A Hierarchical Object-Based Approach
for Urban Land-Use Classification
from Remote Sensing Data

WAGENINGEN UNIVERSITY
University for Life Sciences

Zhan, Qingming

A Hierarchical Object-Based Approach for Urban Land-Use Classification from Remote Sensing Data

Zhan, Qingming

Promotor: Prof. Dr. Ir. Martien Molenaar,
Professor in Geoinformatics and Spatial Data Acqusition
Co-Promotor: Dr. Klaus Tempfli,
Associate Professor in Earth Observation Science

Examining Committee:

Prof. Dr. Ir. Arnold Bregt	Wageningen University
Prof. Dr. -Ing. Manfred Ehlers	University of Vechta, Germany
Prof. Dr. Peter Fisher	University of Leicester, UK
Prof. Dr. Henk Ottens	Utrecht University

A Hierarchical Object-Based Approach for Urban Land-Use Classification from Remote Sensing Data

Zhan, Qingming

Thesis

to fulfill the requirements
for the degree of doctor
on the authority
of the Rector Magnificus of
Wageningen University,
Prof. Dr. Ir. L. Speelman,
to be publicly defended
on Wednesday 29 October 2003,
at 15:00 hrs in the auditorium
of ITC, Enschede

To my wife and my son
&
To my parents

Summary

Zhan, Q., 2003. A Hierarchical Object-Based Approach for Urban Land-Use Classification from Remote Sensing Data. PhD Dissertation

Land-cover and land-use data are essential for urban planning and management. Traditional land-use mapping by visual image-interpretation is expensive, time-consuming and often subjective. Researchers have been searching for automatic or semi-automatic approaches for many years. The combination of airborne LIDAR data with high spatial resolution and multi-spectral images such as IKONOS, QuickBird and SPOT 5 offers great opportunities, especially for application in urban areas. The second generation of airborne scanners with the capacity to acquire simultaneously range and multi-spectral intensity data, makes it possible to extract many meaningful features for land-use classification. The overall objective of this research is the development of a semi-automatic approach for land-cover and land-use classification, based on laser scanning data and multi-spectral images and the development of methods for the consistent aggregation of elementary objects to composite objects at higher abstraction levels.

In this research, several new sensor data have been used and examined for urban land-cover and land-use classification. We have taken the most popular pixel-based classifier, the maximum likelihood classifier (MLC), as an example of traditional classifiers and applied it to high-resolution data. A number of problems have been observed and highlighted, and several remedial measures have been proposed and tested. Land-cover classification accuracy can be improved by modelling the decision surface in the feature space and by selecting samples from both pure pixels and mixed pixels. Spatial partitioning of decision surfaces based on samples of end-member classes is the key to the proposed solutions. The experimental results have confirmed the effectiveness of the proposed class integration method, which uses pure and mixed samples.

Despite the improved land-cover classification accuracy of MLC, we consider the attainable results insufficient for a detailed urban land-use classification. The key features for image-interpretation (size, shape, colour, orientation, pattern, association) are characteristics for certain

types of objects and are only relevant for abstraction levels higher than the pixel level. These features play a key role in image analysis and land-use classification. Object-based image processing techniques are considered for image analysis at a supra pixel level. Therefore, an object-based image analysis approach has become the main focus of this research.

This research develops a hierarchical object-based approach for urban land-use classification. The proposed method consists of three steps: land-cover classification, the definition and delineation of land-use units and land-use classification. It incorporates pixel-based image processing techniques and object-based techniques at different stages. Various techniques have been proposed and tested for object extraction at different aggregation levels.

Several concepts and methods have been proposed and discussed to extract image objects and object properties, and to identify explicit topological relations between image objects. We have elaborated and applied the hybrid-raster data model to explicitly identify topological relationships between image objects. We have tested these concepts and methods for urban land-cover and land-use classification on two test sites. The test results demonstrate the effectiveness of the proposed per-object approach.

Structural information derived from hierarchical image objects plays an important role in land-use classification of urban areas. Delaunay triangulation has been successfully applied to spatially disjoint objects to obtain spatial adjacency relationships and proximity measures; these provide essential information for spatial clustering of objects that form spatial land-use units. Several measures have been proposed and tested for the evaluation of the similarity of buildings. These similarity measures in combination with the spatial adjacency relationships and proximity measures provide information for spatial clustering of land-cover objects, which form spatial land-use units.

Several object properties have been proposed and extracted as attributes of land-use objects for our two test sites. Fuzzy membership functions have been designed to establish the relationships between extracted land-use object properties and designated land-use classes. A fuzzy classifier has been applied for per-object classification based on extracted land-use units and their object properties. The obtained results show that the proposed object-based land-use classification approach is promising. The extracted properties of land-use objects are also important information for urban studies, planning and management.

A united framework for quality assessment has been proposed and tested, based on similarity measures between classified data and reference data. This utilises per-object and per-pixel measurements. The proposed per-object quality measures provide possibilities for obtaining additional quality assessment based on various object properties. The proposed uncertainty measures for extracted land-cover objects and

classified land-use objects have been tested, and we expect them useful in controlling the classification process.

The developed concepts and methods have been implemented by programming in Matlab. The implemented system allows different users to specify characteristics of information that need to be extracted from laser data and spectral data, in order to obtain the desired results. This feature offers planners and other users the opportunity to produce results according to their specific wishes and application requirements from a detailed data set. The multi use of such detailed data sets is important because of the relative high costs of acquiring high-resolution laser data and spectral data. The experimental results show the relevance of hierarchical object modelling in combination with structural image analysis techniques for urban land-cover and land-use classification.

Buildings, green spaces, water surfaces and sealed-ground surfaces have been successfully extracted at the land-cover level. Spatial land-use units have been obtained by aggregation of the extracted land-cover objects. The high quality of the per-object land-use classification has been established by comparing it with the results of visual interpretation.

Keywords

Acknowledgments

I would like to take this opportunity to thank many people for their support and contributions, in one way or the other, during this research project.

First of all, I would like to express my gratitude to my supervisor, Professor Martien Molenaar, for guiding me through the PhD research. His insight has always enlightened me. From him I learned how to think critically, how to select problems, how to solve them, and how to present their solutions. He gave me energy to tackle new problems. I appreciate very much his open mind and patience toward an unknown and uncertain world.

I am indebted to Dr. Klaus Tempfli. He became my second supervisor in the second half of my PhD research. He has spent a lot of his time in discussions, reading my drafts and giving comments and suggestions. From him I learned how to organise ideas and write them down clearly. His sharp eyes helped to detect ambiguities and inconsistencies as they appeared in my earlier drafts.

I would like to thank Dr. Ben Gorte. He was my second supervisor in the early part of my PhD research. He helped me understand many technical problems, especially in image processing. It was a pity that he left ITC in 2000.

I would like to thank the Netherlands government, particularly the Netherlands Development Cooperation (DGIS), for continued funding of the cooperation projects between Wuhan Technical University of Surveying and Mapping (WTUSM), China, which became part of Wuhan University in 2000, and ITC, the Netherlands, in the period from 1986 to 2002. My PhD research was funded by DGIS as part of the third cooperation project and partially by ITC. Many thanks also go to people who have contributed to these projects at various stages. Among them, Professor Morris Juppenlatz and Professor Wang Zhi-Zhuo passed away during my time at ITC. However, they are still living in my memory as my respected teachers, senior scientists and models from many perspectives. Many thanks go to Professor Ning Jing-sheng, Professor Li De-ren, Professor Yang Ren, Professor Lan Yun-chao, Mr. Zhang Run-quan, Professor Yang Hong-yi, Professor Lu Jiang-bin, Professor Yang Kai, Mr. Bai Bo, etc. from the Chinese side and Professor Klaas Jan Beek, Professor Karl Harmsen, Professor Martien Molenaar, Mr. Sjaak Beerens, Mr. Paul Schoonackers, Professor John van Genderen, Dr. Jan Turkstra, Ms. Marjan Kreijns from the ITC side as well as Mr. Oostra, Mr. van den Akker, Mr. Borris Dongelman from the Netherlands embassy in Beijing and many others whose names are not mentioned in this short list. I would like to thank the leadership of the School of Urban Studies(SUS), Wuhan

My appreciation also goes to my office mates, Dr. Cheng Tao, Dr. Ale Raza, Dr. Jose Laurido Campos dos Santos and Ms. Gabi Zimmermann for having a good time working together. Other ITC PhD fellows also gave me a lot of support and help.

Encouragement and energy have been received constantly from my Chinese colleagues and friends. I consider myself fortunate that I have had such an opportunity to know them and share happiness together. Many thanks go to Sun Yu-xian, Dr. Zheng Ding, Wang Li-chun , Dr. Liu Yao-lin, Dr. Liu Xue-hua, Tang Xin-ming, Wu Lan, Zhu Shi-cai, Yao Qing-yan, Zhou Yue-qing, Dr. Yang Hong, Zhang Jian-zhong, Song Ai-hong, Wu Guo-feng, Liu Guang-zheng, Yuan Guo-hua, Wang Chun-qing, Wang Tie-jun, Sun Zhong-wei, Liu Gang, Ren Jin-ming, Shen Xin - a short list. Thanks go to Dr. Huang Zheng-dong and Dr. Cheng Jian-quan as the members of the 'gang of SUS'. Three of us came to ITC together for PhD research and have shared many good moments together at ITC during the four years.

Last but not least, I would like to thank my wife Ying-hui and my son Sheng-an for their love and constant support and for accompanying me to Enschede for three years. My parents are always my source of energy. I dedicate this thesis to my wife, my son and my parents.

Acknowledgments

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Overview of this research

Land-cover and land-use information is essential for urban planning and management. The terms 'land cover' and 'land use' are often confused. Land use can be defined as the use of land by humans, usually with an emphasis on the functional role of land in economic activities. Land use is an abstraction not always directly observable under even the closest inspection. In contrast, land cover designates the visible evidence of land use, or aspects of it such as roads, buildings, parking lots, forest, rivers. Whereas land use is abstract, land cover is concrete and therefore is subject to direct observation. Another distinction is that land cover lacks the emphasis on the economic function that is essential to the concept of land use (Campbell, 1996, 2002).

Many computer-aided classification methods have been developed since the early stages of remote sensing application in 1970s (Curran, 1985; Schowengerdt, 1997; Richards and Jia, 1999; Mather, 1999; Tso and Mather, 2001; Campbell, 2002). Most existing approaches are pixel-based, using multi-spectral data alone, and aim at land-cover mapping, since the spectral information contained in remote sensing images consists of electromagnetic reflections of the physical properties of terrain features. Many existing classifiers fail to produce high-accuracy results because of the existence of mixed pixels caused by the limited spatial resolution of sensors. Therefore high spatial resolution images will be tested in this research to find out if such data can be used for producing better land-cover maps. Please note that in this dissertation high spatial resolution or high resolution refers to images with a spatial resolution from 0.5 m to 4 m produced by sensors such as IKONOS, QuickBird, TopoSys.

Currently, urban land-use mapping is still largely based on visual interpretation using aerial photographs or satellite images, owing to the complexity of urban patterns and the lack of tools for automatic solutions. Human knowledge plays an important role in delineating different land-use units in space and identifying the land-use type of each spatial unit. This is a labour-intensive approach and land-use

classification results produced manually can be variable and inconsistent as regards delineating spatial units and assigning proper land-use types, because of the complexity of urban environments and different understanding of individuals. Therefore an automatic or semi-automatic land-use mapping approach would be preferred. To support feature extraction and land-cover and land-use classification, an object-based image analysis approach is developed and investigated in this research, where image objects are defined based on the hybrid-raster data model. Topological relations between image objects at different abstraction levels are defined and extracted based on image regions (representation of objects in a 2D image). In turn, structural analysis and spatial clustering can be implemented and spatial clusters or spatial units of land use can be extracted in the object-based approach, which is essential to accomplishing land-use classification.

Because of the hierarchical nature of urban planning, land-cover and land-use mapping has to be produced with certain amounts of detail at certain abstraction levels. Single-product approaches may be inconsistent and expensive and should be avoided, a series production approach is likely to be a more efficient way of producing several products at different abstraction levels, based on one set of high-resolution image data. Since data and updating are expensive, the average costs could be reduced if we managed to produce or update several products based on one detailed data set, so that land-use data at a higher abstraction level could be extracted based on land-cover data extracted at a lower abstraction level. Therefore in the context of planning and management it is worthwhile to find out the relationships between different abstraction levels, in terms of scale, contents, the minimum size of spatial units, etc., so that consistent land-cover and land-use maps at different abstraction levels can be produced. In addition, consistent and comparable land-use maps are expected to be produced by applying the same process and rules to image data acquired in the past, present or in the future, because human influences will have been eliminated to a great extent. This feature is crucial for change analysis, since owing to human influences such as diverse understanding and assorted backgrounds, different people often produce different land-use maps (different boundaries, different codes) based on visual interpretation, even when using the same set of images.

The proposed concepts and approaches will be tested on two case study areas. The first test site is in Amsterdam. The data for this densely built-up urban area are an IKONOS image and a digital surface model (DSM) obtained from laser scanning. The second case considers a low-density sub-urban area, the city of Ravensburg, Germany, where we have laser data and high-resolution multi-spectral (MS) images acquired simultaneously. These two different cases were also selected with a view to including different land-use types and spatial patterns in the investigation and examining the effectiveness of different data combinations.

1.2 Research objectives and motivation

Land-cover and land-use data are essential for urban planning and management, based on the roles they play in the planning process, as shown in Figure 1.1. Land-use data are fundamental sources for problem identification and goal formulation
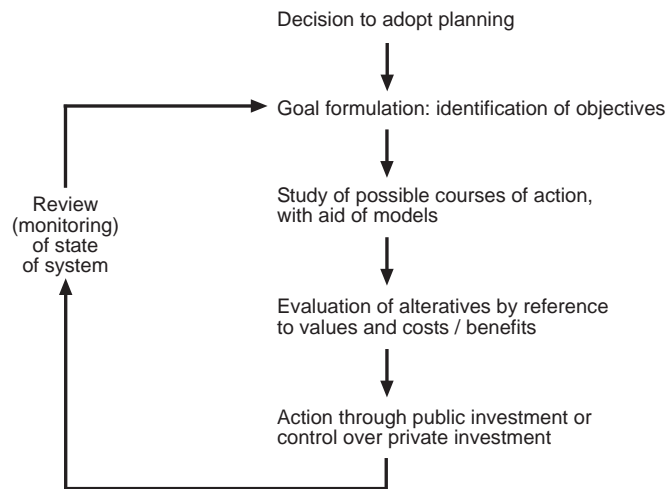
Figure 1.1: Planning process (McLoughlin, 1969; Hall, 2002).

at the initial stage of planning. Land-use data are key factors in planning formulation and forecasting since land-use types and their spatial arrangement are the core business of physical planning. Land-use data play an important role in land suitability evaluation and demand-supply analysis. Land-cover data are fundamental sources for reasoning on land use and for detailed planning. Moreover, land-use planning may be the sole purpose of a planning task in hand. Therefore automatic or semi-automatic land-cover classification and land-use classification, based on remote sensing images and consistent aggregation from lower abstraction levels to higher abstraction levels, are the overall objectives of this research.

To achieve these objectives, the following technical issues are formulated in relation to finding solutions to our research objectives.

Land-cover classification based on high-resolution data

Detailed data are essential for feature extraction, feature handling and the representation of detailed geo-spatial information. Remote sensing technology provides timely available information from spaceborne earth observation systems and airborne laser scanning and imaging systems for a wide range of applications at different scale levels. High-resolution (0.5 m to 4 m) images and airborne laser altimetry data offer exciting possibilities for feature extraction and spatial analysis in urban areas. A combination of IKONOS images and airborne laser scanning data can be one choice. An even more promising data source has become available recently from the second generation of airborne laser scanners combined with a multi-spectral scanner, e.g. TopoSys, which provides high-accuracy 3D data of the earth's surface

and image data simultaneously. However, there are a number of technical issues that have attracted the attention of researchers because of the complexity of the real world and the problems brought about by these newly available data sets. The following are some of the issues that became components of this research.

With the use of high-resolution data the problem of mixed pixels is reduced but the internal variability and noise within land-cover and land-use classes is increased (Cushine, 1987). As a consequence, traditional classification methods such as the maximum likelihood classifier (MLC) method are producing too many classes or classes that are not well defined. Standard techniques have to be augmented for an appropriate analysis because the necessary pixel homogeneity can no longer be achieved by the integration effort of large pixel sizes (e.g., 10 m to 80 m). Because of their high spatial resolution the information content of the high-resolution data in such heterogeneous regions is very complex (Ehlers et al., 2002). In cases of coarse spatial resolution, each pixel may consist of different features appearing in the spatial coverage of a pixel, thus showing a mixed spectral value in each band. Spectral information is used for identifying features by comparison with parameters derived from samples. Comparison is carried out on a basis where sample pixels may be mixed. A pixel of coarse resolution contains a lot of contextual information associated with adjacent pixels. Spectral values between neighbouring pixels are often quite similar, owing to the nature of mixed pixels (smooth transition between neighbouring classes). Sample pixels selected to represent a class are likely to show similarity. With high-resolution images, a pixel will contain only one relatively pure terrain feature in most cases. Pixels as parts of an object may have different spectral values due to the different materials they represent or their orientation toward sunlight. For instance, the roof of a building may be constructed of different materials, say, concrete and asphalt, or, in the case of a gable roofs, the parts of the roof under direct sunlight may have spectral values different from those of the parts on the dark side. Sample pixels selected from different parts of the same roof to represent the roof class may appear in several clusters for each end-member class in the feature space and make a pixel-based classifier such as ML biased. On the other hand, pixels from different objects may have the same or very similar spectral values. For example, as roads and roofs may contain the same or similar material (e.g. asphalt), a pixel from a pixel from a road and a pixel from a roof may have very similar spectral values. In principle, it is impossible for a pixel-based classifier to distinguish them explicitly by using spectral information alone.

In addition, structural and topological information such as the adjacency relationship between buildings is important information for image understanding. Such information requires the detection of individual features and is unlikely to be derived from adjacent pixels by pixel-based approaches, since each pixel and it adjacent pixels will have a relatively small spatial coverage (10 m^2 to 150 m^2) in these images. Meaningful structural and topological information will have to be derived based on the structural analysis of adjacent features (objects) rather than from adjacent pixels. Therefore, conventional pixel-based approaches are not expected to produce good results for high-resolution data. What improvement can be made to enable existing pixel-based classifiers to work with this type of data? Can the object-based approach do a better job here?

Object-based data model for handling images and raster field data

Digital surface models acquired by a laser scanner are becoming increasingly available. DSM provides information on the elevation of terrain features above a well-defined datum, including man-made features such as roofs of buildings. Modelling a surface by elevation values is referred to as 'representation by field data'. Multi-spectral remote sensing images are also field data. Land-cover and land-use classification based on remote sensing images can be treated as mapping or transformation from field data to land-cover and land-use objects. Image objects are conceptualised and can be represented by image regions in a 2D image space. These provide a better representation than individual pixels and are much closer to the human perception of entities such as buildings and residential areas, which we use in planning and many other disciplines. A field model is one of many conceptual models of geographical variation and a basis for much scientific and geographical modelling. In the field model, every location in a spatial framework is associated with a set of attributes measured on a variety of scales. Fields are spatially continuous by definition, but 'continuous' might also refer to the measurement scale (z value). Variables z can be any data type: binary, nominal, ordinal, interval or ratio (Goodchild, 1992, 1997; Cova and Goodchild, 2002). A field can be viewed as a mapping between a locational reference frame and an attribute domain (Worboys, 1995). Representation of fields must always be approximate, as we cannot store an infinite number of locations. Spatial tessellation (regular, irregular or hybrid) is the means most used for representing field-based models. Common operations on fields include interpolation, classification, filtering, spatial overlay, statistical analysis, map algebra, spread functions, corridor analysis, terrain analysis, and many others (Goodchild, 1997; Cova and Goodchild, 2002).

There are many entities in geographical reality that are readily perceived as objects, such as lakes, rivers, buildings. Object representation in a database of real world entities such as buildings and lakes or conceptualised entities such as commercial and residential is considered more natural and logical. In this thesis, we use 'object' to refer to object representation and 'entity' to refer to an entity in reality. From an object perspective, space is viewed as a container populated by these entities, each with an identity, spatial embedding and attributes. Natural language is much more suited to describing objects than fields (Cova and Goodchild, 2002). Molenaar (1998) proposes a theory for spatial object modelling in GIS that provides a theoretical framework for object-based spatial data modelling. Couclelis (1992) and Worboys (1995) note that the field and object conceptual perspectives should not be considered mutually exclusive. The field and object perspectives can be used in conjunction, as well as derived from one another (Cova and Goodchild, 2002). Therefore, we introduce 'image object' (IO) to represent objects extracted from field data (images) according to their definitions and meaning in natural language (semantic). Image objects are regarded as representations of real world entities or conceptualised entities in 2D image. An image object is a spatial container that relates locations in a field space to objects in an object space similar to an object field as described in Cova and Goodchild (2002). The main difference between an image

object and an object field is that an image object emphasises the spatial extent of an object (image region) while an object field is defined as a continuous field in which locations are mapped to spatial objects (Cova and Goodchild, 2002). Both of them share qualities of the field and the object conceptual perspective of geographical phenomena. How can we define a spatially embedded object so that topological relations between objects as represented in a raster can be extracted[1]? What are the roles for object extraction from images? How can image objects be mapped from a field space to an object space? How can we define and identify topological relationships between objects based on a raster data model?

Multi-scale/multi-level aggregation

Urban planning and management use a variety of data in their different stages (Le Clercq, 1990). Many are geo-spatial data. Planning products are hierarchically associated so that planning at a higher level will be used as a guide for planning at the lower levels. For instance, the regional plan will guide the master plan, the master plan will serve as a guide for the district plan or detailed plan, and so on. Planning at each level is an attempt to solve particular problems at an adequate scope or scale. The required degree of detail in geo-spatial data is also quite different at each level. The degree of detail is directly associated with the scale of the geo-data or maps used for analysis and planning formulation. Therefore, aggregation is involved at different levels of the planning hierarchy in order to provide a reasonable amount of information (degree of detail) and a suitable scale at each level. Land-cover and land-use objects obtained from images such as buildings or residential areas are a better form of representation than pixels for human perception. Multi-scale/multi-level aggregation will be based on such objects. A syntax has to be defined to support such multi-scale/multi-level aggregation in the context of urban planning and management. This syntax should be able to take into account the geometric, thematic and semantic attributes of objects in multi-scale/multi-level aggregation. What rules may be relevant for such aggregation with respect to geometric, thematic and semantic attributes?

Semantic and imprecision issues

Planning at different levels will have to look at different aspects of urban reality such as social aspects (e.g. population, education), economic aspects (e.g. industry, employment), environmental aspects (e.g. pollution, green space). Such thematic differentiation of planning looks at the problems from different perspectives, depending on the planning task and the disciplines involved. This may lead to different interpretations of the same feature presented in geo-databases or maps. It implies the need to represent geo-spatial features differently according to different

---

[1]Please note that we use the term 'object' to refer to a spatially embedded unit, which is different from the 'objects' often used in object-oriented programming in computer science. To avoid further confusion, we will use 'object-based' instead of 'object-oriented' in this dissertation, although we may share many similar concepts developed for the object-oriented framework in computer science.

disciplines, but based on a fundamental geo-database. Different representations of geo-spatial features require semantic modelling with respect to corresponding disciplines. What semantic modelling techniques can be applied in representing geo-spatial features?

Land-cover and land-use classification systems (see Appendix A) are defined using linguistic terms such as 'containing multi-story residential apartment buildings with good environment and public facilities available at close range'. Therefore we have to apply certain measurements in order to check whether an object belongs to the defined class as described in a linguistic form. This is called semantic fuzziness. Another issue is raised pertaining to the spatial coverage of land-use classes: when we have extracted a group of buildings, what are believed to belong to one land-use class, say residential, the question arises as to where the spatial boundaries of this class are, since often no such physical boundaries between different land-use classes can be found on the ground or in images. This is called fuzziness of conceptualised boundaries. These types of semantic and imprecision issues exist throughout land-cover and land-use classification. Can fuzzy set and fuzzy logic play a role here? How does the fuzziness of a semantic definition influence the geometric and thematic components of geo-spatial objects? What kinds of measurements can be used for delineating land-use units?

Based on the above discussion, the research objectives can be summarised as follows:

- To examine the main problems in land-cover classification, using pixel-based classifiers based on high-resolution data, and provide potential solutions to these problems, using pixel-based classifiers, and evaluate effectiveness.

- To provide a conceptual framework and formalism of an object-based approach to image analysis and land-cover and land-use classification.

- To provide object-based methods and operations for feature (land-cover objects) extraction.

- To provide spatial aggregation methods for structural analysis toward determining spatial units of land-use classes based on the spatial distribution of land-cover objects.

- To provide land-use classification schema that can deal with different types of information extracted for objects such as thematic, geometric and structural information.

1.3 Main technical problems and proposed solutions

Before exploring a new approach to the defined problems, we need to investigate the problems of using existing approaches and find out if they can cope with new types of data such as high-resolution images. What kind of improvement could be made

using existing approaches in such cases? Three end-member land-cover classes, – built-up area, green space (vegetation) and water surface – are proposed for land-cover classification based on high-resolution multi-spectral data, using a pixel-based approach, the maximum likelihood classifier. Using high-resolution data, we may encounter some new problems because of the existence of many pure pixels; using coarse-resolution data, the problem may be less because the existence of many mixed pixels. One problem is due to the existence of sub-clusters made of pure-samples of sub-classes in the feature space when using high-resolution images, for instance, the dark, medium, light and very light tones of the roofs of buildings and other concrete surfaces are sub-classes of the built-up class. Samples selected from these sub-classes to represent the built-up class will form several sub-clusters, which will then violate the requirement of normal distribution in the case of using the maximum likelihood classifier. As a consequence, a large deviation obtained may be estimated as a parameter for the maximum likelihood classifier. As a result, the decision surface (probability density function) will not be estimated correctly, especially in the margin area of each class in the feature space. The proposed measures focus on spatial modelling in the feature space. We introduce a solution called spatial partitioning for modelling the decision surface in the feature space in order to improve the classification accuracy of a pixel-based classifier and take the maximum likelihood classifier as an example.

Geographical information is the representation of an abstract view of reality. It provides digital data that fit the specifications fixed by the modelling approach. One notable question remains: does the geographical information represent reality? This can be split into two rather different components. Is the modelling approach relevant to the observed phenomenon? Do the data meet the specifications? Together, the answers to these questions allow us to evaluate the most important user requirement: fitness-for-use. As feature extraction is treated as transformation or mapping from field data to objects, what operations or methods are suitable for this type of transformation or mapping? Do these mapped objects represent the abstract view of planning and management in reality? We intend to make a comparison from the object perspective, which is per-object-based rather than per-pixel-based.

The urban land-use classification system is designed hierarchically to correspond with hierarchical levels of planning. What are the spatial data requirements at different aggregation levels for urban planning and management, such as class definition, degree of detail, minimum spatial unit, concerning map scale? What factors should be considered and what kinds of rules could be applied so that data required at high aggregation level can be derived from lower levels through aggregation steps? Can we disaggregate results obtained at high aggregation level to low aggregation level?

In urban areas most human activities are organised through certain forms of spatial arrangement, conscious planning or interactions of different activities. There must be ideas or patterns behind the physical appearances on the ground, such as location, closeness, alignment, spatial clustering. If we could summarise and describe the nature of human activities in urban areas, we would be able to understand them better from what we see on the ground or in images. How can hierarchically structured land-use information be extracted from images and aggregated hierarchically?

What quality level can we reach for extracted objects and in spatial data aggregation? What quality measures should we use? What factors will lead to uncertainty and how can these factors be quantified?

To answer the above questions, three types of objects, elementary objects (pixels), land-cover objects and land-use objects, are defined hierarchically according to the hierarchical layers in which they are located, the roles they play and their geometric properties (see Figure 1.2). Pixels are regarded as elementary objects with properties of uniform size and shape, and fixed adjacency relations. Representations of land-cover entities and land-use entities are image regions (image objects) formed in the land-cover layer and the land-use layer respectively. Land-cover objects are the representations of physical entities such as buildings and lakes. Land-cover objects will be extracted based mainly on physical properties and the spatial distribution of elementary objects (pixels) (Zhan et al., 2002b). Land-use objects, however, are the representations of land-use units that are abstracted or conceptualised in terms of social-economic functions, and often there are no physical boundaries between different land-use types. Land-use objects will be extracted by reasoning based on the various properties and spatial distribution of land-cover objects (Zhan et al., 2002c).

In our three-stage approach we proceed from pixels to land-cover objects in the first stage, to reasoning of spatial coverage for land use based on types and spatial distribution of land-cover objects in the second stage, to reasoning and identifying the land-use type based on thematic and structural information spatially embedded in a spatial unit in the third stage (Zhan et al., 2002d).

In the land-cover object extraction stage, image objects are extracted based on their physical properties, both per-pixel such as electromagnetic reflectance of individual pixels and per-object such as size and height, and similarity or homogeneity measures. Buildings, green spaces, water surfaces and open surfaces (parking space, squares and other paved or bare surfaces such as construction sites) will be extracted in this stage, as four end-members of land-cover classes. To this end the geometric properties such as size, shape, and orientation, as well as thematic information, are acquired based on multi-spectral and laser scanning data.

Reasoning of spatial extent for land use will be based on the local spatial arrangement of land-cover objects. A number of spatial indicators will be extracted to represent thematic, geometric and structural information for land-use classification. Land use will be identified for each spatial land-use unit, based on categorical, geometric and structural indicators extracted from images. Image objects are fuzzy objects in the sense that objects are linguistically defined in planning and management, and the spatial extent of objects (image regions) will be extracted or reasoned based on membership values and the spatial distribution. The fuzziness of the spatial extent of an object is determined by the compactness of membership values in the feature space and geometric space (2D image plane). Therefore, semantic issues and related fuzzy membership functions play important roles in various stages of this research (Zhan et al., under peer review (1)). The work flow of the proposed land-use classification schema is presented in Figure 1.2. The object-based conceptual and computational modelling steps are illustrated in Figure 1.3 (modified after Worboys (1995)).
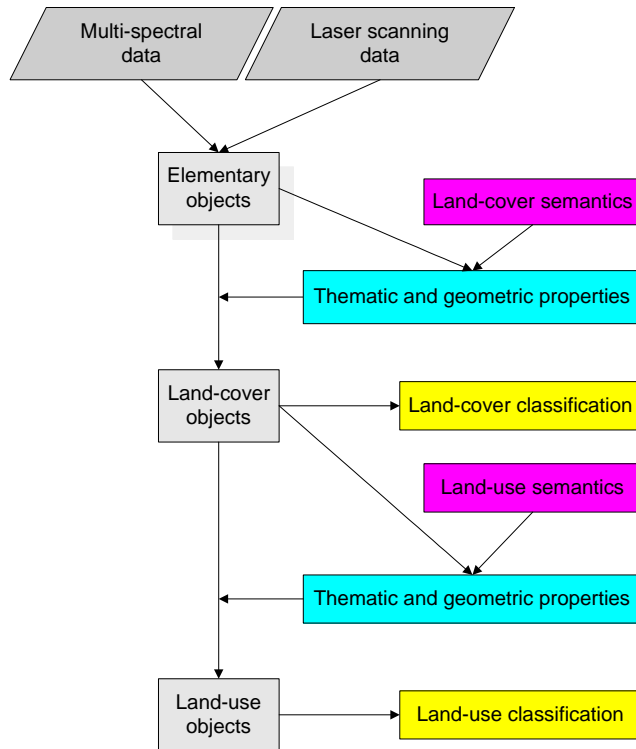
Figure 1.2: A scheme for object-based land-cover and land-use classification.
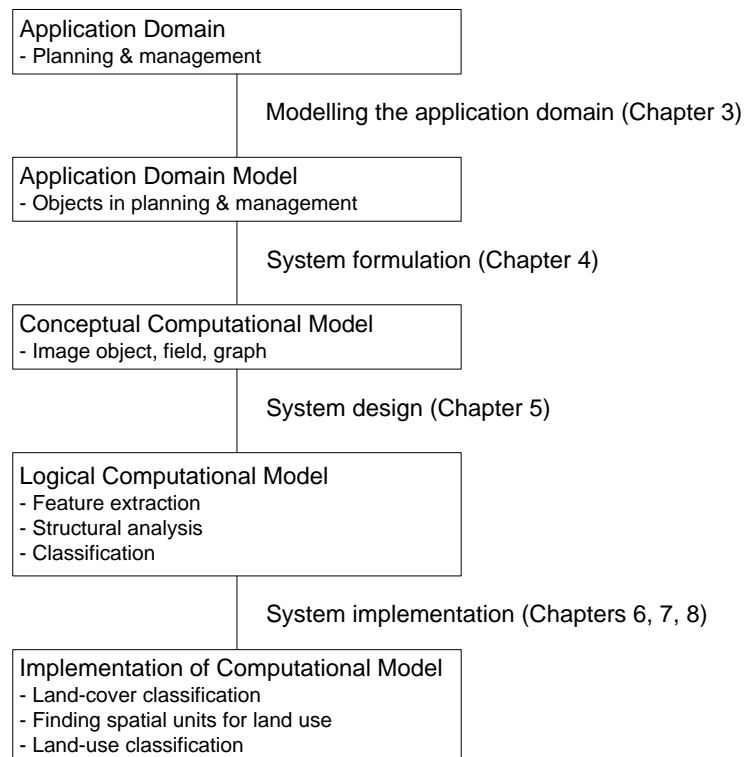
Figure 1.3: Conceptual and computational modelling of the proposed object-based approach (modified after Worboys (1995)).

1.4 Structure of thesis

A review of the existing problems and existing approaches in determining urban land cover and land use from images is presented in Chapter 2. The two study areas and the data used in this research are also introduced in Chapter 2. An investigation is made using the widely used maximum likelihood classifier on high-resolution data. Problems and their causes are highlighted and proposed remedial measures are examined in Chapter 2. In order to apply an object-based approach to obtaining land-cover and land-use data for urban planning and management, an object-based conceptual model for urban planning and management is outlined in Chapter 3. A formalism for the image-object data model is proposed in Chapter 4, which provides a theoretical framework for the logical design and implementation of the identification of topological relations between image objects, based on the hybrid-raster data model. A logical design and an implementation schema for object-based land-cover and land-use classification is presented in Chapter 5. In Chapter 6 object-based methods for land-cover feature extraction are applied and the experimental results are illustrated. Methods for object-based structural analysis and for finding spatial units of urban land use are proposed and elaborated in Chapter 7. Object-based land-use classifications are tested and discussed in Chapter 8. The logical links running from Chapter 3 to Chapter 8 can be seen in Figure 1.3. Assessment of the quality and uncertainty of the experimental results are analysed and reviewed in Chapter 9. Finally, conclusions and future research are outlined in Chapter 10.

Chapter 2

High spatial resolution data and pixel-based classification

2.1 Introduction

In this chapter, a general description of the study areas is provided to illustrate the problems this research is going to tackle. These problems can be found in other urban areas as well. In this chapter, which deals mainly with pixel-based approaches, built-up area, green space and water are considered as three end-member land-cover classes when using multi-spectral data alone, because they are comparatively separate in feature space, based on spectral reflectance. Built-up area is further subdivided into two sub-classes, building and open-surface, when laser data are used to extract buildings from built-up area. To test the proposed approach in different settings, two study areas are selected to represent built-up areas in a large city and in a small town respectively. Next the data used in this research and the preprocessing steps are introduced. A detailed explanation of the problems and comments on existing approaches follow. In the remainder of this chapter, an example of applying one of the existing pixel-based approaches using high-resolution data is presented and a number of modifications are proposed to the existing approach. A short summary is given at the end of this chapter. In general, this chapter explains the existing problems, existing approaches to these problems, and problems remaining even with improved approaches. The shortcomings support the motivation to investigate an object-based approach in an attempt to solve the identified problems. The object-based approach will be presented in the subsequent chapters.

Table 2.1: Data used in the case study Amsterdam

Type of data	Sensor	Date	Band/Colour	Resolution or scale
Satellite imagery	SPOT	13 Oct. 1996	Multi-spectral	20 m
	IKONOS	8 June 2000	Multi-spectral	4 m
Airborne laser data	TopoSys I	28 Mar. 1998	True height value	1 m
Aerial photographs	Optical camera	19 May 1997, 30 May 1999	True colour	1:10,000
Topographic maps		1996, 1997, 1999	Black-and-white	1:1,000

2.2 General description of the study areas

2.2.1 The study area in Amsterdam, the Netherlands

A study area of 3 km × 3 km, southeast of Amsterdam, was selected for the experiment (see Figures 2.1 and 2.2). Approximately 200,000 people live in this suburban district. Several types of residential areas, commercial areas, as well as more natural features such as parks, lakes and canals, can be found in this study area. The landscape of this test site is generally flat, but elevated roads obstruct the straightforward feature extraction of buildings from laser data or images when using existing approaches. Many existing approaches for building extraction are based on analysing profiles derived from laser data. Elevated roads have a similar profile to buildings, which makes it difficult to separate them from buildings when using these approaches. In this regard, the site provides a good opportunity to test whether the proposed object-based land-cover classification approach is more robust than other per-pixel based approaches such as the maximum likelihood method. A list of data used for this area is presented in Table 2.1. Detailed descriptions are given in Section 2.3.

2.2.2 The study area in the city of Ravensburg, Germany

The second test site is an area of 1 km × 1 km in the southwest of Ravensburg, Germany (see Figures 2.3, 2.4, 2.5 and 2.6). This is a difficult area for building extraction because there are many small one- to two- storey houses, often with gable roofs. Some tall trees are very close to the buildings, and the site is situated in a hilly area with various types of vegetation. Both urban and rural land-use types can be found in this area. The advantage is that we are able to use high-resolution data produced simultaneously by a laser scanner and a four-channel multi-spectral scanner. The data used for the experiment are DSM1 (digital surface model acquired

Table 2.2: Data used in the case study Ravensburg

Type of data	Sensor	Date	Band/Colour	Resolution or scale
Airborne imagery	Laser scanner	23 April 2001	First and second pulse	1 m
	Digital line scanner	23 April 2001	Multi-spectral	0.5 m

from the first pulse of the laser beam; see Figure 2.3), DSM2 (digital surface model acquired from the second pulse of the laser beam; see Figure 2.4), colour infrared image (see Figure 2.5) and real-colour image (see Figure 2.6). DSM1 is useful for building extraction and for the delineation of trees from other vegetation. DSM2 can be used as the basis of a digital terrain model (DTM) and provides a reference for identifying high objects above ground. Combining DSM1 and DSM2 provides a good means of checking whether feature surfaces are solid or not. Four bands of images from the multi-spectral scanner are useful information for detecting materials that feature surfaces contain. A list of data used for this area is provided in Table 2.2. A detailed description is given in Section 2.3.

2.3   Data

Several types of data have been used in this research, as listed in Table 2.1 for the Amsterdam test site and Table 2.2 for the Ravensburg test site. However, laser data and IKONOS imagery are the main data sources for this research.

2.3.1   Laser scanning data

Laser scanning is an airborne elevation mapping method that is characterised by a largely automated measuring procedure, where fully digital data collection is followed by a computer-based data evaluation. It is performed with a multi-sensor system with the following main components: laser rangefinder, GPS receiver, and the inertial measurement unit (IMU) recording devices.

In laser scanning, the scanner deflects the laser beam across the flight line; as a result, a swath of ground along the flight line is sampled. The resulting sampling pattern depends on the scanning device (parallel lines, zigzag lines, ellipses, etc). The distance to the earth's surface is determined by measuring the pulse return time. The position and altitude of the sensor are calculated from GPS, IMU and calibration data. The proprietary software of laser scanning system providers (TopoSys, in our cases) is commonly used to calculate the X, Y, Z of terrain points.

The density and distribution of 'points' hit by the laser is determined by the laser system parameters of pulse frequency, scan frequency and scan angle, in com-

bination with the flight parameters of flying height, aircraft speed, and the distance between flight lines. The roughness of terrain relief is another factor influencing the sampling ratio.

Thanks to its variable system parameters (see Tables 2.3 and 2.4), the system offers a wide range of mapping options, from longitudinal profiles to transverse profiles to even spot distribution, and thus a high degree of flexibility with respect to different requirements (TopoSys, 2002).

The laser data of the study area in Amsterdam were captured and processed by TopoSys, as a contractor of the AHN (Actual Height model of the Netherlands) production, using the first generation of laser scanner TopoSys I (Geo-Loket, 2002). The laser data of the study area in Ravensburg were captured and processed by TopoSys as testing data using the second generation of laser scanner TopoSys II (TopoSys, 2002).

At a maximum flying height of 1000 m, above the ground, TopoSys generates its standard product, the 1 m raster elevation model – other raster sizes can be produced by changing the survey height. Elevation accuracy is the order of 0.15 m (in the local coordinate system).

At a flying height of 1000 m the spatial resolution of multi-spectral images acquired by the TopoSys II multi-spectral line scanner is about 0.5 m. Images of the line scanner, which are recorded simultaneously with the laser data, are delivered as true-colour or infrared images.

2.3.2   IKONOS imagery

An IKONOS image is the product of the first commercial satellite remote sensing company, Space Imaging. Images are offered at various levels of processing, based on 4 m resolution for multi-spectral bands and 1 m resolution for the panchromatic band. The first satellite was launched on 24 September 1999. Some of the specifications can be found in Table 2.5. The geo-referenced IKONOS image of the study area is shown in Figure 2.2.

2.4   Data pre-processing

2.4.1   Geo-referencing

The Dutch coordinate system with stereographic projection was selected for geo-referencing all data. Large-scale (1:1,000) base maps were scanned and registered as the base map for other data. Aerial photographs of scale 1:10,000 were scanned and geo-referenced with the base map. The IKONOS image was geo-referenced using the base map and the registered aerial photographs. The laser data had already been processed in the same coordinate system by the Survey Department, Ministry of Transportation and Public Works, the Netherlands. The accuracy for this project, where the entire country was surveyed by laser altimetry, depends strongly on the type of vegetation and topography in the area. Here, a standard deviation of 15 cm

Table 2.3: System parameters of TopoSys I and TopoSys II

Parameters	TopoSys I	TopoSys II
Sensor type	Pulsed fibre scanner	Pulsed fibre scanner
Range	< 1000 m	< 1600 m
Wave length	1.54 μm	1.55 μm
Pulse length	5 nsec	5 nsec
Scan frequency	650 Hz	650 Hz
Pulse repetition rate	83000 Hz	83 000 Hz
Resolution of distance measurements	0.06 m	0.02 m
FOV	14° (± 7°)	14° (± 7°)
Swath width (at maximum range)	220 m	390 m
Av. measurement density (at max. range)	5 meas./m^2	3 meas./m^2
Measurement possibilities	First or last pulse	First and last pulse simultaneously
Intensity measurements	None	Possible

Source: www.toposys.de accessed on 10 September 2002

Table 2.4: Parameters of the Digital Line Scanner

FOV	21°
Pixel per line	682
Resolution (at 1000 m survey height)	0.55 m
4 spectral channels	(1) 440 - 490 nm (2) 500 - 580 nm (3) 580 - 660 nm (4) 770 - 890 nm

Source: www.toposys.de

Table 2.5: Technical specifications of IKONOS sensors

Imagery spectral response	Panchromatic: 0.45 - 0.90 microns
	Multi-spectral:
	Band 1: Blue 0.45 - 0.52 microns
	Band 2: Green 0.52 - 0.60 microns
	Band 3: Red 0.63 - 0.69 microns
	Band 4: Near IR 0.76 - 0.90 microns
Swath widths and scene sizes	Nominal swath width: 11 km at nadir
	a nominal single image at 13 km $\times$ 13 km
Metric accuracy	12 m horizontal and 10 m vertical accuracy with no ground control
	2 m horizontal and 3 m vertical accuracy with ground control
Altitude	681 km
Inclination	98.1 degrees
Speed	7 km per second
Descending nodal crossing time	10:30 a.m.
Revisit frequency	2.9 days at 1 m resolution
	1.5 days at 1.5 m resolution
Orbit time	98 minutes
Orbit type	sun-synchronous

Source: www.spaceimaging.com accessed on 15 September 2002

maximum, with a systematic error of 5 cm maximum, applies to the accuracy of 'solid topography' (such as roads and parking lots) as well as 'flat or soft topography' (such as beaches and grass fields). In wooded areas the accuracy is lower. In this case specifications accept a minimum point density of one point per 36 m^2, a standard deviation of 20 cm maximum and a systematic error of 10 cm maximum (Geo-Loket, 2002). For the test site of Ravensburg, the 'geo-referencing' of both laser data and multi-spectral data was done by TopoSys, achieving an accuracy similar to that of AHN data. Therefore, no additional geo-referencing is necessary for these data.

2.4.2   Data correction

There were several problems with the laser data provided, which had to be corrected as described below.

Flight gaps

A few small gaps were found in the original raster data of the Amsterdam site. To have a complete coverage of the area, manual editing was done using aerial photographs and the large-scale base map.

Missing data

The grided laser data from both sites contained pixels with no values. This may be caused by several factors. The first is the so-called 'mirror reflection'. When the laser beam hits a smooth surface situated near the end of the scan-line, the reflection will be cast off in another direction and the laser rangefinder will receive no signal. This happens often on still water surfaces. Correction can be made by finding the lowest height value in the surrounding region and replacing the missing pixels with this value. Another mirror reflection type was found on top of several high-rise buildings. The highest value in the region will be used to replace the missing pixels in this case. The second cause of false pixels is mainly due to mixed types of vegetation canopy, where the laser signal is weakened or missed because of a mixed type of surface. A few pixels of this type are found in wooded areas.

Noise

Some false pixels were found at random positions. This type of 'missing pixel' can be treated as 'noise'. The average values of surrounding pixels were taken instead.

The laser data used for this research after the above-mentioned corrections had been made are shown in Figures 2.1, 2.3 and 2.4.

Figure 2.1: DSM from laser scanning of the study area in Amsterdam with 1 m resolution (with correction).

2.5 The problems and existing approaches to the problems

2.5.1 Aerial photo-interpretation

Aerial photo-interpretation is still a practical way of obtaining land-use classification results, given the lack of automatic or semi-automatic solutions. However, it is labour-intensive and subjective. Different results may be obtained by different interpreters owing to differences in understanding. Therefore, researchers are searching for automatic or semi-automatic approaches to land-use mapping in various disciplines (Barr and Barnsley, 1997; Barnsley and Barr, 1997; Aplin et al., 1999a,b; Zhan et al., 2002c,d). Size, shape, colour, orientation, pattern and association are some of the cues used in photo-interpretation. These cues should continue to play a key role in potential automatic solutions.

2.5.2 Per-pixel based approaches for land-cover classification

With the use of high-resolution remote sensing data, the problem of mixed pixels is reduced but the internal variability and the noise within land-use classes are increased (Cushine, 1987; Ehlers et al., 2002). Most conventional pixel-based classifiers, such as minimum distance and maximum likelihood assume parametric sta-
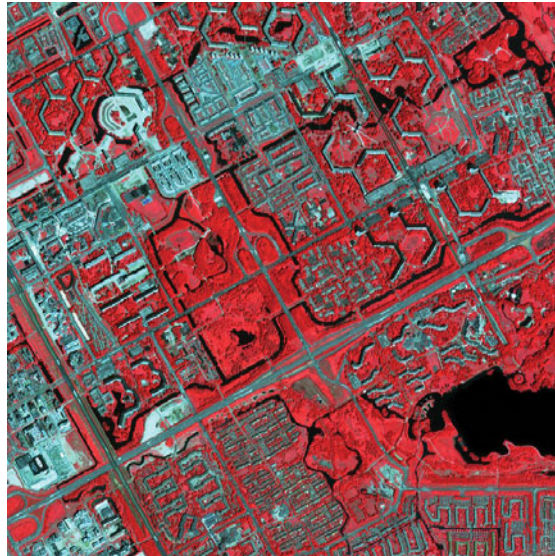
Figure 2.2: Geo-referenced IKONOS image of the study area in Amsterdam (false colour, 4 m resolution).

tistical models, such as the Gaussian distribution (Curran, 1985; Campbell, 1996; Richards and Jia, 1999). These methods are not designed to handle data from different sources or of varying accuracy and they cannot cope with non-numerical data. In practice, the data do not usually obey the conditions imposed by these conventional methods that classify pixels via crisp rules (Mertikas and Zervakis, 2001). Urban areas are complicated because of the mix of man-made features and natural features. Among the cues for photo-interpretation, only colour can readily be extracted from images for land-cover identification. However, it is difficult to be determine size and shape by using per-pixel approaches. In addition, pattern and association are higher-level structural and topological information that is also difficult to extract by using pixel-based approaches, but they are useful in land-use classification. For instance, an isolated building surrounded by woods or other vegetation is likely to be a farm house in a rural setting but a facility or shop in a recreational area in an urban setting. A group of buildings of similar size and regularly spaced or orientated in an urban area is likely to be a residential area. Such higher-level structural information should play an important role in the land-use classification of an urban area. Additional spatial indicators should be extracted, also based on structural analysis, in order to understand and identify spatial patterns or the spatial organisation of features, especially man-made features.
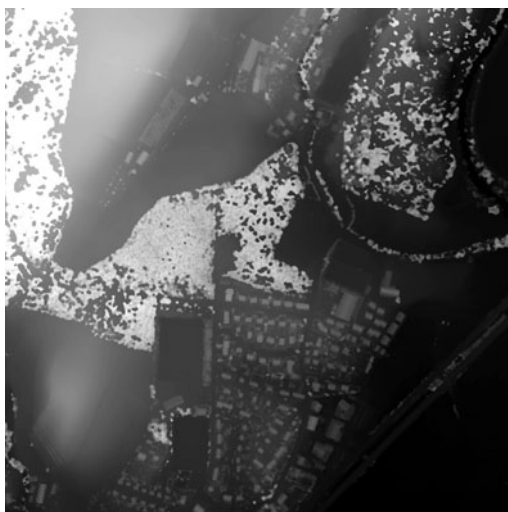
Figure 2.3: Digital surface model from the first pulse of laser beam (DSM1), Ravensburg, Germany (gray tone is proportional to terrain relief: lighter tone refers to higher terrain, not cloud).
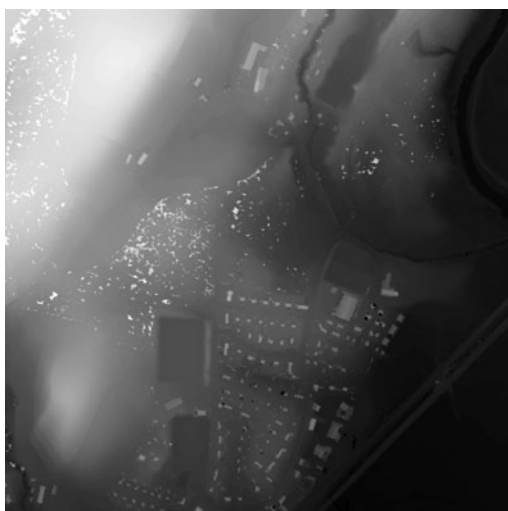
Figure 2.4: Digital surface model from the second pulse of laser beam (DSM2), Ravensburg, Germany (gray tone is proportional to terrain relief: lighter tone refers to higher terrain, not cloud).

Figure 2.5: Colour infrared composite, Ravensburg, Germany.

Figure 2.6: True-colour composite, Ravensburg, Germany.

2.5.3   Per-field approach using vector data

A per-field approach uses vector data to extract image regions (pixels inside a polygon, a spatial unit such as an agricultural field or a parcel) and classify these regions as a whole to improve classification accuracy (Aplin et al., 1999a,b; Zhan et al., 2000; Aplin and Atkinson, 2001). The per-field approach is good for extracting and analysing of structural information. In many cases, however, accurate and up-to-date vector data sets are rarely available (Tatem et al., 2001b). Feature boundaries may have changed between the time of producing vector data and the time of acquiring new image data (Zhan et al., 2002a).

2.6   Improvement of pixel-based land-cover classification

Pixel-based approaches have been developed and are widely used in remote sensing image processing and classification. Since the early stage of remote sensing applications, simple geometric properties such as the uniform shape and size of pixels have attracted many researchers to using pixel-based approaches. Examples of pixel-based classifiers include the maximum likelihood classifier (MLC), the fuzzy classifier, the tree-based approach and the neural network approach. A number of issues have to be taken into consideration when selecting a suitable classifier. A classifier such as the MLC requires normally distributed data. In practical multi-class problems it is rather difficult to guarantee normal or even symmetric distributions with similar covariance matrices for all the classes (Marques de Sa, 2001). The requirement may be further violated when we apply this type of classifier to high-resolution data. In this section, we would like to examine a number of issues raised by the newly available high-resolution images in urban land-cover classification. We take the MLC as a representative of pixel-based approaches, based on the following considerations: widely used in remote sensing, comparatively robust as regards marginal samples, and high classification accuracy can be reached. We will illustrate the normal distribution problem by analysing the decision surfaces in the feature space estimated by the MLC, using given samples, and show that the MLC would be improved by introducing modifications.

2.6.1   Maximum likelihood classifier (MLC)

The MLC is a parametric statistical classifier that estimates mean and standard deviation of each class based on samples. Classification is made by computing the probability of each predefined class for each pixel, according to the probability density function (pdf) derived from estimated parameters with the assumption of the normal distribution (Richards and Jia, 1999). The probability for class ω_i can be calculated for N bands as

$$p(\mathbf{x}|\omega_i) = (2\pi)^{-N/2} |\Sigma_i|^{-1/2} \exp\{-\frac{1}{2}(\mathbf{x} - \mathbf{m}_i)^t \Sigma_i^{-1}(\mathbf{x} - \mathbf{m}_i)\}$$
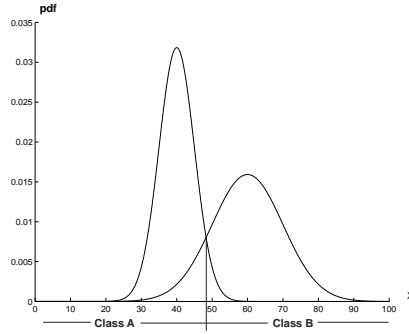
Figure 2.7: 1D pdf of two classes, μ_A=40, ρ_A=5, μ_B=60, ρ_B=10.
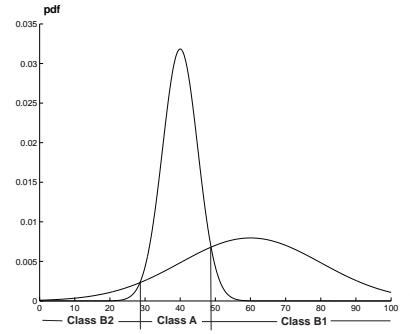
Figure 2.8: 1D pdf of two classes, μ_A=40, ρ_A=5, μ_B=60, ρ_B=20.

where $\mathbf{m}_i$ and Σ_i are the mean vector and covariance matrix of the data in class ω_i.

2.6.2 Known problems with the MLC

The MLC can achieve relatively high accuracy in remote sensing classification when the classes are well defined and samples are selected that meet or nearly meet the requirement of normal distribution, as shown in Figure 2.7 for the one-dimensional case.

Problems may occur when samples of one class have a small deviation while samples of another class have a large deviation, as illustrated in Figure 2.8 for the one-dimensional case and in Figure 2.9 for the two-dimensional case. In this example, class **A** and class **B1** as shown in Figure 2.8 will be classified properly. Pixels in class **B2**, however, will be classified as class **B** instead of the more reasonable class **A** as, using the standard MLC, pdf values computed for class **B** will be higher than those for class **A** in this range. Such a problem may happen when we use high-resolution images, particularly in urban areas, where a class such as water will have a small deviation, green space will have a large deviation, and built-up areas will have a very large deviation (as shown in Figures 2.10 and 2.11). The classification result will be wrong because many water pixels will be wrongly classified as belonging to the built-up class. This occurs because the pdf of the built-up class has been overestimated, owing to the existence of several different types of sub-clusters (newly built-up, buildings with light tone, buildings with medium tone and buildings with dark tone), which causes a large deviation in the built-up class.
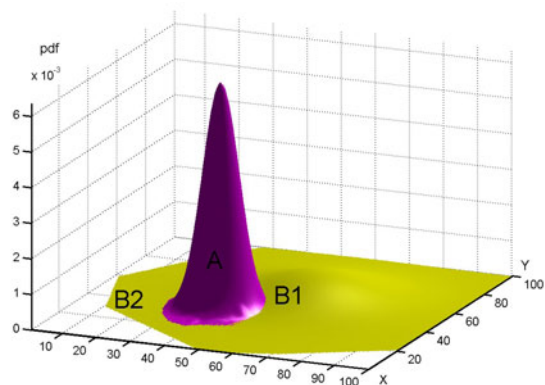
Figure 2.9: 2D pdf of two classes, μ_1=(40, 40), ρ_1=(5, 5), μ_2=(60, 60), ρ_2=(15, 15).
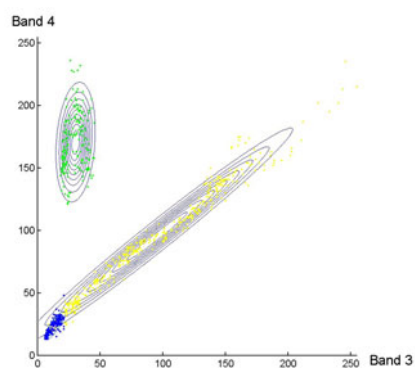
Figure 2.10: 2D plot of pdf based on pure samples of three land-cover classes: built-up (yellow), vegetation (green) and water (blue).
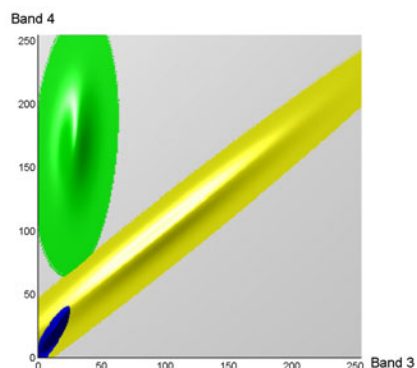
Figure 2.11: 3D draped pdf based on pure samples of three land-cover classes: built-up (yellow), vegetation (green) and water (blue).

2.6.3 Potential solutions to the identified problems

Classification can be treated as partitioning in a feature space. The decision boundaries are formed by the intersection of the pdf obtained for all classes when using the MLC or by the intersection of the fuzzy membership functions in fuzzy classification. Based on the above observation, we consider that three solutions may contribute to improving the maximum likelihood classification. The first measure is to obtain independent samples (Gong and Howarth, 1990). The second and third measures are based on sample and parameter manipulation in the feature space. We might obtain good classification results if we could delineate good decision boundaries in the feature space.

Single-pixel training approach

The single-pixel training approach is a sampling strategy where sample pixels are selected individually instead of using image regions or block training, and each pixel has to be at least several pixels away from any other selected pixel. The single-pixel training approach has proved capable of improving the classification accuracy (Gong and Howarth, 1990). This requirement is meant to reduce positive spatial autocorrelation that may exist among pixels that are spatially contiguous or close together (Campbell, 1981; Lobovitz and Masuoka, 1984). The traditional block training method violates the independent sampling requirement and makes the training signatures for each class less representative. The block training is even worse for high-resolution images. Therefore the single-pixel training approach was chosen in our research without further testing.

Sample selection from central (pure) pixels or boundary (mixed) pixels

The sample selection strategy may have to be adapted to include not only pure pixels but also mixed pixels that can be found along feature boundaries. The reasons for this are the existence of sub-clusters in feature space for candidate end-member classes and the increasing number of pure pixels existing in high-resolution images. Mixed pixels here refer to pixels whose dominant cover classes can be easily defined in visual interpretation although they are located along feature boundaries. We have observed that including such mixed pixels provides more representatives than taking pure pixels only, and the mixed pixels can provide more evidence for determining probability surfaces in the transit regions between different clusters in the feature space, as shown in Figures 2.12 and 2.13. Both the pure-pixel sampling strategy and the pure-plus-mixed pixel sampling strategy will be tested in Section 2.6.4.

Sample integration or class integration

The previous solutions focus mainly on sampling strategies. There is another issue that has to be taken into account when using high-resolution images. Some of the end-member classes may contain a number of sub-clusters in the feature space because of the increasing number of pure pixels. When we select sample pixels for these classes, we have to select samples from each sub-cluster. For instance, we may
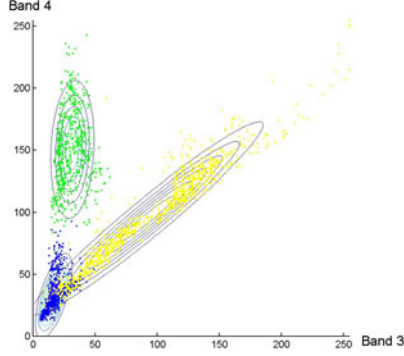
Figure 2.12: 2D plot of pdf based on pure and mixed samples of three land-cover classes: built-up (yellow), vegetation (green) and water (blue).
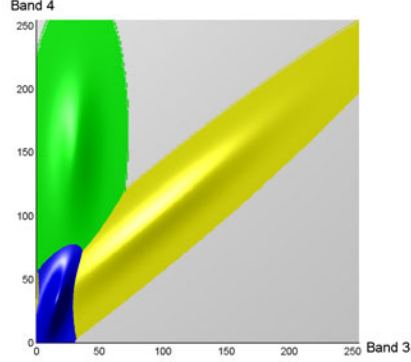
Figure 2.13: 3D draped pdf based on pure and mixed samples of three land-cover classes: built-up (yellow), vegetation (green) and water (blue).

need to select sample pixels from buildings with light tone as well as from buildings with dark tone for the built-up class. The way of combining samples that were selected for each sub-cluster in order to represent a class is called sample integration. Problems may occur in that the built-up class may contain samples from different sub-clusters and the deviation may be overestimated in the feature space with parametric MLC, as shown in Figures 2.10 and 2.11. This problem has been reduced by including mixed pixels as shown in Figures 2.12 and 2.13. However, as you may discover, a small yellow area still exists in the left side of the water cluster in Figure 2.13. Our proposed solution to this problem is using the class integration method instead of sample integration. Unlike sample integration, which simply combines samples of different sub-clusters for parameter estimation of a certain class, class integration estimates a pdf for each sub-cluster in a class and then creates a united pdf for this class by applying a maximum function

$$pdf(\omega_i) = max(pdf_1, pdf_2, \ldots, pdf_k)$$

where class ω_i contains k sub-clusters. The class integration method provides a better solution to this problem, as shown in Figure 2.14 for pure samples and Figure 2.15 for pure and mixed samples. Based on the above testing and observations, we observe that by combining the efforts of single-pixel training, the involvement of boundary or mixed pixels, and class integration, classification accuracy may be improved for high-resolution images.
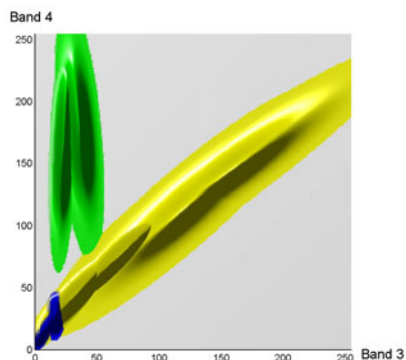
Figure 2.14: 3D draped pdf based on class integration method with pure samples of three land-cover classes: built-up (yellow), vegetation (green) and water (blue).
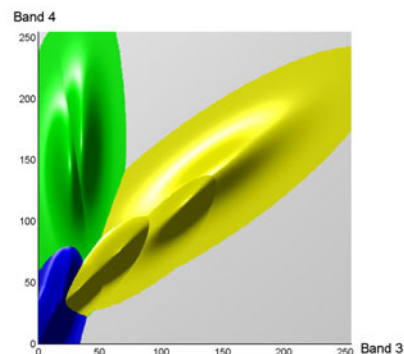
Figure 2.15: 3D draped pdf based on class integration method with pure and mixed samples of three land-cover classes: built-up (yellow), vegetation (green) and water (blue).

2.6.4 Effectiveness of the proposed modifications

To examine the effectiveness of the proposed modifications in a real situation, some 300 samples were manually selected from an IKONOS image for each individual cluster, using the single-pixel approach. According to the spectral values of the four bands of the IKONOS image, the built-up class was subdivided into four clusters, from very light tone to very dark tone; vegetation was subdivided into two clusters, trees and lawn; water was subdivided into two clusters, lake and canal. The first 100 samples were selected from pure pixels and the following 200 samples were selected from boundary or mixed pixels for each cluster. All samples had been divided into two groups by taking every two samples alternately, so that samples with an odd number in the list were used for training and those with even numbers were used for quality assessment. The experimental results of four tests are presented in Tables 2.6, 2.7, 2.8 and 2.9. They refer to sample integration with pure samples, sample integration with pure and mixed samples, class integration with pure samples, and class integration with pure and mixed samples. Comparison of the approaches in terms of overall accuracy and the Kappa coefficient is shown in Figures 2.16 and 2.17. Based on the test results, we can conclude that in general the class integration method achieves a higher quality than sample integration. The highest classification quality was obtained by using the class integration method with pure and mixed samples. Comparing Table 2.6 and Table 2.7 as a result of applying the sample integration method, we can see that the classification quality is reduced when mixed samples are included. This is mainly because the pdf values of classes with small deviations (e.g. water class) are much more sensitive in response to an increasing number of mixed samples in sample integration than classes with

Table 2.6: Quality assessment of ML classification based on sample integration of 50 pure training samples and 150 test samples for each cluster

		Reference Data		
		Built-up	Vegetation	Water
Classified Data	Built-up	573	0	21
	Vegetation	4	299	9
	Water	23	1	270

Overall accuracy: 95.17 %, Kappa coefficient: 92.28 %

Table 2.7: Quality assessment of ML classification based on sample integration of 150 pure and mixed training samples and 150 test samples for each cluster

		Reference Data		
		Built-up	Vegetation	Water
Classified Data	Built-up	553	1	14
	Vegetation	2	298	5
	Water	45	1	281

Overall accuracy: 94.33 %, Kappa coefficient: 91.03 %

large deviations (e.g. built-up class), as is obvious from comparing Figure 2.11 and Figure 2.13. It confirms our earlier observation, and proves that our proposed measures can improve classification quality. One suggestion is to select, in addition to pure samples, samples near the boundaries with other classes and to use the class integration method instead of the sample integration method when sub-clusters exist in end-member classes. A small part of the test site is selected for close observation of the classification results obtained by applying different combinations, as shown in Figures 2.18 and 2.19. The classification results of these different combinations are illustrated in Figures 2.20, 2.21, 2.22 and 2.23. We can observe that many pixels in the dark roofs had been classified as water in Figure 2.20. The situation is even worse in Figure 2.21 because of disproportional changes in the MLC parameters owing to the involvement of mixed samples. The pdf values increase faster for the water class than for the built-up class in the boundary region between the water and built-up classes in the feature space, because the built-up class has a much larger deviation than the water class when sample integration is applied (see the differences between Figure 2.11 and Figure 2.13). From Figure 2.22 we can see that the result is better when using the class integration method. However, the best result is obtained by using pure and mixed samples and applying the class integration method (see Figure 2.23). Figure 2.24 show the improved classification result obtained by using the proposed method.

Table 2.8: Quality assessment of ML classification based on class integration of 50 pure training samples and 150 test samples for each cluster

		Reference Data		
		Built-up	Vegetation	Water
Classified Data	Built-up	571	0	17
	Vegetation	7	298	8
	Water	22	2	275

Overall accuracy: 95.33 %, Kappa coefficient: 92.56 %

Table 2.9: Quality assessment of ML classification based on class integration of 150 pure and mixed training samples and 150 test samples for each cluster

		Reference Data		
		Built-up	Vegetation	Water
Classified Data	Built-up	584	1	21
	Vegetation	4	298	3
	Water	12	1	276

Overall accuracy: 96.50 %, Kappa coefficient: 94.39 %
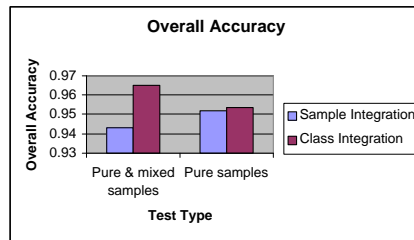
Figure 2.16: Comparison of overall accuracy based on 150 test samples for each cluster.
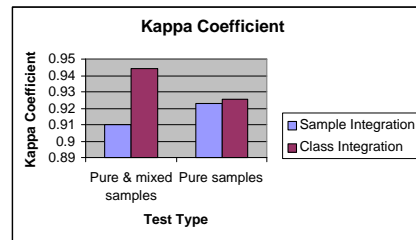
Figure 2.17: Comparison of Kappa coefficient based on 150 test samples for each cluster.
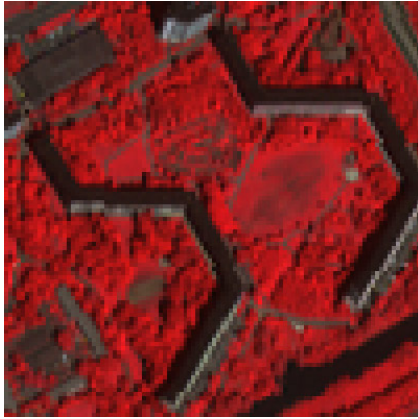
Figure 2.18: A small portion of IKONOS image for close observation. This is a complicated part of the test site due to similarity of spectral features of dark roof, shadow and water.
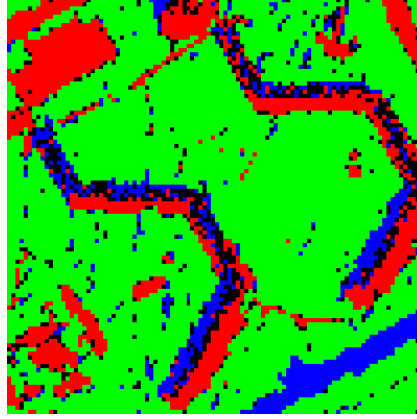
Figure 2.19: The differences between classification results of different combinations (black pixels indicate the different classification results of different combinations).
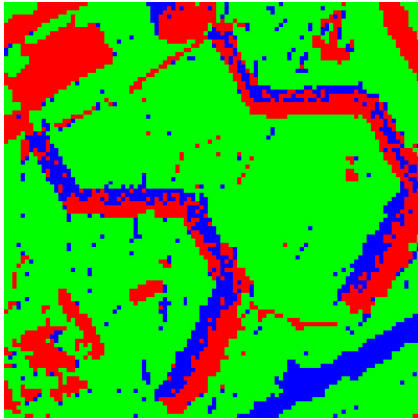
Figure 2.20: Classification result of applying the sample integration method and using the pure samples.
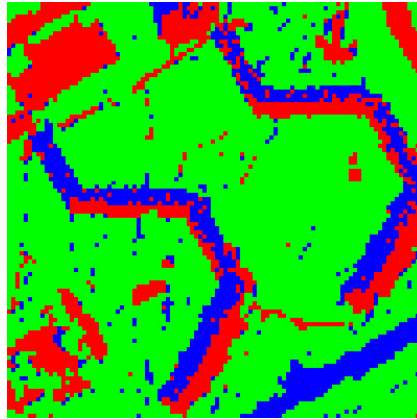
Figure 2.21: Classification result of applying the sample integration method and using the pure and mixed samples.
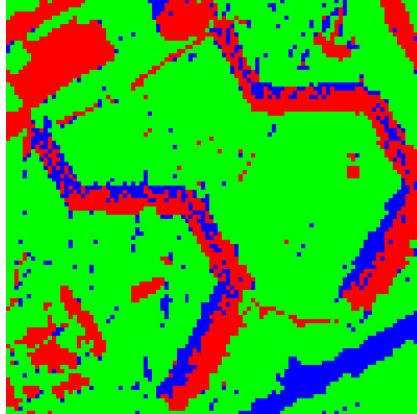
Figure 2.22: Classification result of applying the class integration method and using the pure samples.
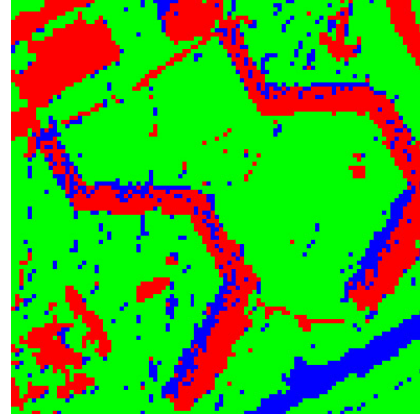
Figure 2.23: Classification result of applying the class integration method and using the pure and mixed samples.
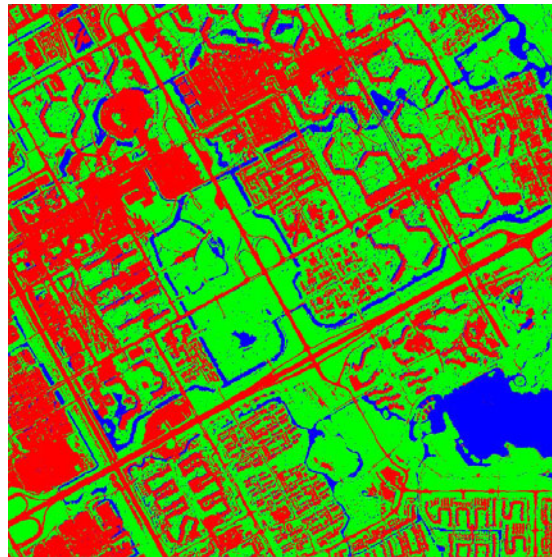
Figure 2.24: Classification result of applying the class integration method and using the pure and mixed samples.

2.7 Proposed object-based approach for land-cover and land-use classification

2.7.1 Complexity of land-use classification

Urban land-use in an urban planning context refers to certain functions with related social-economic characteristics. For instance, a residential area consists of a number of physical features such as residential buildings, parking spaces, footpaths, green space, and maybe canals. Quite often, these features are targets of land-cover classification. Physical features in general have certain associations with spectral features, so they can be identified by using multi-spectral information from remote sensing images. However, land use cannot be determined directly from land-cover information (Barr and Barnsley, 1997; Zhan et al., 2002d). This is because land use is an abstract concept – an amalgam of economic, social and cultural factors – one that is defined in terms of function rather than physical form. Urban land use might be distinguishable in terms of the morphological properties of, and the spatial relations between, their component land-cover parcels (regions) (Barnsley and Barr, 1997). It is possible, for example, that different sample areas of the same nominal land use might exhibit somewhat different morphological properties and/or spatial relations in terms of their component land-cover regions. If the within-class variation (i.e. within a single land use) is greater than the between-class (i.e. between different land uses), then it will not be possible to identify and to distinguish urban land use consistently on the basis of these structural measures (Barnsley and Barr, 1997). Besides the morphological properties and the spatial relations in terms of their component land-cover regions, a number of other indicators are needed in order to identify land use, such as proportion of areas covered by different types of land cover, building density, floor area ratio, or evidence derived from other sources. In addition, correct delineation of the spatial extent of a land-use unit is a crucial factor, and many land-use-related measurements (such as building density) may be influenced by delineating different land-use units, possibly leading to incorrect identification. Therefore, in this research an intermediate stage is proposed for finding spatial units where certain functions are held spatially. This intermediate stage between land-cover classification and land-use classification reflects what is happening in the human vision system. The indicators, as mentioned above, are supposed to be extracted from laser data or multi-spectral images. Many features essential for land-use classification are per-object based. They are difficult or impossible to extracte by per-pixel approaches. Therefore, the object-based approach is needed.

2.7.2 Object-based approach for land-cover and land-use classification

A hierarchy of three levels is proposed to achieve land-use classification at three levels, namely pixel level, land-cover level and land-use level (Figure 2.25). At each level we create image objects in order to represent spatial coverage and respective thematic information. For maintaining logical consistence and uniform expression
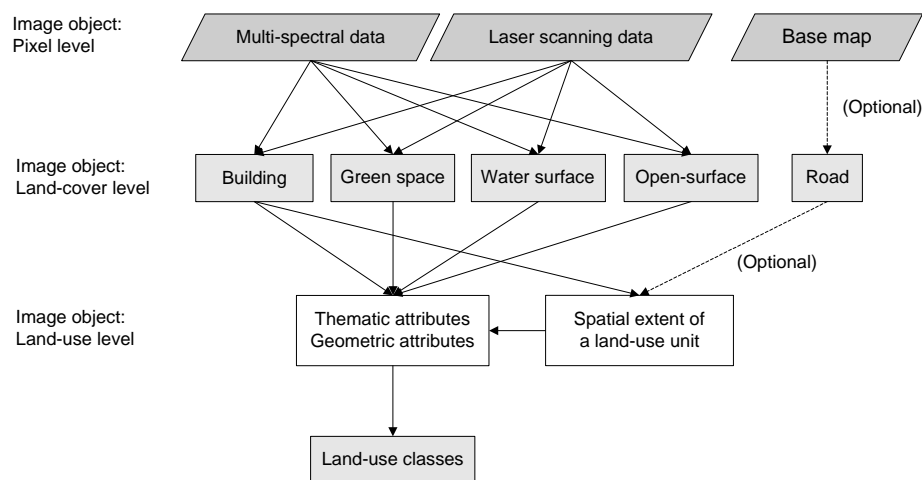
Figure 2.25: Hierarchy of image objects and work flow.

in object-based modelling, each pixel is treated as an object despite the holding of some unique characteristics such as uniform size, shape, function of spatial location. Detailed descriptions of the proposed approach are presented in coming chapters.

2.8   Discussion and outlook

Two test sites with corresponding data have been introduced in this chapter. This is not because we are going to solve problems raised by these particular sites, but because we wish to illustrate typical problems immediately. We shall use them for the practical testing of methods proposed and developed in this research. By introducing the test sites and data preparation in this chapter, we are able to use the data for test purposes and show the test results in the following chapters without any further explanations about the test sites and their data.

Triggered by the potential problems of high-resolution data, we have examined the most popular pixel-based classifier, the MLC, as to its suitability for this type of data. A number of problems have been observed and highlighted, such as the existence of sub-clusters. Several remedial measures for the problems observed have been proposed and tested. Quality improvement has been achieved by modelling the decision surface in the feature space, which aims at obtaining better spatial partitions for each end-member class in feature space, based on selected samples. The experimental results have confirmed the effectiveness of these measures. However, other issues such as pixels in shadow areas and relief displacement caused by non-vertical observation remain untouched by the proposed improvements. These problems will be revisited by using an object-based approach in Chapter 6.

Although the proposed modifications have raised the classification accuracy of

the MLC, we consider the attainable results insufficient for a detailed urban land-use classification. Based on our knowledge of, and experiments in, visual image-interpretation, the key features for image-interpretation, such as size, shape, colour, orientation, pattern, association. are directly associated with explicit objects, which are at higher abstraction levels than pixels. These key features should continue to play a key role in image analysis and land-use classification. For instance, we need to check how buildings are spatially distributed in space in order to find out if they belong to a residential area. We need to know the number of floors of a building in order to achieve better understanding and classification. We need to know if buildings are similar in size, height, orientation, etc. We need to explore the surrounding features of specific objects. Such information cannot be acquired by per-pixel approaches because they are directly associated with objects, not pixels. Therefore, object-based image processing techniques, which provide additional tools and methods for dealing with higher abstraction levels, are considered for higher levels of image analysis. Thus an object-based image analysis approach becomes the main focus of this research. It will be investigated in the coming chapters.

Chapter 3

An object-based conceptual model for urban land-cover and land-use classification

3.1 Introduction

Geographical information systems (GIS) were developed in the late 1960s but very few places installed them because of their expensive hardware and limited software. Since the early 1980s, there has been a marked increase in the installation of GIS at different levels and in different departments of urban and regional governments in the developed countries. with the introduction of microcomputers in the late 1980s, GIS became increasingly being used in planning agencies in the developing countries. The inventory, analysis and mapping capacities of GIS are the main functions that had wide applications in urban and regional planning at this stage (Marble and Amundason, 1988; Yeh, 1988; Chen et al., 1989; Yeh, 1991). The large-scale implementation of urban (planning) information systems indicates the success at this stage. The main characteristics at the time were the practical uses of GIS and remote sensing on a daily-routine basis. The developments during this stage have led the planning discipline to move from analogue to digital. The entity-relationship approach is the prime tool for semantic data modelling. However, experience has shown that for many systems the initial set of modelling constructs (entity, attribute and relationship) is inadequate. There are many phenomena in geographical reality that are readily perceived as entities and represented by objects such as lakes, rivers, roads, buildings. In the planning discipline, many geo-spatial entities such as residential, commercial and industrial in land-use mapping are perceived as en-

tities and are represented by objects as well, despite the fact that no physical boundaries may exist between different land-use objects. Natural language is much more suited to describing objects and fields (Cova and Goodchild, 2002). Many phenomena and concepts in planning are described by way of natural language, such as built-up areas, a good living environment, a walking distance. Object-oriented analysis (OOA) is a method of analysis that examines requirements from the perspective of the classes and objects found in the vocabulary of the problem domain (Booch, 1993). Many principles and techniques for managing complexity have been collected in the OOA, such as abstraction, encapsulation, inheritance, association. (Coad and Yourdon, 1990). The OOA approach is applied in object-based spatial data modelling in the planning and management domain in order to support object-based image analysis toward land-cover and land-use classification in urban areas. Considering the implementation of the OO models in a raster environment, these objects are formulated based on the raster representation. In the following sections, a layered sequence is followed for the OOA – from Subject layer, Class-&-Object layer, Structure layer, Attribute layer to Service layer (Coad and Yourdon, 1990). Object-based conceptual analysis in the planning context follows the same line as OOA and is presented toward urban land-cover and land-use classification as a subject layer. A brief introduction of the OO approach and tools will follow. In the coming sections, conceptual modelling is elaborated by using an adapted syntax based on the notions from both UML (United Modelling Language) and OMT-G (Object Modelling Technique for Geographic Application), and on the notion presented by Molenaar (1998), corresponding to Class-&-Object layer, Structure layer, Attribute layer and Service layer. A number of diagrams are produced concerned with class, structure, attribute and operation. Please note that the conceptual modelling presented in this chapter focuses on the problem domain of urban land-cover and land-use classification based on high-resolution data and does not aim at a general model for GIS in urban planning and management. Therefore, structures, attributes and operations, such as ownership, address and building materials of a building, which rely on additional data rather than remote sensing data, are not considered in this modelling.

3.2 An object view on geo-spatial data used in urban planning and management

3.2.1 Objects and their behaviours in the urban planning context

A starting point for defining objects in an urban planning context may be to look at what features are currently stored and processed in existing GISs. We may find settlements, buildings, roads, commercial districts, green areas, water bodies, industrial areas, etc. These can be treated as objects. However, some of them are not likely to appear on a map at the same time.

By taking a close look at these objects, we can see that some of them are physical entities that have physical properties and physical boundaries, such as buildings, green spaces, water bodies. And there are conceptual entities that consist of other

physical entities, often with fuzzy boundaries, such as residential areas, commercial areas and industrial areas, which directly relate to land use. A residential area may consist of buildings, gardens, footpaths, small lakes or canals, etc. We may notice that certain entities can be treated as physical entities in some cases but considered as conceptual entities in others. For instance, a lake classified as water surface in land-cover classification may, in land-use classification, be identified as recreational land use if it is located in a park, or as a fishing pool if it is located outside the built-up area. Therefore, an entity may 'behave' differently in different circumstances.

3.2.2 Hierarchy of planning

Regional planning

Regional planning aims at the reasonable structure and spatial distribution of production elements at the regional scale. It deals mainly with abstract entities such as human settlement, industrial zones, transportation networks. The central location, physical size and the spatial coverage of its influence zone on the surrounding regions are the main features to be modelled in regional planning.

Master planning

Master planning aims at the sound spatial and sectional distribution of land in urban built-up areas and surrounding regions.

Detailed planning

Detailed planning deals mainly with organic spatial arrangement at the neighbourhood level to meet the certain functions assigned to each neighbourhood block.

3.2.3 Objects at different hierarchical levels of urban and regional planning

Objects at the regional planning level

There are three main types of objects at the regional planning level: point objects, line objects and area objects, corresponding to settlements, transportation networks and influence zones.

Objects at the master planning level

There are three types of objects at the master planning level: point objects and line objects, which are mainly for providing spatial references, and area objects (land use), which are the spatial partitions of major land-use classes as represented in a 2D space.

Objects at the detailed planning level

The objects at the detailed planning level are very similar to the objects at the master planning level. The main differences are small spatial units and greater specificity in land-use functions. Therefore, land-use objects at the master planning level can be generated from land-use objects at the detailed planning level by merging objects from specific land-use classes into major land-use classes and dissolving the boundaries between neighbouring objects within the same major land-use class. Land-use objects at the detailed planning level can be disaggregated or specified from land-use objects at the master planning level.

Objects at the land-cover level

Physical entities such as buildings, roads, green spaces and water bodies are represented by objects at the land-cover level. This is a fundamental base for land-use classification. Each object can be an element of a land-use class. Land-cover objects can be used as an indication for determining the spatial extent of a land-use unit.

3.2.4   Object types

Objects with similar properties or similar behaviours are organised into types. Similar behaviours can be identified according to various criteria or perspectives such as spatial extent and abstraction level. Types of objects will be discussed in the following subsections, according to different perspectives.

3.2.5   Types of objects concerning different abstraction levels

Elementary objects (images or field data)

Pixels are regarded as elementary objects that have uniform geometric properties. Elementary objects share many methods or operations of pixel- or raster-based processing, such as filtering, convolution, classification.

There are two types of elementary objects, one relating to images or field data and one relating to object fields. There is one main difference between the two. The images or field data take data take digital number (DN) values from pixels. These DN values usually range from 0 to 255 or actual height values of laser data. The DN values of an object field are taken mainly from membership functions according to the characteristics of the object. The DN values of an object field are taken from the Boolean value 0 (false) or 1 (true) for crisp objects or a real value from 0 to 1 for fuzzy objects.

Objects at the land-cover level

Objects at the land-cover level have object ID, geometric properties such as location, size, shape, orientation, as well as class-related attributes such as class name or class ID, mean value and standard deviation of membership functions in this class.

Objects at the land-use level

Objects at the land-use level have object ID, geometric properties such as location, size, shape, orientation, as well as class-related attributes such as class name or class ID, composition and proportion of land-cover types contained, number of buildings held by a land-use object.

3.2.6 Types of objects concerning their spatial extent

Objects with fixed boundaries and adjacency relationship (O_{Pixel})

Pixels or O_{Pixel} are elementary objects in a raster data model. A pixel is the smallest spatial unit of uniform size. Pixel values are spectral reflectance in multi-spectral images, height values in laser data, or membership function values in object fields represented in a regular grided space. In general all image processing operations are applicable to this class.

Objects with physical boundaries or physical indications for their spatial extent (O_{PB})

O_{PB} are representations of physical entities existing in the real world, such as buildings, roads, rivers. Their spatial extent can often be determined by their physical boundaries. These types of objects are often considered in land-cover classification. Many land-cover objects such as buildings, roads and water belong to this class. Edge detection and regional growth methods are common tools for delineating this class. These tools are often not sufficient for object extraction. For instance, many edges and regions may be extracted even from the same roof, since it is difficult to identify whether an edge is part of an object or the boundary of the object or whether two regions belong to the same object or to different objects. Additional measures such as size and shape can be introduced by the object-based approach for extracting objects of this class (as discussed in Chapter 6).

Objects with fuzzy boundaries or fuzzy indications of their spatial extent (O_{FB})

O_{FB} are representations of physical objects existing in the real world, such as green spaces in the planning context or vegetation (trees or lawns). It is often difficult to extract the spatial extent with clear sharp boundaries. For instance, it is difficult to determine where a dense wood ends and where grassland begins in an area where tree density is changing gradually in a transition zone. These types of objects such as green space are relevant in both land-cover and land-use classification and are often used in environmental studies. Their spatial extent will have to be determined based on a decision surface, according to fuzzy membership functions.

Objects without physical boundaries, or their existence and spatial extent are conceptualised (O_C)

O_C are representations of conceptualised entities, in the sense that their existence has to be referenced based on the existence of a number of key features and their spatial extent has to be inferred based on the spatial distribution of these key features. Most objects in land-use classification are of this type, such as residential, commercial, industrial areas. The existence and spatial extent is determined based on decision surfaces. The decision surfaces might have to be generated based on a number of indicators, according to definitions such as building density and floor area ratio.

3.2.7 Land-cover and land-use classes

Land-use classes are determined based on economic functions and can be reasoned according to evidence of the ground activities associated with a place. Many classification systems have been developed for different purposes in different countries. For instance, the land-based classification standard was developed by the American Planning Association (American Planning Association, 2001). The first and second levels of land-use classes in the function dimension are presented in Appendix A, Table A.3. The National Land Use Database (NLUD) Land Use Classification version 3.3 was proposed by the British authority (Harrison and Garland, 2001). Its 13 divisions and 51 classes are provided in Appendix A, Table A.4. The National Standard for Urban Land Use Classification was presented by the Chinese authority (The Ministry of Construction P. R. China, 2001). Its first and second levels of land-use classes can be found in Appendix A, Table A.5. The land-use classification that we are talking about in our research concerns the need for investigating and monitoring how urban land is actually used for various purposes in an urban space, and provides planners and decision makers with quantitative measurements of how urban land is used and changed in terms of physical space, time and different categories. The purpose is different from that of land-use codes used in zoning, where particular land-use codes are used to specify the preferences for certain uses of a particular land parcel, although the same or similar land-use classification systems are applied in both cases. In general, these types of classification systems are based mainly on the needs of planners and other related businesses, without much consideration as to whether remote sensing or visual aerial photo-interpretation can yield the wanted data. This type of classification system is too detailed (considering both the first and second levels), thus requiring additional information even if manual visual interpretation is applied. A general land-use classification system was proposed in 1976 considering the use of satellite images (Anderson et al., 1976). Anderson et al. (1976) had manual interpretation in mind and the main data sources available at that time were the Landsat MSS data with a spatial resolution of 79 m.

To make use of high-resolution images (0.5 m to 4 m) in an automated approach, the above-mentioned classification systems need to be adapted and merged. Since we are interested in land-use classification in urban areas, we consider only urban-related land-use types. The proposed urban land-use classification system for the

automatic classification experiments in this research is presented in Table A.6.

3.3   Object-based analysis and modelling

Object-based modelling follows the line of the object-oriented (OO) approach, due consideration to the spatial components in this research. Since many issues regarding the OO approach have been discussed in existing literature (Booch, 1993), here we will mention only some key concepts and provide some examples in applying the OO approach in the urban planning and management domain.

Notions are taken from both UML and OMT-G. Although these tools were originally developed for modelling toward OO programming in software development, they are also important tools for object-based analysis and modelling, and the results can be easily understood and implemented by following these notions and diagrams.

3.3.1   Key elements of an object-based approach

From the object perspective, there are several key elements such as abstraction, encapsulation, modularity, hierarchy. The following definitions are given by Booch (1993).

Abstraction

An abstraction denotes the essential characteristics of an object that distinguish it from other kinds of objects and thus provide crisply defined conceptual boundaries, relative to the perspective of the viewer. For instance, land-cover classes are an abstraction of physical features of both natural and man-made objects in reality. Land-use classes are an abstraction of conceptualised features in terms of human activities, which are usually indicated by certain types of physical features in reality.

Encapsulation

Encapsulation is the process of compartmentalising the elements of an abstraction that constitute its structure and behaviour; encapsulation serves to separate the conceptual interface of an abstraction and its implementation. For instance, an index to pixels that are parts of an image object are encapsulated in the spatial embedding of an image object. Thus other geometric properties such as the centre of mass, size, shape and orientation can be derived from these pixels and can be encapsulated as well.

Modularity

Modularity is the property of a system that has been decomposed into a set of cohesive and loosely coupled models. Image processing operations, which are applicable to elementary objects (images or field data) will be collected in an image processing model, while additional operations applicable to image objects will be the components of an image-object processing model.

Hierarchy

Hierarchy is a ranking or ordering of abstractions. Pixel, land cover and land use form a hierarchy in terms of spatial coverage: a land-cover object is a collection of elementary objects (pixels) and a land-use object is a collection of land-cover objects.

3.3.2 Class

A class is a set of objects that share a common structure and behaviour. A single object is simply an instance of a class. In a land-cover or land-use classification system, each class is readily treated as a class by its definition as it is in the OO domain. However, the definitions provide mainly the indications of the chief characteristics of classes. Their explicit attributes and behaviour are often not provided. Class attributes and operations will have to be defined in the OO modelling.

3.3.3 Relationships among classes

In an OO perspective, several versions of relationships have been proposed for describing relationships among classes (Worboys et al., 1990; Booch, 1993; Fowler and Scott, 1999). The Worboys version used in this research includes generalisation, specialisation, aggregation and association. The following definitions are proposed by Worboys et al. (1990). Examples are given of potential applications to land-cover and land-use classification.

Generalisation

Generalisation is the construction that enables groups of similar types to be considered as a single higher-order type. Land-use object types at the master planning level can be generalised from land-use object types at the detailed planning level.

Specialisation

Specialisation is the construction that enables the modeller to define possible roles for members of a given type. A residential class may be subdivided into sub-classes according to the number of floors of residential buildings or building density, etc. A land-use type consists of a number of land-cover types. Under certain circumstances, some land-cover types should not be included in this particular type. For instance, when a footpath, originally part of a residential area, has been expanded to become a main road, it should be included in the transportation class instead of the residential class.

Aggregation

Aggregation is the construction that enables types to be amalgamated into a higher-order type, the attributes of whose objects are a combination of the attributes of the constituent types. Mapping from land-cover types to land-use types can be treated as

aggregation, so that some attributes of land-cover types become attributes of land-use types.

Association

Association or grouping is the construction that enables a set of objects of the same type to form an object of higher-level type. Hierarchically formed land-use classification systems readily provide associations between classes at different levels. Spatially adjacent residential areas of different types can be merged into a single residential area at a higher level where only the major land-use classes are used.

3.3.4 Object-based modelling tools

United Modelling Language (UML)

UML is a standard tool for OO modelling and OO programming (Fowler and Scott, 1999; OMG, 2001). There have been many publications over the past decades in the fields of OO modelling and UML. For further information, a UML bibliography is available online (UML Bib, 2002).

Object Modelling Technique for Geographic Application (OMT-G)

OMT-G is an object-oriented data model for geographical applications, which was initially based on the classic OMT class diagram notion, and later adapted to approach the concepts and notion of UML. OMT-G offers primitives that provide the means of modelling the geometry and topology of geographical data, making the modelling of geographical applications easier (Borges et al., 2001).

An extension of OMT-G

Since many primitives of geographical objects are defined based on the vector data model in OMT-G, an extension is proposed to include geographical objects with primitives based on the raster data model in this research. Geo-field or image, object field and image object are three types of objects defined based on the raster data model.

Geo-field or image is called an elementary object, which is the fundamental base for object-based image analysis and object-based land-cover/use classification.

Object field is a field but with field values reflecting certain semantic meanings regarding certain features. Such field values are often used to represent fuzzy membership functions for semantic mapping from field data to spatial embedding of objects or desired features in the thematic domain.

Image object is a conditional object in raster form, which is often derived from the object field by image segmentation with certain thresholds. Graphic notation for the extended OMT-G classes is shown in Figure 3.1.

▦ Class name	☁ Class name	▦ Class name
Attributes	Attributes	Attributes
Operations	Operations	Operations
a	b	c
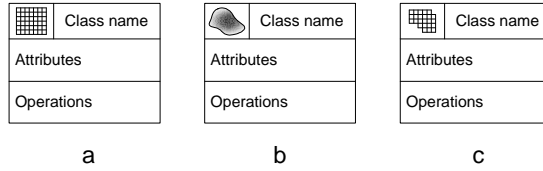
Figure 3.1: Graphic notation for geo-field (a), object field (b) and image object (c).

3.3.5   A syntax for object-based modelling in the context of urban planning and management

Object and object class

We use the bold italic font to represent a class and use an index for an object. For instance, we use ***Building*** to denote the class building and use ***Building***$_i$ to denote a building object.

Method or operation

We use the italic font to represent a method that is applicable to a class, which is quoted in brackets. For instance, we use *Topo*[***Object***$_i$,***Object***$_j$] to denote the method or operation to obtain the topological relationship between ***object***$_i$ and ***object***$_j$.

Notions for spatial relations

The sign $\stackrel{s}{=}$ denotes a relationship in the spatial domain. The symbol $\cup_s$ denotes the union in the spatial domain or spatial partition.

3.4   Object classes in raster for land-cover and land-use classification

3.4.1   Object classes concerning their geometric and topological characteristics

Elementary object

As an elementary object, a pixel in a geo-field or image has a uniform size and an adjacency relationship with surrounding pixels. It also inherits properties and operations of images or field data. Many image or raster processing operations are applicable to this class, such as filtering and convolution.

An elementary object will have following attributes:

- ID: X, Y or row, column
- Geometric properties: uniform size and shape (pixel)
- Thematic attributes: DN values of images and field data (e.g. multi-spectral, elevation values of DSM1 and DSM2, etc.)

Remote sensing image data normally take 8-bit or 11-bit (IKONOS) gray values from 0 to 255 or 0 to 2047 for each spectral band. Laser scanning data usually record height information in real values. Laser scanning data are usually represented in regular grid or raster form and can be processed like images. Variables in field data can be of any data type: binary, nominal, ordinal, interval or ratio.

Object field

An object-field generally has all the same properties and operations as an elementary object. It usually takes values from 0 to 1 as a membership value. In a crisp case, an integer 0 or 1 is the attribute value for each pixel, while a real number between 0 to 1 will be the attribute value in a fuzzy case.

An object field will have the following attributes:

- ID: X, Y or row, column
- Geometric properties: uniform size and shape (pixel)
- Thematic attributes: membership function values of corresponding features of an object

Image object

Image objects will be determined based on a number of conditions according to their definition. The spatial embedding of image objects is similar to image regions, so that geometric attributes include size, shape and orientation. Topological relations are determined based on surrounding objects (or image regions) rather than on neighbouring pixels. Both land-cover objects and land-use objects inherit attributes and operations from image objects. Due to the conceptual nature, land-use objects are at a higher hierarchical level and will be determined based on land-cover objects at a lower level. A number of attributes and operations of land-cover objects can be inherited by the land-use objects through spatial aggregation, so that the land-use object will have additional attributes, such as structural attributes, which represent the spatial distribution of land-cover objects in the spatial embedding of a land-use object.

An image object will have the following attributes:

- ID
- Geometric properties: size, shape, orientation, etc.
- Thematic attributes: land-cover or land-use class
- Semantic attributes: typicality or membership function values to designated classes
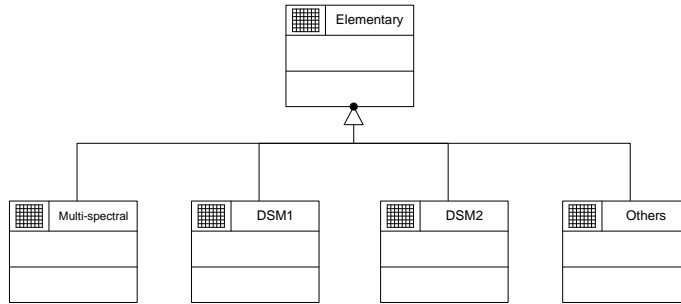
Figure 3.2: Components of elementary class.

3.4.2 Object classes and their sub-classes at different hierarchical levels

At the pixel level

Field data obtained by airborne or satellite-borne sensors provide multi-spectral and/or height information, which is recorded data describing the earth's surface. Images and fields are components or a sub-class of the elementary class.

Multi-spectral imagery, LIDAR data (DSM1, DSM2) etc. are used as attributes of the elementary class in this research.

The relationship between the elementary class and its components is presented in Figure 3.2.

At the land-cover level

Buildings, roads, green spaces and water surfaces are the main land-cover sub-classes that can be extracted from remote sensing images. The land-cover class can be generalised from these sub-classes. The land-cover sub-classes are specialised from the land-cover class. Their attributes and operations can be inherited from the land-cover class. The relationship between the land-cover class and the land-cover sub-classes is presented in Figure 3.3.

Besides the attributes that are inherited from the image object, there are more specific attributes that can be extracted and added to the individual land-cover sub-class. Some examples are proposed:

Building: number of floors, etc.

Green space: average height, etc.

Water: water table, etc.
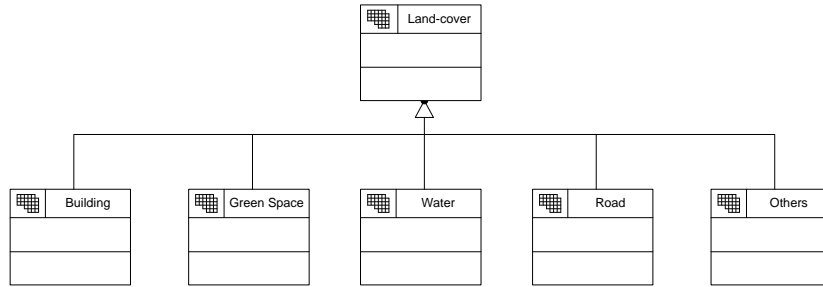
Figure 3.3: The relationship between land-cover class and land-cover sub-classes.
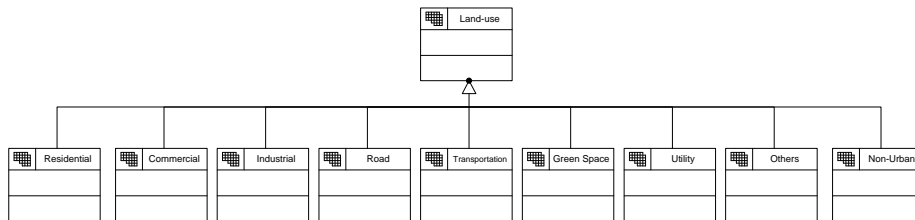
Figure 3.4: The relationship between land-use class and land-use sub-classes.

Road: road width, etc.

At the land-use level

Objects at the land-use level form the land-use class. Residential, commercial and service, industrial and warehouse, main road and main street (including main roads and main streets inside the built-up areas), transportation (including main roads that connect other cities, railway lines and stations as well as airports and harbours), public green space (large green spaces for protectional use, such as trees and lawns planted along river banks and in areas between residential areas and industrial areas, as well as large green spaces for leisure use), utility, others in urban areas and non-urban are some of the major land-use sub-classes (see detailed descriptions and definitions in Appendix A). The land-use class can be generalised from these sub-classes. The land-use sub-classes are specialised from the land-use class. Their attributes and operations can be inherited from the land-use class. The relationship between the land-use class and the land-use sub-classes is presented in Figure 3.4.

An object class at a higher level can be aggregated from a lower level. The object
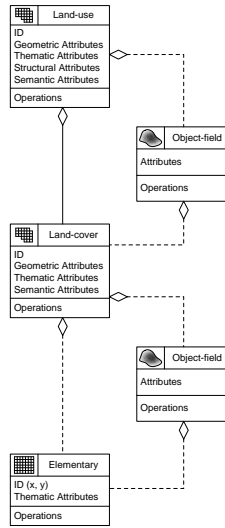
Figure 3.5: Hierarchical aggregation of classes and their attributes.

field plays an important role and acts as a mediator for aggregations or mapping between an elementary object and a land-cover object and between a land-cover object and a land-use object. The aggregation hierarchy of the elementary, land-cover and land-use classes is illustrated in Figure 3.5.

3.5 Methods needed for land-cover and land-use classification

A number of methods are needed to support land-cover and land-use classification and must be attached to applicable objects and classes. The following are some key methods as proposed in the conceptual modelling stage. They will play an important role in various aspects in this research, such as feature transformation, object formation, structural analysis, classification.

3.5.1 Standard pixel-based methods

There are many standard image processing methods or operations such as morphological operations, which can be found in the literature (van der Heijden, 1994; Parker, 1997; Richards and Jia, 1999; Sonka et al., 1999; Tso and Mather, 2001) and are ready to be applied to images and field data. This type of operation can be described using the following syntax:

$$\mathbf{Image}_{\text{out}} = Operation(\mathbf{Image}_{\text{in1}}, ...)$$

Convolution

Convolution is a commonly used operation in image processing. There are several possible notations to indicate the convolution of 2D image to produce an output image. The most common are:

$$\mathbf{Image_{out}} = Convolution(\mathbf{Image_{in}}, \text{Kernel}) = \text{Kernel} \otimes \mathbf{Image_{in}}$$

Morphological operations

Morphological operations are useful tools in image processing. Commonly used operations include dilation, erosion, closing and opening. A more detailed description of morphological operations can be found in Parker (1997).

Dilation operation

A dilation of set A by set B is:

$$A \oplus B = \{a + b, a \in A, b \in B\}$$

where A represents the image being operated on, and B is a second set of pixels, a shape that operates on the pixels of A to produce the result; set B is called a *structuring element (SE)*, and its composition defines the nature of specific dilation. Dilation can be used to acquire the surrounding pixels of an image object (e.g. building), which are then used to obtain a feature (e.g. ground elevation of a building). Dilation can be used to generate the solid core of a land-use object (a residential area), based on clustered land-cover objects (buildings).

Erosion operation

An erosion of set A by set B is:

$$A \ominus B = \{z \mid (B)_z \subseteq A\}$$

Opening operation

The application of an erosion immediately followed by a dilation using the same structuring element is referred to as an opening operation. The name 'opening' is a descriptive one, describing the observation that the operation tends to 'open' small gaps or spaces between touched objects in an image (Parker, 1997). The opening operation is an efficient tool for the removal of noises (small clutters) and for splitting touched image objects.

Closing operation

A closing is similar to an opening except that the dilation is performed first, followed by an erosion using the same structuring element. A closing operation is often

used to fill small holes in an image object.

These morphological operations can be expressed as:

$$\textbf{Image}_{\text{out}} = Dilate(\textbf{Image}_{\text{in}}, SE)$$

$$\textbf{Image}_{\text{out}} = Erode(\textbf{Image}_{\text{in}}, SE)$$

$$\textbf{Image}_{\text{out}} = Open(\textbf{Image}_{\text{in}}, SE)$$

$$\textbf{Image}_{\text{out}} = Close(\textbf{Image}_{\text{in}}, SE)$$

Examples of applying these morphological operations to a test image are shown in Figure 3.6.
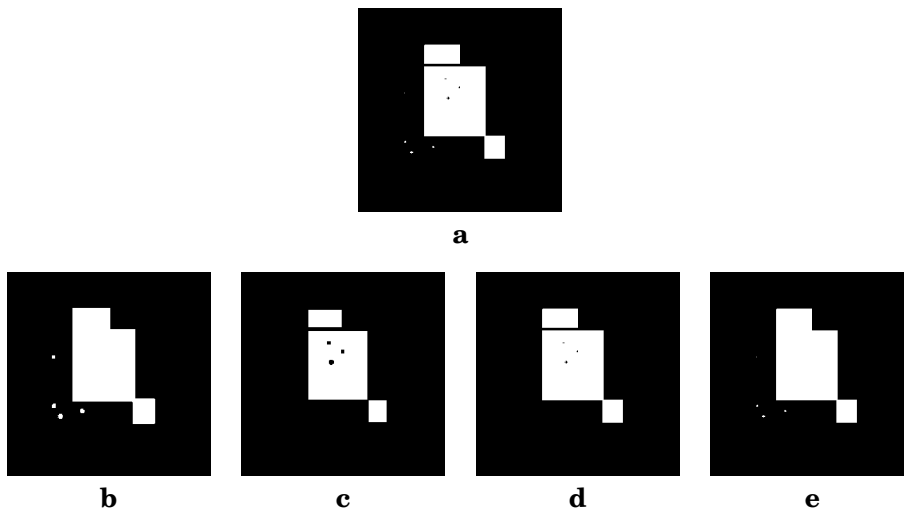
Figure 3.6: Examples of morphological operations: original test image (a); result of dilation (b); result of erosion (c); result of opening (d); result of closing (e).

3.5.2 Transformation for multi-spectral images

Transformation is needed to acquire data that provide indications for desired features, such as the normalised difference vegetation index (NDVI) and the normalised difference water index (NDWI).

Normalised difference vegetation index (NDVI)

The NDVI can be obtained by transformation using the RED band and the NIR band of multi-spectral images.

$$NDVI = \frac{NIR - RED}{NIR + RED}$$

Normalised difference water index (NDWI)

The NDWI can be obtained by transformation using the GREEN band and the NIR band of multi-spectral images.

$$NDWI = \frac{GREEN - NIR}{GREEN + NIR}$$

3.5.3 Methods of sub-pixel analysis

A proposed sub-pixel method is used for image resampling from a coarser resolution to a finer resolution by the spatial modelling of the probability density surface or the surface made by fuzzy membership function values. A detailed description will be presented in Chapter 5.

$$\mathbf{Image}_{sub} = Subpixel(\mathbf{Image}_{in}, factor)$$

3.5.4 Methods of mapping from semantic domain to feature domain

Fuzzy membership function, or mapping from semantic domain to feature domain

In many cases, fuzzy membership functions have to be extracted by semantical modelling, since land-use classes are conceptualised in terms of human activities over the earth's surface in relation to certain social-economic functions that are defined in linguistic terms, and often no physical boundaries can be found for many classes in reality. Therefore methods are needed for constructing fuzzy membership functions or for mapping from the semantic domain to the feature domain.

$$\mathbf{Object\text{-}field} = FuzzyMF(\mathbf{Field}, [Parameter1, ...])$$

3.5.5 Methods of obtaining image objects

Formation of objects

The formation of objects will be based on their membership function according to their characteristics modelled by their definitions. The thematic and geometric components will have to be considered alternately while forming an object. Therefore methods are required for alternately checking the thematic and geometric components for the formation of objects.

$$\mathbf{O}_{bw} = Binary(\mathbf{Object\text{-}field}_{OA}, Threshold)$$

Labelling of objects

When image regions are formed, a method is needed to assign an object ID (label) to each region, using the 4-connection criteria. These IDs are used as identifiers for image objects.

$$\mathbf{O}_{id} = Bwlabel(\mathbf{O_{bw}}, 4)$$

Reasoning on object size

Some features such as vertical walls for building extraction, are based on the third dimension, which can be better assessed by reasoning vertically based on size changes along image regions obtained by segmentation applied to laser data.

$$\Delta_{Size_i} = \frac{Size_i - Size_{i+1}}{Size_i}$$

i and $i+1$ denote objects obtained at elevation layer i and layer $i+1$ respectively.

Reasoning on object location

Vertical walls for building extraction can also be assessed by reasoning based on location changes (centres of mass of image objects) along image regions obtained by segmentation applied to laser data vertically.

$$\Delta_{Loc_i} = \sqrt{(\overline{x}_{i+1} - \overline{x}_i)^2 + (\overline{y}_{i+1} - \overline{y}_i)^2}$$

i and $i+1$ denote objects obtained at elevation layer i and layer $i+1$ respectively.

Building reasoning based on changes in object size and location

$$S_i = \begin{cases} Building, \ \Delta_{Size_i}, \Delta_{Size_{i+1}} < T_{Size} \wedge \Delta_{Loc_i} < T_{Loc} \wedge S_i \in [10, 5000] \ \text{m}^2 \\ Else. \end{cases}$$

i and $i+1$ denote objects obtained at elevation layer i and layer $i+1$ respectively.

3.5.6 Methods of measuring feature similarity

The feature similarity relationship describes the degree of similarity between two objects in geometric feature space, thematic feature space or others. Similarity is measured in a scalable metric space based on certain features. Similarity plays an important role in classification and clustering. In object-based classification or clustering, similarity will have to be measured in feature space based on scalable metric distances as membership functions to desired classes, and in the geometric space based on Euclidean distances. Methods for similarity measurements in geometric space and feature space are required, as well as the integration of similarity measurements from both spaces.

$$Similarity(\mathbf{O}_i, \mathbf{O}_j) = Sim(\mathbf{O}_i, \mathbf{O}_j)$$

3.5.7 Methods of identifying topological relationships

Topological relationship plays an important role in land-use reasoning. The adjacency relationship at pixel level includes two options, 4-connection and 8-connection, between the central pixel and neighbouring pixels. The adjacency relationship at the land-cover level will be determined by close neighbours between an object and its surrounding objects. Close neighbours can be derived by deploying the Delauney triangulation to objects at the land-cover level. Other topological relations between objects will be needed at the land-use level. Therefore methods for extracting topological relationships are useful.

$$\text{Topology}(\mathbf{O}_i, \mathbf{O}_j) = Topo(\mathbf{O}_i, \mathbf{O}_j)$$

3.5.8 Methods of extracting proximity relationships

Delaunay triangulation is applied to obtain proximity relationships between image objects. The shortest distance is extracted by comparing the lengths of the triangle edges that link pixels representing different objects. The result is then used for spatial clustering in order to find spatial units for land-use objects.

3.5.9 Methods of clustering analysis

Fuzzy c-means is selected for finding clusters in the feature space and finding clusters in histogram space.

3.5.10 Methods of spatial clustering

To find land-use spatial units, methods of spatial clustering are needed. The spatial clustering operation has to consider both spatial closeness and feature similarity in order to obtain spatial clusters based on land-cover objects. Such a spatial clustering operation can be regarded as an operation for hierarchical spatial aggregation.

3.5.11 Methods of classification

Classification is often used in remote sensing and planning, while clustering has often appeared in computer sciences and data analysis, although both of them share a core function of grouping or assigning objects into different categories or clusters. Classification is often used when categories are defined such as land-cover or land-use classes. Classification can be applied when data and the required classes are presented. Quite often training samples are needed for the construction of a classifier, either parametric or non-parametric. This is called supervised classification. When training samples are not provided or are unknown, the classifier will have to use data to determine likely clusters. The actual meaning of each cluster will have to be determined by checking the characteristics of each cluster before a class name is assigned. This is called unsupervised classification. Clustering is often used when clusters, or even the numbers of clusters, are unknown. To avoid confusion, we will

use the term 'classification' when the desired classes are defined, and the term 'clustering' when clusters are unknown or are to be formed based on features contained in data. Both classification and clustering methods are required. Classification methods will be used in land-cover and land-use classification. Clustering methods will be needed in finding spatial units or the spatial extent of land-use classes.

3.6 Summary

An object-based conceptual analysis for land-cover and land-use classification is presented. Concepts and tools for object-based modelling are briefly introduced with respect to applications in urban planning and management. Classes, structures, attributes and operations are proposed, which will be further developed and implemented in the following chapters. The image-object data model and image-object fundamentals are presented in Chapter 4. The logical design for implementation is provided in Chapter 5.

Chapter 4

The image-object fundamentals

4.1 Introduction

The choice of a conceptual model determines how information can later be derived (Burrough and McDonnell, 1998). Currently, such a choice has to be made between the vector data model and the raster data model. Since remote sensing images and field data are mainly represented in a regular grid space, image objects are naturally defined based on cells with a regular shape in a discrete 2D space which shares the simplicity of the raster model. To derive meaningful image objects and classify them according to their physical properties, geometric forms and spatial relations, an object-based data model is proposed in this research that uses an image object as a container to check the semantic, thematic and geometric components of a geospatial feature. It will take advantage of both the vector data model and the raster data model. To support hierarchical object-based image analysis and land-cover and land-use classification, a formalism is necessary that can be regarded as an extension of current raster data models. Many concepts used in this formalism are then taken from the vector data model, such as topological relationships (Egenhofer, 1989; Egenhofer and Franzosa, 1991; Egenhofer and Herring, 1991; Molenaar, 1998). By using the concept of object, the dispute between the vector model (which represents objects in $\mathbb{R}^2$) and the raster model (which represents objects in $\mathbb{Z}^2$) can be eliminated to a certain degree. In this chapter, we introduce a raster data model called a hybrid-raster data model, developed to represent image objects. We expect that the proposed object-based model with a raster-based representation will enable smooth and consistent transformation from the object-based conceptual model to the hybrid-raster data model, as well as from the data model to the logical model (Array-table) for implementation. The next section explains the hybrid-raster data model.
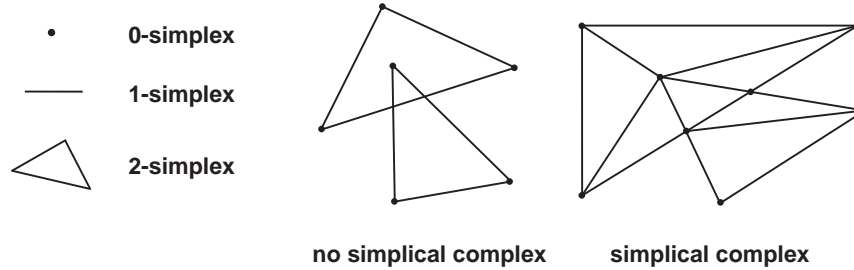
Figure 4.1: Examples of simplices and a simplicial complex in the vector model.

4.2 Elementary objects represented in a raster

In the two-dimensional (2D) discrete space $\mathbb{Z}^2$, pixels are the elementary objects in a raster model. There are three elementary objects in the hybrid-raster model: cell, edge and node (Winter, 1995). In a vector data model, points can be considered as 0-simplices, straight line segments as 1-simplices and triangles as 2-simplices in a 2D space $\mathbb{R}^2$, as shown in Figure 4.1 (Molenaar, 1998). Similarly cells can be treated as 2-simplices, edges as 1-simplices and nodes as 0-simplices in a 2D discrete space $\mathbb{Z}^2$, as shown in Figure 4.2.

The representation of objects in a raster can best be done in a cell raster. Because the cells represent area segments, this geometry is most suitable for the representation of area objects. Rasters are less suitable for representing point objects or line objects. It is possible to indicate in which cell a point object falls and with which cell a line object intersects, but this has to be approximated with an accuracy that is determined by the resolution of the raster (Molenaar, 1998). Despite its weakness in representing point objects or line objects – using high-resolution data can partially compensate for this – the raster data model is still considered a model well suited to land-cover and land-use classification based on multi-spectral images and laser data. There are two reasons for this: both input data, such as images and grided laser data, and output data, such as image objects as presentations of extracted land-cover and land-use objects, can be processed based on the same data model and the extension to the hybrid-raster data model based on the cellular decomposition (Kovalevski, 1989; Winter, 1995; Winter and Frank, 1999, 2000).

The coordinate system is defined by row (i) and column (j). According to tradition, in image processing, rows are counted in from top to bottom and columns are counted from left to right. For graphic presentation or the simultaneous display of raster and vector, a transformation can be applied by flipping the image upside down and shifting the original point to an appropriate location if necessary.
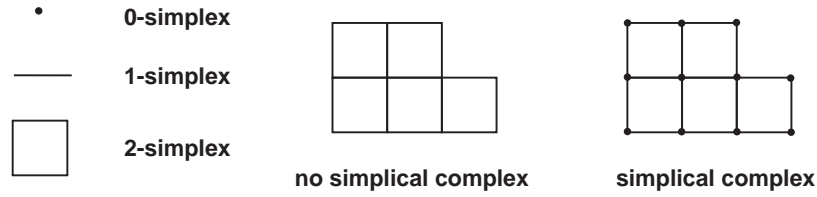
Figure 4.2: Examples of simplices and a simplicial complex in the raster model.
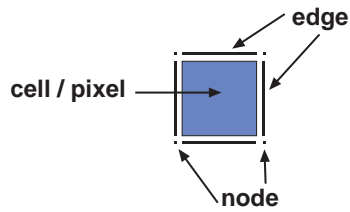
Figure 4.3: A fundamental image object defined in the hybrid-raster data model (cell or pixel: interior; edges and nodes: boundaries).
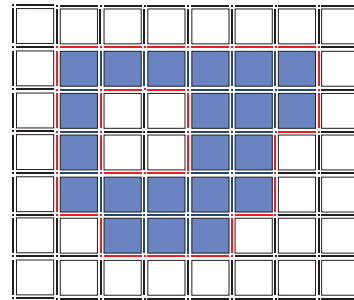
Figure 4.4: A general form image object defined in the hybrid-raster data model (blue: interior; red: boundaries).

4.2.1   The hybrid-raster data model

The hybrid-raster data model is applied in this research to represent objects in a raster-based model for the convenience of direct use of multi-spectral images and rasterised laser data, as well as to enable explicit representation of the topological relationships between two objects in a raster model. In the hybrid-raster data model, a pixel is equivalent to a cell and a cell is an open set. So a pixel or a cell can be treated as the interior of a closed set (a fundamental image object) bounded by its four edges and four nodes, as shown in Figure 4.3. A more general form of an image object, which consists of several pixels and may have a hole inside, can be found in Figure 4.4. Obviously an image object is a simplicial complex in the hybrid-raster data model.

To ensure a partition in a raster as a bounded 2D discrete space, we make the following definitions. A *Cell* is an open set bounded by its four edges and four nodes. An *Edge* is an open set bounded by two cells and two nodes. A *Node* is an open

set bounded by four cells and four edges. Therefore the bounded universe $\mathbb{U}$ of an image in a 2D discrete space $\mathbb{Z}^2$ is the collection of a cell set, an edge set and a node set, and an image object is a subset of $\mathbb{U}$ (Winter, 1995; Egenhofer and Franzosa, 1991; Kainz et al., 1993; Molenaar, 1998; Winter and Frank, 1999). Consequently, topological relationships between two image objects can be identified explicitly.

$$\mathbb{U} = \{Cell, Edge, Node\}$$

$$Cell = \{cell_i \mid cell_i \in \mathbb{U}\}$$

$$Edge = \{edge_i \mid edge_i \in \mathbb{U}\}$$

$$Node = \{node_i \mid node_i \in \mathbb{U}\}$$

4.2.2 Image objects

The geometry of an instance of an image object is represented in a raster by a contiguous region consisting of one or more adjacent cells, as well as edges and nodes lying between these cells.

An image object (O or O_I) can be defined as:

$$O = \{(\mathbf{x}, f(\mathbf{x})) \mid \mathbf{x} \in \mathbb{U}^2, f(\mathbf{x}) \in \{0, 1\}\}$$

where $\mathbf{x}$ is a location vector (a list of pixel locations or location index) and f is a function that defines whether the object is present (1) or absent (0) at a given location for crisp objects. In the case of fuzzy spatial objects, f is determined based on a degree of presence, with a membership value $\mu(\mathbf{x})$ taken from a continuous range between 0 and 1. Its geometry, such as spatial extent, size, shape, orientation, is taken based on a defined threshold, say 0.5, to transform an object field (fuzzy image object) into a crisp image object.

$$f(\mathbf{x}) = \begin{cases} 1, & \mu(\mathbf{x}) \geq Threshold, \ \mu(\mathbf{x}) \in [0, 1] \\ 0, & \mu(\mathbf{x}) < Threshold, \ \mu(\mathbf{x}) \in [0, 1] \end{cases}$$

In practice, if a $cell_{i,j}$ is part of an image object O_k, this will be represented by (Molenaar, 1998):

$$Part_{22}[cell_{i,j}, \ O_k] = 1$$

or we can simply use following expression:

$$p_{i,j}^k = 1$$

to indicate that pixel $p_{i,j}$ is part of an image object O_k.

If an image segment S_i or an image region is part of an image object O, this will be represented by:

$$Part_{22}[S_i,\ O] = 1$$

An image segment S_i is an instance of a potential image object in a binary image derived according to certain attributes of this image segment, which are associated with the definition of the object type. A number of segments may be suitable for presenting an object. In practice, there may be more than one segment that can potentially belong to one object class. They have to be reasoned based on their properties and within a certain application context. For instance, segments have to meet two conditions to qualify as trees: they are 1 or 2 m higher than the surrounding area and obtain higher NDVI values. Only if their size is larger than a defined threshold, can they be further identified as forest.

In the object-based model, an image or a field can be regarded as a partition in a raster in a bounded 2D discrete space $\mathbb{U}$. Considering the implementation in a raster model, an image object (O_k) can be defined as:

$$O_k = \{p_{i,j} \mid \forall\, p \in O_k \overset{s}{=} 4 - \text{connection},\ \ p_{i,j} \in O_k,\ O_k \in \mathbb{U}\}$$

where $p_{i,j}$ is a pixel with location index: ith row and jth column; O_k is an image object; $\mathbb{U}$ is the universe of a given image; $O_k \overset{s}{=} 4 - \text{connection}$ denotes adjacent connection between pixels by at least one edge. Pixels connected only by node(s) in the 8-connection case are not considered as elements of the same object.

In the hybrid-raster data model, an image or a field can be regarded as a partition by the three subsets $Cell$ (pixel), $Edge$ and $Node$ in a raster in a bounded 2D discrete space $\mathbb{U}$. Considering the implementation of the hybrid-raster model for identifying topological relationships between objects, an image-object (O_k) can be defined as:

$$O_k = \{Cell_k, Edge_k, Node_k\}$$

This implies:

$$Cell_k = \{cell_{i,j} \mid cell_{i,j} \in O_k\}$$

$$Edge_k = \{edge_{i,j} \mid edge_{i,j} \in O_k\}$$
$$Node_k = \{node_{i,j} \mid node_{i,j} \in O_k\}$$

$$Cell_k \cup_s Edge_k \cup_s Node_k \overset{s}{=} O_k,\ O_k \in \mathbb{U}\}$$

where $cell_{i,j}$, $edge_{i,j}$ and $node_{i,j}$ are a pixel, an edge and a node, respectively, with location index: ith row and jth column; $Cell_k$, $Edge_k$ and $Node_k$ are a collection of cells, edges and nodes, respectively, in O_k; O_k is an image object; $\mathbb{U}$ is the universe of an given image; $Cell_k \cup_s Edge_k \cup_s Node_k \overset{s}{=} O_k$ denotes that $Cell_k$, $Edge_k$, $Node_k$ are the spatial partition of O_k (spatially connected).

4.3 Fundamental components of an image object

For reasoning topological relations between two image objects, it is necessary to explicitly define the interior (O^o), boundary (∂O) and exterior (O^-) as fundamental components of an image object. Since each object consists of cells, edges and nodes in the hybrid-raster data model, these elementary components will have to be determined explicitly as to the fundamental object component they belong to: interior, boundary or exterior. In the hybrid-raster data model, an image object (O) can be defined by cells, edges and nodes. For a region (image object) in a binary image, the interior (O^o), boundary (∂O) and exterior (O^-) of an image object can be defined based on configurations of surrounding pixels (p) in a 2D binary image.

4.3.1 The interior of an image object

The interior of an image object (O^o) consists of inside cells (^oCell), inside edges (^oEdge) and inside nodes(^oNode).

$$O^o = \{^oCell, {}^oEdge, {}^oNode\}$$

All pixels with value 1:

$$^oCell = \{cell_{i,j} \mid p_{i,j} = 1\}$$

All horizontal edges for which the upper and lower pixels have the value 1:

$$^oEdge = \{edge_{i,j;i+1,j} \mid p_{i,j} = 1 \text{ and } p_{i+1,j} = 1\}$$

All vertical edges for which the left and right pixels have the value 1:

$$^oEdge = \{edge_{i,j;i,j+1} \mid p_{i,j} = 1 \text{ and } p_{i,j+1} = 1\}$$

All nodes for which the four surrounding pixels all have the value 1 at the same time:

$$^oNode = \{node_{i,j} \mid \sum_{i}^{i+1} \sum_{j}^{j+1} p_{i,j} = 4\}$$

4.3.2 The boundary of an image object

The boundary of an image object (∂O) consists of boundary edges ($^\partial Edge$) and boundary nodes ($^\partial Node$).

$$\partial O = \{^\partial Edge, {}^\partial Node\}$$

All edges for which the pixel has the value 1 on one side and 0 on the other side.

Horizontal edges:

$$^{\partial}Edge = \{edge_{i,j;i+1,j} \mid p_{i,j} = 1 \text{ and } p_{i+1,j} = 0 \text{ or } p_{i,j} = 0 \text{ and } p_{i+1,j} = 1\}$$

Vertical edges:

$$^{\partial}Edge = \{edge_{i,j;i,j+1} \mid p_{i,j} = 1 \text{ and } p_{i,j+1} = 0 \text{ or } p_{i,j} = 0 \text{ and } p_{i,j+1} = 1\}$$

All nodes for which one or two or three pixels out of the four surrounding pixels have the value 1 and other pixels have the value 0. Nodes with four surrounding pixels all having the value 1 or all having the value 0 will not be considered as boundary nodes.

$$^{\partial}Node = \{node_{i,j} \mid \sum_{i}^{i+1} \sum_{j}^{j+1} p_{i,j} = 1 \text{ or } 2 \text{ or } 3\}$$

4.3.3 The exterior of an image object

The exterior of an image object (O^{-}) consists of outside cells ($^{-}Cell$), outside edges ($^{-}Edge$) and outside nodes ($^{-}Node$).

$$O^{-} = \{^{-}Cell, \ ^{-}Edge, \ ^{-}Node\}$$

For all pixels with the value 0:

$$^{-}Cell = \{cell_{i,j} \mid p_i = 0\}$$

For all horizontal edges with upper and lower pixels that have value the 0 at the same time:

$$^{-}Edge = \{edge_{i,j;i+1,j} \mid p_{i,j} = 0 \text{ and } p_{i+1,j} = 0\}$$

For all vertical edges with left and right pixels that have value the 0 at the same time:

$$^{-}Edge = \{edge_{i,j;i,j+1} \mid p_{i,j} = 0 \text{ and } p_{i,j+1} = 0\}$$

For all nodes with four surrounding pixels that all have value the 0 at the same time:

$$^{-}Node = \{node_{i,j} \mid \sum_{i}^{i+1} \sum_{j}^{j+1} p_{i,j} = 0\}$$

Based on the above definition, the topological relations hold by excluding the left edge and the top edge of an image, which makes it more efficient in implementation by using the schema proposed below.

4.3.4 A schema for implementation

For the extraction of explicit topological relations between two image objects derived from different sources and the efficient implementation of the hybrid-raster model, an index schema proposed by Winter et al. (Winter, 1995; Winter and Frank, 1999) can be applied here for each cell, edge and node. $cell_{i,j}$ represents a cell in i row and j column; $node_{i,j}$ represents the node in the lower-right corner (node surrounding by $cell_{i,j}$, $cell_{i,j+1}$, $cell_{i+1,j}$ and $cell_{i+1,j+1}$) of $cell_{i,j}$. $edge_{i,j;i,j+1}$ indicates the vertical edge between $cell_{i,j}$ and $cell_{i,j+1}$ and $edge_{i,j;i+1,j}$ indicates the horizonal edge between $cell_{i,j}$ and $cell_{i+1,j}$. The derived cells, edges, and nodes of an image object are illustrated in Figure 4.7. Compared with the standard raster model, the hybrid-raster data model and the index method provide powerful tools for exploring additional topological relations between two image objects in raster data or field data.

A tube structure

For efficient computation, we make the following proposal. Instead of checking neighbouring pixels using index i and i+1 or j and j+1 in a loop each time, it seems more efficient to create a four-layer tube in the third dimension of a 2D image, i.e. T(1:m, 1:n, 1:4) for I(1:m, 1:n). The first layer T(1:m, 1:n, 1) is used to store all pixels in the current image (T(1:m, 1:n, 1)=I(1:m, 1:n)). The second layer T(1:m, 1:n, 2) is for a shift image with one pixel upward (T(1:m, 1:n, 2)=I(2:m+1,1:n)). The third layer T(:, :, 3) is for a shift image with one pixel to the left (T(1:m, 1:n, 3)=I(1:m,2:n+1)). The fourth layer T(:, :, 4) is for a shift image with one pixel upward and with one pixel to the left (T(1:m, 1:n, 4)=I(2:m+1,2:n+1)). Having created such a tube means that pixel p(i, j) in T(i, j, 1) has its lower pixel p(i+1, j) in T(i, j, 2), its right pixel p(i, j+1) in T(i, j, 3), and its lower-right pixel p(i+1, j+1) in T(i, j, 4). (A graphic illustration of such a tube structure is given in Figure 4.5.) By creating the tube, the loop operation for finding neighbouring pixels is avoided, which makes it computationally more efficient at the cost of using extra memory space. This tube structure will be generated temporarily for extracting the fundamental components of image objects ($Cells$, $Edges$ and $Nodes$).

The same arrangement can be used for another tube to store cells, edges and nodes in the hybrid-raster model. The first layer of the tube H(1:m, 1:n, 1) is used for cells. The second layer H(1:m, 1:n, 2) is for horizontal edges (H-edges) on the lower side. The third layer H(1:m, 1:n, 3) is for vertical edges (V-edges) on the right side. The fourth layer H(1:m, 1:n, 4) is for nodes at the lower-right corners. This tube establishes a primary data structure for representing image objects for the hybrid-raster model. An example is shown in Figure 4.6.

An encoding schema

The following encoding schema is proposed for recording the different types of cells, edges and nodes of an image object derived from images in the implementation of the hybrid-raster model. Based on this information, topological relations between two image objects are extracted explicitly.
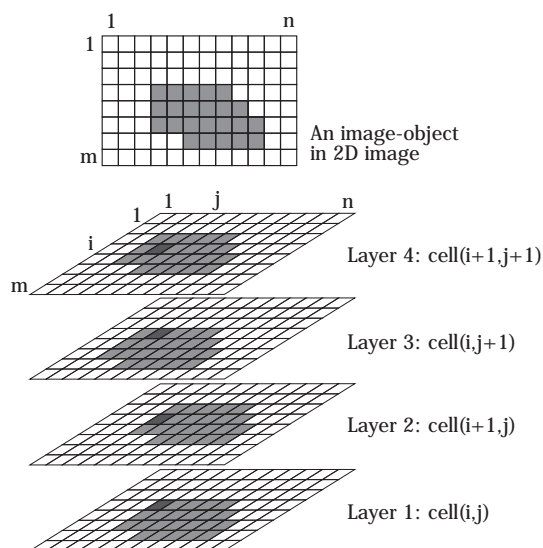
Figure 4.5: A tube structure for representing pixels and their neighbouring pixels.

$$Cell = \left\{ \begin{array}{l} 1, \; inside \; cells \\ 0, \; outside \; cells \end{array} \right.$$

$$Edge = \left\{ \begin{array}{l} 2, \; inside \; edges \\ 1, \; boundary \; edges \\ 0, \; outside \; edges \end{array} \right.$$

$$Node = \left\{ \begin{array}{l} 2, \; inside \; nodes \\ 1, \; boundary \; nodes \\ 0, \; outside \; nodes \end{array} \right.$$

4.3.5   Implementation

Working on the definitions and schema described in the previous sections, we are now ready to derive, store and present all necessary information to implement the
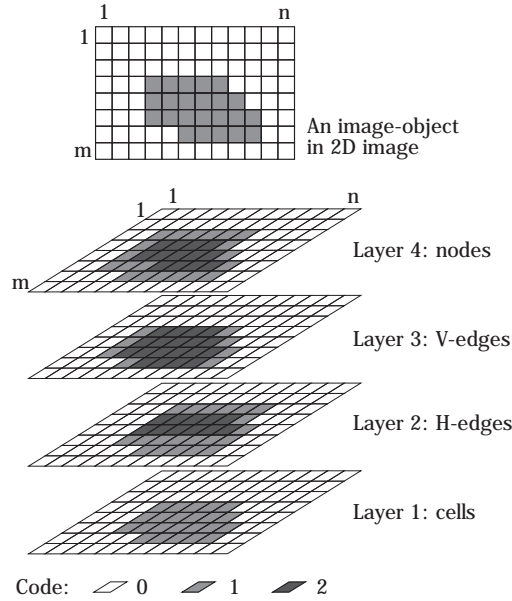
Figure 4.6: A tube structure for recording extracted cells, edges and nodes.

hybrid-raster data model based on a binary image. Examples of a derived image object without holes and an image object with holes are illustrated in A and B of Figure 4.7. Two special cases can be presented as the full region $\mathbb{U}$ and the empty region $\emptyset$, which can be defined as

$$O_k = \{\mathbb{U}\}$$

and

$$O_k = \{\emptyset\}$$

as shown in Figure 4.8 A and B respectively. The current implementation is made in Matlab.

4.4 Identification of topological relationships between image objects

With the definitions and the implementation concept described in the previous sections, we are able to determine the topological relationships between two image objects, based on the 9-intersection matrix (Egenhofer, 1989; Egenhofer and Franzosa, 1991; Egenhofer and Herring, 1991; Egenhofer, 1993; Egenhofer and Sharma, 1993; Molenaar, 1998).
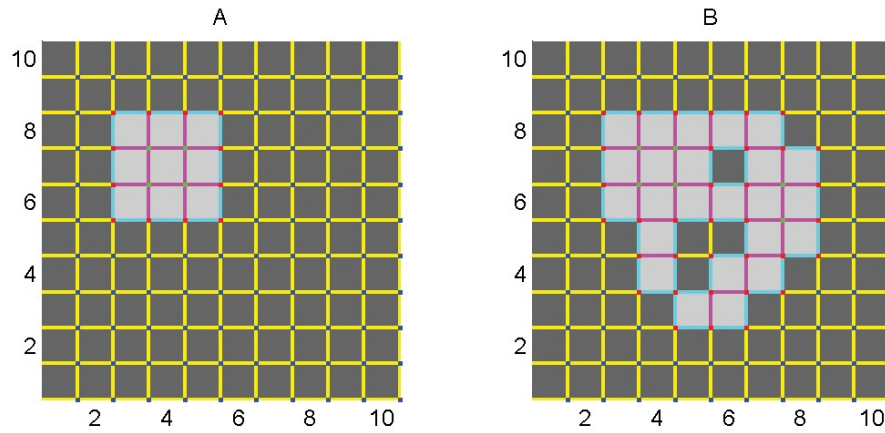
Figure 4.7: The derived cells, edges and nodes of an image object based on a binary image (A: a simple region; B: a region with holes). Light pixels represent the inside cells, dark pixels represent the outside cells; magenta lines represent inside edges, cyan lines represent boundary edges, yellow lines represent outside edges; green dots represent inside nodes, red dots represent boundary nodes, blue dots represent outside nodes of an image object.
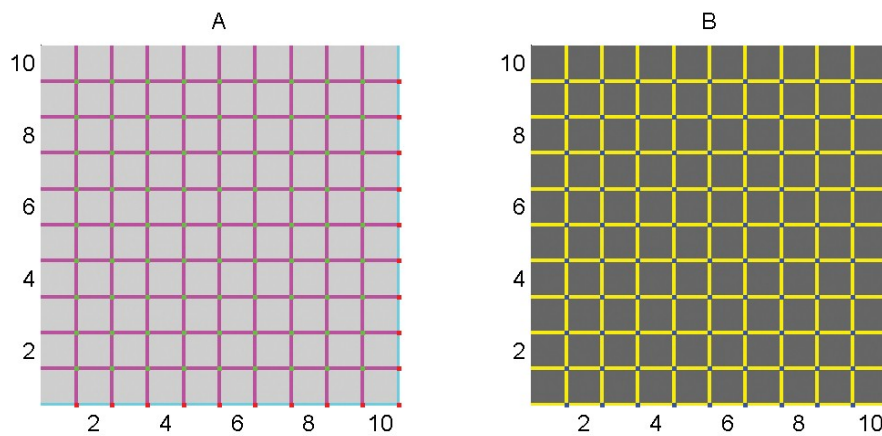
Figure 4.8: The derived faces, edges and nodes based on a region (A: full region; B: empty region).

$$R(A, B) = \begin{pmatrix} A^o \cap B^o & A^o \cap \partial B & A^o \cap B^{\text{-}} \\ \partial A \cap B^o & \partial A \cap \partial B & \partial A \cap B^{\text{-}} \\ A^{\text{-}} \cap B^o & A^{\text{-}} \cap \partial B & A^{\text{-}} \cap B^{\text{-}} \end{pmatrix}$$

The derived topological relations between two objects are consistent regardless of whether the objects have holes or not, because holes in an image region are regarded as the exterior of an object, as shown in Figure 4.7 B. To examine whether the proposed schema works well for identifying the topological relationships of two image objects, defined based on the hybrid-raster data model, we create a number of different cases and check whether the right topological relationship is identified in each case. In the following, the values in the 9-intersection matrix are presented for each type of topological relationship in order to check if the results obtained by using the given image objects match the corresponding values.

4.4.1 Disjoint

Two objects are disjoint if

$$R_{Disjoint}(A, B) = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 1 \\ 1 & 1 & 1 \end{pmatrix}$$

An example of 'disjoint' objects is shown in Figure 4.9, where two objects are spatially separate. Another example of 'disjoint' objects is shown in Figure 4.10, where object A is inside a hole of object B but remains 'disjoint'. In both cases, the same values in the 9-intersection matrix are obtained, which indicate the 'disjoint' topological relationship.

4.4.2 Equal

Two objects are equal if

$$R_{Equal}(A, B) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Two examples of 'equal' objects are shown in Figures 4.11 (simple objects) and 4.12 (objects with holes). In both cases, the same values in the 9-intersection matrix are obtained, which indicate the 'equal' topological relationship.

4.4.3 Contain

Object A contains object B if
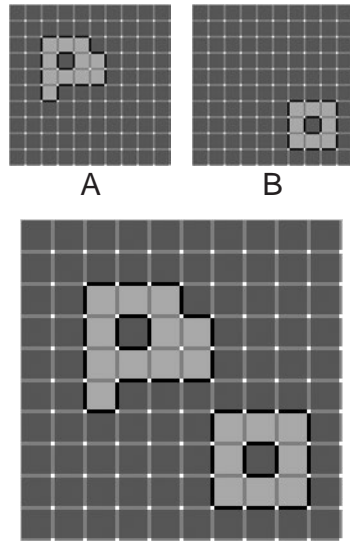
Figure 4.9: A 'disjoint' B (A and B are spatially separate).
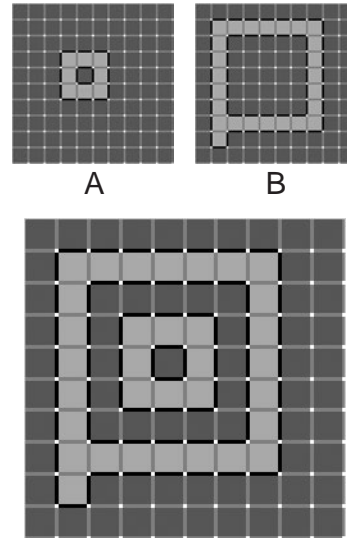
Figure 4.10: A 'disjoint' B (A in a hole of B).
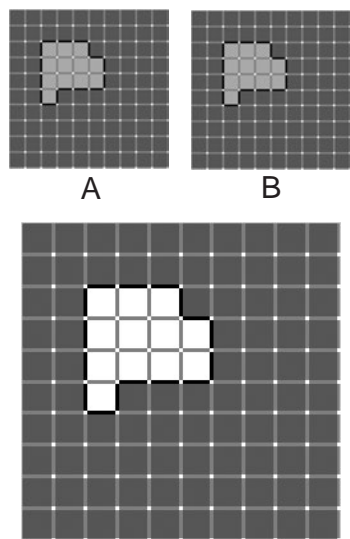
Figure 4.11: A 'equals' B (without a hole).
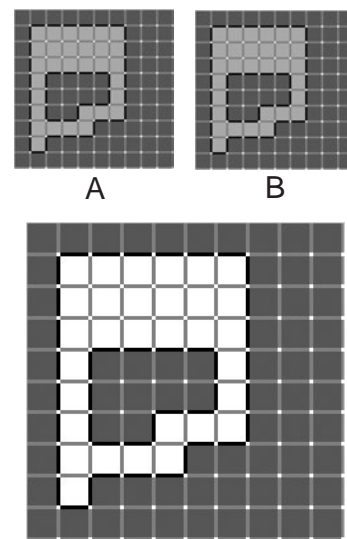
Figure 4.12: A 'equals' B (with a hole).
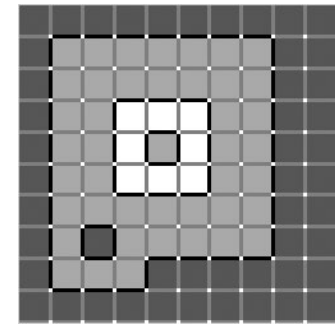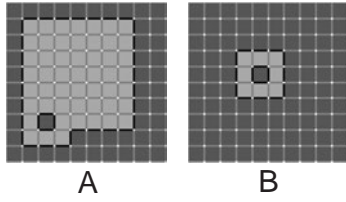
Figure 4.13: A 'contains' B.   Figure 4.14: A 'contained by' B.
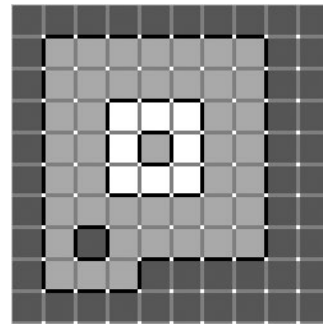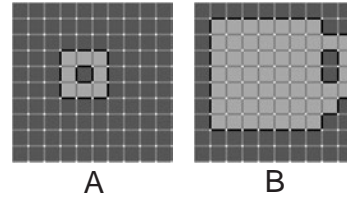
$$R_{Contain}(A, B) = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{pmatrix}$$

An example of 'contain' objects is shown in Figure 4.13. In this case, the same values in the 9-intersection matrix are obtained, which indicate the 'contain' topological relationship.

4.4.4 Contained by

Object A is contained by object B if

$$R_{ContainedBy}(A, B) = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \end{pmatrix}$$

An example of 'contained by' objects is shown in Figure 4.14. In this case, the same values in the 9-intersection matrix are obtained, which indicate the 'contain by' topological relationship.

4.4.5 Meet

Two objects are meet if

$$R_{Meet}(A,B) = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$$

There are two situations that can be equally treated as 'meet': the connection is either by touched boundary edge(s) or by touched boundary node(s). An example of two objects 'meet' by a touched boundary edge is shown in Figure 4.15. An example of two objects that 'meet' by a touched boundary node is shown in Figure 4.16. Figure 4.17 shows two objects that 'meet' from inside by a touched boundary edge. Figure 4.18 shows two objects that 'meet' from inside by a touched boundary node. In these four cases, the same values in the 9-intersection matrix are obtained, which indicate the 'meet' topological relationship.

4.4.6 Cover

Object A covers object B if

$$R_{Cover}(A,B) = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$$

There are two situations that can be equally treated as 'cover': the connection is either by boundary edge(s) or by boundary node(s). An example where object A covers object B, with a connection by boundary edge(s), is shown in Figure 4.19. Figure 4.20 shows object A covering object B, with a connection by boundary node(s). In both cases, the same values in the 9-intersection matrix are obtained, which indicate the 'cover' topological relationship.

4.4.7 Covered by

Object A is covered by object B if

$$R_{CoveredBy}(A,B) = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix}$$

There are two situations that can be equally treated as 'covered by': the connection is by boundary edge(s) or by boundary node(s). Figure 4.21 shows that object A is covered by object B and connected by a boundary edge(s). Figure 4.22 shows object A covered by object B and connected by a boundary node(s). In both cases, the same values in the 9-intersection matrix are obtained, which indicate the 'covered by' topological relationship.
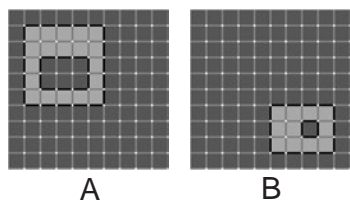
Figure 4.15: A 'meets' B by edge.
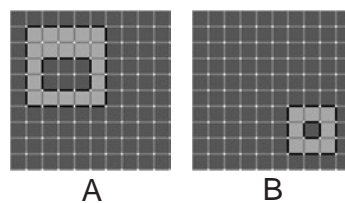
Figure 4.16: A 'meets' B by node.
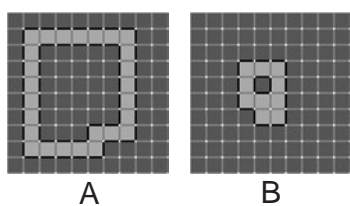
Figure 4.17: A 'meets' B by edge from inside.
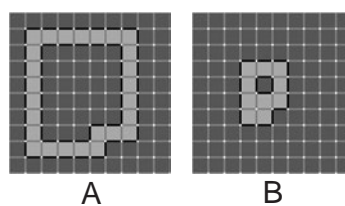
Figure 4.18: A 'meets' B by node from inside.
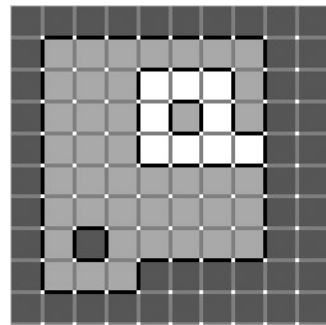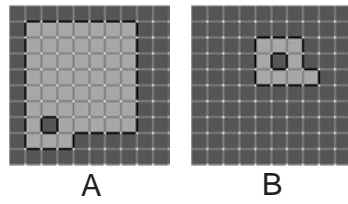
Figure 4.19: A 'covers' B with connection of boundary edge.
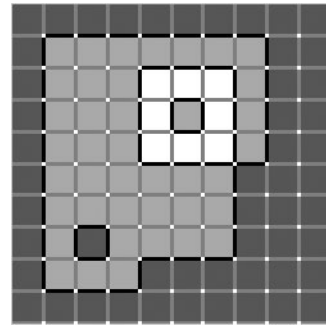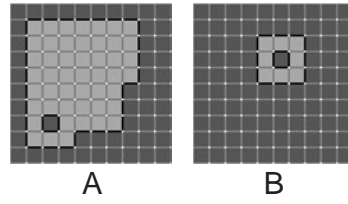
Figure 4.20: A 'covers' B with connection of boundary node.
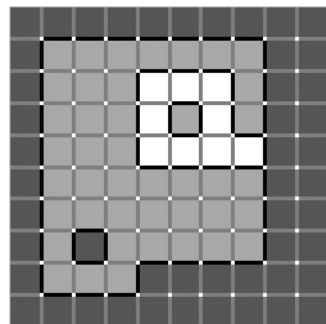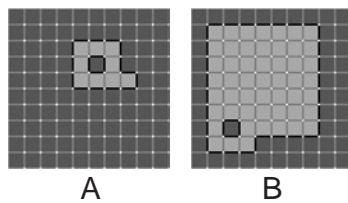
Figure 4.21: A 'covered by' B with connection of boundary edge.
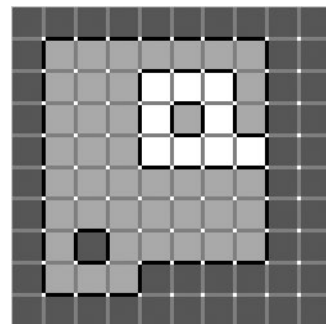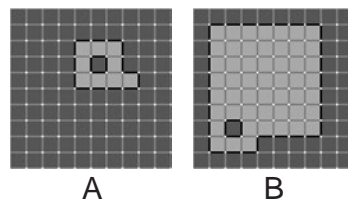
Figure 4.22: A 'covered by' B with connection of boundary node.
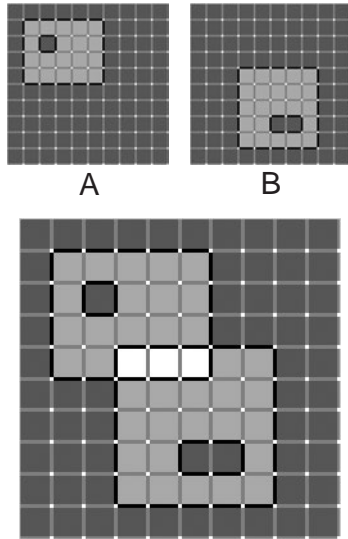
Figure 4.23: A 'overlaps' B with boundary nodes.
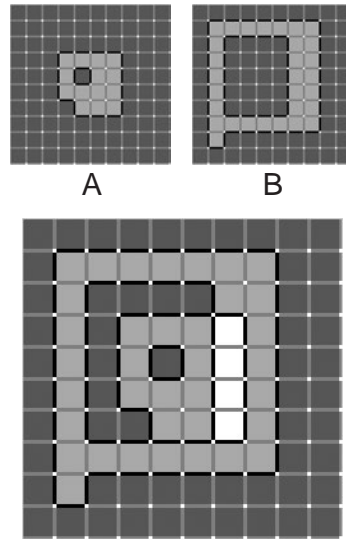
Figure 4.24: A 'overlaps' B with boundary edges and nodes.

4.4.8 Overlap

Two objects are overlapping if

$$
R_{Overlap}(A, B) = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}
$$

There are two situations that can be equally treated as 'overlap': (a) overlapping cell(s) and boundary node(s), or (b) overlapping cell(s), boundary edge(s) and boundary node(s). An example of overlapping objects with overlapping cell(s) and boundary node(s) is shown in Figure 4.23. An example of overlapping objects with overlapping cell(s), boundary edge(s) and boundary node(s) is shown in Figure 4.24. In both cases, the same values in the 9-intersection matrix are obtained, which indicate the 'overlay' topological relationship.

4.5 Geometric properties of an image object

The geometry of a planar region comprises the following aspects: size, position, orientation and shape. Many of these aspects are covered by a family of parameters called *moments*. In probability theory, moments are used to characterise probability density functions, e.g. expectation (first-order moment), variance, covariance

(second-order central moments) (van der Heijden, 1994). In our 2D image case, we use the same definitions but replace the density function with a binary function,

$$p_{i,j}^k = \left\{ \begin{array}{ll} 1, & p_{i,j}^k \in O \\ 0, & p_{i,j}^k \notin O \end{array} \right.$$

in a 2D image I(1:M,1:N). The moments of order $p + q$ of a region represented by the bitmap $p_{i,j}$ are:

$$M_{p,q} = \sum_{i=1}^{M} \sum_{j=1}^{N} i^p j^q p_{i,j} = \sum_{p_{i,j} \in O_k} i^p j^q$$

Size of an image object

The size of an image object is the total number of pixels that belonging to this image object ($M_{0,0}$) in units of pixel area (Δ^2).

$$Size(O_k) = M_{0,0} = \sum p_{i,j}^k \tag{4.1}$$

Location of an image object

The first-order moments $M_{1,0}$ and $M_{0,1}$ are related to the balance point ($\overline{x}$, $\overline{y}$) of the region. The location of an image object can be represented by the balance point ($\overline{x}, \overline{y}$).

$$\overline{x} = M_{1,0}/M_{0,0} \quad \text{and} \quad \overline{y} = M_{0,1}/M_{0,0} \tag{4.2}$$

This point (x,y) is also called the *centre of gravity*, or *centroid* (centre of mass), and can be used to determine the position or location of the region given in units of pixel period (Δ).

$$Location(O) = (\overline{x}, \overline{y}) \tag{4.3}$$

In order to make the description independent of position, moments can be calculated with respect to the centroid. The results are the so-called *central moments*:

$$\mu_{p,q} = \sum_{i=1}^{M} \sum_{j=1}^{N} (i - \overline{x})^p (j - \overline{y})^q p_{i,j} = \sum_{p_{i,j} \in O_k} (i - \overline{x})^p (j - \overline{y})^q \tag{4.4}$$

If the ordinary moments are known, it is less expensive to derive the central moments from the ordinary moments than to evaluate expression (4.4) directly (van der Heijden, 1994).

For instance:

$\mu_{0,0} = M_{0,0}$
$\mu_{0,1} = \mu_{1,0} = 0$
$\mu_{0,2} = \mu_{2,0} - \overline{x} M_{1,0}$
$\mu_{1,1} = M_{1,1} - \overline{x} M_{0,1}$
... ...

The second-order central moments exhibit a number of properties that are comparable to the covariance matrices in probability theory and the moments of inertia associated with rotating bodies in mechanics. The principal axes of a region are spanned by the eigenvectors of the matrix (van der Heijden, 1994):

$$\begin{bmatrix} \mu_{2,0} & \mu_{1,1} \\ \mu_{1,1} & \mu_{0,2} \end{bmatrix}$$

The principal moments are the corresponding eigenvectors.

$$\lambda_{max} = \frac{1}{2}(\mu_{2,0} + \mu_{0,2}) + \frac{1}{2}\sqrt{\mu_{2,0}^2 + \mu_{0,2}^2 - 2\,\mu_{0,2}\,\mu_{2,0} + 4\,\mu_{1,1}^2}$$

$$\lambda_{min} = \frac{1}{2}(\mu_{2,0} + \mu_{0,2}) - \frac{1}{2}\sqrt{\mu_{2,0}^2 + \mu_{0,2}^2 - 2\,\mu_{0,2}\,\mu_{2,0} + 4\,\mu_{1,1}^2}$$

The direction of the largest principal moment is:

$$\theta = tan^{-1}(\sqrt{\frac{\lambda_{max} - \mu_{2,0}}{\mu_{1,1}}})$$

Orientation

The direction of the largest principal moment (θ) is often used to specify the orientation of a region. When orientation (θ) is used as a feature for comparison, we have to keep in mind that a disk-like region ($\lambda_{max}/\lambda_{min} \approx 1$) does not make sense for quantitative comparison based on θ. Therefore the ratio ($\lambda_{max}/\lambda_{min}$) should be consulted before making orientation matching between objects. However, the ratio can be used as an uncertainty measurement for matching.

Shape of an image object

Exact geometric shape descriptions of an image object are difficult to derive because of the existence of vague boundaries and uncertainty in an image in many cases. On the other hand, it is not necessary to use exact descriptions since it will be difficult to make a robust comparison between objects according to exact shape matching anyway. However, many shape descriptors can be derived based on the statistics of the spatial distribution of the pixels that form an image object. These descriptors are considered robust for matching similar shapes and can be used as indicators for similarity comparison of object shape. Below we describe a number of such indicators that can be directly derived from image regions. Many of them are provided by the Image Processing Toolbox, Matlab (The MathWorks Inc., 2001).

- Length of the major axis and the minor axis
  The principal moments λ_{max} and λ_{min} can be used to represent the lengths (in pixels) of the major axis and the minor axis of the ellipse that has the same second moments as the region.

- Eccentricity
  The eccentricity of a region can be defined as the ratio between the square roots of the two principal moments:

$$\text{Eccentricity} = \sqrt{\frac{\lambda_{\max}}{\lambda_{\min}}}$$

- Convex area
  The number of pixels in the convex hull, the smallest convex polygon that can contain the region.

- Solidity
  The proportion of the pixels in the convex hull that are also in the region. Computed as:

$$\text{Solidity} = \frac{\text{Area}}{\text{Convex area}}$$

- EquivDiameter
  The diameter of a circle with the same area as the region. Computed as:

$$\text{EquivDiameter} = \sqrt{\frac{4 * \text{Area}}{\pi}}$$

- Hole size
  Hole size can be computed as:

$$\text{Hole size} = \text{Filled area} - \text{Area}$$

- Hole ratio
  Hole ratio can be computed as:

$$\text{Hole ratio} = \frac{\text{Filled area} - \text{Area}}{\text{Area}}$$

4.6   Thematic attributes of an image object

The thematic attributes of an image object can be derived directly from images and other field data, usually by taking the average value of all pixels belonging to this image object.

For instance, multi-spectral values of an image object:

$$\text{Band } 1(O_k) = \frac{\sum_{i=1}^{M} \sum_{j=1}^{N} p_{i,j} \cdot \textbf{Band 1}}{\sum_{i=1}^{M} \sum_{j=1}^{N} p_{i,j}}, \ p_{i,j} \in O_k$$

$$\text{Band } 2(O_k) = \frac{\sum_{i=1}^{M} \sum_{j=1}^{N} p_{i,j} \cdot \textbf{Band 2}}{\sum_{i=1}^{M} \sum_{j=1}^{N} p_{i,j}}, \ p_{i,j} \in O_k$$

$$\ldots\ldots$$

The same principle can be applied with other data, such as:

$$\text{DSM } 1(O_k) = \frac{\sum_{i=1}^{M} \sum_{j=1}^{N} p_{i,j} \cdot \textbf{DSM 1}}{\sum_{i=1}^{M} \sum_{j=1}^{N} p_{i,j}}, \ p_{i,j} \in O_k$$

$$\text{NDVI}(O_k) = \frac{\sum_{i=1}^{M} \sum_{j=1}^{N} p_{i,j} \cdot \textbf{NDVI}}{\sum_{i=1}^{M} \sum_{j=1}^{N} p_{i,j}}, \ p_{i,j} \in O_k$$

4.7 Semantic component of an image object

The semantic component of an image object includes features used to extract image objects, methods and parameters applied for obtaining these objects, as well as quality and uncertainty assessment figures.

This type of information is very important to users, in that it allows users to decide whether extracted image objects meet their requirements, and what kinds of changes may be needed in order to acquire image objects that meet their specific needs and satisfy quality aspects.

If this type of information is made transparent to users, problems that arise when users from different disciplines interpret the same scene differently can largely be avoided. Users are allowed to choose different models or modify related parameters in order to obtain the image objects they desire. Since the whole system is intended to be semi-automatic, any user can specify his/her requirements by choosing the desired features and models, as well as parameters, to produce the desired results based on one fundamental data set. Such a set-up provides the possibility for different users to share the relatively high costs of acquiring high-resolution data and to produce the data they expect.

4.8 Summary

A model for image objects is proposed in this chapter. The hybrid-raster data model based on a regular cellular decomposition of 2D space is applied in the research as an extension of the simple raster data model in order to derive explicit topological relations between two image objects. This data model has been implemented in Matlab, thus enabling the extraction of topological relationships between image objects derived from image and laser data. The topological relations provide useful information for reasoning in urban land-use classification. It is impossible to derive such information by per-pixel approaches. The presented considerations and implementation schemes focus on topological relationships between image regions. The relationships between image regions are sufficient for the current research since both land cover and land use are represented as image regions. Topological relationships for linear features and point features are not included in this research. This chapter has also introduced measures for the geometric properties and attributes of an image object. These per-object properties provide useful information, which can only be derived by the object-based approach. Object-based measures are considered robust to noise and other high-frequency signals existing in high-resolution data, and hence can reduce complexity in classification compared with per-pixel approaches.

Chapter 5

Logical design for object-based land-cover and land-use classification*

5.1 Introduction

Urban land-cover classification and land-use classification are processes of partitioning urban space into discrete spatial units that hold certain physical features or certain functions. Man-made features are easily conceived of as determinate objects because they have discernible boundaries. The physical evidence of the building materials used and the simple constructs, as compared with natural features such as trees, imply very little fuzziness in determining the boundaries of man-made features. Natural features share with man-made features the property of an easily defined physical interior at the abstraction level of reality we are interested in. The boundary of a natural river, however, gives far more rise to fuzziness than the boundary of a paved road. If the boundary of a feature is indiscernible or, although discernible, is not the boundary of interest (e.g. the boundary of leaves, branches, or trunk of a tree), we may refer to such a feature as an indeterminate object. In landcover classification we attempt to delineate determinate and indeterminate objects, such as buildings, water surfaces, green space. They have in common the fact that they are physical features of urban space, a fact we fairly successfully make use of in semi-automatically producing land-cover maps from remote sensing data.

Land-use classification, however, aims at delineating regions that have a func-

---

tional meaning. The boundary of a residential area is indiscernible. The object 'residential area' is indeterminate on solely physical evidence as provided by remote sensing data. Only if we can relate land-use classes to physical appearance – containing certain types of features in certain constellations – can we have a chance of delineating such indeterminate objects from remote sensing data. 'Delineate' implies the determination of the boundaries with a quantifiable degree of certainty. We aim at determining the boundaries of land-use classes based on probability surfaces derived from identified feature constellations. The indicant features and their spatial distribution will be reasoned on and extracted, based on high-resolution multi-spectral and laser data. The probability surfaces of land-use objects must be modelled based on available data and their thematic associations with their definitions concerning the certain functions they fulfil. The intersection of the probability surfaces of neighbouring objects will need to be modelled and extracted as the boundary of objects. Therefore, we have to collect detailed information about land-cover features and their properties before we can classify land-use. We expect the combination of high-resolution multi-spectral image and laser data to be able to provide such detailed information and allow us to infer both land cover and land use.

In Chapter 2 we have shown that we cannot expect pixel-based classification approaches to yield land-use identification. In Chapter 3 we have presented a conceptual framework for object-based image analysis. In Chapter 4 we have presented the concept of the image object and its representation based on the hybrid-raster data model. We shall now elaborate on the logical design of how to extract image objects and how to accomplish land-cover and land-use classification in urban areas. We shall lay down the methods of dealing with image objects. The mapping from field space to object space is done in three steps: feature extraction toward land-cover classification; finding spatial units for land-use classes based on land-cover features; land-use identification for each spatial unit. The overall scheme is described in Figures 1.2 and 2.25. A number of issues will be discussed in the following sections of this chapter, such as the spatial representation of image objects, the hierarchy of image objects, image object extraction at the land-cover level, finding spatial units for land-use types, extracting image-object properties at the land-use level, and classifying land-use objects.

5.2   Spatial representations of image objects

5.2.1   Homogeneity and semantic description of image regions

Homogeneity

Image objects are represented by image regions. Image regions are usually obtained by segmentation. Most of the existing segmentation algorithms developed in image processing, computer vision and other related communities are based on homogeneity measures and certain constraints based on digital numbers (DN) in an image. They are based on the assumption that no a *priori* knowledge is available about

objects in the scene (Ballard and Brown, 1982; Beaulieu and Goldberg, 1989). Segmented regions should be uniform and homogenous with respect to some characteristic such as gray tone or texture. Region interiors should be simple and without many holes. Adjacent regions of segmentation should have significantly different values with respect to their homogeneity characteristics. Boundaries of each segment should be simple, not ragged, and must be spatially accurate (Haralick and Shapiro, 1985). Image segmentation often constitutes the low-level processing stage of an image analysis system, while the high-level stage is then devoted to the interpretation of these regions (Ballard and Brown, 1982; Beaulieu and Goldberg, 1989). In the past most researchers concentrated on the stage of low-level processing of images. The 'region-growing' technique, the 'split-merge' approach, as well as the adaptive threshold or optimisation approach, have been applied in many areas of image segmentation and analysis. Many good applications can been found in literature. But the fact is that one homogeneity measure that may work well for certain features may not be good for other features. A more comprehensive homogeneity measure may be able to cope with some more features, but it may increase vagueness in what should be included in the segments. Again it cannot always be suitable for all features. On the other hand, filtering and kernel-based approaches are sensitive to the size of the window. Therefore in this research we propose to initiate image segmentation at the high-level stage, based on meaningful features, and apply the low-level image analysis techniques at a later stage to refine these regions.

Semantic description of image regions

The DNs of an image are recorded according to the electromagnetic reflection or emission from locations on the earth. Despite bias or errors made during data collection, due to atmospheric impacts, limited accuracy of equipment involved etc., these DNs often have meaningful associations with the physical environment. For instance, in the infrared band, the DN values are associated with the amount of vegetation and its growth status. Therefore, knowledge of physical features can play a very important role in image segmentation. Such knowledge can be derived based on the semantic description of desired features and by the quantitative measurement of the desired features and corresponding DN values based on samples, as we did by using the MLC. Knowledge of both definitions and their associations with features that can be extracted from various remote sensing data is used in this research for designing an image segmentation schema and for image region refinement by using fuzzy membership functions to represent such associations. The basic idea is to identify ranges of DN values that can be associated with certain land-cover classes. For instance, vegetation pixels have higher NDVI (normalised difference vegetation index) values, while most built-up and water pixels have lower NDVI values. A number of techniques can be applied to identify membership functions between DN values and classes, such as fuzzy logic, probability theory, rule-based reasoning. The fuzzy membership functions obtained by semantic analysis can be implemented at the pixel level, as well as the land-cover level and the land-use level, when we regard each image object as a unit like a pixel.

Some semantic descriptions are, however, directly related to image regions or the

geometric properties of image regions. For instance, one characteristic of a region belonging to the building class is that buildings should have vertical walls. This is a characteristic that refers to the geometric properties of an object. In this case, we need to look at LIDAR data for solutions. We may find a solution to extracting this 3D feature based on LIDAR data (2D data containing height information). It can be measured by comparing the sizes of image regions that are obtained by a 'vertical segmentation' of the DSM. The size differences between the image regions of a building at several elevation slices should be small with regard to the vertical wall characteristic, as shown in Figures 6.1 and 6.2. A detailed description of the vertical segmentation and its utilisation can be found in Zhan et al. (2002b) and in Section 6.2. This type of information cannot be derived by per-pixel approaches because the height value for each pixel in LIDAR data can vary. Thus it is difficult to form regions according to these height values alone, for example, we can extract edges by edge detection algorithms. To detect a region based on these edges is often problematic owing to the lack of evidence as to how these edges are interrelated. Therefore, the object-based approach can play a better role than pixel-based approaches in region-based feature extraction.

5.2.2   Extraction of image objects

In current geo-information systems, a spatial entity can be described by the geometric properties and thematic properties that are linked with it as attributes. Two commonly used data models in geo-information communities and related fields are vector and raster. Using a vector data model, the geometric boundaries have to be measured first, often manually on the ground or in images, according to the individual's understanding of the requirements. A unique identifier and a class name and attributes will be attached to this spatial entity later on. The raster data model plays an important role in remote sensing image processing. We look at the thematic information contained in each pixel first, and then assign class labels to each pixel according to likelihood value in the case of maximum likelihood classification. Geometric boundaries will be determined in a post-processing stage if necessary. In general, a geo-spatial object has geometric and thematic components (Molenaar, 1998). However, in the proposed image-object approach, we try to include semantic components in addition to geometric and thematic components in image objects. Semantic components provide information as to how image objects are formed by indicating the characteristics used in determining these objects, the membership values corresponding to these characteristics, and uncertainty assessment, etc. We start with semantic analysis based on the semantic components of class definition and features that can be derived from images. In some cases, ancillary data such as population, infrastructure and cadastre may be useful, especially for land-use classification. In this research, we try to use features that can be derived from images or laser data without relying on additional ancillary data. Quite often these ancillary data are not available or there may be changes taking place between the acquisition of images and the acquisition of ancillary data. A logical design can be carried out by using characteristics that meet the requirement of classes in the semantic description and can be derived from images. Image objects are determined based on

GIS/Vector RS/Raster   Image objects

| Geometric (Boundary) | Geometric (Pixel) | Semantic (Application context) |

$\Downarrow$ $\Downarrow$ $\Downarrow$

| Thematic (Attributes) | Thematic (Classes) | **Thematic**
- Values of pixels,
- Values of regions
- Geomatric properties of regions |

$\Downarrow$

Geometric – Image objects
- Pixel-based attributes
- Region-based attributes
- Geometric properties of regions
- Thematic attributes
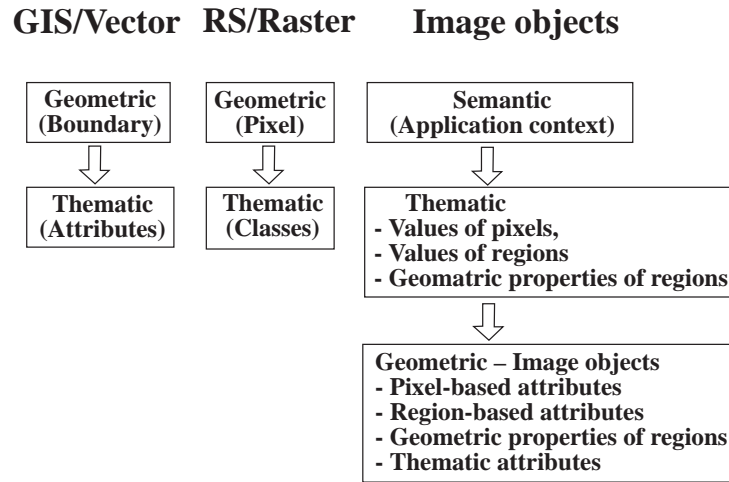
Figure 5.1: Different ways of extracting geo-spatial features, using vector, raster and image-object data models.

the membership values of pixels and refereed based on the image segments derived from images. Ways of extracting object by using vector, raster and image-object data models are compared and illustrated in Figure 5.1.

Image objects (O_I *or* O) or regions contain thematic information from images. Once an image object is formed, a number of geometric properties can be derived such as location, size, orientation, shape (see Zhan et al. (2002d)). As described in the previous section semantic information will be included as well. Therefore, three components can be included in an image object.

$$O\{Thematic, Geometric, Semantic\}$$

For feature extraction based on image objects, we can define several conditions that image objects have to meet in order to be classified. Because of the limitation of the spatial resolution of images and data collection errors attributable to atmospheric impacts, limited accuracy of equipment involved, etc., these conditions may be satisfied to only a certain degree. Membership functions are needed to check the degree to which these conditions are satisfied using the minimum function. The following expressions are based on notions of the Formal Data Structure (FDS) proposed by Molenaar (1998). In the following text an object (O_i) is equivalent to an image object (O_I).

If an object O_i passes the test formulated in a decision function for a class C_j, then it will be a member of that class.

$$MF[O_i, C_j] = \begin{cases} 1, & \text{if } O_i \text{ is a member of } C_j \\ 0, & \text{otherwise} \end{cases}$$

For each object O_i we define the function specifying the class C_j to which the object belongs:

$$CLASS(O_i) = \{C_j \mid MF[O_i, C_j] = 1\}$$

In a multi-condition case, the minimum function is used:

$$CLASS(O_i) = \{C_j \mid min(MF_k[O_i, C_j]) = 1, k = 1, ..., n\}$$

where k denotes a condition.

As many membership functions are fuzzy ($0 \leq MF(O_i, C_j) \leq 1$), an overall membership function ($MF_{OA}[O_i, C_j]$) is required for the final decision, which is obtained by applying the minimum function to all related membership functions to meet these conditions:

$$MF_{OA}[O_i, C_j] = min(MF_k[O_i, C_j]), k = 1, ..., n \tag{5.1}$$

or by using the normalised Euclidean distance when MF_k are considered as measures in a metric space:

$$MF_{OA}[O_i, C_j] = \sqrt{\frac{\sum_{k=1}^{n} MF_k(O_i)^2}{n}}, \quad k = 1, ..., n \tag{5.2}$$

A final decision can be made by choosing a threshold (T) for the overall membership function. When geometric conditions are also included, the final decision will have to be made by reasoning between the overall membership function and geometric properties of potential image objects.

$$CLASS(O_i) = \{C_j \mid MF_{OA}[O_i, C_j] > T\} \tag{5.3}$$

5.3 Hierarchical image objects and hierarchical aggregation

As discussed in Chapter 3, objects may have different types of characteristics and behaviours at different abstraction levels.

5.3.1   Image objects at the pixel level

Each pixel in an image can be regarded as an image object at the lowest level of reasoning. Its spatial coverage is a square covered by a pixel in an image. Its attributes are values from image sources, i.e. intensity of each spectral band of multi-spectral images, or the height value from laser scanning data. The 4-connection and 8-connection adjacency relationships, which respectively consider four and eight directly connected pixels as the adjacent neighbours of a pixel, are commonly used for its spatial relation in a 2D image space.

5.3.2   Image objects at the land-cover level

An image-object at the land-cover level is a group of adjacent (4-connection) pixels that are likely to have the same or similar values (homogeneity) based on certain characteristics or membership functions of a certain class. Its spatial coverage is derived by image analysis and meaningful image segmentation based on image objects (pixels) at the pixel level (Zhan et al., 2001). Its attributes are the average value of pixels forming the object from different image sources. There are two types of objects at the land-cover level, as discussed in Chapter 3 and in the previous section. The first type of object has a determinate interior and describable boundaries, such as buildings and water surface. The second type has a determinate interior, but fuzzy boundaries, such as green space. Topological relations between objects can be identified based on the approach introduced in Chapter 4. Image objects at the land-cover level have to be extracted or aggregated based on image objects (pixels) at the pixel level. Detailed design is presented in Section 5.4.

5.3.3   Image objects at the land-use level

An image object at the land-use level is a spatial unit that contains a number of land-cover objects and represents a particular type of land use. Image objects at the land-use level are often conceptualised entities, such as residential areas. The concepts used for extracting land-use objects may vary depending on the understanding and perception of a specific discipline. Therefore, many image objects at the land-use level are indeterminate objects and have to be modelled based on probability surfaces created based on image objects obtained at the land-cover level. Topological relations between objects can be identified based on the approach introduced in Chapter 4. Detailed design is given in Sections 5.6, 5.7 and 5.8.2.

5.4   Formation of objects at the land-cover level

5.4.1   Extraction of buildings

Before a building can be extracted, we have to answer the following questions and try to specify them semantically and logically:

What is a building? How can we describe a building from different perspectives?

We could try to describe a building by using common-sense knowledge (semantic analysis) from the following perspectives. These features are considered to be extracted based on remote sensing data.

From an application perspective

From an application point of view, different disciplines may have different meanings or understandings regarding these spatial entities. In such cases, additional specifications are required at this stage. In an application where building forecourts should be included as part of the buildings, they can impose a loose condition as regards checking the size differences between two height layers of the same building. In another application where the upper parts of buildings are essential, the forecourt of buildings may be excluded, for instance in assessing the number of dwellings. They can impose more restricted conditions by checking the size differences between the higher layers.

From the geometric perspective

Size

The size of a building as projected in a plane should be larger than 10 m^2 and should be smaller than 5000 m^2.

Height

Buildings should usually stand at least 3 m or more above the surrounding ground and have vertical walls.

From the building material perspective

Building roofs are constructed using various building materials, such as concrete, asphalt, iron and steel, wood, glass. However, most building roofs are built using concrete or tiles. This implies that building roofs are solid and that building roofs are usually not vegetation or water.

5.4.2   Extraction of green spaces

Green space in urban areas usually includes trees and lawns, both having features of vegetation. Vegetation has a unique feature that absorbs electromagnetic waves in most of the panchromatic range, particularly in the RED band range, and emits strong electromagnetic waves in the near infrared (NIR) range, as shown in Figure 5.2. The normalised difference vegetation index (NDVI) is a transformation

designed to enhance such characteristics. Therefore, green space can be extracted based on NDVI values calculated using the NIR band and the RED band from multi-spectral images, using the formula:

$$\text{NDVI} = \frac{\text{NIR} - \text{RED}}{\text{NIR} + \text{RED}}$$

NDVI values can be best extracted based on original multi-spectral images (IKONOS for Amsterdam test site). NDVI values are transformed to fuzzy membership values via an 'S-shape' fuzzy membership function toward two classes, vegetation and non-vegetation, using two parameters obtained by a fuzzy clustering approach called fuzzy c-means to specify the start- and end-points of the S-shape spline curve. A sub-pixel spatial modelling approach is implemented to convert the fuzzy membership value from 4 m resolution to 1 m resolution for integration with laser data. The proposed sub-pixel spatial modelling approach is presented in Section 5.5.

5.4.3   Extraction of water surface

A water body absorbs the full range of electromagnetic waves in different degrees, as shown in Figure 5.2. The normalised difference water index (NDWI) was proposed to reflect such characteristics (McFeeters, 1996). Therefore, water surfaces can be extracted based on NDWI values calculated using the NIR band and the GREEN band from multi-spectral images using the formula:

$$\text{NDWI} = \frac{\text{GREEN} - \text{NIR}}{\text{GREEN} + \text{NIR}}$$

Water extraction by using NDWI alone may be good enough for remote sensing images with coarse spatial resolution and in natural areas or rural areas where not many man-made objects exist. For a high-resolution image and in much more complicated urban areas, using the standard NDWI alone may not be sufficient to separate water from other objects that have very similar spectral features to water pixels, such as dark shadow, dark road, dark building roof. To illustrate the problem, we manually pick up a number of typical samples from an IKONOS image (see Figure 2.2) to represent land-cover features typical of an urban scene, as shown in Figure 5.2.

Enhance normalised difference water index (eNDWI)

Based on Figure 5.2, we can observe that lake water, canal water, shadow and dark building are very similar in all four bands of the IKONOS image. We can also observe that both normal and new buildings show a similar trend to water if we choose Band 2 (GREEN band) and Band 4 (NIR band) to acquire NDWI. Only vegetation can be easily separated from others by using NDWI. Therefore, we propose another formula in which all bands are used for water extraction. To make it different from the existing NDWI, we called this index the enhanced NDWI and use the notion eNDWI. The eNDWI intends to enhance the difference between a shallow water
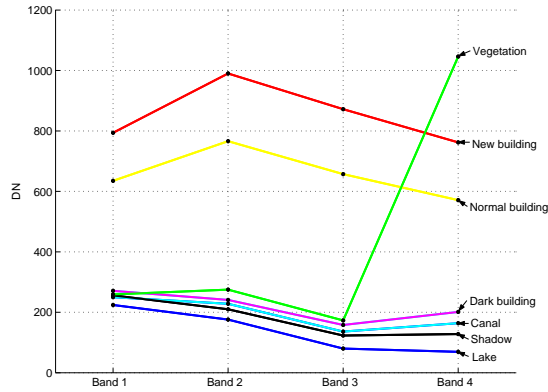
Figure 5.2: Spectral reflectance of typical urban features, based on IKONOS image.

body such as a canal and shadow area, as well as the difference between water and light building.

$$eNDWI = \frac{BLUE + GREEN - RED - NIR}{BLUE + GREEN + RED + NIR}$$

Even when using eNDWI, we may still not be able to completely separate water, shadow and dark building, but these water-like features are better separated from buildings, particularly new buildings. By comparing two histograms, NDWI and eNDWI, obtained from the same image (see Figures 5.3 and 5.4), it can be seen that, despite the slight shrinkage in contrast, eNDWI separates a significant feature (canal indicated by second small peak to the right in Figure 5.4), that was previously mixed with building and concrete in NDWI.
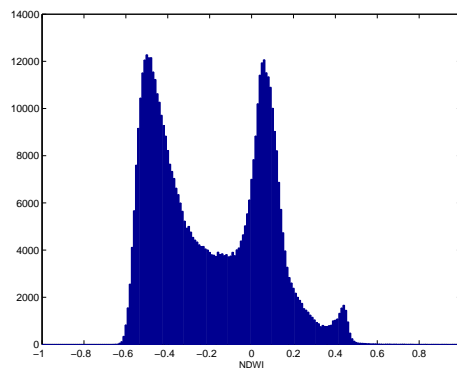
Figure 5.3: Histogram of NDWI based on IKONOS image.
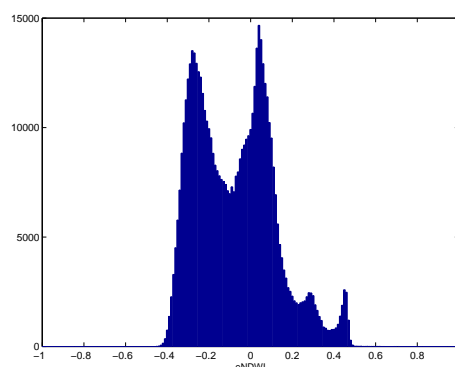
Figure 5.4: Histogram of eNDWI based on IKONOS image.

Remove other features from water surfaces

Although the proposed eNDWI performs better than NDWI, it is still not able to clearly distinguish shadow and dark building from water. Since we have detailed laser data, and buildings are extracted as described earlier, we could consider using this information to mask building-related features such as shadow and dark building.

Using laser data and meta data of the multi-spectral image, we can simulate shadow areas caused by buildings and the relief displacement of buildings by applying the hillshade algorithm. Water surfaces can be refined by masking shadow areas and buildings, including displaced roofs caused by the slightly oblique viewing of the sensor.

5.4.4 Derivation of image-object properties at the land-cover level

The extracted image regions obtained by segmentation based on the overall membership function or reasoning are represented in the form of a binary image. Pixels belonging to such regions will have the value 1 and other pixels will have the value 0.

Such a binary image is further labelled with a unique ID for each region so that all pixels in a region should be connected by the 4-connection. The unique ID is used as the identifier for each image region.

A number of geometric properties such as size, shape and orientation can be derived for each image object as described in Sections 4.5 and 4.6.

5.5 Spatial modelling for pixel interpolation from a coarser resolution to a finer resolution

Urban features often have sharp boundaries. Because of the limited spatial resolution of remotely sensed images, pixels containing boundary elements will contain a mixture of the spectral responses from different features. Among the four causes of mixed pixels described by Fisher, and as shown in Figure 5.5 (Fisher, 1997), the 'Sub-pixel' and 'Linear sub-pixel' cases are the most difficult, owing to lack of information regarding their existence and their spatial extents for objects smaller than the pixel size. Therefore the 'Boundary pixel' and 'Intergrade' cases are the main targets of our sub-pixel approach. However, the proposed approach aims at solutions to the 'Boundary pixel' and 'Intergrade' cases without neglecting the potential existence of the other two cases.
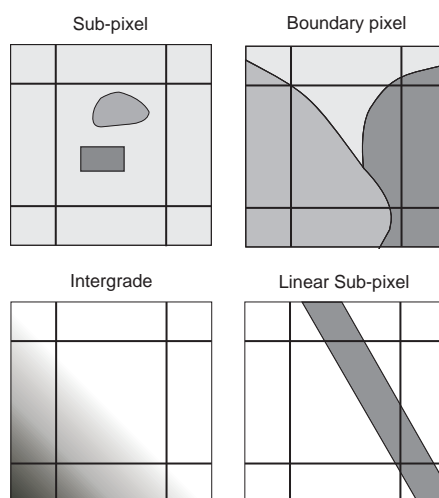
Figure 5.5: Four causes of mixed pixels (Fisher, 1997); reproduced from the International Journal of Remote Sensing by permission of Taylor and Francis Ltd, http://www.tandf.co.uk/journals

There are a number of techniques proposed for dealing with the pixel unmixing issue (Atkinson, 1997; Foody et al., 1997; Schowengerdt, 1997; Foody, 1998; Steinwendner, 1999; Tatem et al., 2001a,b). As the value of each pixel is the composite spectral signature of the land-cover types present, these approaches were applied based on spectral pixel unmixing and spatial unmixing respectively, or based on both aspects.

Mixture modelling, neural networks and fuzzy c-means classifier are currently available for estimating the proportions of different classes that a pixel may rep-

resent (Foody, 1996; Atkinson, 1997; Atkinson et al., 1997; Bastin, 1997). Conventional classifiers such as the maximum likelihood classifier (MLC) are based on the spectral signatures of pure pixels and do not recognise spatial patterns in the same way that a human interpreter does (Gong and Howarth, 1990). On the other hand, the MLC generates a substantial amount of information on the class membership properties of a pixel, which provides valuable information on the relative similarity of a pixel to the defined classes (Foody et al., 1992). The objective of our approach is to incorporate probability measures derived from the MLC or fuzzy membership values, together with spatial information at the pixel scale to increase the accuracy classification and to produce finer classification maps or interpolation results at the sub-pixel scale.

A spatial dependence model was applied in mapping the location of the land-cover proportions estimated from the mixture model at the sub-pixel scale (Atkinson, 1997). The algorithm was iterated several times through all the pixels at the pixel and sub-pixel scales to avoid a 'hole' at the centre and to smooth the surface. However, such an arrangement conflicts with our understanding of what constitutes 'maximum spatial order' as the authors indicated (Atkinson, 1997). The proposed approach therefore intended to define the contributions from the central pixel and the neighbouring pixel to spatial allocation at the sub-pixel scale, based on the assumption that the land cover is spatially dependent both within and between pixels, as our aim is to respond to the 'Boundary pixel' and 'Intergrade' cases (Atkinson, 1997; Verhoeye and De Wulf, 2000).

Some promising results have also been achieved for pixel unmixing by using a neural network approach (Tatem et al., 2001a,b). However, neural network approaches are sensitive to samples used in training and testing phases. It may cost a lot of effort to select a sufficient number of good samples, with due consideration to representatives of the spectral and spatial aspects, to ensure the ability of neural networks in pixel unmixing. This training procedure has to be repeated again when it applies to other data sets. However, in our approach only spectral information of samples is used. Spatial allocation at the sub-pixel scale will be processed in a separate stage.

A per-field approach using detailed vector data can improve classification accuracy (Aplin et al., 1999a,b; Aplin and Atkinson, 2001). In most cases, however, accurate vector data sets are rarely available (Tatem et al., 2001b). Feature boundaries may have changed as well if image data and vector data have been captured at different periods.

In the proposed two-stage (spectral-spatial) approach, we implemented fuzzy classification, using an S-shape fuzzy membership function in the first stage. In the second stage, we applied the inverse distance weighting predictor to interpolate a membership surface at the sub-pixel scale with the *MF* values of the central pixel and neighbourhood at the pixel scale, and classified the image according to the interpolated *MF* values at sub-pixel scale.

The proposed approach aims at increasing the accuracy of land-cover classification and producing finer classification maps, especially for boundary pixels. The results will be used later in inferring urban land use based on land cover classification. We have already demonstrated the ability of the proposed sub-pixel approach

to improve classification accuracy, based on probability derived by maximum likelihood classifier (see Zhan et al. (2002a)). In this case, we apply this spatial modelling approach for pixel interpolation from coarse resolution (IKONOS, 4 m) to finer resolution (Airborne LIDAR, 1 m), based on fuzzy membership values.

5.5.1 Proposed sub-pixel methods

After implementing the maximum likelihood classification or fuzzy classification, the proportion of each class in a pixel is given by the pixel's probability or membership vector. Each pixel is split up by the zoom factor. We use zoom factor 4 to illustrate that each pixel is split up into 4 x 4 = 16 sub-pixels, as shown in Figure 5.6. To determine the class probability or membership value of a sub-pixel for each endmember class, a new probability vector or membership value is calculated, based on the probability or membership value vectors of the central pixel and its eight neighbouring pixels. The inverse distance weighting predictor was used in computing a new probability value or membership value for each sub-pixel. The assumption is that the value of an attribute z at an unvisited point is a distance-weighted average of data points occurring within a neighbourhood or window surrounding the unvisited point (Burrough and McDonnell, 1998).

$$\hat{Z}(x_0) = \frac{\sum_{i=1}^{n} Z(x_i) \cdot d_{ij}^{-r}}{\sum_{i=1}^{n} d_{ij}^{-r}} \tag{5.4}$$

$\hat{Z}(x_0)$: is the value of the attribute at an unvisited location
$Z(x_i)$: z is the known value of the attribute at location xi
d_{ij} : d is the distance between the unknown point x_j and a neighbour x_i
r : is a distance weight factor
n : is the number of neighbours

For a given sub-pixel the distances to the nearest edges or corners of the neighbours are calculated. These distance measures d_{ij} are used to calculate the new probability vector of the sub-pixel $z(x_0)$ by taking the distances from a given sub-pixel to the edges of the N, E, S and W neighbours and to the corners of the NW, NE, SE and SW neighbours (see Figure 5.6). The effect of these distance measures on the interpolation result is tested. The distance weight factor r is set to 1.0. An important factor to consider is how to incorporate the probability or membership value vector of the current pixel itself in the interpolation. One option is to leave the centre probability vector out. In this case, only the neighbouring probability vectors are used (neglecting the existence of the 'Sub-pixel' and 'Linear sub-pixel' cases in Figure 5.5). Another option is to choose a distance value for the central pixel in the interpolation (considering the potential existence of the 'Sub-pixel' and 'Linear sub-pixel' cases in Figure 5.5). For example, the distance from each sub-pixel to the centre pixel could be set at 1.0 to give this pixel a large weight.
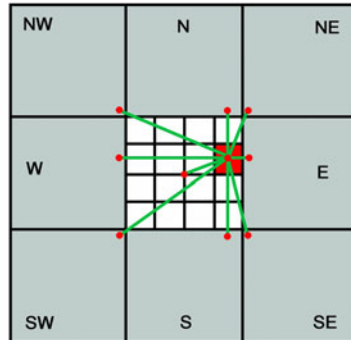
Figure 5.6: Inverse distance interpolation used to compute sub-pixel probability vectors. Distances are taken from sub-pixel to corners or edges of neighbouring pixels.

5.5.2 Experimental testing of the proposed sub-pixel method

To verify the proposed sub-pixel approach, a number of controlled tests were implemented. Four test images were created manually as truth at the sub-pixel scale (200 × 200 pixels), with two classes with values 1 and 0 (white and black respectively as shown in Figures 5.7-c, 5.8-c, 5.9-c and 5.10-c). Simulated probability images (50 × 50 pixels) were generated from four test images, using the averaging aggregation method in order to maintain the statistical and spatial properties of the simulated data (Bian and Butler, 1999). Simulated images are presented in Figures 5.7-a, 5.8-a, 5.9-a and 5.10-a (50 × 50 pixels). Each pixel of a simulated image covers 4 × 4 pixels at the sub-pixel scale (truth images), corresponding to the same spatial aggregation scale of IKONOS imagery (4 m) and LIDAR (1 m). Classification results obtained at pixel scale are provided in Figures 5.7-b, 5.8-b, 5.9-b and 5.10-b. Probability results at the sub-pixel scale by applying the 'bilinear', the 'bicubic' and the proposed approaches are shown in Figures 5.7-d,f,h; 5.8-d,f,h; 5.9-d,f,h and 5.10-d,f,h respectively. The classification results at the sub-pixel scale by using these three approaches are shown in in Figures 5.7-e,g,i; 5.8-e,g,i; 5.9-e,g,i and 5.10-e,g,i.

The proposed sub-pixel interpolation results are compared with the results obtained at pixel scale, as well as the standard 'bilinear' and 'bicubic' interpolation approaches, as shown in Figures 5.7, 5.8, 5.9, and 5.10. Pixel-to-pixel comparison with corresponding 'truth' images shows that the proposed approach gains the highest overall accuracy and Kappa coefficient in the four tests, as indicated in Table 5.1.
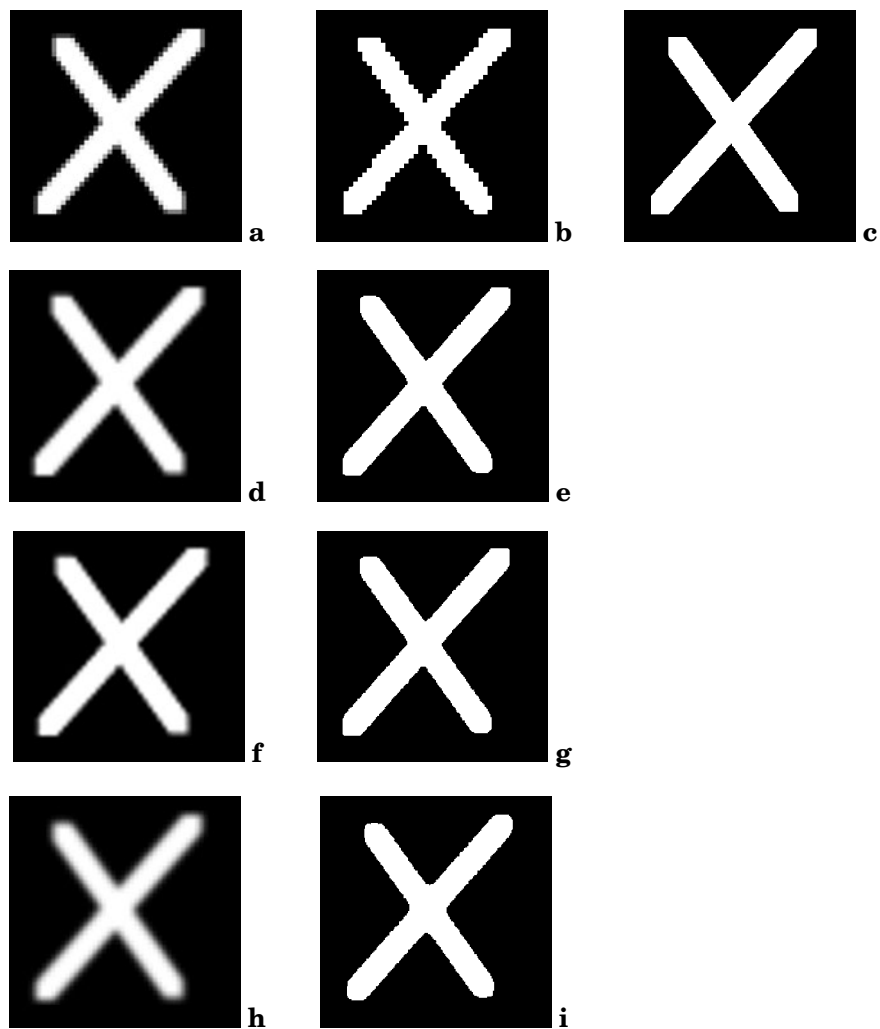
Figure 5.7: Sub-pixel test 1: a - simulated image at the pixel scale based on 'truth' image (c); b - classified image based on image (a) at the pixel scale; c - 'truth' image prepared at the sub-pixel scale; d - result of 'bilinear' interpolation based on image (a); e - classified image based on image (d); f - result of 'bicubic' interpolation based on image (a); g - classified image based on image (f); h - result of proposed sub-pixel approach based on image (a); i - classified image based on image (h).

Figure 5.8: Sub-pixel test 2: a - simulated image at the pixel scale based on 'truth' image (c); b - classified image based on image (a) at the pixel scale; c - 'truth' image prepared at the sub-pixel scale; d - result of 'bilinear' interpolation based on image (a); e - classified image based on image (d); f - result of 'bicubic' interpolation based on image (a); g - classified image based on image (f); h - result of proposed sub-pixel approach based on image (a); i - classified image based on image (h).
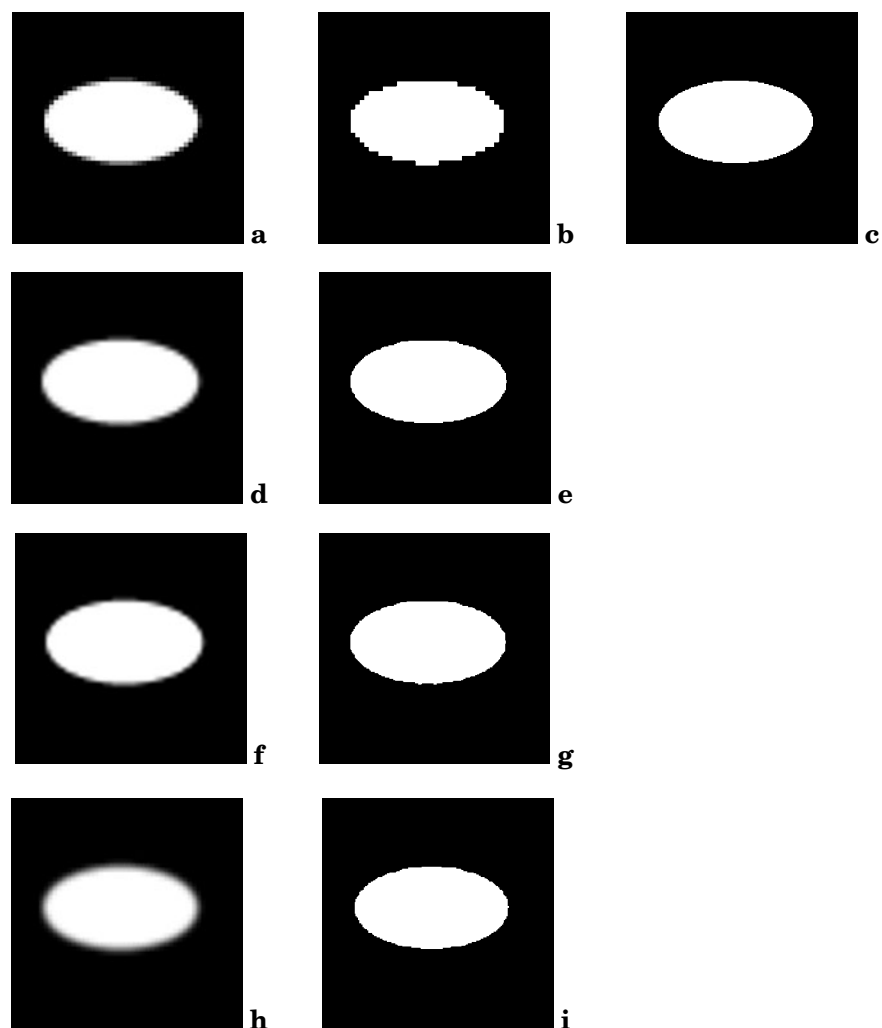
Figure 5.9: Sub-pixel test 3: a - simulated image at the pixel scale based on 'truth' image (c); b - classified image based on image (a) at the pixel scale; c - 'truth' image prepared at the sub-pixel scale; d - result of 'bilinear' interpolation based on image (a); e - classified image based on image (d); f - result of 'bicubic' interpolation based on image (a); g - classified image based on image (f); h - result of proposed sub-pixel approach based on image (a); i - classified image based on image (h).

Figure 5.10: Sub-pixel test 4: a - simulated image at the pixel scale based on 'truth' image (c); b - classified image based on image (a) at the pixel scale; c - 'truth' image prepared at the sub-pixel scale; d - result of 'bilinear' interpolation based on image (a); e - classified image based on image (d); f - result of 'bicubic' interpolation based on image (a); g - classified image based on image (f); h - result of proposed sub-pixel approach based on image (a); i - classified image based on image (h).

Table 5.1: Test results using overall accuracy and Kappa coefficient

Test image	Pixel scale		Sub-pixel scale					
			Bilinear		Bicubic		Proposed	
	Overall accuracy	Kappa coefficient	Overall accuracy	Kappa coefficient	Overall accuracy	Kappa coefficient	Overall accuracy	Kappa coefficient
Test1	98.21%	94.33%	99.42%	98.19%	99.40%	98.12%	99.51%	98.44%
Test2	98.73%	95.95%	98.73%	95.97%	98.67%	95.77%	99.62%	98.78%
Test3	99.22%	97.42%	99.52%	98.42%	99.48%	98.31%	99.80%	99.35%
Test4	98.53%	93.60%	99.12%	96.17%	99.09%	96.03%	99.57%	98.09%

5.6 Finding spatial units for land-use types

To avoid confusion caused by using different terms, we would like to give more explicit explanation of the terms used in this section. We use 'spatial pattern' to refer to both spatial arrangement and spatial distribution. 'Spatial arrangement' means spatial alignment such as along the road, or shapes such as rectangular and circular. 'Spatial distribution' indicates the spatial extent of certain features that can be bounded by convex hulls. 'Adjacency' refers to a qualitative spatial relationship where no other object exists between the adjacent objects, but the adjacent objects may not necessarily 'touch', i.e. the adjacent objects can be disjointed. 'Proximity' refers to a quantitative spatial relationship that indicates closeness, considering the distance between objects.

Since an individual land-cover feature does not tell much about the spatial extent of a land-use unit, we need to look for the spatial arrangement and spatial distributions of land-cover features. Certain land-use types often show certain spatial patterns or even unique patterns in urban areas. On the other hand, some indicators as to what a spatial unit contains can be derived according to feature type, number and their distribution. For instance, we can use building as a feature for reasoning land-use units, and use the number, average size of buildings and building density as properties or attributes of such units (land-use units) to classify the use of such units (land-use types). Therefore, finding reasonable spatial units and boundaries is the key to land-use classification.

Buildings are land-cover objects with certain shapes and sharp boundaries. It is relatively easy to extract buildings from image and laser data of good quality. The spatial distribution and spatial arrangement of buildings are often the results of conscious planning and development to serve certain functions, and thus can provide indications for land-use reasoning. For instance, a residential area normally consists of closely situated houses of similar size, shape and orientation (parallel or perpendicular). Gardens or green space can often be found in open space between residential buildings. Therefore, spatial clustering of buildings plays an important role in finding spatial units of land uses. In this section we will introduce a method of extracting land-use spatial units based on the spatial distribution and feature similarity of buildings.

5.6.1 Cluster analysis and spatial clustering

A spatial unit is a spatial cluster of land-cover objects that serves certain economic functions and contains certain types of physical features. Physical features such as buildings in a spatial unit often show similar characteristics, such as size, shape, orientation, and these features are often located within a close range in space, forming a certain spatial pattern. Cluster analysis techniques can help to find clusters based on similarity of features in a feature space. k-means, k-nearest neighbour (k-NN), fuzzy c-means etc. are some well-known approaches for clustering.

Spatial clustering is the process of grouping a set of objects into classes or clusters so that objects within a cluster are highly similar to one another, but are dissimilar to objects in other clusters (Han et al., 2001). Partition methods such as the k-means algorithm (MacQueen, 1967), the expectation maximisation (EM) algorithm (Dempster et al., 1977; Bradley et al., 1998; Yu et al., 1998) and the k-medoid algorithm (Kaufman and Rousseeuw, 1990) tend to find only spherical-shaped clusters and encounter difficulty in discovering clusters of arbitrary shapes (Han et al., 2001). Unfortunately, land-use spatial units often have arbitrary shapes and the number of clusters k is often unknown.

Han et al. (2001) provided a survey on spatial clustering. Spatial clustering is a special type of clustering that takes into account both similarity of features in a feature space and spatial distance in a 2D or 3D physical space.

Therefore, we need to find suitable solutions to our specific problems by investigating the following aspects:

- Distances between objects

- Similarity measures of features

- Spatial arrangement and/or spatial distribution

- Spatial partitioning based on closeness (distance), and spatial comparability that takes feature similarity into consideration (i.e. features that are more similar based on certain characteristics are more likely to belong to the same cluster)

5.6.2 Distances between objects

Euclidean distance plays a very important role in spatial clustering. It remains a key player in our case since proximity often denotes a similarity of use. Euclidean distance is used to determine how close objects are situated in 2D space.

The distance between two objects can be defined in several ways: the distance between the centres of two objects (the centre of gravity), or the shortest distance between two objects (i.e. the shortest distance between pixels that belong to two objects). To illustrate the different distance measures and their clustering results, a small sample area is selected, as presented in Figure 5.11. Buildings shown in Figure 5.11 should be clustered in three groups, as shown in Figure 5.12, by visual interpretation, according to spatial pattern and our planning knowledge of spatial arrangement. The same result may not be achieved by using distances between the centres of objects, as illustrated in Figure 5.13 where distance **a** is obviously larger
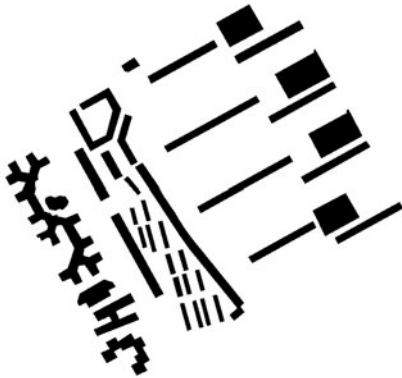
Figure 5.11: Buildings in an urban area where buildings vary in size, shape, orientation, and perhaps in height etc.
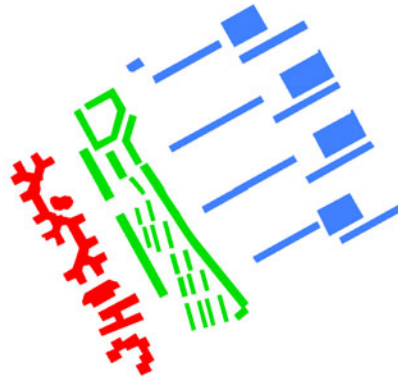
Figure 5.12: Manual clustering of buildings (red, green and blue) according to our planning knowledge.

than **b**, and **c** differs slightly from **d**. Therefore, for spatial clustering based on the distribution of buildings, we consider the shortest distance between objects a better measure than the distance between object centres, as shown in Figure 5.14.

In general, the shorter the distance between two buildings, the higher the possibility that the two buildings belong to the same cluster. Buildings may vary in size and shape. The distance between small and low-rise residential buildings tends to be small, and the distance between large and high-rise residential buildings has to be large in accordance with the planning requirement in many countries that residential dwellings need direct sunlight, even during the season with the lowest sun angle, to ensure a healthy living environment. Such a requirement can be achieved by considering the longest shadow a building may project in direction of sunlight, and the terrain relief, as shown in Figure 5.15. The minimum distance between residential buildings is often enforced by planning regulations governing geographical locations and building height. The required minimum distance may be smaller when a location is closer to the equator, when building height is constant. Such a requirement and association are valid for residential areas. Rules may be different for other land-use types and between different land-use types. In addition, real situations are very complicated, especially in urban areas, where similarity in size, shape, orientation and building height should have an impact on spatial clustering as well. Therefore, for spatial clustering it may be necessary to consider and incorporate similarity measures based on these features, in addition to the shortest distance.
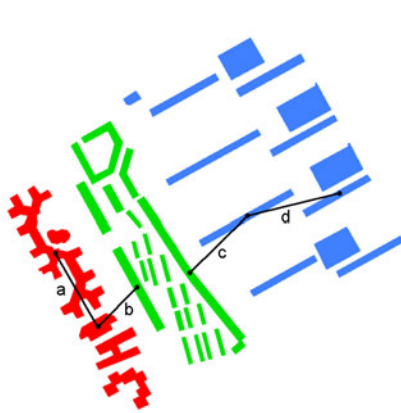
Figure 5.13: Distances between centres of objects (a, b, c and d).
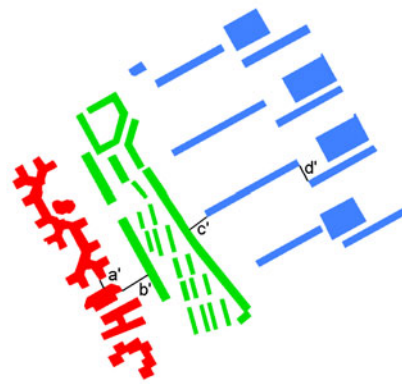
Figure 5.14: The shortest distances between objects (a', b', c' and d').
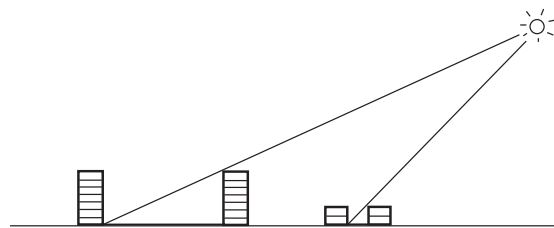
Figure 5.15: Relationship between the minimum distance to avoid shadow, building height and sun angle of a location with flat terrain.

5.6.3 Similarity measures in feature space

Many similarity measures have been proposed in literature for various applications, such as feature-based similarity (Tversky, 1977), similarity for case-based reasoning (Osborne and Bridge, 1996, 1997a,b), similarity measures for content-based image retrieval (Santini and Jain, 1995, 1999) and semantic similarity evaluation for categorical data (Bishr, 1997; Liu et al., 2002; Rodríguez and Egenhofer, 2003), just to mention a few. It is extremely difficult to identify a single all-encompassing 'best' similarity measure (Zobel and Moffat, 1998). The following are a number of indicators and features that are considered useful for reasoning spatial units of land-use in this research.

We observe that many features are not linear or not even monotonic in feature space when we use these features to determine whether two objects should be in the same spatial cluster or in different spatial clusters, i.e. we may not be able to determine whether or not two objects belong to the same spatial cluster according to the individual values of these features. For example, two objects of similar size are likely to belong to the same cluster, but we are not so sure because of the complexity in reality. In such cases, a distance measure based on object features may not be a metric one, because it may not follow the three axioms required for metric space: minimality, symmetry and triangle inequality.

Tversky proposed the well-known contrast model and the ratio model which are based on a feature matching function using a set-theoretic model (Tversky, 1977):

$$\text{Similarity } Sim(\mathbf{a}, \mathbf{b}) = F(\mathbf{A} \cap \mathbf{B}, \mathbf{A} - \mathbf{B}, \mathbf{B} - \mathbf{A}).$$

The similarity Sim of object $\mathbf{a}$ to object $\mathbf{b}$ is expressed as a function F of three arguments: $\mathbf{A} \cap \mathbf{B}$, the features that are common to both $\mathbf{a}$ and $\mathbf{b}$; $\mathbf{A}$-$\mathbf{B}$, the features that belong to $\mathbf{a}$ but not to $\mathbf{b}$; $\mathbf{B}$-$\mathbf{A}$, the features that belong to $\mathbf{b}$ but not to $\mathbf{a}$. One representation called the contrast model was proposed under certain assumptions (see details in Tversky, 1977):

$$Sim(\mathbf{a}, \mathbf{b}) = \theta f(\mathbf{A} \cap \mathbf{B}) - \alpha f(\mathbf{A} - \mathbf{B}) - \beta f(\mathbf{B} - \mathbf{A}), \; for \; \theta, \alpha, \beta \geq 0.$$

This is a linear combination of the measures of the common and the distinctive features. For example, if $\theta = 1$, α and β vanish, then $Sim(\mathbf{a}, \mathbf{b}) = f(\mathbf{A} \cap \mathbf{B})$; that is, the similarity between objects is the measure of their common features. If, on the other hand, $\alpha = \beta = 1$, and θ vanishes then $-Sim(\mathbf{a}, \mathbf{b}) = f(\mathbf{A}-\mathbf{B})+f(\mathbf{B}-\mathbf{A})$; that is, the dissimilarity between objects is the measure of the symmetric difference between the respective feature sets. Note that in the former model ($\theta = 1$, $\alpha = \beta = 0$), it is determined by their common features, whereas in the latter model ($\theta = 0$, $\alpha = \beta = 1$), it is determined by their distinctive features. The contrast model expresses similarity between objects as a weighted difference of the measures of their common and distinctive features, thereby allowing for a variety of similarity relations over the same domain (Tversky, 1977). The contrast model intends to obtain an absolute value to indicate the degree of similarity between two objects. In practice, however, it is quite difficult to verify such a function and determine the three parameters θ, α and β. The ratio model, which is presented in the same paper, provides another option:
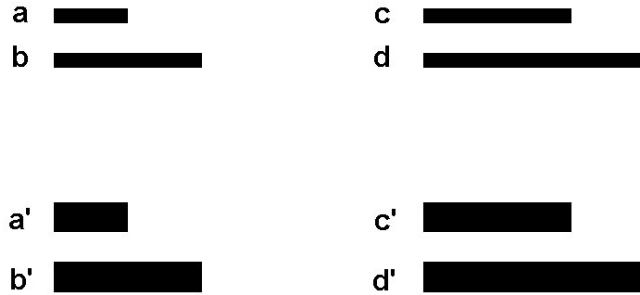
Figure 5.16: Similarity comparison between the absolute difference and the proportional difference.

$$Sim(\mathbf{a}, \mathbf{b}) = \frac{f(\mathbf{A} \cap \mathbf{B})}{f(\mathbf{A} \cap \mathbf{B}) + \alpha f(\mathbf{A} - \mathbf{B}) + \beta f(\mathbf{B} - \mathbf{A})}, \; \alpha, \beta \geq 0$$

where similarity is normalised so that Sim lies between 0 and 1. The ratio model gives a relative value to the degree to which features are similar.

5.6.4   The ratio model

Before choosing a similarity measure, we would like to do a simple experiment and use object size as a feature for similarity comparison. As shown in Figure 5.16, we can observe that the degree of similarity between **a** and **b** is more or less equivalent to that between **a'** and **b'** in terms of size because the ratio of size difference between **a** and **b** and the ratio between **a'** and **b'** are the same (**a** is half the size of **b** or **b** is two twice the size **a** and the same ratio is applicable to **a'** and **b'**). Similar observations can be made between the pairs **c**, **d** and **c'**, **d'**. When we compare the degree of similarity between **a** and **b** and the degree of similarity between **c** and **d**, we can observe that **c** and **d** are more similar than **a** and **b** (i.e. $Sim(\mathbf{c}, \mathbf{d}) > Sim(\mathbf{a}, \mathbf{b})$) because **c** is two-thirds of the size **d** and **a** is half the size of **b**. The first conclusion that can be made is that the ratio of size between two objects is proportional to the degree of similarity in terms of size. We should notice that the absolute size difference between **a** and **b** (i.e. $|\mathbf{a} - \mathbf{b}|$) is the same as that between **c** and **d** (i.e. $|\mathbf{c} - \mathbf{d}|$). The second conclusion is that the absolute size difference is not a suitable measure for similarity assessment, since **c** is more similar to **d** than **a** is to **b** although $|\mathbf{a} - \mathbf{b}| = |\mathbf{c} - \mathbf{d}|$. A similar observation can be made for similarity comparison in terms of building height. Therefore the ratio model (proportional) is better than the absolute measure in these cases.

Table 5.2: Similarity in size of each pair of 'squares' in Figure 5.17

A, B	a, b	b, c	c, d	d, e	e, f	f, g	g, h	h, i
$\mathbf{A} \cap \mathbf{B}$	1	4	16	25	36	49	64	81
$|\mathbf{A} - \mathbf{B}|$	3	12	9	11	13	15	17	19
$\mathbf{A} \cup \mathbf{B}$	4	16	25	36	49	64	81	100
Similarity	0.25	0.25	0.64	0.694	0.735	0.766	0.79	0.81
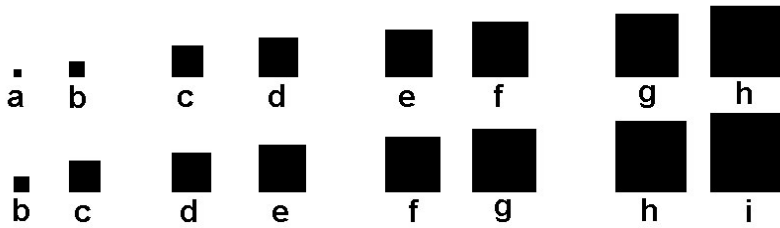
Figure 5.17: Similarity test in terms of size.

To determine the choice of the two parameters of the ratio model α and β, we consider that the similarity function is symmetric when we compare two objects in terms of building size or building height in our case, since $|\mathbf{a} - \mathbf{b}| = |\mathbf{b} - \mathbf{a}|$. Therefore we use $\alpha = \beta = 1$ in the ratio model and construct our similarity measures for building size, building height etc. as follows:

$$Sim(\mathbf{a}, \mathbf{b}) = \frac{f(\mathbf{A} \cap \mathbf{B})}{f(\mathbf{A} \cap \mathbf{B}) + f(\mathbf{A} - \mathbf{B}) + f(\mathbf{B} - \mathbf{A})} = \frac{f(\mathbf{A} \cap \mathbf{B})}{f(\mathbf{A} \cup \mathbf{B})} \tag{5.5}$$

In a test using the above formula, the similarity value is calculated for each pair of 'squares' (see Figure 5.17), as shown in Table 5.2. The test results show that the similarity values fit well with our visual observation, as discussed earlier. The similarity values of pair (**a**, **b**) and pair (**b**, **c**) are the same and fit well with our visual observation. This is because the larger object is four times bigger than the corresponding smaller object in both pairs, despite the fact that the absolute size difference between **b** and **c** is four times bigger than that between **a** and **b**.

This formula will be applied in spatial clustering for similarity assessment of building size and height. Similarity measures for shape and orientation are much more complicated because the similarity can hardly be measured by a single similarity measure. Additional investigation has to be made in this regard. The similarity measures for building size and height will be used to determine whether or not two neighbouring buildings belong to the same spatial cluster, in addition to the distance measure (i.e. the shortest distance between these objects).

5.6.5 Spatial distribution of buildings as indication of the spatial extent of land-use units

Although we believe in the fundamental assumption that the shorter the distance between two objects, the more similar the instances are, the actual spatial patterns are function-based, in that some facilities (such as schools) are built to serve residents at the local neighbourhood level, while other facilities (such as shopping centres) serve the local communities or even for the whole city. These types of functional spatial associations are much more complicated for spatial modelling than are other physical constraints such as terrain relief, rivers, canals and lakes, or social and political constraints such as administrative boundaries. There are still some rules that can be applied at the local level. For instance, in a residential neighbourhood, houses are likely to be built by a developer at a certain time and in a similar fashion with regard to building size, shape, number of floors and orientation, etc. These buildings are also likely to be situated at close range in an urban space. Therefore, additional similarity measures such as size and height should be involved in spatial clustering, besides the shortest distance between adjacent objects. This type of information can be extracted from airborne laser data and multi-spectral data by using the object-based approach.

Many social-economic functions in urban areas are proposed and organised in the planning and implementation phases. Urban space used for certain social-economic functions is conceptualised and continuous in nature, and thus often has vague boundaries. This means that we may not be able to find sharp spatial boundaries for different functions. For instance, vegetation may be found in a residential area and continuously distributed to a park (recreational use) next to the residential area when no physical boundaries can be found in the transit zone between the residential area and the park. However, a number of physical components that provide these functions may be found in a spatial extent, such as buildings for residential use. Therefore, we may use the distribution of physical features (land-cover objects) for reasoning the spatial extent of land-use functions (i.e. land-use objects).

Indicators for spatial clustering

- Adjacency relationship
  The adjacency relationship provides a meaningful spatial relation between two objects. The adjacency relationship can be obtained by the Delauney triangulation. In our approach, adjacent objects are only considered if they belong to the same cluster, while property similarity of adjacent objects is also considered. Non-adjacent objects will not be checked. Two non-adjacent objects may be grouped in the same cluster only if adjacency links hold (i.e. two adjacent objects connected by the adjacency link meet the criteria for belonging to the same cluster) and such links pass from one object to another non-adjacent object via other adjacent objects. For instance, if object **a** and its adjacent object **b** are considered as belonging to the same cluster, and **b** and its adjacent object **c** are also in the same cluster, then non-adjacent objects **a** and **c** are considered to be in the same cluster (as shown in Figure 5.18).

- The shortest distance between two adjacent objects
  The distances between two adjacent objects can be measured from many pixels that belong to these two adjacent objects. The shortest distance between two adjacent objects is obtained by comparing the length of all triangle edges (Delauney triangulation) that link two adjacent objects. The shortest distance between two adjacent objects is a quantitative proximity measure for how close two adjacent objects are.

- Feature similarity
  The more properties found similar (i.e. similar size, shape, etc.), the more likely it is that two adjacent objects are grouped in the same cluster.

- Density and other area-related measures
  Density and other area-related measures such as building density and floor area ratio (FAR) are spatially related measures in a continuous or discrete space. Similar density or other measure values provide indications for the possible merging of clusters. Building density and FAR are useful measures in built-up areas, not only for checking whether to combine two clusters, but also for cluster identification (land-use classification).
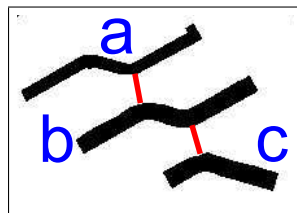
Figure 5.18: Clustering of adjacent and non-adjacent objects.

The adjacency relationship can be derived by applying Delaunay triangulation or its dual graph, the Voronoi diagram, since no other point can be found along any edges of Delaunay triangulation or two points share a common boundary in the Voronoi diagram, which ensures the two points are adjacent. When Delaunay triangulation is applied, pixels as parts of image objects at the land-cover level can be treated as points and used for creating Delaunay triangulation. We can then determine the shortest link between two adjacent objects by comparing the lengths of all the Delaunay triangulation edges that link pixels that belong to the two objects. Such shortest links and their lengths are stored in a table, similar to the region adjacency graph (RAG), as an indicator for checking if these two objects belong to the same spatial cluster (land-use unit), along with other indicators such as similarity in size and height. The shortest links provide information that indicates that these objects are spatially adjacent and the degree to which these objects are close.

Feature similarity provides additional measures for decision making on merging or splitting in spatial clustering. As mentioned earlier in this section, the ratio model is applied as a similarity measure for features such as building size and

building height. Similarity measures will be based on the average per cluster when comparing two clusters, while each of these clusters contains more than one object.

5.6.6 Partitioning according to similarity of features and the shortest distance

Owing to the complexity of the spatial arrangement and spatial distribution of features in urban areas, we may not be able to find these spatial clusters by using similarity measures of features or spatial distances alone. It has to be clear that similarity measures of features are derived from a feature space whereas distances are derived from the physical space (i.e. Euclidean distance). These two types of measures are independent in nature. To find reasonable spatial units for land-use classes in a space partitioning process, we need to integrate both feature similarity between objects and the distance between them.

Similarity of features and proximity

Studies using texture stimuli have found that the human visual system can quickly group similar colour and shape features into global spatial regions and then rapidly segregate them at their boundaries or edges in order to begin establishing figure-ground relations within a scene (Cook, 2001). Similarly, in our cases many factors may have to be considered in terms of similarity and spatial closeness in order to define and delineate spatial clusters. The adjacency relationship, the shortest distance, feature similarity and density are considered important factors for reasoning in finding spatial units of land use in this research. The effectiveness of these factors will be discussed in detail in Chapter 7.

Delineation of land-use units

As discussed in Chapter 3, there is often no direct physical evidence to be found to indicate the spatial extent of a land-use unit, and some land-use types are conceptualised. Therefore we proposed a surface modelling approach to delineate boundaries for such land-use units when spatial clusters are determined. To illustrate the approach step by step, we use the example shown in Figure 5.19. A morphological closing operation is applied to create the solid core of a spatial cluster, as shown in Figure 5.20. The distance transformation is then applied to generate a transit zone around the core of each land-use unit and an S-shape fuzzy membership function is used to create a simulated surface represented by fuzzy membership values which gradually decrease from the edge of the core toward the neighbouring clusters, as shown in Figures 5.21, 5.22 and 5.23. The integrated surface as shown in Figure 5.24 is used to determine the boundaries of land-use units by taking the local minima as the boundary between different clusters, as illustrated in Figure 5.27. This can be treated as the watershed when we reverse the surface. In the final stage, the watershed algorithm (Vincent and Soille, 1991) is used to obtain the boundaries of land-use objects, as shown in Figure 5.25. The land-cover objects superimposed with land-use objects are shown in Figure 5.26.

For application of the surface model, we can specify the maximum width of the transit zone (e.g. 80 m) and obtain surface values by applying the Z-shape fuzzy membership function, as shown in Figure 5.27. In this model, we consider the same width for all land-use types since we have not identified the land-use type yet. It is possible to specify different transit zone width for different land-use types (e.g. an industrial area may have a larger width than a residential area) when related information is available. Another option is to specify the transit zone width proportional to the average building size of the specific clusters.

The above-mentioned approach is based on the assumption that only multi-spectral data and laser data are available. However, often a road map is available, which may help in obtaining better spatial partitioning results. If so, land-use boundaries could be improved by incorporating the road network in the spatial partitioning process. A corresponding surface model would be as shown in Figure 5.28. To consider roads as a separate land-use class and prevent the possible mixing of roads and clusters, we can specify a narrow transit zone along both sides of roads (e.g. 5 m). To avoid the transit surface across the road of a large transit zone from a nearby cluster, the solid cores of clusters are subtracted first by means of a mask. This mask is created by applying a morphological dilation operation to road pixels, using a circular structuring element with a reasonable radius (e.g. 40 m when the transit zone width is 60 m). The relationship between road width ($Width_{Road}$), transit zone width for roads ($Width_{Road\_Transit}$), radius of the circular structuring element ($Radius_{SE}$) and transit zone width for clusters ($Width_{Cluster\_Transit}$) should be:

$$Width_{Road} + 2 \times Width_{Road\_Transit} + Radius_{SE} \leq Width_{Cluster\_Transit}$$

When the minimum road width is 15 m, the proposed surface model is as presented in Figure 5.28.

5.7 Extraction of image-object properties at the land-use level

When spatial units for land use are extracted, a number of land-use-related properties can be derived accordingly based on these spatial units. The following properties are considered to be land-use-related features that play an important role in land-use reasoning and identification. For instance, building density, floor area ratio and green coverage ratio can be derived for each land-use image object; these are directly associated with definitions of several land-use types. These and other properties can be derived from image objects when their spatial units are determined.

5.7.1 Numerical and categorical properties

In the following, several meaningful numerical and categorical properties are listed, which can be derived for each land-use image object from image and laser data and can play an important role in land-use classification.
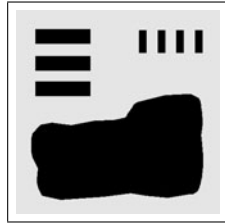
Figure 5.19: Three clusters.
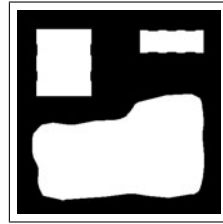
Figure 5.20: Cluster cores.
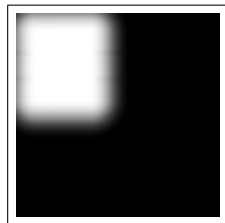
Figure 5.21: Transit zone: cluster 1.
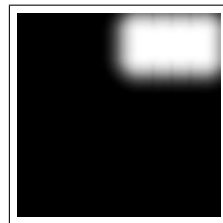
Figure 5.22: Transit zone: cluster 2.
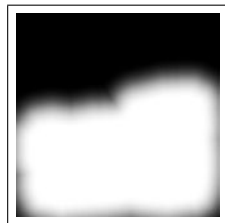
Figure 5.23: Transit zone: cluster 3.
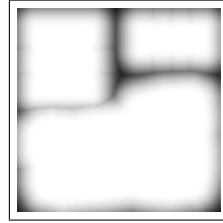
Figure 5.24: Transit zone: integrated.
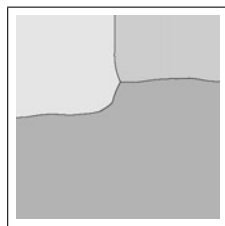
Figure 5.25: Spatial partitioning using the watershed algorithm.
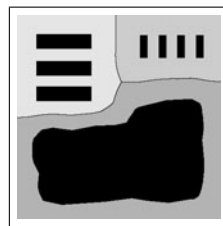
Figure 5.26: Three clusters and the corresponding partitions.
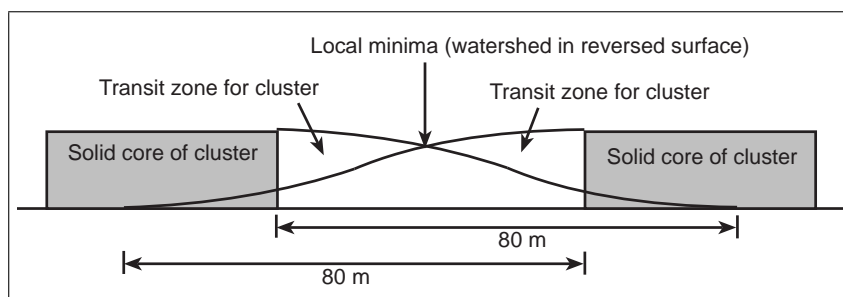
Figure 5.27: Profile of a proposed surface model for delineation of cluster land-use units without using road map.
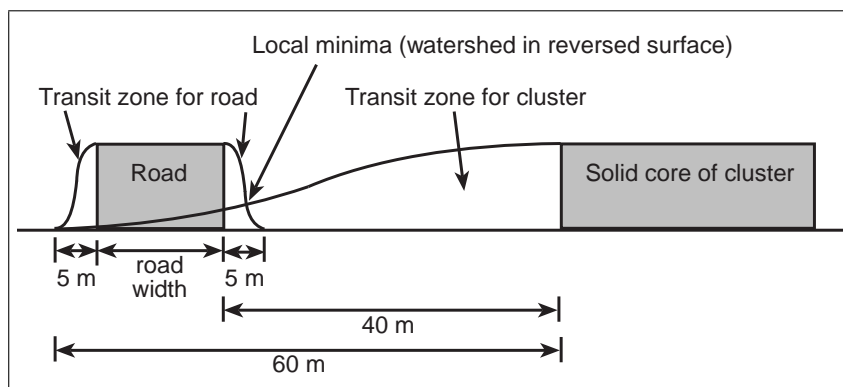
Figure 5.28: Profile of a proposed surface model for delineation of cluster land-use units by using road network.

- Type and proportional composition of land-cover objects a land-use object contains
- Number of buildings
- Average building size
- Average building height
- Building density

$$\text{Building density} = \frac{\text{Total area of buildings}}{\text{Size of the spatial unit}}$$

- Floor area ratio (FAR)

$$\text{FAR} = \frac{\text{Total area of building floor space}}{\text{Size of the spatial unit}}$$

- Green coverage ratio (GCR)

$$\text{GCR} = \frac{\text{Total area of green space}}{\text{Size of the spatial unit}}$$

- Water coverage ratio (WCR)

$$\text{WCR} = \frac{\text{Total area of water surfaces}}{\text{Size of the spatial unit}}$$

- Open-space coverage ratio (OCR)

$$\text{OCR} = \frac{\text{Total area of built-up area} - \text{Total area of building footprints}}{\text{Size of the spatial unit}}$$

(Note: built-up area is the complement of vegetation and water in a spatial unit)

5.7.2 Geometric properties

Location, size, shape and orientation are the geometric properties of an image object. These geometric properties can be described by several indicators (van der Heijden, 1994; Shufelt, 2000) as introduced in Section 4.5.

5.7.3 Structural properties

The spatial distribution of land-cover objects over the space of a land-use object is an essential element that can be derived based on the geometric properties of land-cover objects, as presented in Section 4.5. These geometric properties can also be used to identify land-use objects and can be treated as structural properties of land-use objects.

Several structural indicators, which are useful measures for the spatial distribution of specific land-cover objects in a land-use unit and can be extracted directly from images, are proposed as follows. Spatial coverage ratio (SCR), spatial mixture ratio (SMR), and spatial bias ratio (SBR) are useful for characterising the spatial distribution of features and determining if further subdivision of a land-use unit is required.

- Spatial coverage ratio (SCR)

$$\text{SCR} = \frac{\text{Size}_{\text{Convexhull}}}{\text{Size}_{\text{Spatial unit}}}$$

SCR indicates the spatial distribution of a certain land-cover feature in a spatial unit.

- Spatial mixture ratio (SMR)

$$\text{SMR} = \frac{\cap_s \,{}^n_{i=1}(\text{Convex hull}_i)}{\cup_s \,{}^n_{i=1}(\text{Convex hull}_i)}$$

SMR measures the degree of overlay in the spatial distribution of different features and can be used to check out if different types of features are mixed in a space.

- Spatial bias ratio (SBR)

$$\text{SBR} = \frac{2 \cdot |\text{Centre}_{\text{LC-O}} - \text{Centre}_{\text{LU-O}}|}{\text{EquivDiameter}_{\text{LU-O}}}$$

SBR is calculated as the distance between the gravity centre of land-cover objects ($\text{Centre}_{\text{LC-O}}$) and the gravity centre of the land-use object ($\text{Centre}_{\text{LU-O}}$) divided by the equivalent radius of a spatial unit, and can be used to check out if land-cover objects are equally distributed over a space or concentrated in certain parts of space.

5.8 Land-use classification

Land-use classification is based on the delineated spatial units and their properties, as described in the previous sections. The possibilities of a spatial unit belonging to defined classes are evaluated based on fuzzy membership functions or probabilities of each end-member class – a method called per-object fuzzy classification. Fuzzy membership functions are often used to express existing knowledge on relationships between selected features and defined classes. Probabilities can be obtained by training or learning from samples. Since not all features can be derived from images and laser data, in this research fuzzy membership functions and probabilities are constructed based on features that are extracted from these data. Many social-economic features such as population may be useful and can be included by integrating remote sensing and GIS. In this research, only those features and measures that can be derived from high-resolution multi-spectral images and laser data are under investigation.

5.8.1 Selected features and their associations with end-member classes

Features that make each land-use class distinguishable from other classes vary, and are very much related to the local settings. Therefore it is necessary to select a number of indicators that are robust in a general sense, such as building density and percentage of green space, and can be derived from image analysis. A list of such indicators and their associations with land-use classes is presented in Table 5.3.

5.8.2 Land-use classification

Based on rich features that can be derived from image and laser data as described earlier in this chapter, many classifiers can be applied in land-use classification. For instance, the fuzzy classifier, maximum likelihood classifier and neural network classifier are suitable classifiers. For unsupervised classification, the tree-classifier, k nearest neighbour, fuzzy c-means etc. are candidate classifiers. To incorporate knowledge in land-use identification and make it more robust in different urban areas, the fuzzy classifier is used, since the general knowledge used for designing the fuzzy membership functions is not too sensitive to different locations. Fuzzy membership functions can be adjusted by local knowledge obtained using samples from the particular site.

5.8.3 Summary

In this chapter, a logical design for object-based land-cover and land-use classification is proposed and discussed. It consists of three steps: land-cover classification, land-use unit reasoning and delineation, and land-use classification. The proposed approach incorporates per-pixel image processing techniques and per-object techniques in different stages. Many per-object features can be derived in addition to those that can be extracted by per-pixel approaches. Extracted objects can be directly compared with definitions of land-cover classes and land-use classes based on the characteristics they share. The spatial extent of such objects can readily be presented in a raster format in a GIS and can be converted to vector representation if necessary. Another feature of this approach is that all types of information are transparently preserved for each object, for example the average DN values for each type of raw data such as spectral values, height data, fuzzy membership value for each feature, characteristics used in land-cover classification, spatial clustering and land-use classification. Such detailed information provides useful sources for quality assessment and uncertainty analysis and allows different users to choose different features and modify related parameters to run the system again in order to acquire desired outputs. Based on the object-based land-cover and land-use classification schema proposed in this chapter, detailed case studies for land-cover classification are given in Chapter 6, structural analysis and the extraction of spatial units of urban land use are presented in Chapter 7, land-use classification is provided in Chapter 8, and quality assessment and uncertainty analysis are discussed in Chapter 9.

Table 5.3: A list of indicators and their associations with land-use classes

Indicators	Residential	Commercial	Industrial	Transportation	Recreational	Non-urban
Number of buildings	⊕	⊕	⊕	⊙/⊖	⊙/⊖	⊙/⊖
Type and proportion	B: ⊕ G: ⊙/⊕ W: ⊙	B: ⊕ G: ⊙ W: ⊙/⊖	B: ⊕ G: ⊙ W: ⊙	B: ⊙/⊖ G: ⊙/⊖ W: ⊖	B: ⊙/⊕ G: ⊕ W: ⊙/⊕	B: ⊙ G: ⊙/⊕ W: ⊙/⊕
Average building size (m^2)	50 - 2000	200 - 2000	> 2000	0 - 500	0 - 2000	0 - 2000
Building density	med. - high	high	med.	very low	low	low
Floor space ratio	0.5 - 10	1 - 20	0.2 - 2	< 0.1	< 0.2	< 0.5
Green coverage ratio	0.1 - 0.8	0 - 0.3	0 - 0.5	0 - 0.3	> 0.3	0 - 1

Note: ⊕ indicates positive proportional

⊖ indicates negative proportional

⊙ indicates the feature can be absent or discarded

B: building

G: green space or vegetation

W: water surfaces

Chapter 6

Object-based land-cover feature extraction*

6.1 Introduction

Exploiting remote sensing in urban areas has been a challenge for quite some time because of the complexity and fragmentation of objects and the combination of man-made features and natural features. Airborne laser altimetry data offer possibilities for feature extraction and spatial modelling in urban areas. There are many approaches for deriving buildings and other features reported in literature (Brunn and Weidner, 1997; Hug and Wehr, 1997; Lemmens et al., 1997; Axelsson, 1999; Haala and Brenner, 1999; Haala and Walter, 1999; Morgan, 1999). However, there are many cases, where it is still difficult to extract particular features by using these approaches – for instance, in an urban area where many roads are raised above ground level, with special characteristics similar to those of buildings (surface profile and spectral reflection, etc.). The reported approaches all seem to have their short-comings for building extraction in such a complicated urban context. The proposed object-based approach that we have developed tries to extract buildings through reasoning in a slice-based layer space. In the proposed approach, the DSM from the laser scanning in raster format is segmented into slices at 1 m increments in elevation. The resulting image regions of each slice are then labelled and treated as image objects. Hence, a number of properties can be derived based on labelled segments (image objects) such as location (centroid), size, shape, orientation. These properties are used for reasoning in the layer space. The layer space is defined by elevation at 1 m intervals as the X-axis and the properties of an image region (as they change throughout the slices) as the Y-axis. Image objects are linked and reasoned on vertically. A tree structure is created using links between segments throughout the different layers. Reasoning is based on the patterns of these properties on the paths

---

*This chapter is based on the following papers: Zhan et al. (2002b) and Zhan et al. (under peer review (1)).

of each branch of the search tree in the layer space. Several experiments have been performed in both study areas, southeast of Amsterdam and Ravensburg, Germany, based on the proposed approach. The approach is very promising, also for features other than buildings. Theoretical considerations, the detailed description of the major steps, as well as experimental results, are presented in this chapter. Also the methods of extracting other land-cover classes such as vegetation and water surface are described in this chapter.

6.2 Object-based building extraction

One of the problems in automatically extracting buildings from the DSM lies in discriminating between buildings and other protruding man-made structures such as flyovers and driveways (see e.g. Brunn and Weidner (1997); Hug and Wehr (1997); Axelsson (1999); Haala and Brenner (1999); Shufelt (2000)). Instead of trying to solve the problem by pixel-based analysis of the DSM, we study the change in properties of image objects in elevation slices. We slice the DSM at a fixed vertical interval (1 m in our test data) to obtain image objects at various levels of elevation, which are then subjected to reasoning. The underlying assumption is that for a building certain properties of its image object hardly change from one level to the next (see Figure 6.1). In the present study we detect buildings based on two properties, i.e. vertical change in size of an image segment and shift of its centre of mass. To this end, we have to link the image objects at the different layers by a tree structure. The degree of change from level to level also permits the production of uncertainty estimates of extracted buildings. We have tested the approach using high-resolution laser data of our Amsterdam and Ravensburg test sites.

6.2.1 Semantic and context analysis

Based on the characteristics of a building such as vertical wall, size and building material as discussed in Chapter 5, we could try to describe a building by using common-sense knowledge.

From the geometric perspective

- Size

  Size of a building as projected in a plane should be larger than 10 m^2 and smaller than 5000 m^2:

$$MF_{Size}[O_i, Building] = \left\{ \begin{array}{l} 1, \ O_i \in [10, 5000] \text{ m}^2 \\ 0, \ \text{otherwise} \end{array} \right.$$

- Height

Figure 6.1: Profile of real world (a), laser image (b) and profile of image segments for building reasoning (profile of segments at intervals of 1 m (c).

Buildings should be at least 3 m or more above surrounding ground:

$$MF_{Height}[O_i, Building] = \begin{cases} 1, \ O_i \in [3, 300] \text{ m} \\ 0, \ \text{otherwise} \end{cases}$$

Height information is used intensively in this research. More detailed uses of height information obtained by laser scanning are provided in later sections.

- Vertical wall

Most buildings should have vertical walls to support the building structure. The features used for describing vertical walls can be reasoned based on the size difference between image regions obtained at different elevation slices at the same location. Same size or similar size of such vertically adjacent slices indicates the existence of vertical walls. The locations of the centre of mass of such slices can also be used for reasoning on possible vertical walls. In the case of vertical walls, the centre of mass of vertically adjacent slices should not have moved or should show only a minor movement. The first measure is robust for building extraction. However, there may potentially be exceptions when such slices are obtained at the terrain layers. There the proportional size

difference may also be very small and very similar to the situation found with buildings, because the size of such image regions is very large. This potential problem can be avoided by checking the size of the image region; if the image region is too large, it will be rejected as a building / will not be identified as a building. The second measure (i.e. height) can also be used to largely avoid such cases. The shortcoming of the second measure is that it may not be able to distinguish a tree canopy from a building since the centres of mass of such image regions should not move across slices. Therefore, both measures should be included for building reasoning in order to reduce possible uncertainties, which may be caused by applying one of them alone.

From the building material perspective

Buildings are constructed using various building materials such as concrete, brick, iron and steel, wood. However, most building roofs in our test sites are tiled roofs and corrugated roofs using materials such as concrete, brick and metal. This implies that building roofs are solid and that building roofs should not, in general, contain vegetation or water.

- Normalised difference vegetation index (NDVI)

  NDVI can be used to indicate whether a derived object contains vegetation. A small NDVI value indicates absence of vegetation, which in turn refers to rock, bare soil or concrete, such as roofs and roads. A large NDVI value suggests the existence of vegetation, which indicates the existence of trees or lawn in urban areas.

  $$MF_{NDVI}[O_i, Building] = \left\{ \begin{array}{l} 1,\ NDVI(O_i) \rightarrow low \\ 0,\ NDVI(O_i) \rightarrow high \end{array} \right.$$

  In implementation, we first obtain two cluster centres by using the fuzzy c-means algorithm (fcm) from the histogram of $NDVI$ based on the consideration as discussed earlier.

  $$Cluster\_cen = fcm(NDVI, 2)$$

  A Z-shape fuzzy membership function (zmf) is then constructed by using two cluster centres, $min(Cluster\_cen)$ and $max(Cluster\_cen)$, as parameters to obtain a membership function (MF) value for each pixel, indicating the degree to which a pixel is likely to be part of buildings.

  $$MF_{NDVI}[p, Building] = zmf(NDVI, min(Cluster\_cen), max(Cluster\_cen))$$

- Surface compactness

  Surface compactness can be reasoned based on the difference between the distance measured by the first laser pulse and the second pulse, by subtracting the DSM2 from the DSM1. A smaller dDSM value indicates a 'harder' surface, which in turn refers to compact surfaces such as roofs and roads. A larger dDSM value suggests a 'softer' surface, which likely refers to vegetation, particularly trees. We may expect large dDSM values also at the building edges. Such edges can largely be eliminated by applying gray-scale morphological operations such as opening to the dDSM, since such edges are thinner (width of one or two pixels) than regions containing trees.

  $$dDSM = DSM1 - DSM2$$

  The MF value of an object O_i is generally classified as a building.

  $$MF_{Solidness}[O_i, Building] = \begin{cases} 1, \ dDSM(O_i) \rightarrow small \\ 0, \ dDSM(O_i) \rightarrow large \end{cases}$$

  In our cases, we first obtain two cluster centres by using the fuzzy c-means algorithm (fcm) from the histogram of $dDSM$, which can be regarded as a data-driven approach, in order to find parameters that fit the location situation without human intervention.

  $$Cluster\_cen = fcm(dDSM, 2)$$

  An Z-shape membership function (zmf) is then constructed by using two cluster centres, $min(Cluster\_cen)$ and $max(Cluster\_cen)$, as parameters to obtain a MF value for each pixel, indicating the degree to which a pixel is likely to be part of buildings.

  $$MF_{Solidness}[p, Building] = zmf(dDSM, min(Cluster\_cen), max(Cluster\_cen))$$

From an application perspective

From an application point of view, different disciplines may have different meanings or understanding regarding these spatial entities. In such cases, additional specifications are required at this stage. In applications such as landscape architecture, urban design or transportation engineering, where the building ground floor and forecourt should be included as part of the buildings, users can set a loose condition, checking the size differences between two height layers of the same building. In other applications, such as in the extraction of number of dwellings or population assessment, the upper parts of buildings are essential and the forecourts tend to be excluded. In this case, users can set more restricted conditions, checking the size differences between several layers.
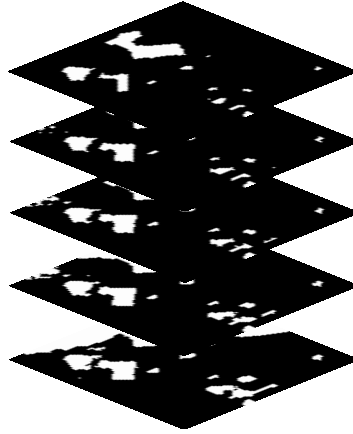
Figure 6.2: Vertical image segmentation of laser data.

6.2.2 Formation of image objects and their properties

The result is a set of binary images as illustrated by Figure 6.2. Next, the image segments (4-connection pixels) are uniquely labelled per image, thus obtaining identifiers of the image objects. To test our concept we consider two properties that we expect to be very relevant: size and location.

The size of an object is calculated as the actual number of pixels of the segment by using Formula 4.1. The location is computed as the centre of mass of the segment by using Formulas 4.2 and 4.3.

The linking of image objects in a tree structure is accomplished in a table. The first column records the identifier (ID) of a segment in a binary image at the lowest layer, and the following columns record labels of its linked segments in the images at higher layers. As shown in Table 6.1, the table provides information about the tree structure and links, as well as recording other properties derived for each path. The columns indicate the layer sequence from lowest to highest (-5 m to 50 m in the case of our Amsterdam test site). A row indicates a branch of search paths. The IDs in the table provide explicit links to corresponding image regions at the specific layers as sequentially sorted by the columns. Extracted sizes of linked image objects are recorded in a table such as shown in Table 6.2. The associated segments are identified by their position in the multi-layer grid. Other tables are generated in a similar manner for the size differences and location shifting of the image regions between adjacent layers vertically at a location (row, column in image). Please note that record nos. 270 and 300 presented as examples in these tables actually correspond to image objects **a** and **b** respectively, as indicated in Figures 6.3 and 6.4. These tables are used for detecting buildings in the 'layer space'.

Table 6.1: Region IDs of the linked regions segmented by using different elevation based on the DSM for each branch of search paths

Search path	Elevation values (in m) used to obtain image regions at different layers by segmentation based on the DSM						
	-5 m	-4 m	-3 m	-2 m	-1 m	0 m	1 m
. . .							
270	1	1	9	1	201	178	190
. . .							
300	1	1	9	1	14	5	5
. . .							

2 m	3 m	4 m	5 m	6 m	7 m	8 m	9 m	10 m	11 m	. . .
. . .										
228	191	177	164	147	133	124	116	0	0	. . .
. . .										
127	118	169	156	142	128	120	0	0	0	. . .
. . .										

Table 6.2: Region sizes (m^2) of the linked regions segmented by using different elevation based on the DSM for each branch of search paths

Search path	Elevation values (in m) used to obtain image regions at different layers by segmentation based on the DSM						
	-5 m	-4 m	-3 m	-2 m	-1 m	0 m	1 m
. . .							
270	9000000	8649993	7445947	4077599	607	567	541
. . .							
300	9000000	8649993	7445947	4077599	1951423	1533481	894627
. . .							

2 m	3 m	4 m	5 m	6 m	7 m	8 m	9 m	10 m	11 m	. . .
. . .										
512	511	503	500	490	466	351	305	0	0	. . .
. . .										
82706	63955	942	844	820	793	511	0	0	0	. . .
. . .										

6.2.3   Reasoning for building extraction

The layer space is defined by a plot of the property of an image object (e.g. size of the object or percentage change of size by going up one layer) against layer altitude. For every vertically linked image object a plot results. The reasoning is then based on the patterns of a property as obtained from all the paths of each branch of the search tree in the layer space.

Building identification

A fair assumption for the majority of buildings seems to be near-vertical walls within a certain height range, and this may help to distinguish them from flyovers, access ramps and the like. Accordingly, a requirement for identifying a building is to find image objects that have little deviation in size and only a small shift in the centre of mass between adjacent layers.

We consider the following indicators computed for layer i and layer $i + 1$:

$$\Delta_{Size_i} = \frac{Size_i - Size_{i+1}}{Size_i}$$

$$\Delta_{Loc_i} = \sqrt{(\overline{x}_{i+1} - \overline{x}_i)^2 + (\overline{y}_{i+1} - \overline{y}_i)^2}$$

The extracted Δ_{Size} (see Table 6.3) and $\Delta_{Location}$ (see Table 6.4) are also recorded in the same manner as Tables 6.1 and 6.2.

Reasoning in finding buildings

To identify a building, we need to define thresholds for the tolerated change between layers.

In a 2D image space Z^2, a segment (S) or an image object (O) can be identified as belonging to a building if it meets the following conditions:

$$S_i = \begin{cases} Building, \ \Delta_{Size_i}, \Delta_{Size_{i+1}} < T_{Size} \wedge \Delta_{Loc_i} < T_{Loc} \wedge S_i \in [10, 5000] \ \text{m}^2 \\ Else. \end{cases}$$

where T_{Size} and T_{Loc} are the thresholds for size difference and location shifting between two objects linked vertically. A small part of the laser range image used in the case study and the extracted buildings are shown in Figures 6.3 and 6.4 respectively. Buildings that are either lower or higher than the elevated roads are extracted properly. The plot of relative size differences versus elevation is shown in Figure 6.5 for two selected buildings (**a** and **b**). Building **a** is located on a lower part of the ground whereas building **b** rises from the level of an elevated road. The curve of building **a** shows a large size change from the bottom layer to the next layer of the segmentation, which reflects the fact that at the bottom layer segments are very

Table 6.3: The size differences of the linked regions segmented by using different elevation based on the DSM for each branch of search paths

Search path	Elevation values (in metre) used to obtain image regions at different layers by segmentation based on the DSM						
	-5 m	-4 m	-3 m	-2 m	-1 m	0 m	1 m
... 270	0.039	0.139	0.452	0.999	0.066	0.046	0.053
... 300	0.039	0.139	0.452	0.521	0.214	0.417	0.908
...							

2 m	3 m	4 m	5 m	6 m	7 m	8 m	9 m	10 m	11 m	...
... 0.002	0.016	0.006	0.020	0.049	0.247	0.131	0	0	0	...
... 0.227	0.985	0.104	0.028	0.033	0.356	0	0	0	0	...
...										

Table 6.4: The location shifting of the linked regions segmented by using different elevation based on the DSM for each branch of search paths

Search path	Elevation values (in metre) used to obtain image regions at different layers by segmentation based on the DSM						
	-5 m	-4 m	-3 m	-2 m	-1 m	0 m	1 m
... 270	12.23	47.86	282.76	1261.20	0.18	0.12	0.31
... 300	12.23	47.86	282.76	353.06	71.14	54.52	1226.9
...							

2 m	3 m	4 m	5 m	6 m	7 m	8 m	9 m	10 m	11 m	...
... 0.05	0.11	0.01	0.06	0.15	0.98	0.26	0	0	0	...
... 48.09	526.44	0.50	0.22	0.18	5.63	0	0	0	0	...
...										
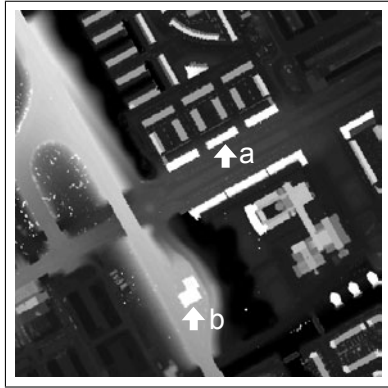
Figure 6.3: Original laser data.   Figure 6.4: Extracted buildings.
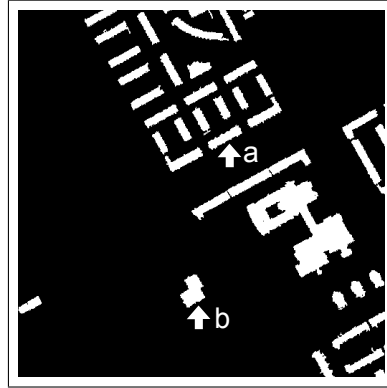
large in the case of horizontal ground (Amsterdam). The same holds for building **b**. For building **a**, the curve then drops to small for the next level and remains stable, indicating the near-vertical walls built on low ground. For building **b**, the decrease in size difference is slow while climbing up from the bottom to the elevated road. Once reached (at 2 m above sea level), the vertical walls cause the curve to remain stable.

In most cases, the above reasoning can differentiate well between buildings and other features. In cases where high trees are close to relatively low-rise houses, the above reasoning is unlikely to differentiate well between buildings and high trees (see Figure 6.1). Other information sources would have to be added (spectral information or first/last return of laser pulse) in order to refine the obtained image regions.

Additional reasoning in a building

- Building ground floor

The lowest segment along the vertical line which meets the criteria of a building will be treated as the ground floor (including forecourt) of the building.

- Building height

When a segment has been identified as the ground floor of a building, the difference between the average height of the ground surrounding the building ground floor and the DSM as masked by the ground floor region can be taken as the height of this building. The average height of the ground surrounding the building ground floor is obtained by computing the average height values of all pixels from the surrounding region of the building. The surrounding region of a building is derived by
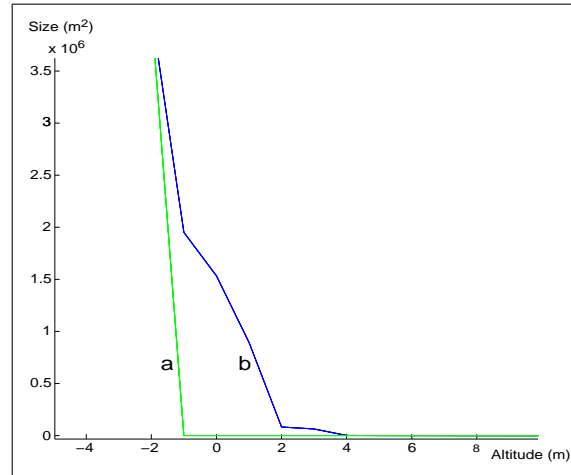
Figure 6.5: Plots of size differences for two buildings (**a** and **b**).
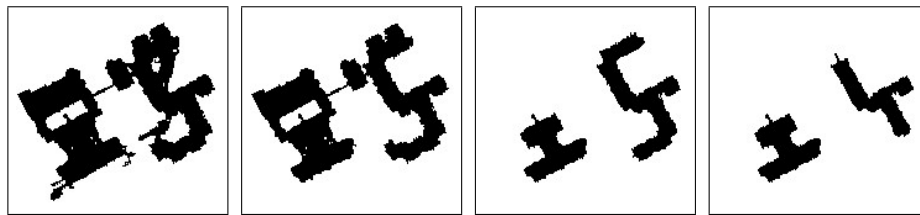
Figure 6.6: Outline differences of a building from ground floor to its upper layers (from left to right).

using a dilation operation with a small structuring element (e.g. a 5×5 'disk-shape' structuring element is applied in our cases) on the building region. A surrounding region of this type can be understood as a 5 m wide ring belt surrounding the specific building. Based on building height we can infer the number of floors, which is a useful property for later land-use identification.

- Outline of a building

Since the lower segments in particular may contain noisy pixels caused by adjacent vegetation or structures in gardens, as shown in Figure 6.6, it is up to the user or application objectives to decide from which layer to extract the outline of the building ground floor and forecourt. For a high-rise building, the upper layer may give the better outline.
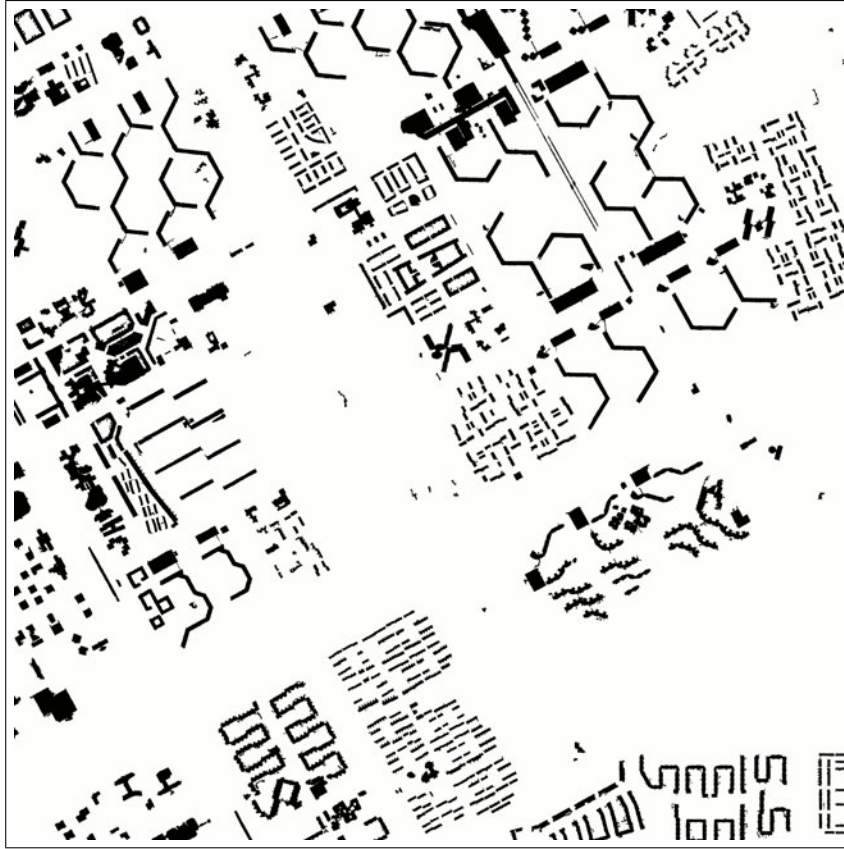
Figure 6.7: Buildings extracted by using 15 % as the threshold for checking the size differences.

6.2.4   Experimental results

Amsterdam test site

Based on DSM data acquired by the first-generation TopoSys laser scanner for the Amsterdam test site, most buildings were successfully extracted by checking only the size differences in the layer space, using 15 % as the threshold. Figure 6.7 shows the result of the building extraction from laser data, i.e. the DSM shown in Figure 2.1. Figure 6.8 shows buildings digitised from the large-scale base map, which was used as reference data. Per-pixel comparison of the extracted buildings and reference data is shown in Figure 6.9. The extracted building heights above ground level are shown in Figure 6.11, based on the building ground floor as shown in Figure 6.7.
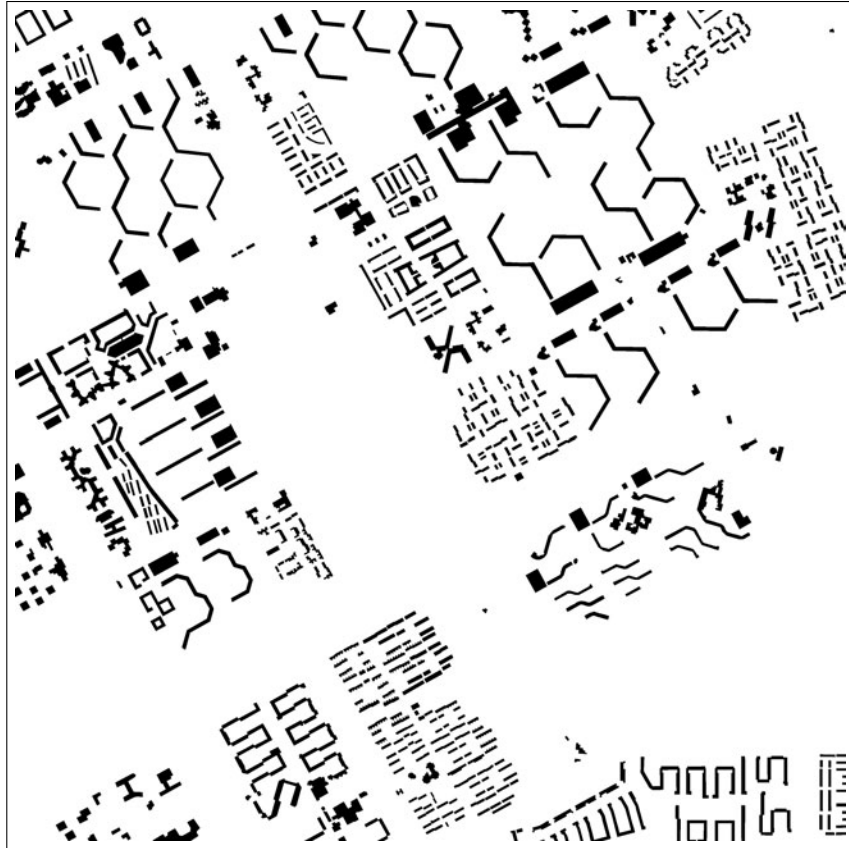
Figure 6.8: Buildings digitised from the base map of scale 1:1,000.
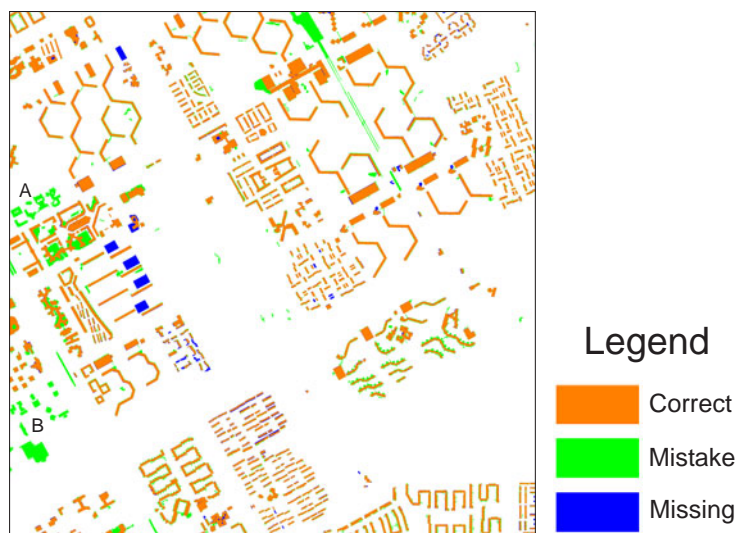
Figure 6.9: Comparison of extracted buildings (Figure 6.7) with the reference data (Figure 6.8).
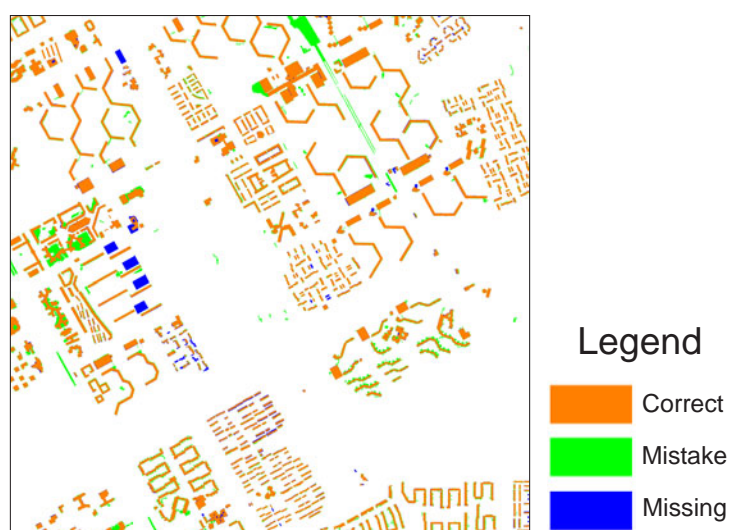
Figure 6.10: Comparison of extracted buildings after removal of changed buildings with the reference data (Figure 6.8).
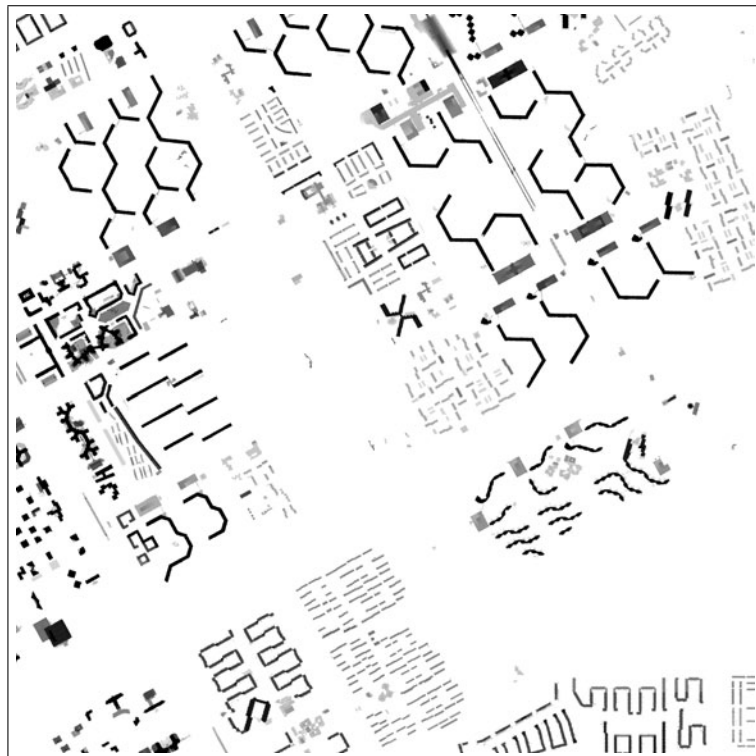
Figure 6.11: Building height above ground level produced from the DSM and outline of building ground floor (darker tone indicates higher building).

Table 6.5: Accuracy assessment of extracted buildings from Amsterdam test site based on the number of objects

	Building (extracted)	Building (from map)
Total Number	727	730
Correct Number	683	704
Mistake Number	44	26
Correct (%)	93.9 %	96.4 %
Mistake (%)	6.1 %	3.6 %

Ravensburg test site

Figures 6.12 and 6.13 show the results of the building extraction from laser data as shown in Figures 2.3 and 2.4, and from multi-spectral data as shown in Figures 2.6 and 2.5, by applying different thresholds. Figure 6.14 shows buildings delineated manually by visual interpretation based on images; This is used as reference data.

6.2.5   Quality assessment

Amsterdam test site

For the sake of comparison, we created a 'reference image', which contained exclusively buildings (derived from image analysis and edited with reference to the 1:1000 scale cadastral maps). Accuracy assessment was made based on image-to-image comparison between the result of building extraction and the ground truth as shown in Table 6.5. The total number of buildings is different owing to different interpretations as to what a building is (e.g. the map did not include the metro stations and some other small buildings, while the extraction result did). On the other hand, several parking garages have not been detected owing to the direct connection with raised roads. In general, high-quality results have been obtained, as can be seen from Table 6.5 and the map showing the extracted buildings, the existing buildings according to the reference map, and the differences between them as shown in Figure 6.9. We can notice that two groups of buildings (19 buildings) have been identified as mistakes, and are indicated by **A** and **B** in Figure 6.9. However, on a field visit these buildings proved to be correctly extracted. The reason is that these are new buildings and are not included in the reference. To make a fair comparison of the proposed method, we mask these new buildings that are not presented in the base map, as shown in Figure 6.10. Based on the modified data as shown in Figure 6.10, the correct figures for accuracy assessment as presented in Table 6.6 are slightly higher than the original assessment as shown in Table 6.5.

Figure 6.12: Building extracted by using 30 % as threshold for checking the size differences and using fuzzy membership functions based on other features, Ravensburg, Germany.

Table 6.6: Accuracy assessment of extracted buildings from Amsterdam test site based on the number of objects with the updated map

	Building (extracted)	Building (from map)
Total Number	708	730
Correct Number	683	704
Mistake Number	25	26
Correct (%)	96.5 %	96.4 %
Mistake (%)	3.5 %	3.6 %

Figure 6.13: Buildings extracted by using 50 % as threshold for checking the size differences and using fuzzy membership functions based on other features, Ravensburg, Germany.

Figure 6.14: Reference data prepared by visual interpretation and manual delineation, Ravensburg, Germany.
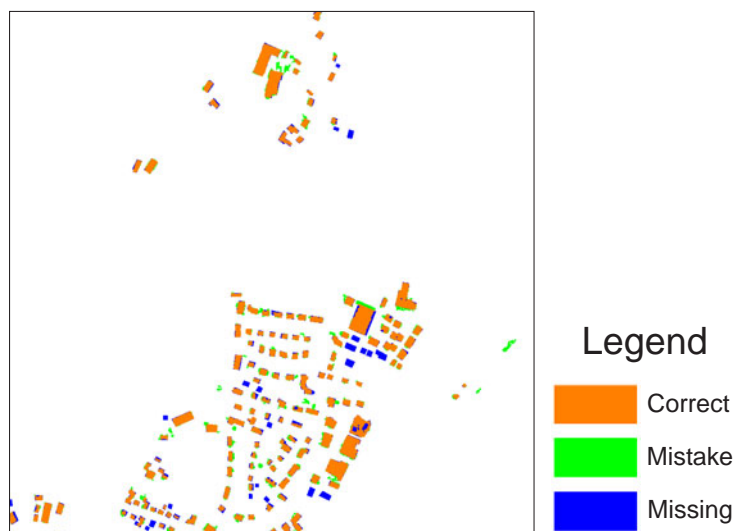
Figure 6.15: Comparison of extracted buildings as shown in Figure 6.12 with the reference data of Figure 6.14, Ravensburg, Germany.
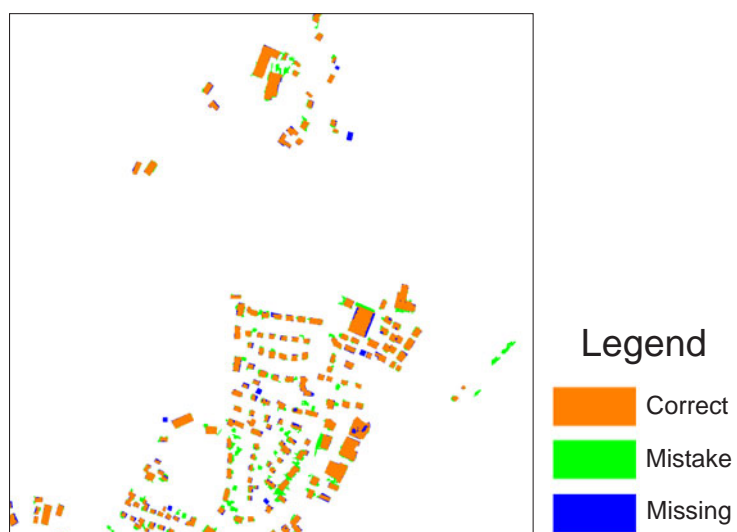
Figure 6.16: Comparison of extracted buildings as shown in Figure 6.13 with the reference data of Figure 6.14, Ravensburg, Germany.

Table 6.7: Quality assessment of extracted buildings from Ravensburg test site based on the number of objects

	Number of extracted buildings		Number of buildings from visual interpretation
Total Number	157	Total number	177
Correct Number	150	Correctly detected	154
Mistake Number	7	Not detected	23
Correct (%)	95.5 %	Correct (%)	87.0 %
Mistake (%)	4.5 %	Mistake (%)	13.0 %

Note: Four buildings are spatially separate in reference data, but they are merged in extracted results.

Ravensburg test site

For quality assessment, we created a 'reference image' containing exclusively buildings (obtained by screen digitising based on a high-resolution image). Accuracy assessment is made based on image-to-image comparison between the result of building extraction and the ground truth, as shown in Table 6.7. The total number of buildings is different because four buildings that are spatially separated in the reference data are merged in the extracted results. In general, high-quality results have been obtained, as can be seen from Table 6.7 and the maps showing the extracted buildings, the existing buildings according to the reference map, and the differences between them as shown in Figures 6.15 and 6.16.

6.2.6 Uncertainty assessment of extracted buildings from the Amsterdam test site

For the uncertainty assessment, we indicate whether the defined criteria for a building are met for each segment and store this in a table with the same structure as mentioned earlier (i.e. Table 6.1). Then we count the number of segments existing above building basements, and the number of segments that met the criteria we established for two adjacent layers. The uncertainty measure is expressed as the percentage of segments that met the criteria from all the segments that exist in the search path. The uncertainty assessment result is presented in Figure 6.17. Should several building branches exist above a basement, the average is computed. By using the proposed uncertainty measure, buildings that have the dominant characteristic (i.e. vertical wall), such as large and high-rise apartment buildings, are more certain; buildings that do not have this characteristic, such as low-rise buildings,
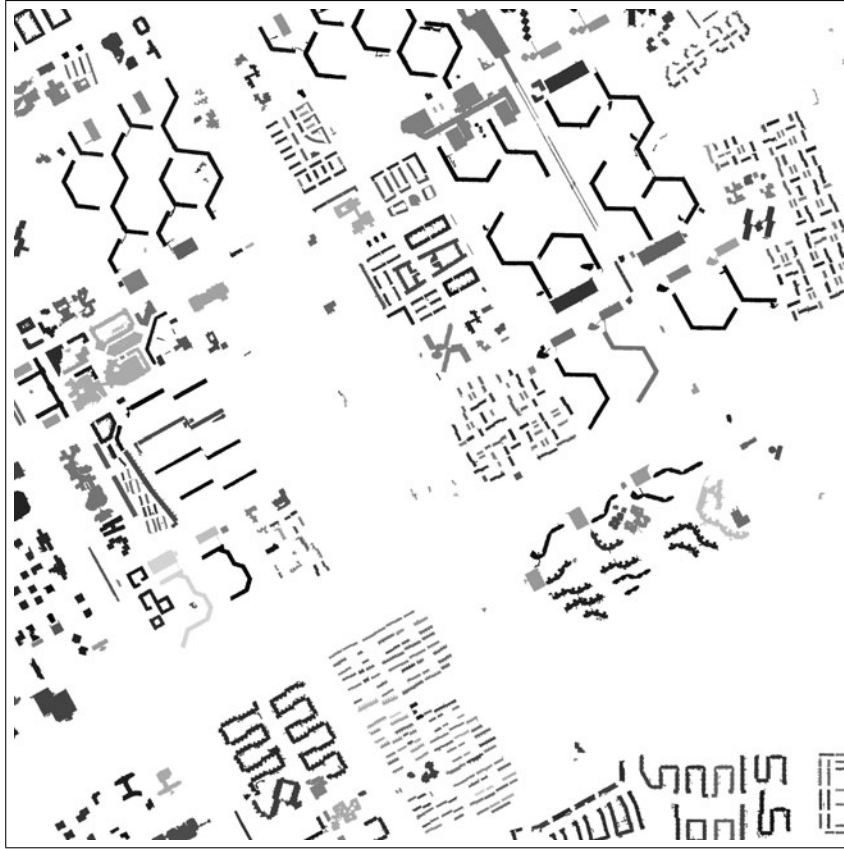
Figure 6.17: Uncertainty assessment result (lower tone indicates higher uncertainty, dark tone indicates higher certainty respectively).

buildings with gable roofs or buildings with multi-branches on the upper layers, are more uncertain, as can be observed from Figure 6.17. By providing such uncertainty assessment results, users can save time on quality inspection by concentrating on uncertain objects only.

6.3 Object-based green space extraction

Green space is extracted based on the normalised difference vegetation index (NDVI) from the IKONOS image, using the formula:

$$\text{NDVI} = \frac{\text{NIR} - \text{RED}}{\text{NIR} + \text{RED}}$$

To automatically or semi-automatically extract green space based on NDVI or in a so-called data-driven approach, we apply the fuzzy c-means algorithm for k=2 to NDVI values in the histogram space and obtain two cluster centres. As we know that vegetation should have higher NDVI values, non-vegetation such as built-up areas and water surfaces should have lower NDVI values. Therefore, two centres (C_{high} and C_{low}) can be used as estimators to represent vegetation and non-vegetation respectively. The fuzzy membership function for vegetation extraction is formulated by the S-shape function, using C_{low} and C_{high} as the bounded points. The NDVI image derived from the 4 m resolution IKONOS image of the Amsterdam test site is shown in Figure 6.18. The histogram of NDVI values and the formulated fuzzy membership function are shown in Figure 6.19. To avoid any additional distortion that may be caused by resampling from 4 m resolution to 1 m resolution to meet the resolution of rasterised laser data, we use the original DN values of the IKONOS image to obtain the NDVI and transform it into fuzzy membership values relating to two classes, vegetation and non-vegetation. Pixel resampling from 4 m resolution to 1 m resolution is then made based on fuzzy membership values by the proposed sub-pixel interpolation approach as introduced in Section 5.5. The extracted green space is presented in Figure 6.20, which is based on the fuzzy membership values at 1 m resolution, using 0.5 as the threshold. The extracted result is good for land-cover mapping and for computing the green coverage ratio. However, many small objects attributable to the fragmental distribution of the vegetation are considered noise, and this may increase the complexity by presenting many small objects such as vegetation in domestic gardens for land-use reasoning later on. When we consider small objects to be noisy or too small to be considered as public green space in land-use classification, we can remove them. Figure 6.21 shows the result obtained after removing objects smaller than 1000 m^2.

6.4   Object-based water surface extraction

In many cases, water surfaces can be extracted based on spectral information. However, when using multi-spectral data such as an IKONOS image alone, it is quite difficult to separate water pixels from pixels falling in the shadow areas of buildings and pixels in dark building roofs because of their similarity in spectral space, as illustrated and discussed in Chapter 5. Therefore we propose an integrated approach to extract water surfaces by using multi-spectral data and laser data to eliminate such mixtures.

6.4.1   Enhanced normalised difference water index (eNDWI)

For the extraction of water surfaces from multi-spectral data, the eNDWI was applied in this research (see details in Chapter 5). The eNDWI image derived from the IKONOS image is shown in Figure 6.22.

Extracted water surfaces as presented in Figure 6.24 were extracted based on the S-shape fuzzy membership function in the range 0 to the highest fuzzy cluster cen-
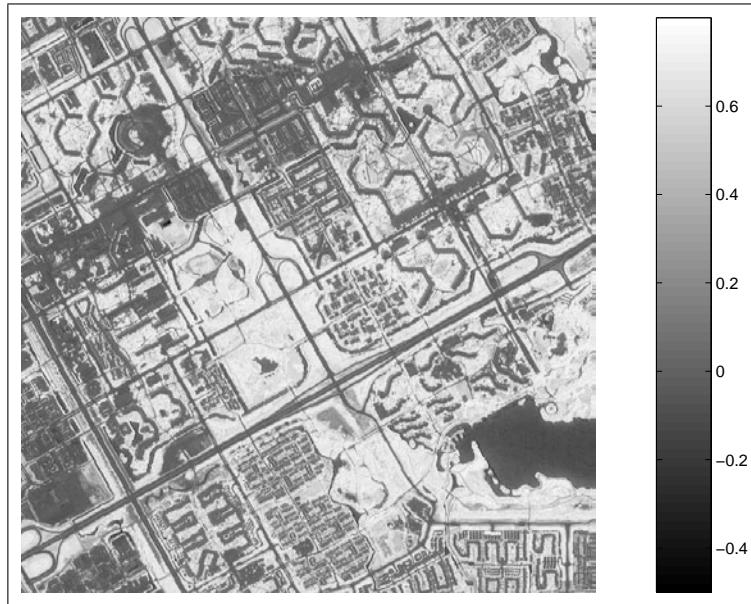
Figure 6.18: NDVI image (4 m resolution) derived from IKONOS image.
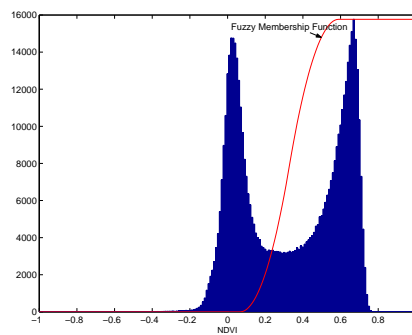
Figure 6.19: Histogram of NDVI image and fuzzy membership function.

Figure 6.20: Green space extracted based on NDVI using fuzzy membership function.

Figure 6.21: Green space after removal of objects smaller than 1000 m^2.
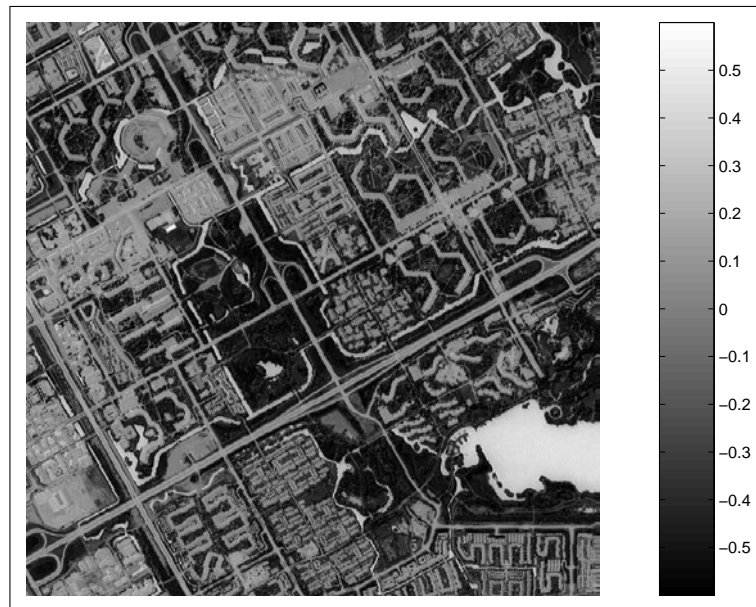
Figure 6.22: eNDWI image derived from IKONOS image.
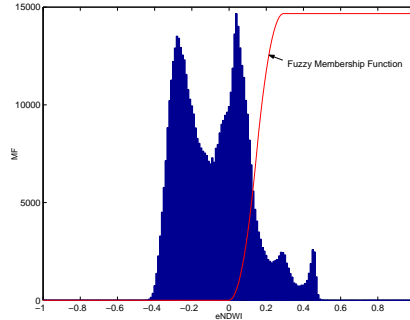
Figure 6.23: Histogram of eNDWI image and fuzzy membership function for water extraction.

tre (three clusters in the histogram using fuzzy c-means), as shown in Figure 6.23, and resampled from 4 m resolution to 1 m resolution using the proposed sub-pixel method. In this result, there may be other objects that have a spectral reflectance very similar to water surfaces, such as shadows and dark buildings. Such non-water objects will be detected and removed by using height information contained in the DSM.

6.4.2 Extraction of shadow areas and building relief displacement

Shadow areas in a scene are extracted from the simulation, using the DSM, sun angle azimuth and sun angle elevation as input. Building relief displacement caused by slightly oblique viewing in imaging is derived from simulation based on the DSM, using nominal collection azimuth and nominal collection elevation. These simulations are implemented in ArcView using the hillshade analysis for shadow and building relief displacement respectively. The simulated image for shadow areas appearing in the IKONOS image is made based on laser data using the sun angle azimuth (150.9920 degrees) and the sun angle elevation (58.17625 degrees), as shown in Figure 6.25. The simulated image for building relief displacement is made based on laser data using the nominal collection azimuth (200.9442 degrees) and the nominal collection elevation (69.52011 degrees), as shown in Figure 6.26. These meta data are attached to the IKONOS image provided by the company Space Imaging. The black areas in these simulated images are shadow areas, and building roofs that are caused by relief displacement in the corresponding images.

6.4.3 Removal of non-water areas

Extracted water objects possibly mixed with other objects are refined by masking shadow areas, building roofs caused by relief displacement, as well as buildings
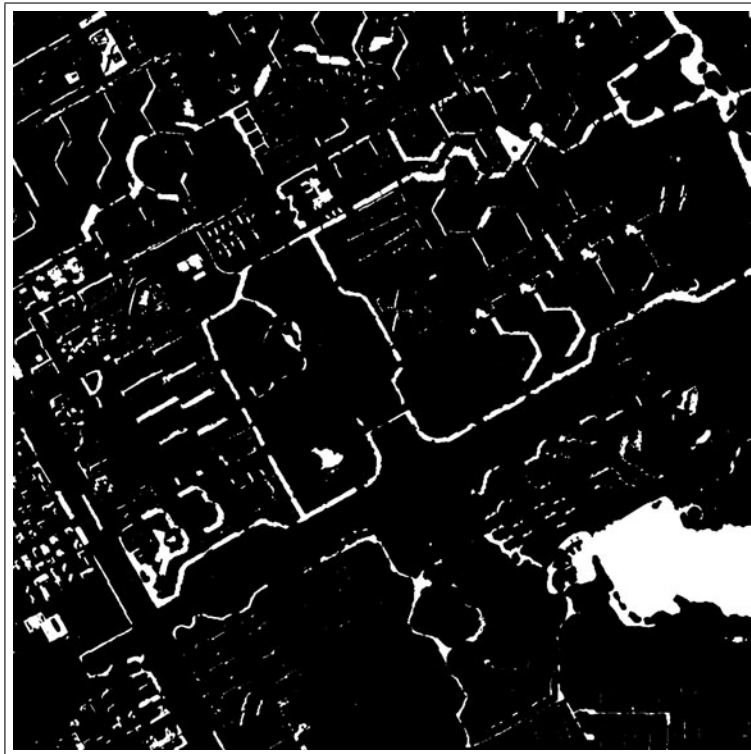
Figure 6.24: Water areas extracted based on fuzzy membership function and sub-pixel interpolation.
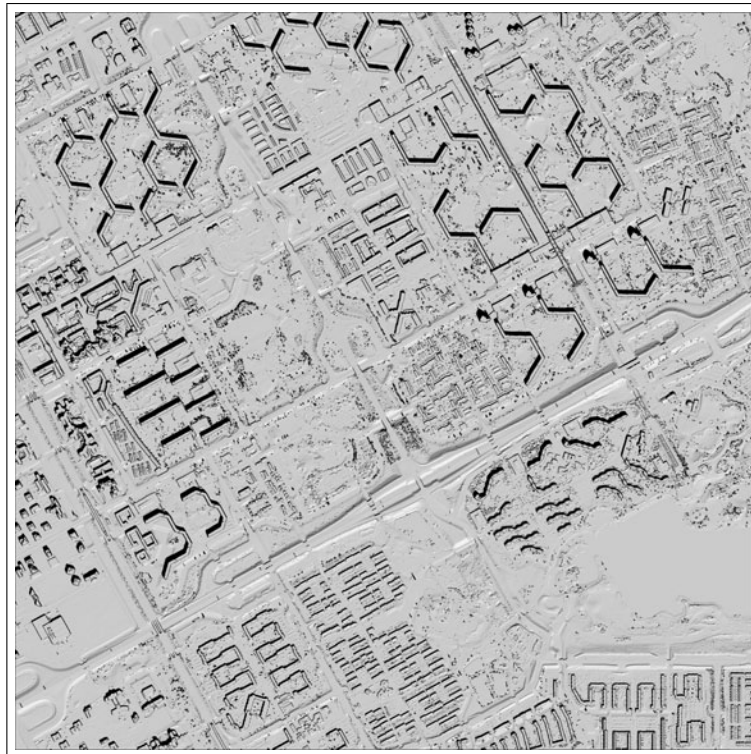
Figure 6.25: Simulated shadow areas of buildings by hillshade analysis based on laser data.
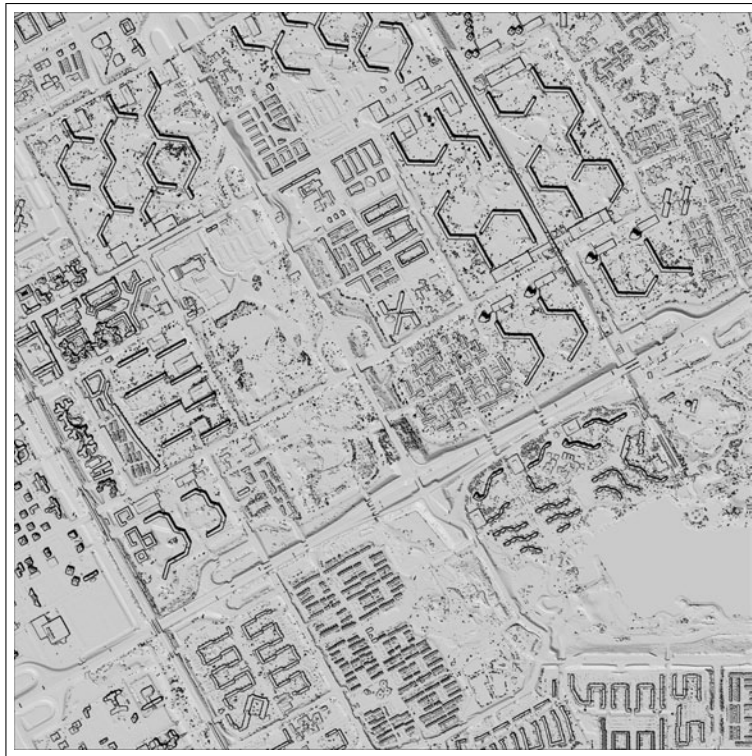
Figure 6.26: Simulated building relief displacement by hillshade analysis based on laser data.

Figure 6.27: Water bodies after removal of objects that are masked by simulated shadow and buildings.

themselves as extracted based on laser data. The refined water objects are presented in Figure 6.27.

There may be some unexpected objects existing among the water objects. The DSM is used again to check if most pixels in water objects are similar in height value. The refined result is used for land-cover mapping and for computing the water coverage ratio. We consider that many small water objects are noise or are too small to be presented as water bodies for land-use reasoning. Therefore we remove such small objects in the final stage. The final result, as shown in Figure 6.28 is obtained by removing objects smaller than 400 m^2 and those that have a standard deviation of height values (DSM) larger than 3.
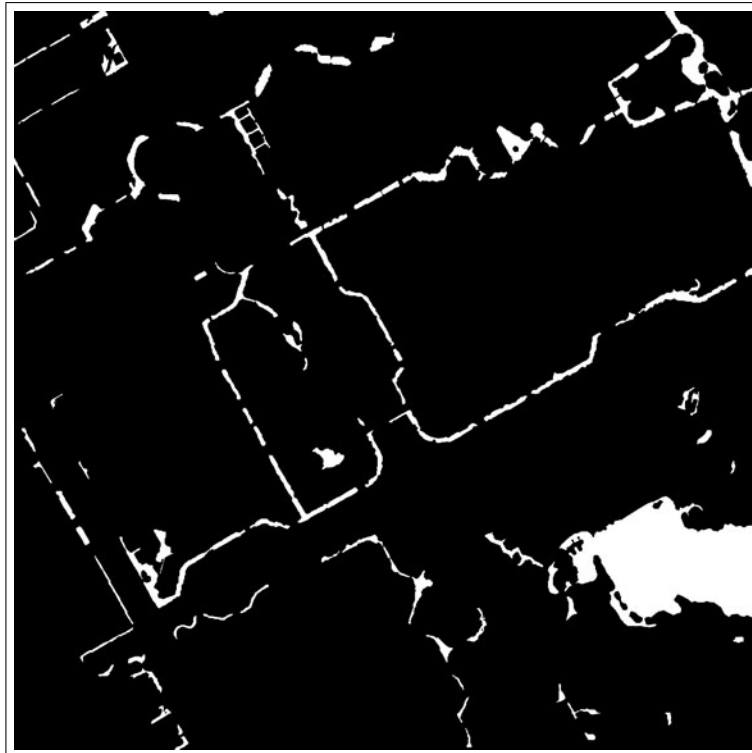
Figure 6.28: Water bodies after removal of objects that are smaller than 400 m^2 and those that have a standard deviation of height values (DSM) larger than 3.

6.4.4 Water surface extraction based on missing pixels from laser data

Water surfaces can also be extracted from laser scanning data. In some cases, smooth water surfaces may cause the mirror reflection of the omitted laser beam, which can result in not receiving a return signal. The water surfaces extracted from laser data are likely to give better identification and sharper boundaries than those based on multi-spectral information, as shown in Figure 6.29. To make sure those 'missing-value' pixels are parts of water surfaces, spectral information should be checked as to whether they have the spectral properties of water. It should be noted that noisy pixels and flight gaps should be removed in the data preparation phase.

The water surfaces extracted from laser data can be treated as reference data. We observe by comparing Figure 6.28 and Figure 6.29 that the water objects extracted from IKONOS and laser data are very similar. This is despite some changes that took place between acquiring the laser data in 1998 and capturing the IKONOS image in 2000 and despite some narrow canals that are missing from the IKONOS image owing to the coarser resolution.

6.5 Summary

The test results show that the proposed image-object-based approach is robust and reliable for building extraction for our purposes. It works well in a complicated urban context, such as the Amsterdam test site, where elevated roads have a similar profile to buildings. It also works quite well in a difficult area, such as the Ravensburg test site, where there are small buildings with gable roofs and high trees very close to buildings and the terrain is undulating. In such cases, multi-spectral data are needed for additional efforts in refining buildings extracted based on laser data.

It is relatively easier to extract the vegetation areas, using the NDVI. Water areas are very difficult to extract using multi-spectral data alone, because of the very similar spectral reflectance of shadow and dark buildings. Thus, an enhanced NDWI is proposed for water extraction in urban areas and this performed better in our test sites. In addition, buildings, the displaced building roofs caused by oblique image acquisition, and shadow areas are derived. These non-water objects are used to refine water objects that are derived using spectral information alone. Such additional data can be extracted based on laser data.

The highly successful extraction of land-cover objects such as buildings, vegetation, water surfaces and other built-up areas provides us with a very promising basis for structural analysis toward extracting spatial units of land-use and land-use classification, which will be presented in the coming chapters.
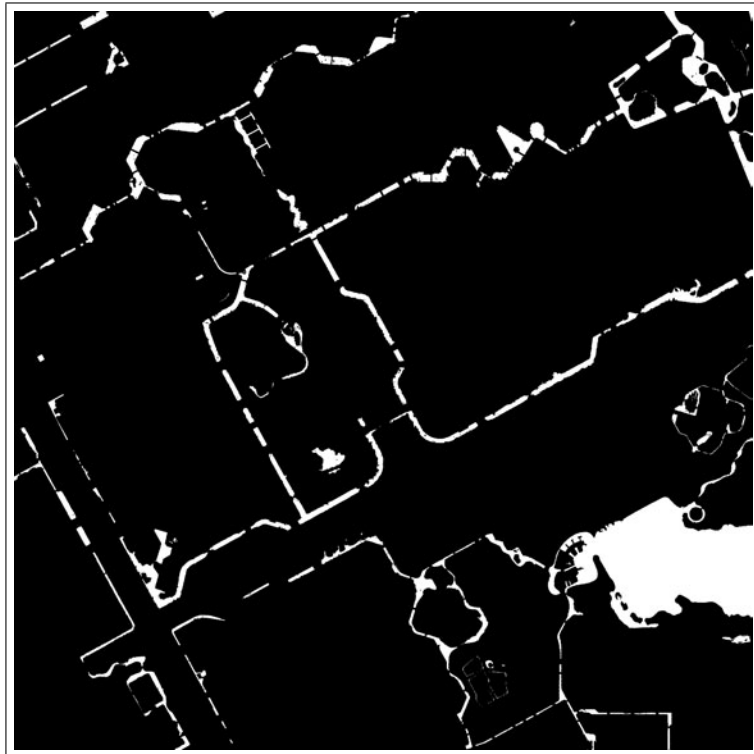
Figure 6.29: Water surface extracted based on missing pixels in laser data.

Chapter 7

Object-based structural analysis and spatial units of urban land use*

7.1 Introduction

Land-use mapping can be understood as the mapping of land features into a spatial partition of categorical land-use units in a 2D space. A land-use map is a spatial tessellation of categorical land-use types in a 2D reference space. This implies that the land-use classification process can be subdivided into two interactive and interrelated components: the spatial tessellation of categorical land-use types appearing in a given 2D bounded space, and the identification of land-use types for each spatial unit. These two components should be processed alternately in a parallel manner. By taking certain land-cover features into account in the first step, an approximate spatial partition can be made. The procedure continues by alternately taking into account other land-cover features and spatially partitioning until the given 2D space has been tesselated. When spatial tessellation is supported by evidence that suggests certain land-use functions may hold in a certain location and indicates their likely spatial extent in the given space, final land-use identification can be made based on the characteristics that each land-use spatial unit contains. Object-based structural analysis and extraction of spatial units are the central issues of this chapter. Following the logical design and discussions of Chapter 5, a number of techniques are investigated for spatial clustering and spatial partitioning in order to find spatial units for land use. Delaunay triangulation is deployed to acquire spatial proximity relationships between land-cover objects in the 2D reference space; mathematical morphology is applied to find the solid core of a spatial unit in 2D space; distance transformation and fuzzy membership function are used

---

*This chapter is based on the paper Zhan et al. (2002c).

together to model and create the spatial transit zones surrounding the solid core of each spatial unit. The watershed algorithm is proposed for deriving explicit boundaries between spatially adjacent land-use units. Experimental results are presented and discussed in this chapter. Land-use identification based on such spatial units will be described in Chapter 8.

7.2 Extraction of proximity relationship and the shortest links between adjacent land-cover objects

Land-cover classification leaves us with buildings, vegetation, water and other open surfaces. Spatial analysis of the land-cover objects is an essential step toward identifying land use. Proximity of objects is an important measure for finding spatial clusters. Delaunay triangulation applied to the raster image of a land-cover type (e.g. buildings) is a good tool for finding adjacent buildings and the shortest distance between them. To do so, we must eliminate triangle edges that link two pixels of one and the same object. The remaining edges indicate adjacent objects. Thus the shortest edge between two adjacent objects can be extracted for representing the proximity relationship and how close these objects are situated (i.e. proximity).

7.2.1 Delaunay triangulation and Voronoi diagram

Delaunay triangulation and its dual Voronoi diagram have been receiving increasingly attention because of the ability to produce tessellation of space (Gold, 1991, 1992; Gold and Edwards, 1992; Pilouk and Tempfli, 1992; Okabe et al., 1994; Li et al., 1999; Li and Huang, 2002; Estivill-Castro and Lee, 2002). Detailed mathematical formulation and description of Delaunay triangulation and the Voronoi diagram can be found in Okabe et al. (2000) and de Berg et al. (2000). Detailed description of the Quickhull algorithm applied in this research can be found in Barber et al. (1996). The Delaunay triangulation links up the natural neighbours in a point set by triangle edges; the edges of such triangles indicate the proximity relationship between linked points. If applied to the centres of pixels that represent image objects, we get triangle edges between adjacent pixels of one and the same image object and triangle edges that link up two pixels of adjacent image objects, the two pixels satisfying the natural neighbour criterion. Thus the shortest links between two adjacent image objects can be derived based on the length of the edges that link two adjacent image objects. In triangulation, we use the row and column number of a pixel to represent a point vector, as shown in Figure 7.1, and use the object ID as the ID of the point. To extract adjacent image objects, we deploy Delaunay triangulation for all pixels that constitute image objects such as buildings. Thanks to the properties of Delaunay triangulation, each triangle edge indicates proximal points (pixels), as shown in Figure 7.2.
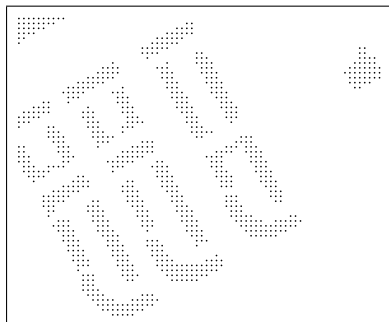
Figure 7.1: Pixels (points) embedded by image objects such as buildings.
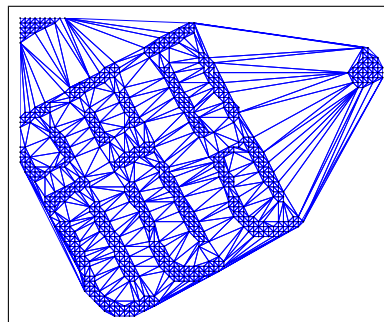
Figure 7.2: Delaunay triangulation deployed in all building pixels.

7.2.2 Extraction of proximity relationship between objects

Since each pixel has an ID that indicates the object to which it belongs (see object-based land-cover feature extraction described in Chapters 5 and 6), the internal edges that link pixels of the same object can be identified and removed. The remaining edges are links between adjacent objects, as shown in Figure 7.3.

7.2.3 Extraction of the shortest links between adjacent objects

By comparing the lengths of all edges linking two objects, we can easily determine the shortest edges that link adjacent objects. Figure 7.4 shows the shortest links between adjacent objects (buildings). The shortest links between adjacent objects provide useful information for the spatial clustering of objects, as discussed earlier in Section 5.6.2. A matrix is created which indicates adjacent buildings and the shortest distance between them. A detailed description of this approach can be found in Zhan et al. (2002c) and Zhan et al. (2002d).

7.2.4 Spatial clustering by checking the shortest links between adjacenct objects

Two simple spatial clustering examples are shown in Figures 7.5 and 7.6. They illustrate the result of clustering when using the shortest links as the only indicator and 40 m and 20 m as the thresholds respectively. The shortest links that are longer than the specified threshold are removed. The remaining edges are then used to indicate that their corresponding objects are believed to be parts of the same cluster.
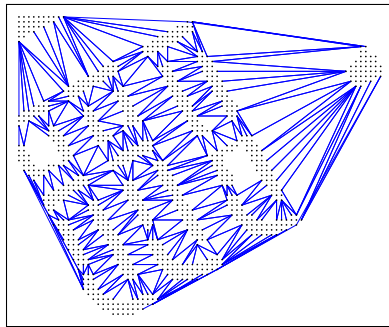
Figure 7.3: Edges between pixels of different buildings.
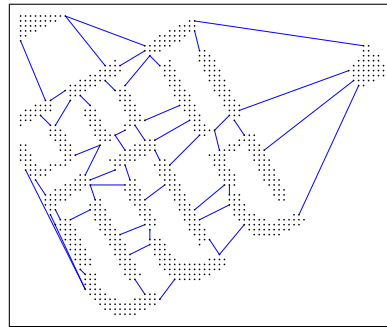
Figure 7.4: The shortest links between different buildings.
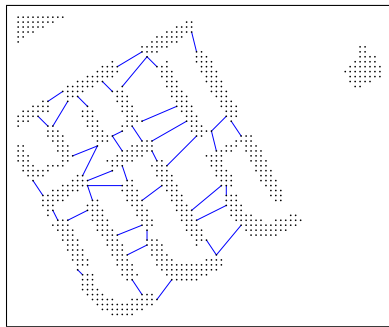
Figure 7.5: Adjacent buildings with links shorter than 40 m.
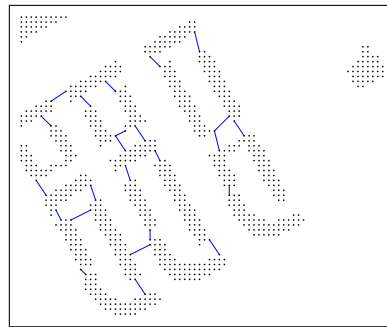
Figure 7.6: Adjacent buildings with links shorter than 20 m.
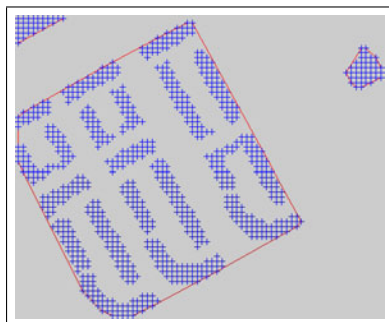
Figure 7.7: Convex hull of clustered buildings.

Figure 7.8: A raster presentation of clustered buildings.

7.2.5 Convex hull and representation of clusters

When objects are spatially clustered, the convex hull of each cluster can be extracted and delineated to indicated objects that belong to this cluster. An example of such a result, using the threshold of 40 m, is shown in Figure 7.7. Objects belonging to a cluster can also be presented by assigning different colours to objects that belong to different clusters in a raster (image). An example of such a result, using the threshold of 40 m, is shown in Figure 7.8.

7.3 Extraction of spatial clusters for land-use classification

To test the ideas of spatial clustering and similarity measures as discussed in Sections 5.6 and 7.2, we studied them in respect to the Amsterdam test site. This test site has one peculiarity. Some corridors that connect buildings in Amsterdam have been extracted as parts of extracted buildings in the previous stage (see Figure 6.7); thus the actual data indicates the spatial relation. To counteract this – as it may not occur in other places – and test our approach in more general circumstances, we use digitised buildings from the base map as test data for spatial clustering.

7.3.1 Extraction of the shortest links between adjacent objects

Buildings extracted in the previous stage, as presented in Chapter 6 (digitised buildings in this case), are used as indicators for spatial clustering in order to find spatial units of land use. The shortest links between adjacent buildings are extracted as shown in Figure 7.9, based on methods discussed earlier. The shortest distances
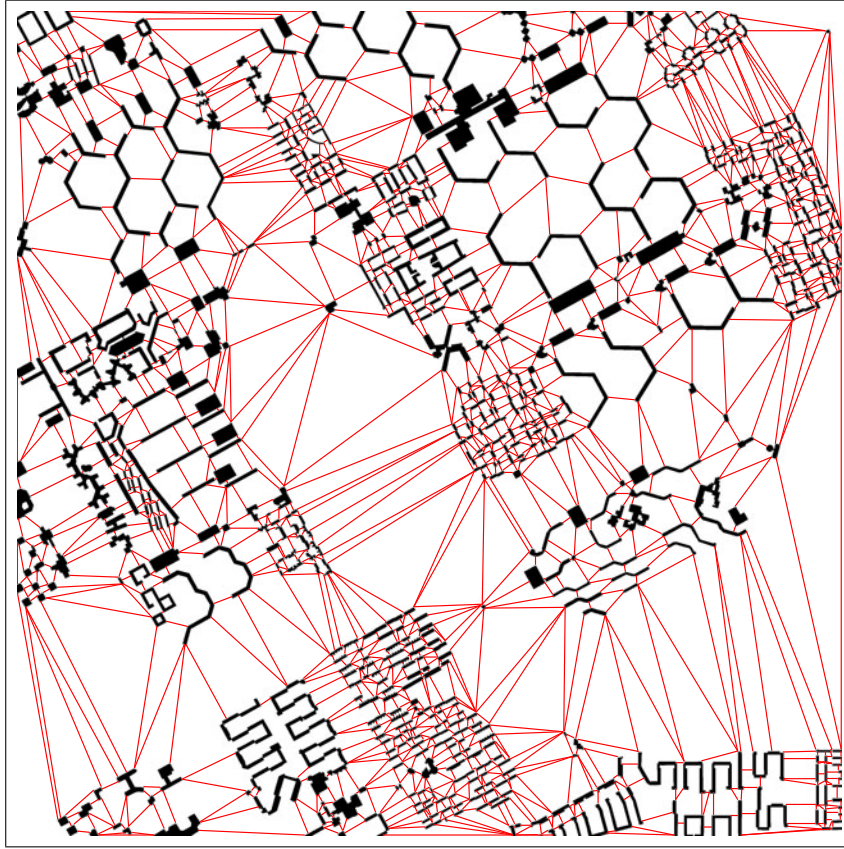
Figure 7.9: Linked buildings by the shortest distance between objects.

and the properties of linked objects are used for reasoning whether linked buildings belong to the same cluster.

7.3.2 Reasoning for spatial clustering based on the shortest links between adjacent objects

Finding spatial clusters in an urban area is much more complicated than illustrated by the simple examples in the previous figures. A crucial decision has to be made as to what is the best threshold that can be used to find spatial clusters, since it will have significant impact on the results. In order to find a way to obtain an optimised threshold, we need to observe a number of proposed cluster parameters and see how they react to a specific threshold. Thus we carry out the following experiments that

use the threshold as a variable in order to find out how good the results we obtain actually are and how cluster parameters may reflect such differences when different thresholds are applied.

The threshold (MaxDist) is defined as the largest possible distance between adjacent buildings considered to be in the same cluster (spatial unit of a land-use type). All links between adjacent buildings will be broken or removed if their length is greater than the threshold. The remaining links indicate that those objects should stay in the same cluster.

In order to find suitable measures that can be used to determine the optimal threshold, a sequential spatial clustering is done in a loop, using trial thresholds from 5 m to 250 m, spaced at 10 m intervals. For the Amsterdam test site, the 5 m threshold will generate a cluster for each building. The 250 m threshold will group all buildings in only one cluster. An optimised threshold must be somewhere between these two extremes. To determine the optimised threshold, we propose the following cluster-related measures: the number of spatial clusters, the maximum number of buildings in a cluster, the minimum number of buildings in a cluster, the number of isolated buildings, and the average number of buildings per cluster. In the following, we define and briefly explain these measures.

- Number of spatial clusters (NoCluster)

The number of spatial clusters is counted according to the number of separated clusters (no link between them) when a threshold is applied. When the threshold increases, the number of spatial clusters decreases.

- Maximum number of buildings in a cluster (MaxNoObjects)

The maximum number of buildings in a cluster indicates the largest number of buildings among all clusters. When the threshold increases, the maximum number of buildings in a cluster increases.

- Minimum number of buildings in a cluster (MinNoObjects)

The minimum number of buildings in a cluster indicates the smallest number of buildings among all clusters. When the threshold increases, the minimum number of buildings in a cluster also increases but not at the same rate as the maximum number of buildings in a cluster.

- Number of isolated buildings (NoIsolatedObjects)

The number of isolated buildings is taken from the number of clusters that consist of only one building. When the threshold increases, the number of isolated buildings decreases.

- Average number of buildings per cluster (AvNoObjectsPerCluster)
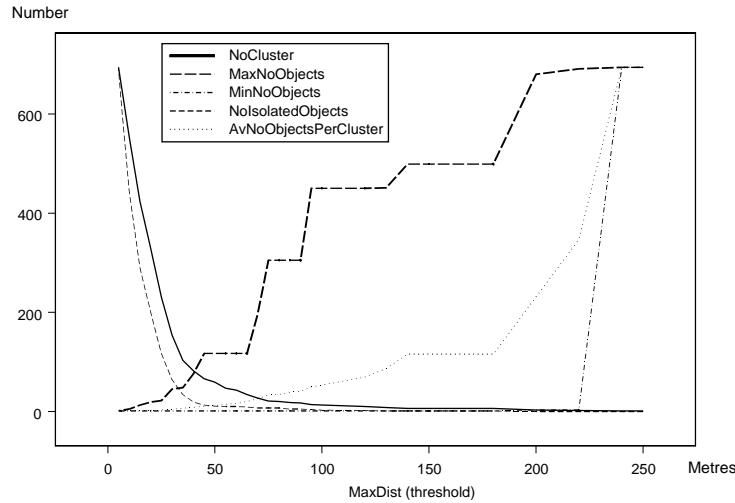
Figure 7.10: Proposed measures change corresponding to different thresholds for reasoning the optimal threshold.

The average number of buildings per cluster is calculated as the total number of buildings divided by the number of clusters. When the threshold increases, the average number of buildings per cluster increases. The rate of increase ranges between the rates relating to the maximum and minimum number of buildings in a cluster.

By comparing the clusters obtained by using a series of thresholds ranging from 5 m to 250 m as discussed earlier, we observed that the result obtained with the threshold equal to 40 m was the best according to our visual interpretation. The results are shown in Figures 7.12, 7.13 and 7.14. Therefore, we consider 40 m the optimal threshold for the test site. So interpreting Figure 7.10, which plots out the above-described properties in the function of the threshold, we discover NoCluster and MaxNoObjects are possibly good measures for optimising the threshold.

Based on this experiment, we propose a model for reasoning the optimal threshold for spatial clustering according to the shortest distances between adjacent objects by using NoCluster and MaxNoObjects as measures.

$$T_{optimal} = arg\ min_{i=5}^{250} |NoCluster_i\ -\ MaxNoObjects_i|$$

In this model, the threshold (MaxDist) is used as a variable (i) for reasoning on the optimal threshold ($T_{optimal}$) by minimising the differences between NoCluster and MaxNoObjects ($|NoCluster_i\ -\ MaxNoObjects_i|$). The optimal threshold, 40 m, is obtained by reasoning based on the proposed model, as shown in Figure 7.11. This approach seeks a global solution using the shortest distance between adjacent objects as a measure and using the optimal threshold to separate different clusters.

Number
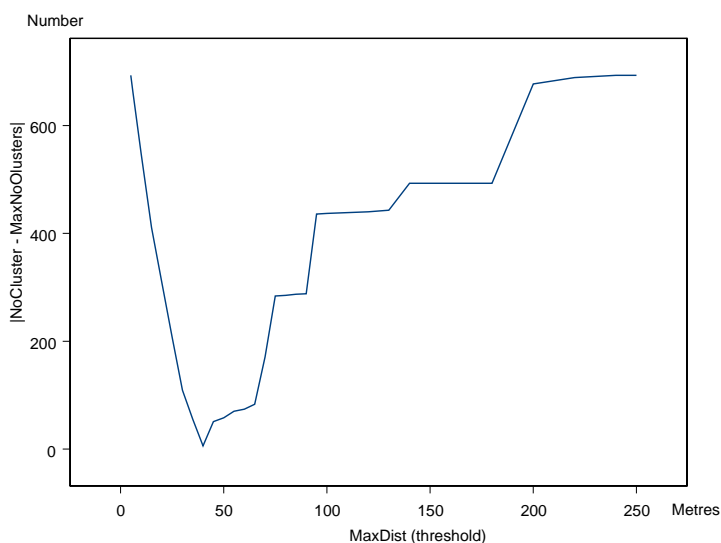
Figure 7.11: Result of proposed model for reasoning the optimal threshold.

7.3.3 Spatial clustering based on the shortest links between adjacent objects and the optimised threshold

Using 40 m as the threshold, the following results are obtained. The linked buildings are shown in Figure 7.12. The clustered buildings are presented in different colours in Figure 7.13. The clustered buildings and their convex hulls are shown in Figure 7.14.

Based on the obtained results, we can observe that most buildings are correctly clustered according to our visual interpretation. However, this method has its limits; for instance, the globally selected optimal threshold may still cause some mistakes at certain locations, as indicated by **A**, **B** and **C** in Figures 7.12, 7.13 and 7.14. Using the shortest links between objects as the only indicator for spatial clustering may be good for regular spatial patterns, but it may not suffice for complicated cases such as our test site. Therefore, additional measures are needed in order to obtain better results in spatial clustering.

7.3.4 Integration of the shortest distance and feature similarity

Buildings may vary in size, height, etc. even within a small neighbourhood. Such internal dissimilarity in a cluster has to be minimised before combining feature sim-
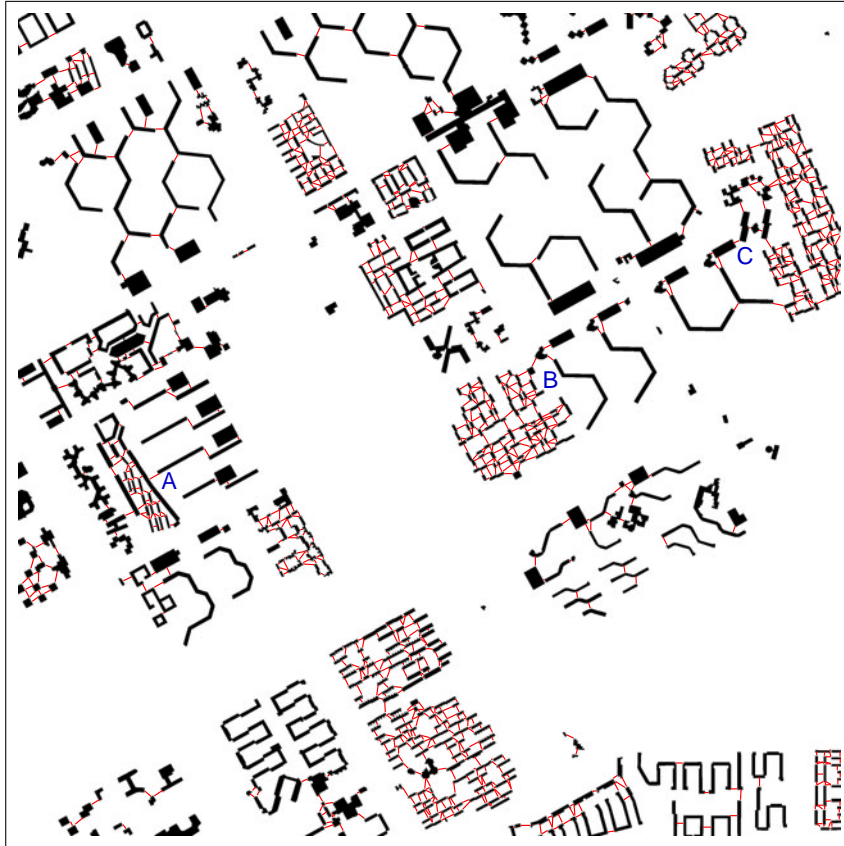
Figure 7.12: Linked buildings where the shortest edges are shorter than 40 m.

Figure 7.13: Clustered buildings where the shortest links are shorter than 40 m.
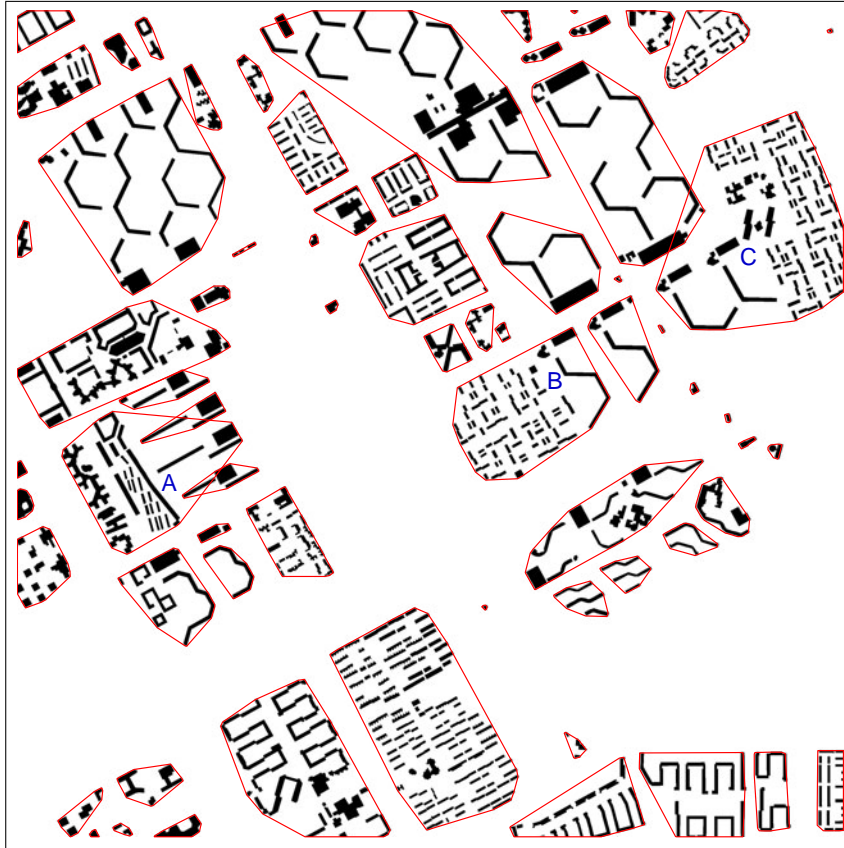
Figure 7.14: Clustered buildings and convex hulls where the shortest links are shorter than 40 m.

ilarity with the shortest distance measure. This is achieved by introducing an intermediate step that produces natural clusters with the nearest neighbours. Clustering with the nearest neighbour is to link an object to its nearest neighbour among all adjacent objects by comparing the edge length created by Delaunay triangulation. Clusters created by the nearest neighbours form natural clusters that can be considered as fundamental clusters for possible further grouping. Further clustering is based on the shortest distance between natural clusters and feature similarity between linked natural clusters. Therefore, the integration of the shortest distance and feature similarity between objects that are linked by Delaunay triangulation edges is implemented in two stages: spatial clustering with the nearest neighbour to create natural clusters, and applying the feature similarity measured between natural clusters in addition to the shortest links between natural clusters.

Natural clusters consisted by the nearest neighbours

The nearest neighbour is detected from Delaunay triangulation edges that link an object with its adjacent objects by checking the edge length. In some cases an isolated object may have a long edge with its nearest neighbour, such as a petroleum station or an individual building in a park. Therefore a threshold is needed to avoid such isolated objects being clustered by the nearest neighbour. To determine a threshold for all the links between objects that should be grouped in a cluster, we compute the average distance and the standard deviation of the distances between the objects that we have identified as the nearest neighbours. We keep those links that are shorter than the average distance plus three times the standard deviation. The histogram of link length for all the nearest neighbours is shown in Figure 7.15. Objects linked by the remaining edges after thresholding are used for finding the natural clusters. Extracted triangle edges that link the nearest neighbours and are shorter than the threshold are shown in Figure 7.16. The natural clusters extracted by this approach are presented in Figure 7.17.

Integration of shortest distance and feature similarity

As we discussed earlier, using the shortest distance alone is not sufficient for spatial clustering in a complicated urban area, so feature similarity is proposed for further consideration. We observe that adjacent buildings of similar size or similar height are more likely to be in the same cluster most, but adjacent buildings of different size or different height may still belong to the same cluster in some cases. A similar observation can also be made based on extracted natural clusters. The feature similarity between linked natural clusters is likely to be higher than the feature similarity between the individual buildings at the cluster links. Thus, first finding natural clusters and then considering the feature similarity is likely to produce the wanted grouping. Therefore, feature similarity is considered after natural clusters are formed. The similarity measures are calculated based on features (average building size and building height in this case) between adjacent natural clusters, using Formula 5.5. Due to the complexity of an urban scene, there is no simple rule that determines whether two objects should be in the same cluster or not. By considering relationships discussed earlier and by taking a limited number of samples
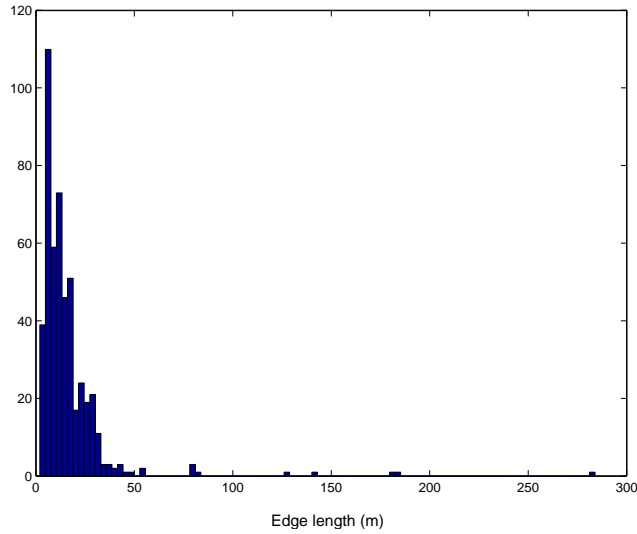
Figure 7.15: A histogram of edge length for all the nearest neighbours.

(edges and linked clusters by these edges) that should or should not be in the same cluster, we derive the following rules based on the shortest distance and the similarity in size and height.

$$
Cluster_{i,j} = \begin{cases}
1, & \begin{array}{l} SD \leq 20 \text{ or} \\ SD \leq 30 \text{ and } Sim\_Size \geq 0.3 \text{ and } Sim\_Height \geq 0.5 \text{ or} \\ SD \leq 40 \text{ and } Sim\_Size \geq 0.5 \text{ and } Sim\_Height \geq 0.6 \text{ or} \\ SD \leq 50 \text{ and } Sim\_Size \geq 0.6 \text{ and } Sim\_Height \geq 0.7 \text{ or} \\ SD \leq 70 \text{ and } Sim\_Size \geq 0.8 \text{ and } Sim\_Height \geq 0.8; \end{array} \\
\\
0, & else.
\end{cases}
$$

where SD denotes the shortest distance or the edge length; Sim_{Size} denotes similarity in building size; Sim_{Height} denotes similarity in building height; code 1 indicates that the linked clusters should be combined; and code 0 indicates that the linked clusters should be separated.

Applying these rules to our test site, we could achieve a significant improvement as compared with the clustering obtained by using the shortest distance alone, as shown in Figures 7.18, 7.19 and 7.20. Objects that appear only partially near the edges of the image may cause dissimilarity problems. We can admire the fact that the improved results are very similar to what we may interpret visually – although

Figure 7.16: Linked buildings by the nearest neighbours closer than the threshold of 78.3 m.
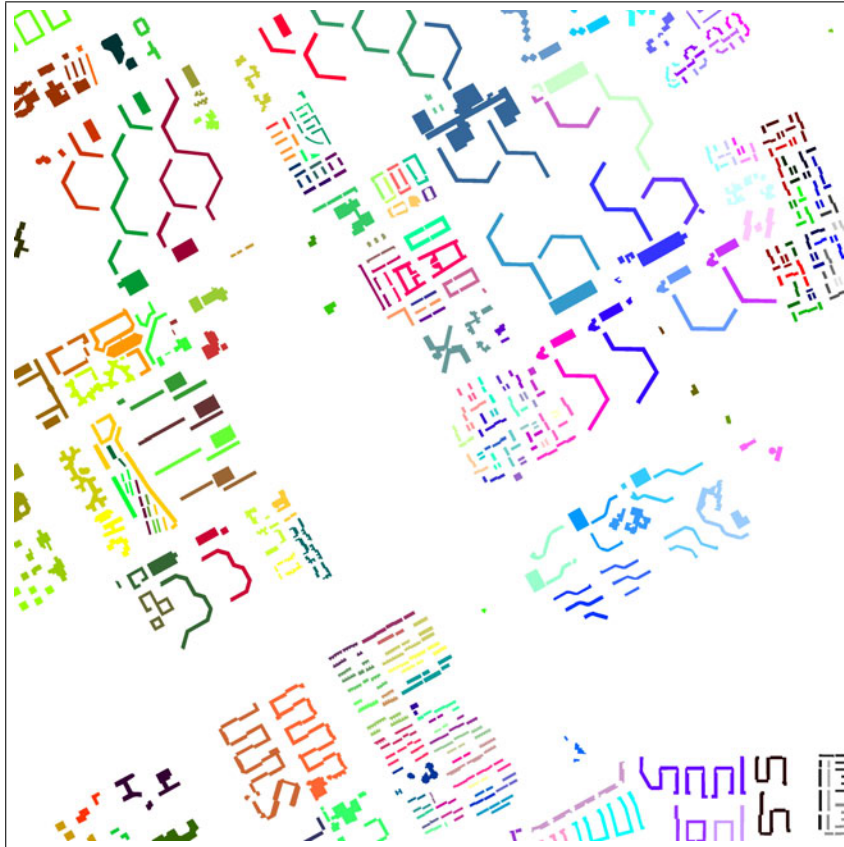
Figure 7.17: The initial clusters created by the nearest neighbours (clusters are presented by colours).

different people may have different interpretations, especially when they have different professional backgrounds.

7.3.5 Quantitative analysis toward rule extraction for spatial clustering based on the shortest distance and feature similarity

As presented earlier, clustering is improved by considering the shortest distance and feature similarity based on rules obtained by visual observation. To confirm such rules quantitatively and obtain explicit relationships between the distance measure and the similarity measures, a quantitative analysis is carried out based on the data obtained in the previous stage. Some 410 edges are extracted from the Amsterdam test data, which link 217 adjacent natural clusters; 302 edges link clusters that should be combined, the other 108 edges should be removed according to the visual interpretation that we consider as reference data. The need for further merging of obtained clusters will be up to users, according to their application requirements.

Relationship between measures proposed for spatial clustering and the decision to combine or separate linked clusters

The relationship between the shortest distance between adjacent natural clusters and the decision to combine or separate linked clusters is presented in a scatter plot that indicates 410 edges includes edge length and the decision (1 for combination, 0 for separation) as well as a linear relationship estimated by linear regression, as shown in Figure 7.21. This figure shows that the possibility of linked clusters that should be combined declines as the edge length increases. The result confirms our observation that the shorter the link edge, the higher the possibility that the linked objects (buildings or clusters) should be combined. The frequency curves (blue and red curves as shown in Figure 7.21) also show this tendency.

The relationship between similarity in building size between adjacent natural clusters and the decision to combine or separate linked clusters is presented in a scatter plot that indicates 410 edges and includes edge length and the decision (1 for combination, 0 for separation) as well as a linear relationship estimated by linear regression, as shown in Figure 7.22. This figure shows that the possibility of linked clusters that should be combined increases as the similarity measure based on building size increases. The result confirms our observation that the more similar the linked objects (buildings or clusters), the higher the possibility that the linked objects (buildings or clusters) should be combined. The frequency curves (blue and red curves as shown in Figure 7.22) also show such a tendency.

The relationship between similarity in building height between adjacent natural clusters and the decision to combine or separate linked clusters is presented in a scatter plot that indicates 410 edges and includes edge length and the decision (1 for combination, 0 for separation) as well as a linear relationship estimated by linear regression, as shown in Figure 7.23. This figure shows that the possibility of linked clusters that should be combined increases as the similarity measure based on building size increases. The result confirms our observation that the more similar

Figure 7.18: Linked buildings by the shortest distance and feature similarity.
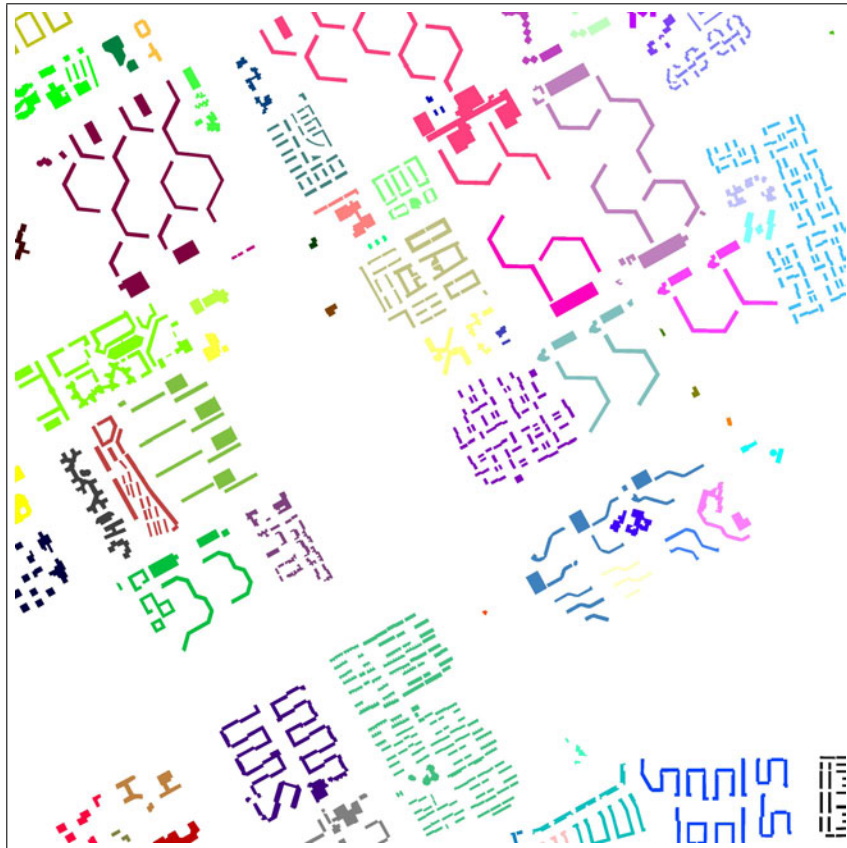
Figure 7.19: Clustered buildings by shortest links and feature similarity.
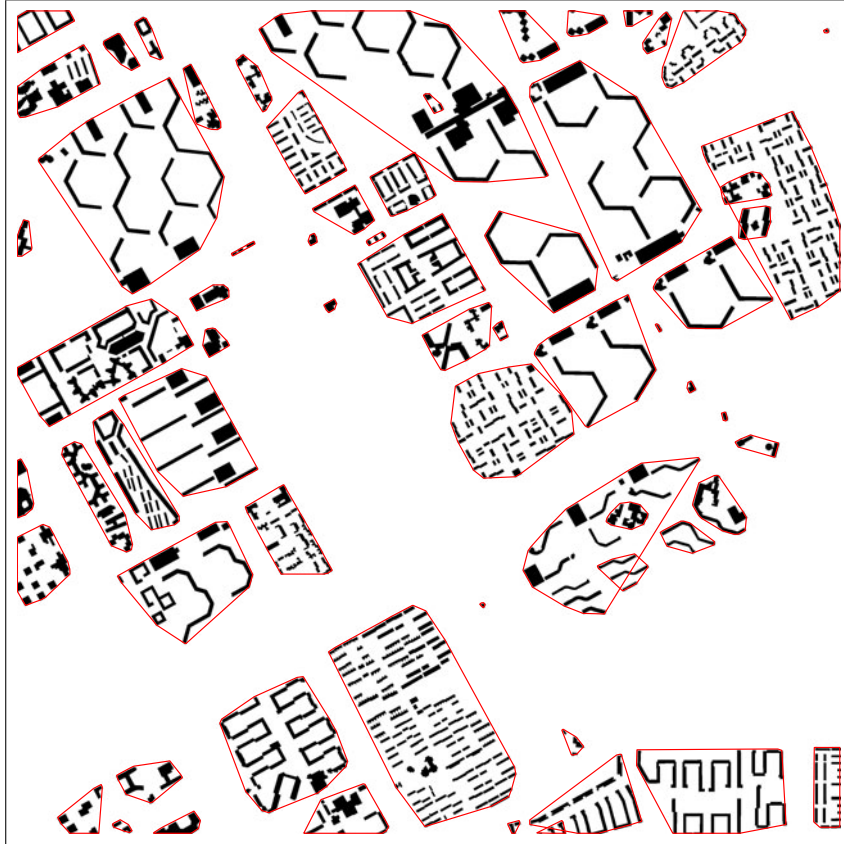
Figure 7.20: Clustered buildings and convex hulls by the shortest links and feature similarity.
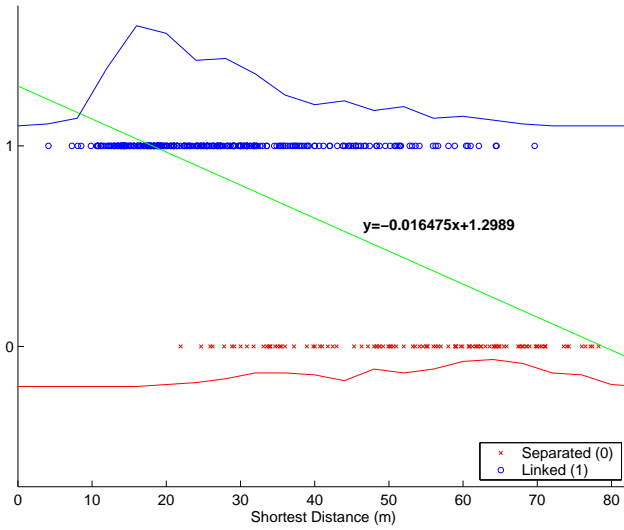
Figure 7.21: Relationship between the shortest distance and the binary decision (1 for combination, 0 for separation).
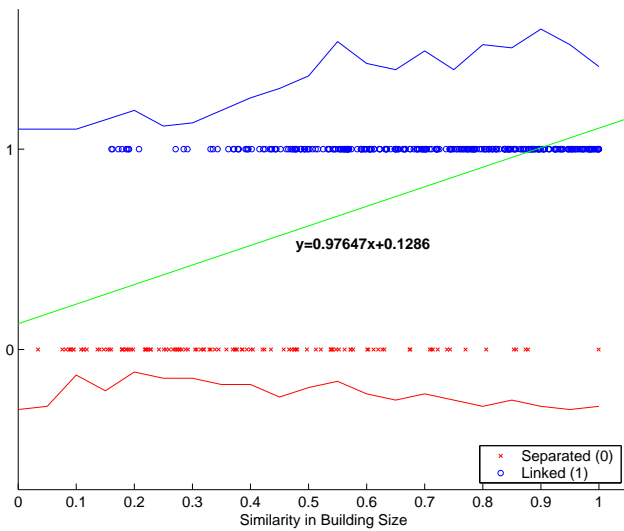
Figure 7.22: Relationship between similarity in building size and the binary decision (1 for combination, 0 for separation).
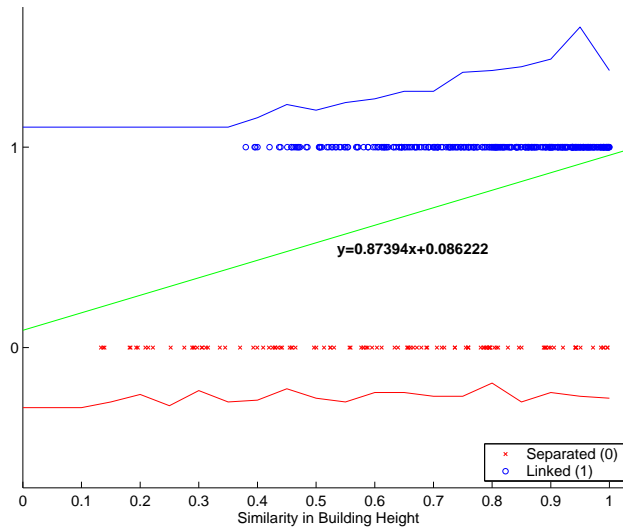
Figure 7.23: Relationship between similarity in building height and the binary decision (1 for combination, 0 for separation).

the linked objects (buildings or clusters), the higher the possibility that the linked objects (buildings or clusters) should be combined. The frequency curves (blue and red curves as shown in Figure 7.23) also show such a tendency.

Relationship between the measures proposed for spatial clustering and the decision to combine or separate linked clusters

Based on 410 edges that link 217 natural clusters, the relationship between edge length and similarity in building size with respect to deciding whether to combine or separate linked clusters is shown in Figure 7.24. We can observe that most edges that should remain are located in the upper-left corner (i.e. shorter distance and more similar in building size) and most edges that should be removed are located in the lower-right corner (i.e. longer distance and less similar in building size). The frequency of edges that should be linked among all edges over the shortest distance declines when the shortest distance increases (see the green curve in Figure 7.24).

The relationship between edge length and similarity in building height is shown in Figure 7.25. We can observe that most edges that should remain are located in the upper-left corner (i.e. shorter distance and more similar in building height) and most edges that should be removed are located in the lower-right corner (i.e. longer distance and less similar in building height). The frequency of edges that should be linked among all edges over the shortest distance declines when the shortest distance increases (see the green curve in Figure 7.25).
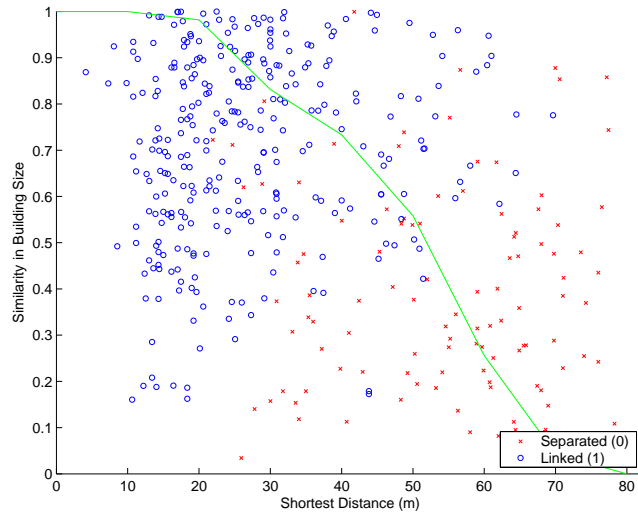
Figure 7.24: Relationship between the shortest distance and similarity in size with respect to the decision to combine or separate linked clusters (the green curve indicates the frequency of edges that should be linked).

The relationship between similarity in building size and similarity in building height is shown in Figure 7.26. We can observe that most edges that should remain are located in the upper-right corner (i.e. more similar in both building size and height) and most edges that should be removed are located in the lower-left corner (i.e. less similar in both building size and height). The frequency of edges that should be linked among all edges over the similarity in building size increases when the similarity in building size increases (see the green curve in Figure 7.26).

Similarity change between linked objects and linked natural clusters with the same links

When we select all edges (410) that link natural clusters and compare the buildings linked by these edges and the natural clusters linked by these edges, we find measured similarity changes as follows.

For all edges that link natural clusters, the average similarity in building size increases from 0.5653 (buildings) to 0.6226 (natural clusters); the average similarity in building height increases from 0.6941 (buildings) to 0.7442 (natural clusters). These changes indicate that the natural clusters created by the nearest neighbour can yield larger similarity values for both building size and building height of the edges that link them.

For all edges that link natural clusters which should remain linked, the average similarity in building size increases from 0.6593 (buildings) to 0.7238 (natural clus-
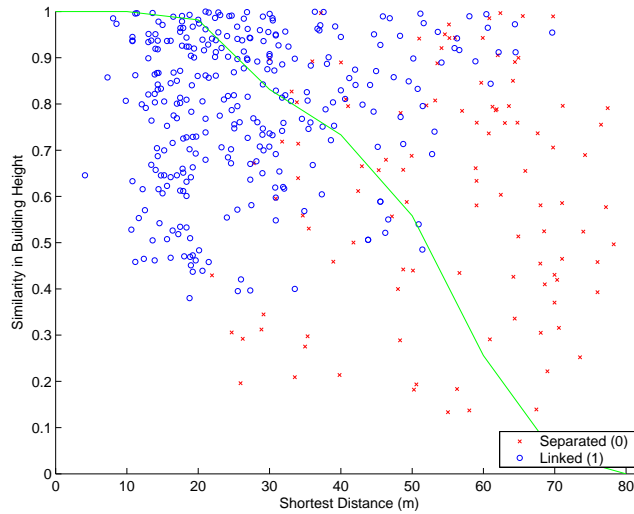
Figure 7.25: Relationship between the shortest distance and similarity in height with respect to the decision to combine or separate linked clusters (the green curve indicates the frequency of edges that should be linked).

ters), and the average similarity in building height increases from 0.7534 (buildings) to 0.8113 (natural clusters).

For all edges that link natural clusters which should be removed, the average similarity in building size increases from 0.3652 (buildings) to 0.4072 (natural clusters) and the average similarity in building height increases from 0.5680 (buildings) to 0.6012 (natural clusters).

Changes in the average similarity of the linked buildings and the linked natural clusters with the same links (410) are compared in Figure 7.27. We can observe that similarity in both building size and building height increases for linked clusters compared with linked buildings (see **a** and **a'** in Figure 7.27). Similarity in both building size and building height of objects (buildings and natural clusters) linked by edges that should remain (see **b** and **b'** in Figure 7.27) is larger than that which should be removed (see **c** and **c'** in Figure 7.27). This confirms that buildings and clusters that should be combined are more similar to each other than those that should be separated.
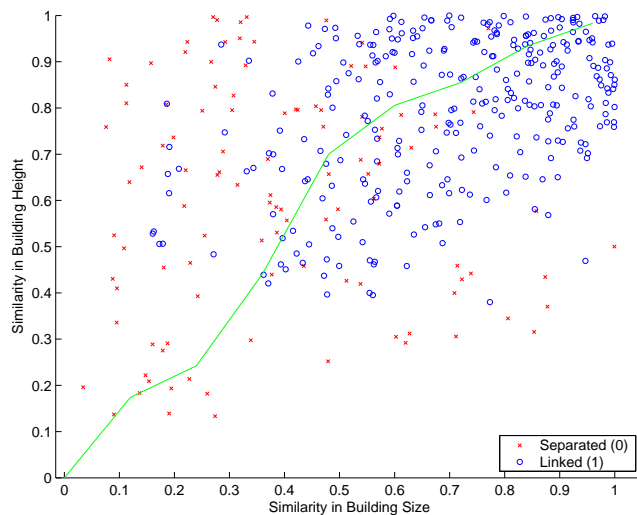
Figure 7.26: Relationship between similarity in size and height with respect to the decision of combining or separating linked clusters (the green curve indicates the frequency of edges that should be linked).
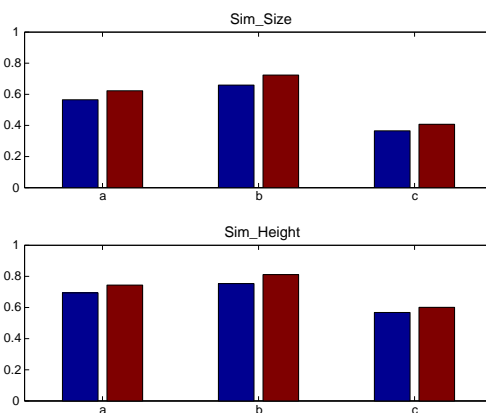
Figure 7.27: Comparison of similarity changes between linked objects and linked natural clusters with the same links. Blue colour indicates the average similarity of linked buildings. Red colour indicates the average similarity of linked natural clusters. a: all edges that linked natural clusters; b: all edges that linked natural clusters and should be retained; c: all edges that linked natural clusters and should be removed.

7.4 Spatial partitioning based on clustered objects

When spatial clusters are extracted by spatial clustering as described in the previous section, a spatial partitioning process has to follow to achieve a tessellation. As discussed earlier in Chapter 5, the spatial extent of a cluster can be regarded as having two components, the core or interior and the transit zone or fuzzy boundary. The core is the space delineated by objects belonging to this cluster. It can be extracted by using a convex hull or the morphological operation 'closing' (erosion after dilation). Boundaries derived by using either the convex hull or the morphological operation closing are considered too sharp and not a fit human perception of space. To obtain a natural transit zone surrounding the clustered objects, we use a distance transformation so that the fuzzy membership values decline toward the neighbouring clusters. Finally, the watershed algorithm is applied to acquire explicit spatial partitions in 2D space based on the simulated surfaces created for each cluster.

7.4.1 Morphological closing and the interior of a cluster

To determine the interior of a cluster, both the convex hull and closing can be used. The convex hull method is faster than the morphological closing, but it is more suitable for convex shapes or compact forms of clusters and not so good for concave shapes. Therefore, we prefer the closing operation. For the closing operation, we use a circular structural element (SE) and apply it to all pixels of the objects of each cluster. The radius of the SE is determined based on the largest distance between clustered objects, to make sure a solid core (without holes inside) is formed. In this case, 100 m is used for the morphological closing operation and the result is presented in Figure 7.28.

7.4.2 Distance transformation

The fuzzy transit zones are modelled by the distance transformation based on solid cores of clusters. A Z-shape fuzzy membership function is formulated in the range of 0 to 80 m in this case. The results of distance transformation based on clustered buildings are shown in Figure 7.29.

7.4.3 Watershed algorithm

The watershed algorithm (Vincent and Soille, 1991) is applied to obtain explicit spatial partitions in 2D space based on the simulated surfaces created for each cluster. The result is shown in Figure 7.30.

Figure 7.28: The solid interiors of clusters created by the morphological closing operation using 100 m as radius for the circular SE.
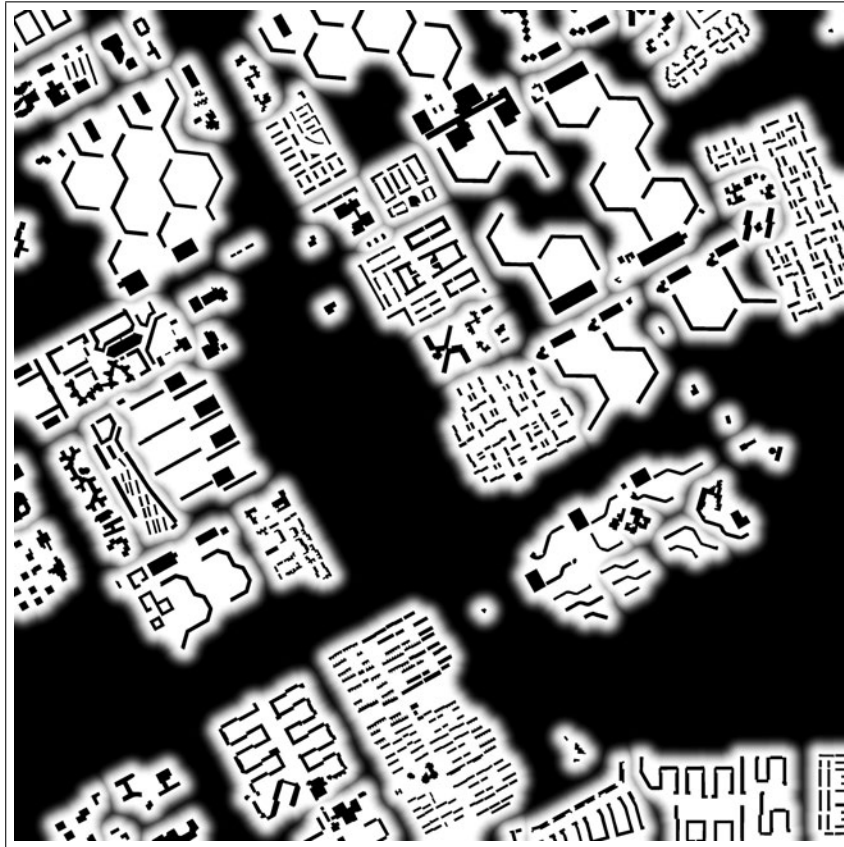
Figure 7.29: Transit zones surface created by distance transformation based on clustered buildings (superimposed on all buildings).
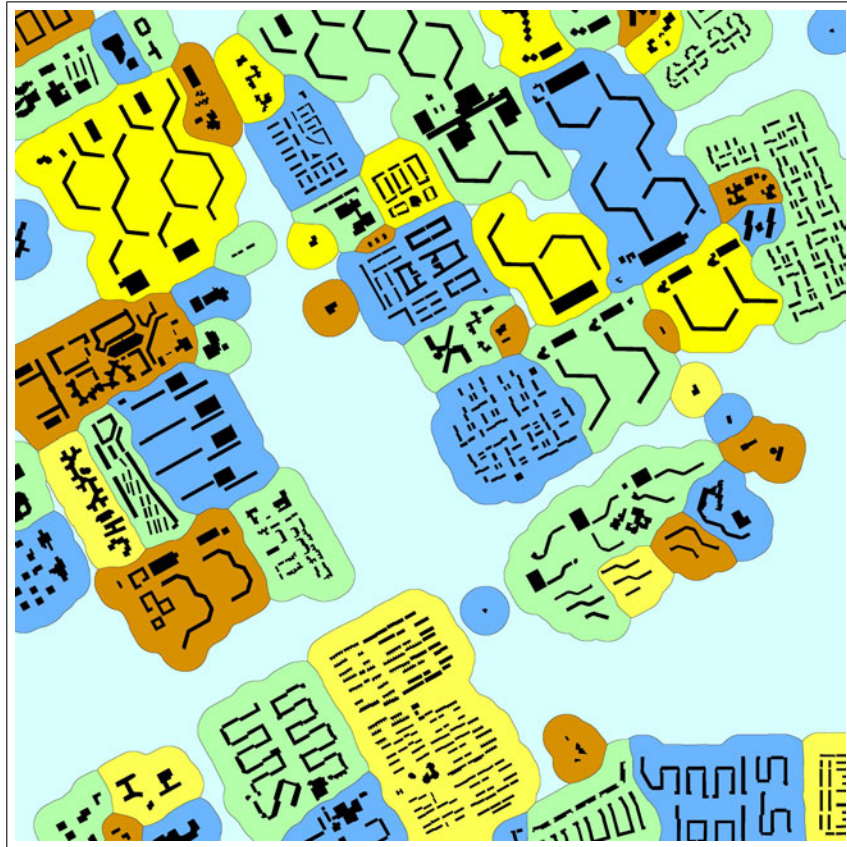
Figure 7.30: Spatial partitioning using the watershed algorithm based on the distance transformation (superimposed on all buildings).

Figure 7.31: Road map of the Amsterdam test site.

7.4.4   Using a road network in spatial partitioning

As discussed in Chapter 5, a road map is often available, which may help in obtaining a better spatial partitioning result. The road map is rasterised and used in our partitioning process as shown in Figure 7.31.

To find space to insert the road network and create the watershed along roads, the solid cores of clusters are subtracted first by a mask created by applying a morphological dilation operation to road pixels, using a circular SE with a radius of 40 m, based on the discussion in Chapter 5. The distance transformation and the Z-shape fuzzy membership function is then applied to the subtracted solid cores and the road network with the transit zone widths of 60 m and 5 m respectively. The transit zone surface is obtained as shown in Figure 7.32. The land-use units are delineated after applying the watershed algorithm to the reversed surfaces as shown

in Figure 7.33.

7.5   Summary

A structural analysis approach proposed in this chapter is tested on the Amsterdam test site. The experimental results show that the proposed object-based approach is powerful for spatial clustering. Delaunay triangulation is a good tool for extracting proximity relations among disjoint objects such as buildings. The shortest links between objects based on Delaunay triangulation allow us to form natural clusters. The natural clusters consist of buildings that are identified as the nearest neighbours. The natural clusters represent the elementary clusters of possibly larger clusters that represent land-use units. Whether adjacent natural clusters should be combined is decided based on the shortest edges (from Delaunay triangulation) that link the natural clusters, by checking the edge length (the shortest distance between clusters), and on similarity measured from features of linked clusters (i.e. building size and building height). The experimental results and acquired relationships show that the shortest distance between clusters and similarity measures in terms of building size and building height are good measures. Rules for the combination of linked clusters can be extracted based on these measures and checking the effect of different settings on a limited number of representative sample edges by visual interpretation. The spatial partitioning achieved by morphological closing, distance transformation, Z-shape fuzzy membership function and the watershed algorithm looks natural, i.e. closely corresponding to what a well trained interpreter is likely to produce. The delineated regions are regarded as spatial units of land use, thus describing the spatial extent of land-use image objects. The extraction of land-use-related properties and land-use classification will be based on these spatial units and will be presented in the next chapter.

Cities and metropolitan areas of all sizes provide many supportive functions for their constituents – the residents, transients and employees that live, visit and gain their livelihood there. In response to these roles, a complex pattern of land use evolves. Generally, these land uses conform to a regular, predictable pattern, but strong historically, culturally and technologically based traditions at work mean that cities around the world exhibit tremendous differences in form (Hartshorn, 1992). We cannot expect spatial clustering to be done best by using a single measure. The proposed approach offers a high degree of automation in delineating land-use units and reduces human intervention to rule specification and manual corrections of the results if necessary. We observe that the proposed method works well in regular scenes (with regular patterns and newly developed areas). For highly complex areas such as city centres or historical urban districts, further investigation is needed.
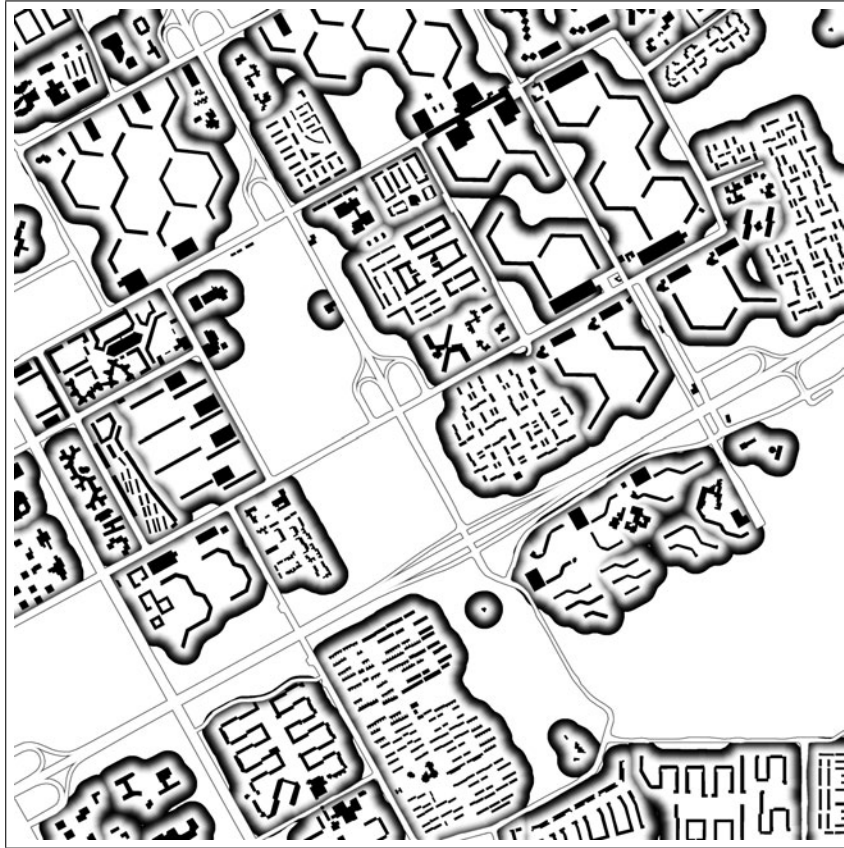
Figure 7.32: Transit zones surface created by the distance transformation based on clustered buildings and the road network (superimposed on all buildings).

Figure 7.33: Spatial partitioning using the watershed algorithm based on the distance transformation and the road network (superimposed on all buildings).

Chapter 8

Object-based land-use classification

8.1 Introduction

Land-use classification is based on the spatial units of land use obtained by structural analysis and spatial clustering, as discussed in Chapter 7. Extracted regions in spatial clustering can be seen as a representation of land-use objects in a 2D image. Such regions are regarded as the spatial extent of land-use objects and make the links between land-use objects and their locations in the image. Land-use classification can be treated as object classification based on object properties. Object properties used for land-use classification are also extracted based on such spatial units.

In a general sense, classification is a broad theme that indicates feature selection and data reduction, feature or property extraction, classifier selection, sample and classifier training, post-classification processing, as well as quality and uncertainty assessment. However, the main focus of this research is the extraction of features (object properties or object attributes) in a hierarchical structure based on high-resolution remote sensing data. 'Object' in this chapter refers to an object at land-use level, so we also call it a land-use object. Land-use mapping is one of many applications in which the extracted land-use objects and properties can be applied. Instead of aiming at a land-use map as an end-product, we rather emphasise in this thesis on the associations between extracted properties and urban land-use classes. We mainly discuss properties derived at the land-use level in this chapter. The land-use spatial units are the spatially bounded masks for property extraction and land-use classification. There are many urban land use classification systems proposed for various applications (see Appendix A). Classification, being a human abstraction process, will depend on the purpose and the techniques applied much more than on the intrinsic properties of the individual objects or object components that are being classified. Thus the result of classification will depend on various factors,

such as discipline, perception, techniques applied and classifier used. Urban land-use classification issues are discussed in this chapter mainly with regards to the needs of urban planning. In the following sections, we will present land-use property extraction, relationships between extracted properties and urban land-use classes, and fuzzy membership functions associated with land-use classification.

8.2 Extraction of object properties for urban land-use classification

To classify an urban area into discrete classes in terms of their use, a number of properties have to be extracted which should provide indications of land usage. In the following we describe various numerical and categorical properties that we consider meaningful, and which can be derived from image and laser data and can play an important role in land-use classification. We consider the following measures to be meaningful indicators for land-use classification. The definitions and formulae of the measures are given and discussed in Chapter 5.

- Type and proportional composition of land-cover objects a land-use object contains
- Number of buildings
- Average building size
- Average building height
- Building density
- Floor area ratio (FAR)
- Green coverage ratio (GCR)
- Water coverage ratio (WCR)
- Open-space coverage ratio (OCR)

8.2.1 Type and proportional composition of land-cover objects which a land-use object contains

For land-use classification, we consider four types of land-cover objects, i.e. building, green space, water, and open-surface (secondary road, footpath, parking space and other concrete open space). These features are extracted from laser data and multi-spectral data, as described in Chapters 5 and 6. The proportional composition of land-cover features is extracted for each land-use unit as shown in Figure 8.1. The proportional composition of land-cover components for each land-use unit is represented in the proportion map using a colour composition where the red colour component represent the built-up proportion (percentage of building and open-surface), the green colour component represents vegetation proportion and the blue colour component represents water proportion. In the proportion map, the reddish colour indicates a higher proportion of buildings and other sealed surfaces; the greenish colour indicates a higher proportion of vegetation (trees and lawns); the bluish colour indicates a higher proportion of water surfaces.
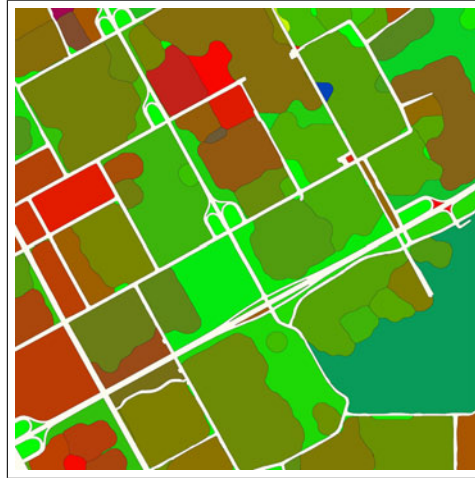
Figure 8.1: Land-cover proportion: reddish, greenish and bluish colours are proportional to the percentage of built-up area, green space and water surface a land-use unit contains.

8.2.2   Size of a land-use unit

The size of a land-use unit is counted as the total number of pixels of a land-use unit (each pixel represents 1 m^2). In land-use classification we consider size as one of the factors for formulating rules. For instance, large green space may be considered as recreational use, but small green space may be merged into one of its adjacent units. Some small objects, similar to sliver polygons in vector representation, produced by distance transformation and the watershed algorithm need to be detected by checking their size. In this case, we consider land-use objects that are smaller than 5000 m^2 as small objects and not as an independent land-use unit, as shown in Figure 8.2. Such small objects will be treated separately. If a small object is adjacent to one or several large land-use objects, it should be merged with its largest neighbour, as this is what people usually do in visual interpretation. If a small object is surrounded only by roads, it will remain as a separate object and labelled as public green space if its vegetation proportion is relatively high, say 50 %; otherwise it will be labelled as others.

8.2.3   Number of buildings

The number of buildings in a land-use unit is an indirect indicator. A residential area usually consists of many buildings of similar size. Few buildings are found in a public green space such as a park. To obtain this measure, we mask all buildings contained in a land-use object, label them, and count the total number of buildings in a land-use unit.
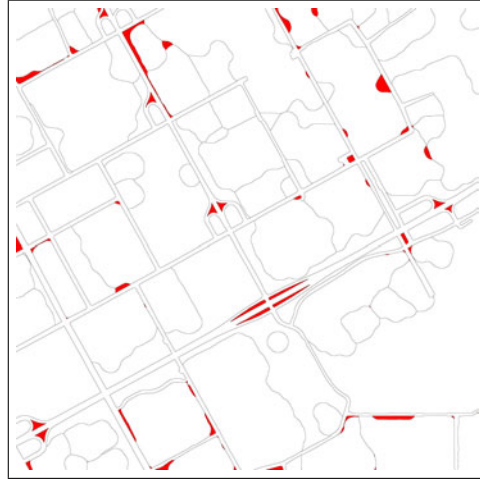
Figure 8.2: Objects smaller than 5000 m^2 and not considered as land-use objects (indicated in red colour).

8.2.4 Average building size

The average building size is a measure for certain land-use classes. For example, commercial and industrial areas often consist of some large buildings, whereas residential areas usually consist of relatively small buildings of uniform size. We first extract all buildings contained in a land-use object, accounting for the total number of building pixels. Average building size is obtained by dividing the total number of building pixels times pixel size (1 m^2 in this case) by the total number of buildings for each land-use object.

8.2.5 Average building height

High-rise buildings are distributed mainly in the central business district (CBD), as well as in some residential areas (apartment buildings). Schools and public facilities are very often low-rise buildings. Average building height is also a measure used to subdivide residential areas into residential sub-classes such as low-rise residential areas, multi-storey apartment areas and high-rise apartment areas. Since different countries, different regions, or even different cultures may have different intentions toward building height, local knowledge is needed for building classification rules based on building height. Average building height for each land-use object is obtained by extracting all buildings with building height information (extracted as described in Chapter 6) contained in a land-use object and taking the average. By dividing the average building height by 3 (average floor height in metres), we can estimate the average number of floors, which is useful information for land-use
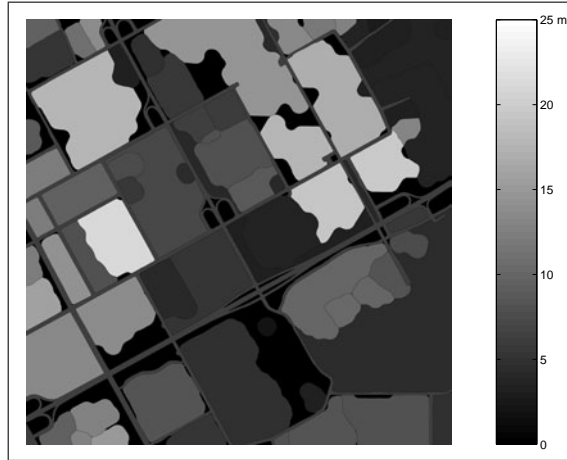
Figure 8.3: Extracted average building height.

classification. An extracted map that shows the average building height is shown in Figure 8.3.

8.2.6   Building density

Building density provides indications on the degree of concentration and intensity in terms of urban development, and is often used as a measure for building control in urban planning and management. In general, building density increases toward the city centre. Density itself is a key identifier of sub-classes of residential area, such as high-density, medium-density and low-density residential areas. Building density is calculated by dividing the total area of buildings by the size of the spatial unit for each land-use object. An extracted building density map is shown in Figure 8.4.

8.2.7   Floor area ratio (FAR)

Floor area ratio is a comprehensive indicator often used by planners to estimate development intensity or intensity of use of urban space and it is often used as a measure for building control in urban planning and management. In general, commercial areas have high FAR values. Parks and other open space have low FAR values. Furthermore, FAR is considered proportional to population density in residential areas. FAR is computed as the total area of building floor space (building size times number of floors) divided by the size of the spatial unit. The extracted result is presented in Figure 8.5.
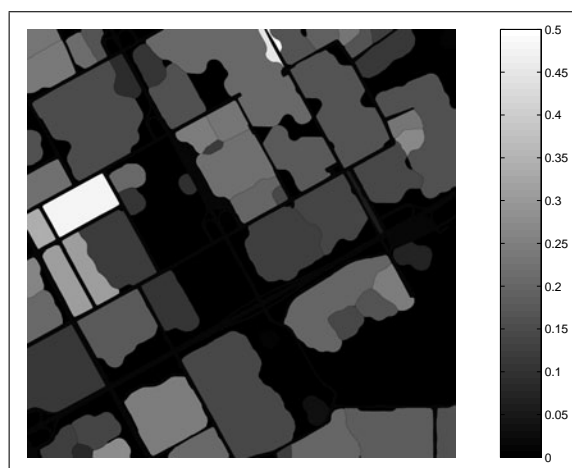
Figure 8.4: Extracted building density.
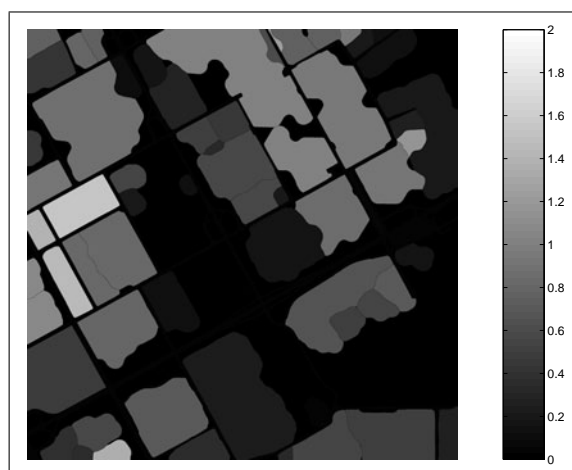
Figure 8.5: Extracted floor space ratio (FAR).
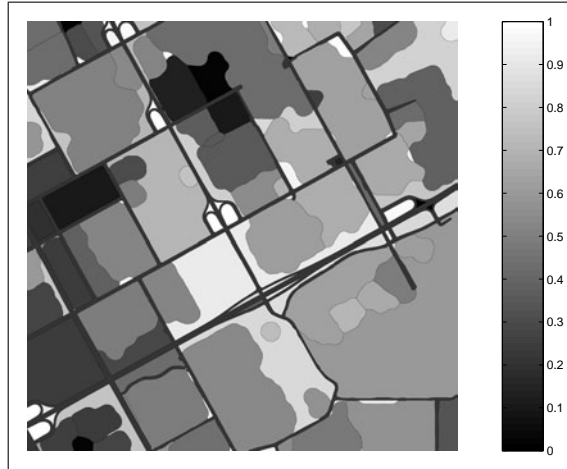
Figure 8.6: Extracted green coverage ratio (GCR).

8.2.8 Green coverage ratio (GCR)

Green space (trees and lawns) is a major indication of environmental quality. A high green coverage ratio indicates a better living environment. This is a good indicator for reasoning whether an area is likely to be a public park. In some cases, the green coverage ratio is combined with the percentage of water surface in order to assess a living environment. In general, parks and other open spaces, as well as low-density residential areas, have high GCR values. Commercial and industrial areas usually have low GCR values. The GCR or the proportion of green space is calculated as the total area of green space divided by the size of the spatial unit. The result of extraction is shown in Figure 8.6.

8.2.9 Open-surface coverage ratio (OCR)

The open-surface coverage ratio or proportion of open space is calculated as the complement of the total of buildings, green space and water surface in the spatial unit, based on spectral information. The major OCR components in an urban area are parking spaces and squares. Construction sites and industrial areas may cause high OCR values as well. This is a good indicator for reasoning whether an area is a public gathering place such as a commercial centre, an office area, public facilities or roads and footpaths inside a land-use unit. This is a key indicator for detecting construction sites or newly developed areas. Residential areas usually have a low OCR ratio. The computed OCR for Amsterdam is shown in Figure 8.7.

The derived object properties are stored in a table for each land-use object, as shown in Figure 8.8. These object properties are used as object attributes and will be used for object-based classification. The location and spatial extent of land-use

Figure 8.7: Extracted open-surface coverage ratio (OCR).

objects are defined by corresponding image regions. Thus the proposed object-based land-use classification is more similar in its (per-object) approach to most vector-based GIS than to the (per-pixel) methods of most remote sensing image processing systems.

8.3 Characteristics of different land-use classes and responses from the extracted properties

To examine and build up quantitative relationships between the proposed measures and land-use classes, we manually select 25 known land-use objects and examine how the proposed measures (land-use object properties) behave for different land-use classes, in order to design the discriminating functions toward designating land-use classes based on the extracted land-use object properties. In this section, we will also examine the robustness and sensitivity of the extracted land-use object properties identifying different land-use classes.

8.3.1 Class discrimination based on extracted properties

To classify land-use objects, we need to explore and establish the relationships between extracted properties and designated land-use classes. Some 25 known samples have been selected manually from different land-use classes of the Amsterdam

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
	ID	Size	No.B	B.Size	B.avSize	B.stdSize	B.Den	B.avHeight	B.stdHeight	FAR	GCR	WCR	OCR	Class	Certainty	Correct
1	1	32492	3	7539	2513.00	2088.58	0.23	9.75	0.53	0.75	0.43	0.00	0.34	1	0.05	0
2	2	81454	10	19595	1959.50	2315.90	0.24	5.44	3.63	0.44	0.43	0.04	0.29	2	0.00	0
3	5	26056	0	0	0.00	0.00	0.00	0.00	0.00	0.00	0.84	0.04	0.11	3	0.42	1
4	6	22994	1	3602	3602.00	0.00	0.16	8.91	0.00	0.47	0.59	0.10	0.16	2	0.06	1
5	7	64475	5	14250	2850.00	2849.74	0.22	12.35	2.81	0.91	0.26	0.03	0.49	2	0.01	1
6	8	33690	4	11348	2837.00	2917.40	0.34	12.34	6.93	1.39	0.19	0.00	0.47	2	0.12	1
7	10	25137	2	6612	3306.00	2141.12	0.26	12.28	3.04	1.08	0.60	0.05	0.09	2	0.11	1
8	11	63239	11	13173	1197.55	869.95	0.21	15.66	3.50	1.09	0.27	0.06	0.46	2	0.05	1
9	12	226519	14	24885	1777.50	3082.72	0.11	13.58	7.69	0.50	0.23	0.03	0.63	2	0.01	1
10	13	7883	0	0	0.00	0.00	0.00	0.00	0.00	0.00	0.99	0.00	0.01	3	0.59	1
11	14	12817	0	0	0.00	0.00	0.00	0.00	0.00	0.00	0.99	0.00	0.01	3	0.59	1
12	15	13774	0	0	0.00	0.00	0.00	0.00	0.00	0.00	0.78	0.00	0.22	3	0.41	1
13	17	364324	18	54155	3008.61	2930.61	0.15	17.68	9.45	0.88	0.50	0.04	0.31	1	0.18	1
14	18	16145	0	0	0.00	0.00	0.00	0.00	0.00	0.00	0.76	0.04	0.20	3	0.41	1
15	22	37996	5	4787	957.40	345.44	0.13	9.57	2.76	0.40	0.25	0.00	0.62	2	0.06	1
16	23	121525	17	58411	3435.94	6210.35	0.48	9.61	3.79	1.54	0.11	0.00	0.41	2	0.01	1
17	24	73297	3	22701	7567.00	7841.74	0.31	14.10	4.21	1.46	0.25	0.02	0.42	2	0.13	1
18	27	63508	0	0	0.00	0.00	0.00	0.00	0.00	0.00	0.75	0.05	0.20	3	0.41	1
19	31	6078	0	0	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.34	0.64	3	0.31	1
20	32	72275	26	22530	866.54	1442.78	0.31	7.86	3.16	0.82	0.38	0.00	0.31	1	0.08	1
21	33	31172	2	6709	3354.50	2575.99	0.22	11.07	1.69	0.79	0.30	0.19	0.30	2	0.08	1
22	37	12199	2	1022	511.00	130.11	0.08	12.88	0.20	0.36	0.03	0.00	0.89	2	0.10	1
23	38	62390	3	8317	2772.33	1733.85	0.13	12.84	1.83	0.57	0.24	0.04	0.58	2	0.09	1
24	39	175330	9	30402	3378.00	3949.82	0.17	13.92	7.00	0.80	0.49	0.09	0.25	1	0.01	1
25	40	171806	8	20342	2542.75	1790.67	0.12	21.23	5.58	0.84	0.52	0.00	0.36	1	0.10	1

Figure 8.8: A table for recording object IDs and attributes (object properties) for land-use image objects.

test site. The relationships between extracted properties and designated land-use classes are presented in a matrix of scatter plots, as shown in Figures 8.9 and 8.10. The difference between Figure 8.9 and Figure 8.10 is that the measured values presented in the former are obtained from extracted buildings based on laser data, whereas the measured values given in the latter are acquired from buildings in the base map. Based on Figures 8.9 and 8.10, we can observe that several land-use classes are well identified by these properties, such as commercial and public green space (see rows and columns associated with building density (B.Den), FAR and GCR in Figures 8.9 and 8.10), whereas facilities are mixed with residential areas because of the existence of different sub-classes such as schools, hospitals and community centres that are included in the class called public facility. We may also note that among these properties, building density, FAR and GCR are better measures for class discrimination in general (see also the enlarged version in Figures 8.11 and 8.12).
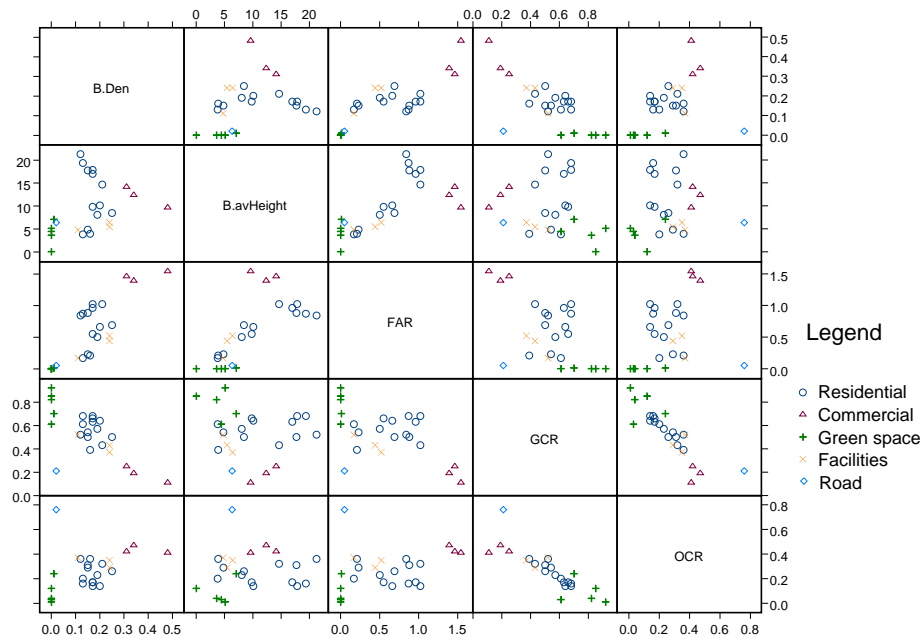
Figure 8.9: Class discrimination based on selected properties (based on extracted buildings). Each scatter shows distribution of samples in terms of two corresponding properties.
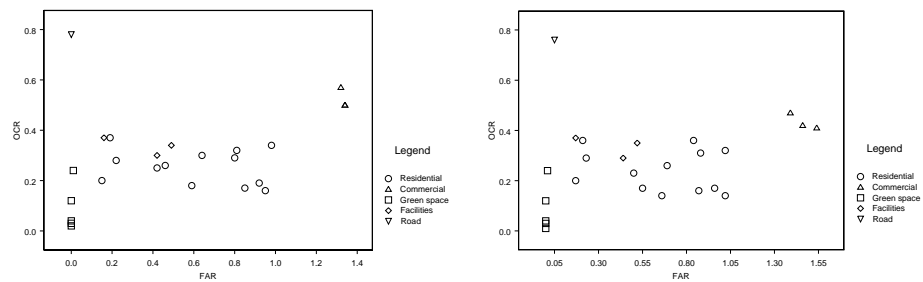
Figure 8.12: Class discrimination based on FAR and OCR using extracted buildings (left) and buildings digitised from map (right).

When we take a close look at these relationships as shown in Figures 8.11 and 8.12, some linear discriminating functions can be derived for certain classes based
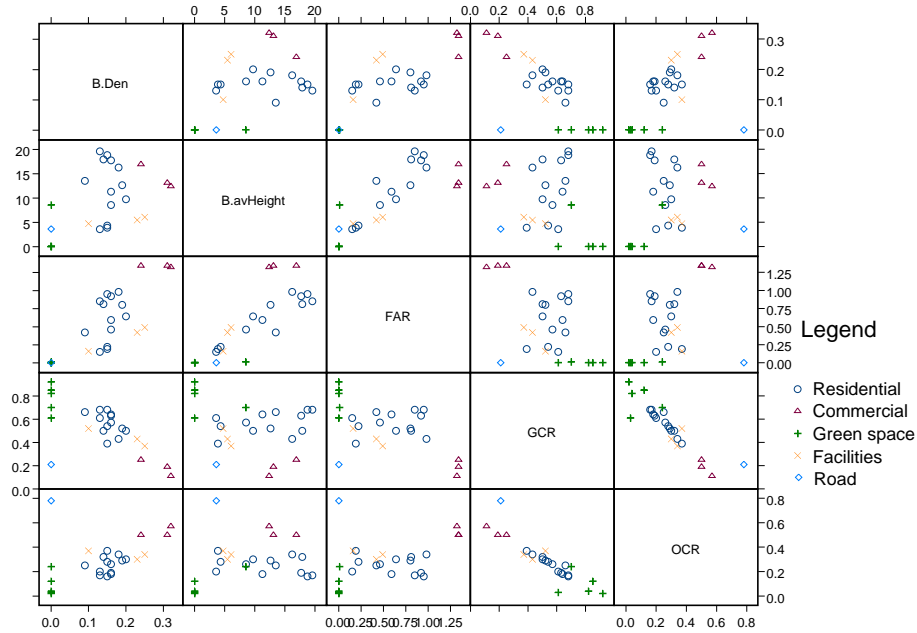
Figure 8.10: Class discrimination based on selected properties (based on buildings from map). Each scatter shows distribution of samples in terms of two corresponding properties.

on these properties, while additional properties may be needed for mixed classes. Object properties obtained at land-use level, such as building density, FAR, GCR, OCR are discriminative features for land-use classification and are also meaningful to urban planners. We may not be able to extract such properties at the land-cover level and certainly not at the pixel-level. Therefore we draw the general conclusion that the proposed hierarchical object-based approach is a better option for urban land-use classification than the conventional per-pixel based approaches.

8.3.2 Robustness and sensitivity of extracted properties

When comparing Figures 8.9 and 8.10, we can observe that the extracted properties are quite robust, despite the fact that the derived values of object properties may be slightly different owing to some errors inherited from land-cover object extraction (see also Figures 8.11 and 8.12). This means that small errors and mistakes made during land-cover object extraction (see details in Chapter 6) do not have a significant impact on the proposed measures for land-use objects.
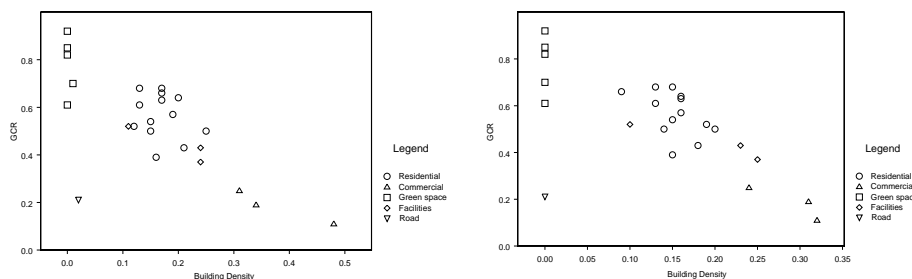
Figure 8.11: Class discrimination based on building density and GCR using extracted buildings (left) and buildings digitised from map (right)

Another interesting issue is how these properties respond to changes in the real world. As we mentioned earlier in Chapter 2, the base map of the Amsterdam test site was updated before 1998 (buildings appearing in laser data and the IKONOS image were not mapped), laser data were acquired in 1998, and the IKOKOS scene was obtained in 2000. Based on our knowledge and indications from the data, we found that this study area was in the process of dynamic innovation during this period. Quite a lot of changes have taken place in certain locations in the study area. We have selected two changed sites and take a close look at them in Figures 8.13 and 8.14. The extracted properties are given in Tables 8.1 and 8.2 respectively. These examples demonstrate that not only can land-cover changes be mapped but also the extracted properties reflect these changes well. Since we have acquired only one IKONOS image, we cannot make a comparison for the GCR.

Table 8.1: Properties derived from changed site 1 (Figure 8.13)

Object properties	Based on buildings digitised from base map	Based on buildings extracted from image
Building density	0	0.22
Average building height (m)	0	12.35
FAR	0	0.91

We are also interested in how these properties respond to areas that have not changed in the real world. We have selected two unchanged sites and take a close look at them in Figures 8.15 and 8.16. The extracted properties are given in Tables 8.3 and 8.4 respectively. These examples demonstrate that these properties are quite stable as regards unchanged sites, as compared with object properties derived based on buildings digitised from the map and buildings extracted from images, despite some noisy pixels that exist along the edges of extracted buildings. A similar
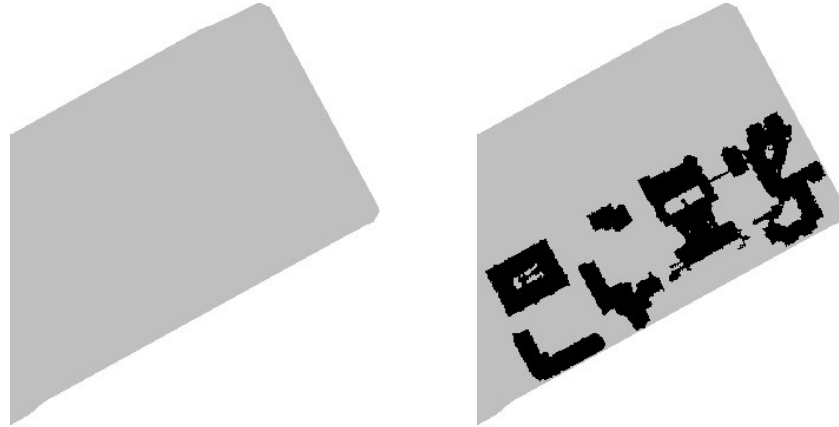
Figure 8.13: Changed site 1: according to the base map (left) and extracted buildings (right).

Table 8.2: Properties derived from changed site 2 (Figure 8.14)

Object properties	Based on buildings digitised from base map	Based on buildings extracted from image
Building density	0	0.11
Average building height (m)	0	13.58
FAR	0	0.50

observation can also be made regarding the cases presented in the next section. These examples show that the proposed properties are sensitive in reflecting big changes that are taking place in reality, and are robust in respect to small errors made during building extraction.

8.4 Characteristics of different land-use classes

To specify discriminating features based on extracted object properties and to design the fuzzy membership functions for land-use classes, we need to investigate the characteristics of each land-use class. In this section, we select several samples of known land-use classes, determine by which object properties they can be identified, and check whether derived properties are robust and reliable in describing class characteristics based on data derived from different sources.
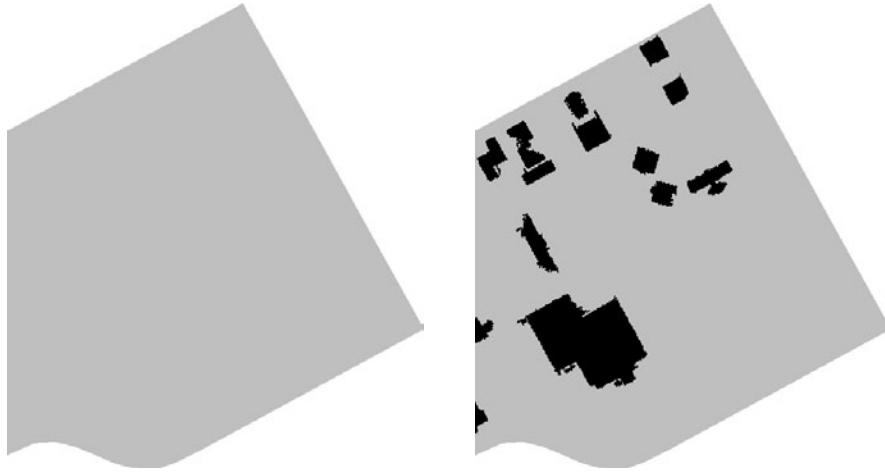
Figure 8.14: Changed site 2: according to the base map (left) and extracted buildings (right).

Table 8.3: Properties derived from unchanged site 1 (Figure 8.15)

Object properties	Based on buildings digitised from base map	Based on buildings extracted from image
Building density	0.13	0.11
Average building height (m)	2.93	3.33
FAR	0.13	0.13

8.4.1 Commercial area

The extracted properties as presented in Table 8.5 are derived for a commercial area on the Amsterdam test site, based on buildings digitised from the map and extracted buildings from the image as shown in Figure 8.17. By comparing these two images, we see that, despite some noisy pixels along building edges, several building forecourts at ground floor have been extracted from the image by using the proposed method, whereas they have not been mapped in the base map. Such building forecourts may have been ignored in the field survey or in the cartographic process of map production. These building forecourts have been confirmed by field visits. They are the main cause of different results being derived from different data sources. Such differences have an impact on some derived properties, such as building density and average building height. We believe that object properties derived from the image are better in describing the reality in this case.
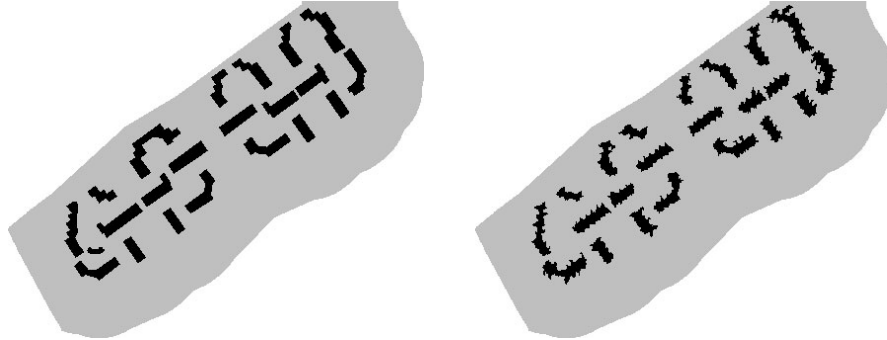
Figure 8.15: Unchanged site 1: according to the base map (left) and extracted buildings (right).
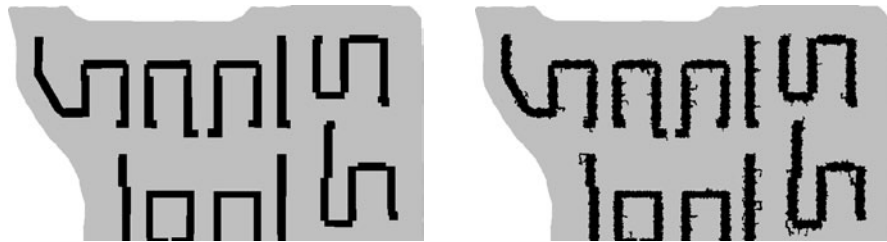
Figure 8.16: Unchanged site 2: according to the base map (left) and extracted buildings (right).

In general, the main characteristics of a commercial area are high building density (> 0.2), high FAR value (>1), but low GCR value (< 0.3) as compared with other land-use classes, as we can also see in Figures 8.10, 8.11 and 8.12. Since office area and other facility areas share some common features that are similar to commercial areas in our test site, we combine them in the commercial class or call them commercial and other facilities.

8.4.2 Residential area

The extracted properties as presented in Tables 8.6, 8.7 and 8.8 are derived from several types of residential areas on the Amsterdam test site, based on buildings digitised from the map and the extracted buildings, as shown in Figures 8.18, 8.19 and 8.20 respectively. The major differences between buildings derived from different sources are that several covered corridors have been extracted as parts of buildings, while they have been removed according to the base map despite some

Table 8.4: Properties derived from unchanged site 2 (Figure 8.16)

Object properties	Based on buildings digitised from base map	Based on buildings extracted from image
Building density	0.16	0.19
Average building height (m)	9.35	8.03
FAR	0.50	0.50

Table 8.5: Properties derived from a commercial area

Object properties	Based on buildings digitised from base map	Based on buildings extracted from image
Building density	0.32	0.48
Average building height (m)	12.38	9.61
FAR	1.32	1.54
GCR	0.11	0.11
OCR	0.57	0.41

noisy pixels existing along the building edges. Such differences may have a slight impact on the average building height for high-rise buildings.

In general, the main characteristics of a residential area are medium building density (0.1 to 0.2), medium FAR value (0.2 to 1.0), but medium GCR value (0.3 to 0.7) as compared with other land-use classes, as we can also see in Figures 8.10, 8.11 and 8.12.

8.5 Land-use classification

Given the information obtained from the processes described earlier, several classification methods can be used in the final stage of land-use classification. The classifiers, however, must not conflict with the nature of the data. Classifiers to consider for land use include fuzzy logic, the nearest neighbour classifier and the tree-based classifier. The relationships between extracted properties and land-use classes are often fuzzy. For instance, a high percentage of green coverage ratio indicates the high probability or possibility that a land-use object is likely to belong to a public park or green space. (Here 'high' is a relative term or fuzzy.) Therefore, fuzzy classification is believed to be a more suitable approach for land-use classification. It has special significance for remote sensing. Fuzzy logic permits partial membership, a property that is especially significant in field remote sensing, as partial membership translates closely to the problem of mixed pixels (Campbell, 2002). The same prop-

Figure 8.17: Buildings in a commercial area based on map (left) and extracted from image (right).

Table 8.6: Properties derived from a residential area (multi-story apartment)

Object properties	Based on buildings digitised from base map	Based on buildings extracted from image
Building density	0.14	0.15
Average building height (m)	19.28	17.68
FAR	0.88	0.88
GCR	0.50	0.50
OCR	0.32	0.31

Table 8.7: Properties derived from a residential area (multi-story apartment)

Object properties	Based on buildings digitised from base map	Based on buildings extracted from image
Building density	0.13	0.16
Average building height (m)	6.62	5.54
FAR	0.30	0.30
GCR	0.13	0.13
OCR	0.63	0.60
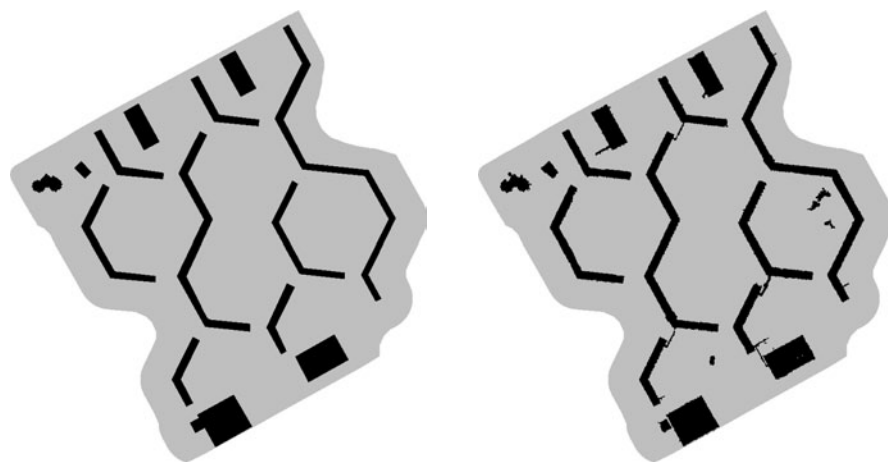
Figure 8.18: Residential area 1: buildings based on map (left) and extracted from image (right).
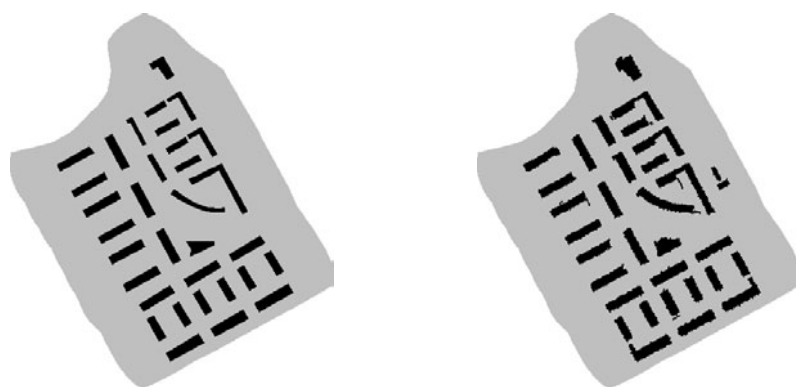
Figure 8.19: Residential area 2: buildings based on map (left) and extracted from image (right).

Table 8.8: Properties derived from a residential area (multi-story apartment)

Object properties	Based on buildings digitised from base map	Based on buildings extracted from image
Building density	0.15	0.15
Average building height (m)	4.52	4.79
FAR	0.23	0.23
GCR	0.54	0.54
OCR	0.28	0.29
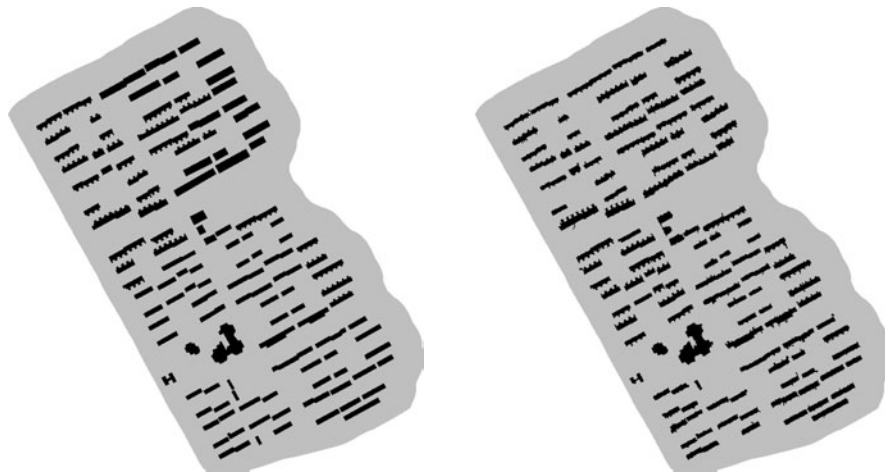
Figure 8.20: Residential area 3 (multi-story apartment): buildings based on map (left) and extracted from image (right).

erty holds in our object-based cases, where problems of mixed objects are similar to problems of mixed pixels. In addition, knowledge and rules can be easily represented by using fuzzy approaches. A land-use classification example based on fuzzy logic can be found in Zhan et al. (2000).

8.5.1 Fuzzy classification

Fuzzy set theory (Zadeh, 1965) and fuzzy logic have been developed and used in many research and application areas. Fuzzy set theory has been widely applied in clustering and classification as well (Foody, 1996; Hoeppner et al., 1999; Zhan et al., 2000; Tso and Mather, 2001). Fuzzy c-means is a well-known unsupervised classifier. A fuzzy classifier assigns membership values to objects based upon a membership function. Membership functions for classes are determined either by general knowledge about the relationships between object properties and land-use classes or by definitional rules describing the relationships between derived measures and classes. Or, as is more likely in the instance of remote sensing classification, membership functions are derived from experimental data for the specific scene to be examined (Campbell, 2002). In our object-based approach, derived object properties rather than the spectral values of a pixel are used for designing the membership functions. These properties are commonly used in planning practice. They are relatively easier to associate with different classes and thus can be directly deployed in forming membership functions.

8.5.2 Fuzzy membership functions for land-use classification

There are many forms of membership functions, such as the triangular membership function, the trapezoidal membership function, the Gaussian curve membership function. In this experiment, we use the S-shaped and the Z-shaped curve membership functions for 'one-end' cases and the 'π-shaped' curve membership function for 'two-end' cases, since the spline membership functions are more natural than the linear membership functions for representing the transit zones. The parameters for the spline membership functions can be easily specified by indicating the starting and ending points of a spline curve.

Based on our observations with sample sites and planning knowledge, we design fuzzy membership functions and use extracted object properties to classify land-use objects. In fuzzy classification, we need to establish specific fuzzy membership functions for each end-member class, based on extracted object properties. Number of buildings, building density, average building height, FAR, GCR and OCR are selected as such object properties for land-use classification in this case. Residential and public green space and a combined class consisting of commercial, office and public facilities (called commercial for short) are our end-member classes in this test. Other land-use types do not exist on our test site.

Fuzzy membership function according to the number of buildings in a land-use unit

The number of buildings in a land-use unit is a weak association regarding land-use classes, since we can hardly specify the actual building numbers a land-use unit should contain. However, we could specify the range that a land-use unit of a particular class is likely to have, based on our observations and planning knowledge. A residential area usually consists of a large number of houses. An area that contains fewer buildings is unlikely to be considered a residential area; thus a fuzzy low bound can be specified from 0 to 5. Commercial and other facilities may consist of several buildings but not too many; thus a fuzzy high bound can be specified from 10 to 20. A public green space should not contain many buildings; thus a fuzzy high bound can be specified from 0 to 5.

Fuzzy membership function according to building density

Building density is a measure of building control for specific land uses, particularly for residential areas, and is usually enforced by the planning acts. This requirement varies from country to country. In a developed country such as the Netherlands, building density for a residential area is more restricted than in a developing country such as China, to ensure the higher quality of the living environment. Thus we can specify a range with a fuzzy low bound of 0 to 0.1 and a fuzzy high bound of 0.3 to 0.4 for the Amsterdam test site. A likely range for the building density of a residential area in China can be specified as 0.1 to 0.3 for a fuzzy low bound and 0.5 to 0.8 for a fuzzy high bound. Therefore local knowledge plays an essential role in designing fuzzy membership functions for land-use classification. The building density requirement for commercial and other facilities is not as restricted as for a residential area. Thus we should specify a large range – for example, with a fuzzy low bound of 0 to 0.1 and a fuzzy high bound of 0.5 to 0.8 in our case. As discussed earlier, a public green space should not contain many buildings, so we give a fuzzy high bound of 0.05 to 0.2 for a public green space.

Fuzzy membership function according to average building height

The building height may vary even within a land-use class such as a residential area. The average building height may vary as well from place to place; thus it is a relatively weak measure. We can design fuzzy membership functions based on the general knowledge that it is unlikely that high-rise buildings will be found in a public green space, and building height for a residential area may have a relatively larger range than commercial and public facilities.

Fuzzy membership function according to floor area ratio (FAR)

Floor area ratio (FAR) is a measure of building control for specific land uses, particularly for residential areas, and is usually enforced by the planning acts. For some areas such as city centres or commercial areas, it is often preferred to take a high FAR to allow efficient use of space. Therefore we specify a fuzzy low bound of 0.05 to

0.15 for commercial and other public facilities use, a range with a fuzzy low bound of 0.05 to 0.15 and a fuzzy high bound of 1.0 to 1.2 for residential areas, and a fuzzy high bound of 0.2 to 0.25 for public green spaces.

Fuzzy membership function according to green coverage ratio (GCR)

The green coverage ratio (GCR) is a key indicator for a living environment. It is obvious that GCR should be high for public green space; thus a fuzzy low bound of 0 to 0.5 is specified. A residential area should have a reasonable range for GCR as a living space and is specified using a fuzzy low bound of 0.05 to 0.10 and a fuzzy high bound of 0.7 to 0.8. GCR for commercial and public facilities may vary over a large range; therefore a weak fuzzy low bound of 0 to 0.3 and a weak fuzzy high bound of 0.5 to 1.0 is given.

Fuzzy membership function according to open-space coverage ratio (OCR)

The open-space coverage ratio (OCR) can be interpreted as the percentage of sealed ground surface, and mainly consists of roads and footpaths at the neighbourhood level, parking spaces and playgrounds. It is also a relatively weak measure for land-use classification. In general, a public green space should not have a high OCR; thus a weak fuzzy high bound of 0.5 to 1.0 is given. Commercial and other facilities are likely to contain large parking spaces and other sealed surfaces such as squares; thus a relatively large range is specified by using a fuzzy low bound of 0 to 0.1 and a weak fuzzy high bound of 0.6 to 1.0. Residential areas are specified by using a relatively smaller range (a fuzzy low bound of 0.5 to 0.1 and a weak fuzzy high bound of 0.4 to 0.8).

The fuzzy membership functions proposed for land-use classification of the Amsterdam test site is presented in Figure 8.21.

8.5.3 Computation of the normalised overall membership values for each end-member land-use class

The normalised overall membership function of a land-use class for each land-use object is proposed and calculated as follows. The normalised overall membership function will have a value in the range from 0 to 1.

$$MF_k(O_i) = \sqrt{\frac{\sum_{j=1}^{n} MF_j^k(O_i)^2}{n}}, \;\; k = 1, ..., m$$

k denotes the kth land-use class designated in classification
i denotes a land-use object (image-object)
j denotes object properties that the fuzzy membership function is based upon
m denotes the total number of land-use classes designated in classification
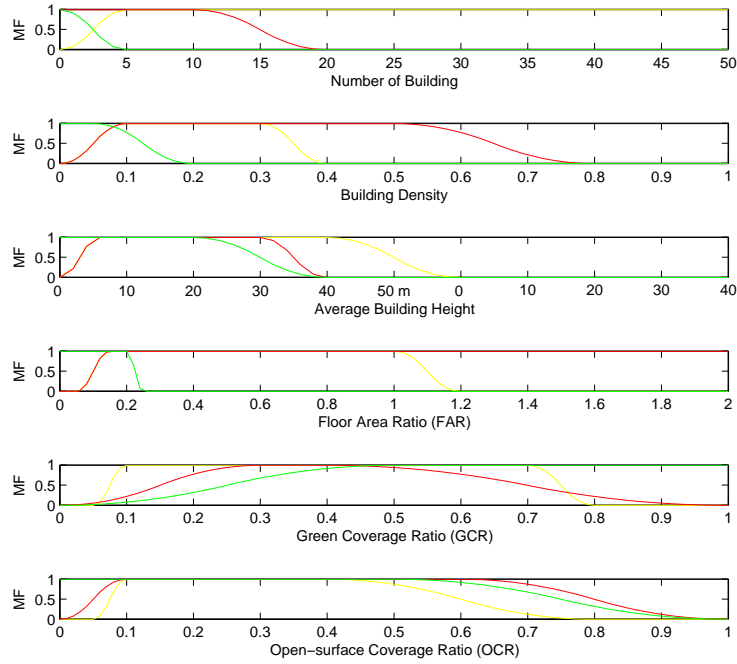
Figure 8.21: Fuzzy membership functions designed for land-use classification of Amsterdam test site (red colour denotes commercial class, yellow colour residential class, green colour green space).

n denotes the total number of object properties involved.

This formula assumes equal weight for each object property that is applied in this case. If the user can specify the weights to reflect the degree of importance among object properties, the following formula should apply.

$$\boldsymbol{MF}_k(O_i) = \sqrt{\frac{\sum_{j=1}^{n} \boldsymbol{W}_j \cdot \boldsymbol{MF}_j^k(O_i)^2}{\sum_{j=1}^{n} \boldsymbol{W}_j}}, \quad k = 1, ..., m$$

W_j denotes the weight for jth object property involved.

The class which receives the highest overall fuzzy membership value for a particular object is assigned to the object to obtain the 'hard' classification result – the land-use map. The difference (subtraction) between the highest and second to highest overall fuzzy membership values is used as an uncertainty measure. More detailed uncertainty assessment will be discussed in Chapter 9.

8.5.4 Land-use classification of Amsterdam test site

Due to the peculiarities of the developed modelling for creating the transit zones and the used watershed algorithm for delineation of land-use objects, some small sliver-like objects were created. In addition, objects smaller than 5000 m^2 were not considered as independent land-use objects. They should be merged with adjacent objects after classification. Therefore, the actual classification was based on 102 objects (larger than 5000 m^2) out of a total of 192 objects. The classification result shown in Figure 8.22 is obtained by applying the above-mentioned membership functions to extracted object properties. By comparing this with the reference land-use map which was prepared by visual interpretation and a field visit (see Figure 8.23), we find that most land-use objects are correctly classified (90 out of 102 objects). Some 12 objects are misclassified. Most of misclassifications (11 out of 12 objects) occur between commercial and residential, owing to the similarity between these two classes. Only one facility was misclassified. This was caused by a mistake made in building extraction, where a metro station had been extracted as a building. We can also observe that of the 12 misclassified objects, seven are located in the edge of image. Partially cut objects are more likely to be misclassified since most object properties may have been seriously damaged by edge-cutting, such as the number of buildings, building density, FAR, GCR, OCR. The remaining five misclassifications were mainly caused by a mixture of residential buildings and facility buildings. Detailed quality assessment of the land-use classification result will be quantitatively measured and analysed in Chapter 9.
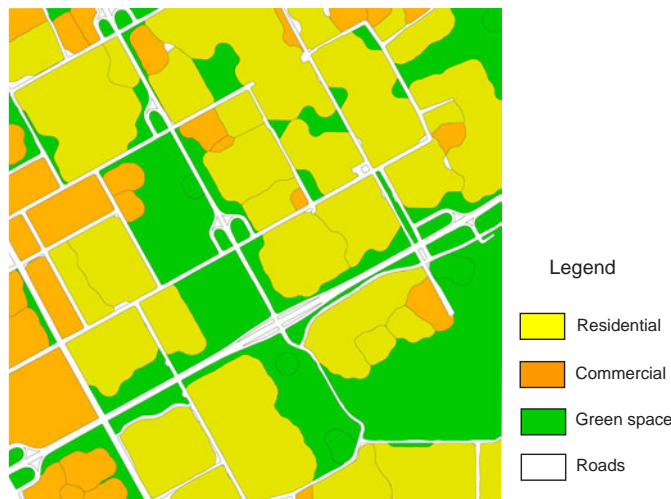
Figure 8.22: Classification result obtained by applying the proposed fuzzy membership functions.
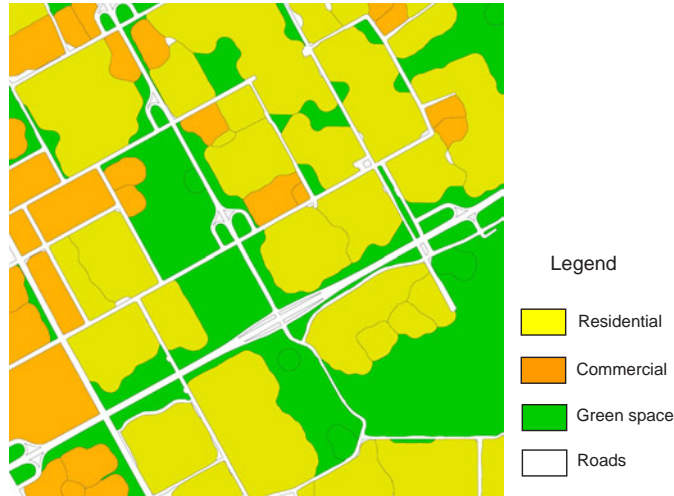
Figure 8.23: Classification based on visual interpretation and field visit.

8.5.5   Summary

In this chapter, we have proposed a number of meaningful object properties that are commonly used in urban planning and can be extracted from laser data and multi-spectral data. Several investigations have been made to find out the relationships between these properties and land-use classes. A fuzzy classification approach is applied, with fuzzy membership function based mainly on object definitions and local knowledge. The classification result obtained by using the proposed approach is promising. Such a result is unlikely to be achieved by the conventional per-pixel-based approaches, as much useful information and many object properties that can be extracted by the per-object approach cannot be derived by the per-pixel approaches. Urban areas are very complicated. The proposed approach has been tested in a relatively easy area where urban patterns are regular, and the test site is comparatively small. Further research is needed for applying the proposed methods in more complicated urban areas. However, this research can be regarded as one step forward toward automatic land-use extraction and classification from images in urban areas.

Chapter 9

Quality and uncertainty assessment [*]

9.1 Introduction

Many per-pixel classification approaches have been developed, such as the MLC, the k nearest neighbour classifier (k-NN), the neural network classifier (NN). In these approaches, the individual pixel is treated as a fundamental unit throughout the whole process: sample selection, classifier training, classification, preparation of reference data, accuracy and uncertainty assessment, etc. Many efforts have been made in accuracy and uncertainty assessment regarding the per-pixel approaches (Congalton and Mead, 1983; Janssen, 1994; Skidmore, 1999; Foody, 2000). In these approaches only uncertainty at a particular location for the variable has been discussed. Many applications, however, require predictions about multi-pixel regions, and issues of uncertainty become more complicated in such circumstances (Dungan, 2002). So far not much attention has been paid to accuracy and uncertainty assessment with regard to features extracted from images by object-based approaches.

The error matrix or confusion matrix is often used to compute quality measures such as user's accuracy, producer's accuracy, overall accuracy and the Kappa coefficient for quality assessment of classification results obtained by visual interpretation or per-pixel approach (Congalton and Mead, 1983; Janssen, 1994; Skidmore, 1999; Foody, 2000). In the case of visual interpretation, operators usually interpret an object and delineate the object boundaries, and then label it according to a designated class. To assess the interpretation quality, a field visit is made to check whether the assigned classes are correct for some randomly selected sample objects, and to count the number of objects that are correctly classified and the number of objects that are misclassified for each class. Quality assessment results are represented as an error matrix in order to compute the mentioned quality measures.

---

[*]This chapter is based on the following papers: Zhan et al. (2003) and Zhan et al. (under peer review (2)).

In the case of digital image classification by computer, quality assessment results are also represented by an error matrix in order to compute the mentioned quality measures, but often using randomly selected sample pixels. In the former case, the obtained quality measures indicate the classification quality in terms of objects (per-object) from the object perspective. In the latter case, the obtained quality measures indicate the classification quality related to object locations (per-pixel).

In our object-based land-cover and land-use classification, acquired objects (image objects) may have both classification errors and location errors. Quality assessment concerning classification errors can be divided into two aspects: correctness and completeness. Correctness measures the percentage of extracted objects that are correctly classified. Completeness measures the percentage of existing objects that are correctly explained by the classification of all existing objects (in the reference data). Location errors can be divided into two categories: errors in terms of object position (per-object location errors) and errors in terms of the spatial extent of an object (per-pixel location errors). Per-object location errors indicate the positional difference between the centre of mass of an extracted object and the centre of mass of the same object in the reference data.

In the following sections of this chapter, we will discuss both the per-object quality measures with regard to different object properties and the per-pixel quality measures concerning the spatial extent of objects. A short review of the error matrix and related measures, as well as certain limitations, is presented and discussed in Section 9.2. The proposed framework that is expected to utilise per-object and per-pixel quality measures is presented and discussed in Section 9.3. The quality assessment of land-cover objects is presented in Section 9.4. The uncertainty assessment of extracted buildings from the Amsterdam test site was presented in Section 6.2.6. Some efforts proposed for acquiring more compact objects and experimental results are presented and discussed in Section 9.5. The quality and uncertainty assessment of derived land-use objects is discussed in Section 9.6. The chapter closes with a general discussion and considers the outlook for the future.

9.2   The known methods of quality assessment

Quality is a very broad issue that may relate to a variety of properties but most frequently the property of interest is map or classification accuracy (Foody, 2000). Classification accuracy is typically taken to mean the degree to which the derived image classification or thematic map agrees with reality (Campbell, 1996). The error matrix or confusion matrix is a popular means for quality assessment of classification results (Congalton and Mead, 1983; Janssen, 1994; Skidmore, 1999; Foody, 2000), as shown in Table 9.1. Based on the error matrix, a number of quality measurements can be derived, such as overall accuracy, user's accuracy, producer's accuracy and the Kappa coefficient.

$$\text{Overall accuracy} = \frac{1}{\mathbf{n}} \sum_{k=1}^{m} \mathbf{n}_{kk}$$

Table 9.1: Error matrix for quality assessment

		Reference data				
		A	B	C	Total	User's accuracy
Classified data	A	$\mathbf{n}_{AA}$	$\mathbf{n}_{AB}$	$\mathbf{n}_{AC}$	$\mathbf{n}_{A+}$	$\mathbf{n}_{AA}/\mathbf{n}_{A+}$
	B	$\mathbf{n}_{BA}$	$\mathbf{n}_{BB}$	$\mathbf{n}_{BC}$	$\mathbf{n}_{B+}$	$\mathbf{n}_{BB}/\mathbf{n}_{B+}$
	C	$\mathbf{n}_{CA}$	$\mathbf{n}_{CB}$	$\mathbf{n}_{CC}$	$\mathbf{n}_{C+}$	$\mathbf{n}_{CC}/\mathbf{n}_{C+}$
	Total	$\mathbf{n}_{+A}$	$\mathbf{n}_{+B}$	$\mathbf{n}_{+C}$	$\mathbf{n}$	
	Producer's accuracy	$\frac{\mathbf{n}_{AA}}{\mathbf{n}_{+A}}$	$\frac{\mathbf{n}_{BB}}{\mathbf{n}_{+B}}$	$\frac{\mathbf{n}_{CC}}{\mathbf{n}_{+C}}$		

$$\text{Kappa coefficient} = \frac{\mathbf{n}\sum\limits_{k=1}^{m}\mathbf{n}_{kk} - \sum\limits_{k=1}^{m}\mathbf{n}_{k+}\mathbf{n}_{+k}}{\mathbf{n}^2 - \sum\limits_{k=1}^{m}\mathbf{n}_{k+}\mathbf{n}_{+k}}$$

Dungan (2002) stated that when new observational evidence is acquired that is incompatible with the results of the currently accepted model, uncertainty will be increased. Uncertainty may change when one is talking about a single pixel or multiple pixels. A confidence statement about the limited area represented by a single pixel may be different from a confidence statement about a large area of which that pixel forms only a part. This statement implies that the current per-pixel-based quality assessment measures may not be sufficient for quality assessment in the case of objects derived from images. This is because our spatial unit has been changed from an individual pixel to an individual object or multi-pixel region, while in image processing the error matrix and related measures are usually location-based (per-pixel). Additional per-object measures are needed to assess the quality and uncertainty of extracted objects from different perspectives such as position, size, shape, correctness, completeness. Detailed discussions on such measures follow in the coming sections. In the remaining part of this section, we will demonstrate the limitation of existing per-pixel quality measures by using the following example.

To examine whether known quality measures are still applicable in our object case, we compute them for the Ravensburg test site, using sampling at random locations. To this end we generate 100,000 samples from random locations by randomly generating a number $\in [0, 1]$, which is then scaled to represent the x coordinate, and randomly generating a number, which is then scaled to represent the y coordinate. We take the extracted buildings as an example of a classification. The elements of the error matrix are computed as shown in Table 9.2. A problem can be observed from the figures presented in this table, in that a very large number of pixels are found in the cell representing non-building in both classified data and reference data (*e* in Table 9.2). This indicates that the objects of interest only cover a small portion

Table 9.2: Error matrix for quality assessment of extracted buildings from the Ravensburg test site, based on 100,000 random samples

		Reference data			
		Building	Not building	Total	User's accuracy
Classified data	Building	3177 (*a*)	479 (*b*)	3656	86.9 % (*c*)
	Not building	699 (*d*)	95645 (*e*)	96344	99.3 % (*f*)
	Total	3876	96109	100000	
	Producer's accuracy	82.0 % (*g*)	99.5 % (*h*)		

Overall accuracy: 98.8 %, Kappa: 83.7 %, Overall quality: 73.0 %

of the scene. Samples falling in areas that do not contain objects will not make much sense for quality assessment, since we are only interested in the extracted objects. As a consequence, the error matrix shows an overestimated user's accuracy and producer's accuracy for non-building (*f* and *h* in Table 9.2), as well as an overestimated overall accuracy. The user's accuracy and the producer's accuracy for building (*c* and *g* in Table 9.2), however, can still be considered as reasonable measures for quality assessment since they have not been corrupted by the large number of pixels which exist in non-object areas. We also consider the Kappa coefficient to be valid since the Kappa coefficient takes into account the agreement contributed by chance. Kappa considers that the frequency of a sample appearing in a class is proportional to the percentage of locations (pixels) this class covers among all possible locations (the total size of the image). Therefore, we need to have a different interpretation of the figures in the error matrix. We can observe that user's accuracy and producer's accuracy for the object-related cells (*c* and *g* in Table 9.2) are calculated based on pixels falling in the object-related cells (*a*, *b* and *d* in Table 9.2). They can be understood as correctly classified (*a* in Table 9.2), wrongly detected (*b* in Table 9.2) and undetected (*d* in Table 9.2). In the coming section, we try to provide a united framework for quality assessment that utilises per-object and per-pixel measures. In the new framework, we expect to be able to suggest a number of new measures, as well as to solve the problems as presented in this section.

9.3 Quality measures for object properties and spatial extent

The generic meaning of uncertainty deals with the subjective. While two individuals may arrive at the same answer to a question, one individual may be more certain than the other about that answer. Given that measures of uncertainty, Bayesian or Frequentist, are important for gauging progress, they should be agreed upon with

some degree of consensus and shown to be used over time (Dungan, 2002). We will explain in the following why a united framework is needed for quality assessment from different aspects.

9.3.1 Initiatives for a united framework for quality assessment from different aspects

Need for quality assessment of the single-class cases

As demonstrated in the previous section, simple pixel-based quality measures are biased because of the existence very large samples that do not make much sense for quality assessment of the single-class cases, such as quality assessment of extracted buildings. A similar problem may also occur in the multi-class cases. For instance, when classification results are obtained by per-pixel classification in a coastal area where the water surface covers a very large portion of the image, an overestimated overall accuracy is likely to obtained. A very large number of samples are likely to be selected from water areas by the random sampling approach. Most of these samples will easily have been classified correctly, based on a clear and compact cluster in the feature space, whereas other classes may be classified with low accuracy because of a certain degree of mixture in the feature space. In such cases, the Kappa coefficient is likely to produce a reasonable figure for a per-pixel classification result. But we still need a per-object overall quality measure for the single-class cases, since we cannot produce the Kappa coefficient when the number of objects not belonging to the designated class is not available or is not of interest to us, as shown in Tables 9.3 and 9.2.

Differences between per-object and per-pixel measures

We explain the difference in quality assessment by per-object and per-pixel measures in the following way: per-pixel measures assess how good the classification is at locations (pixels) while per-object measures assess how good the classification is of a multi-pixel grouping (image object), using the objects as a counting unit. Therefore, they can be seen as measures concerned with different aspects. To apply per-object measures, we must solve the object matching problem. In this research, we consider an extracted object (such as a building) as matching an object in the reference data if the two overlap by at least 50 % and the overlapping part is larger than or equal to 10 pixels. We have chosen these values considering that the ratio criterion of 50 % may not be sufficient for small objects that consist of only a few pixels. Figure 9.1 illustrates various constellations of two matched objects. Objects that match according to the above criteria are then considered as being classified correctly. All four cases, which are shown in Figure 9.1, are considered as being classified correctly irrespective of their spatial extent. The per-object measures, however, should also be able to assess the differences in the spatial extent of matching objects.

Need for quality assessment of other object properties

In many cases, we may be interested in the quality in terms of object properties such as object size and the position of extracted objects, in addition to simply assessing correct object classification. A per-object quality measure related to object size can be used to assess situations, as presented in cases **C** and **D** in Figure 9.1. Case **C** is an example of the extracted object being smaller than the reference object while in case **D** the extracted object is larger than the corresponding object in the reference data. In both cases there is no error of position. A per-object quality measure related to position can be used to assess situations such as that in case **B** in Figure 9.1, where the extracted object is not in the same position as the corresponding reference object, while their sizes are identical. This measure can be associated with a registration error between image and reference map.
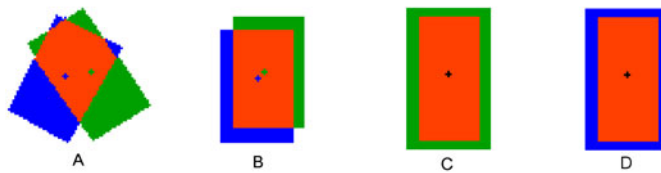
Figure 9.1: Four matched cases of an extracted object (orange colour indicates the matched region; blue colour indicates extracted region that is not explained by the reference data; green colour indicates a region in reference data that was not extracted): A - more than 50 % matched; B - matched, with the same size but different position; C and D - matched, with the same position but different spatial extent.

9.3.2   A united framework for quality assessment based on the feature contrast model

To develop a united framework for quality assessment, we consider the use and extension of Kversky's feature contrast model (Tversky, 1977) to measure the degree of similarity between classification results and reference data from different aspects, and to use them as quality measures suitable for both per-object and per-pixel cases. The more features that match between the classification results and the reference data supposedly representing reality, the higher we consider the quality of those results to be – this also applies when reality is subjectively described by definitions such as land-use classes.

$$Similarity = \frac{f(\mathbf{C} \cap \mathbf{R})}{f(\mathbf{C} \cap \mathbf{R}) + \alpha \cdot f(\mathbf{C} - \mathbf{R}) + \beta \cdot f(\mathbf{R} - \mathbf{C})}$$

The similarity between classified data (**C**) and reference data (**R**) based on a

specific feature is expressed as a function (f) of the three arguments: $f(\mathbf{C} \cap \mathbf{R})$, the features that are common to both $\mathbf{C}$ and $\mathbf{R}$; $f(\mathbf{C} - \mathbf{R})$, the features that belong to $\mathbf{C}$ but not to $\mathbf{R}$; $f(\mathbf{R} - \mathbf{C})$, the features that belong to $\mathbf{R}$ but not to $\mathbf{C}$. α and β denote weights for $f(\mathbf{C} - \mathbf{R})$ and $f(\mathbf{R} - \mathbf{C})$ respectively. $\alpha = \beta$ if $\mathbf{C}$ and $\mathbf{R}$ are symmetric. $\alpha \neq \beta$ if $\mathbf{C}$ and $\mathbf{R}$ are asymmetric. We can relate this model to the error matrix. For an error matrix of those classes (see Table 9.1), $\mathbf{n}_{AA}$ can be regarded as $f(\mathbf{C} \cap \mathbf{R})$. $\mathbf{n}_{AB}$ and $\mathbf{n}_{AC}$ can be treated as $f(\mathbf{C} - \mathbf{R})$. $\mathbf{n}_{BA}$ and $\mathbf{n}_{CA}$ can be treated as $f(\mathbf{R} - \mathbf{C})$. This similarity ratio model can be extended and applied to assess the quality of extracted objects since many features can be selected for such comparison. We will now explain the existing quality measures and propose some new measures within the framework of feature similarity.

9.3.3 Explanation of the existing quality measures in the new framework

The two parameters, α and β, as presented in the feature contrast model can be regarded as weights for two aspects of mismatch. In most cases, we consider $\alpha = \beta = 1$. Within the framework of feature similarity, the figures in the diagonal cells of an error matrix are regarded as matched features, i.e. $f(\mathbf{C} \cap \mathbf{R})$; the figures in off-diagonal cells of an error matrix are regarded as mismatched features, i.e. $f(\mathbf{C} - \mathbf{R})$ and $f(\mathbf{R} - \mathbf{C})$. For the single-class assessment, we introduce the *overall quality* (*OQ*) (please note this is different from overall accuracy as defined before). The overall quality can be understood as a percentage of the number of matched objects among the total number of objects in the classification result and the reference data.

$$OQ_k = \frac{f(\mathbf{C}_k \cap \mathbf{R}_k)}{f(\mathbf{C}_k \cap \mathbf{R}_k) + f(\mathbf{C}_k - \mathbf{R}_k) + f(\mathbf{R}_k - \mathbf{C}_k)}, k = 1, ..., m,$$

where k denotes a designated class (land-cover or land-use) and m denotes the total number of designated classes. Thus the overall quality for both visual interpretation results (per-object) and for computer image classification (per-pixel) can be expressed as:

$$OQ \text{ for class } k = \frac{N(\mathbf{C}_k \cap \mathbf{R}_k)}{N(\mathbf{C}_k \cap \mathbf{R}_k) + N(\mathbf{C}_k - \mathbf{R}_k) + N(\mathbf{R}_k - \mathbf{C}_k)}$$

$$= \frac{\mathbf{n}_{kk}}{\mathbf{n}_{kk} + (\mathbf{n}_{k+} - \mathbf{n}_{kk}) + (\mathbf{n}_{+k} - \mathbf{n}_{kk})},$$

where N is a function of object numbers (number of objects (N_o) in visual interpretation cases, number of pixels (N_p) in computer image classification), $\mathbf{n}$ denotes the actual the number of objects, and k denotes a designated class.

Similarly the user's accuracy (*UA*) and producer's accuracy (*PA*) can be expressed as follows:

$$UA \text{ for class } k \;=\; \frac{N(\mathbf{C}_k \cap \mathbf{R}_k)}{N(\mathbf{C}_k \cap \mathbf{R}_k) + N(\mathbf{C}_k - \mathbf{R}_k)}$$

$$=\; \frac{\mathbf{n}_{kk}}{\mathbf{n}_{kk} + (\mathbf{n}_{k+} - \mathbf{n}_{kk})}$$

$$=\; \frac{\mathbf{n}_{kk}}{\mathbf{n}_{k+}}$$

$$PA \text{ for class } k \;=\; \frac{N(\mathbf{C}_k \cap \mathbf{R}_k)}{N(\mathbf{C}_k \cap \mathbf{R}_k) + N(\mathbf{R}_k - \mathbf{C}_k)}$$

$$=\; \frac{\mathbf{n}_{kk}}{\mathbf{n}_{kk} + (\mathbf{n}_{+k} - \mathbf{n}_{kk})}$$

$$=\; \frac{\mathbf{n}_{kk}}{\mathbf{n}_{+k}}$$

In the proposed framework, the overestimation problem in quality assessment when using the error matrix for a single class (the example presented in Section 9.2) can be solved as follows. Both user's accuracy and producer's accuracy are useful measures for assessing the quality of a single class, and so is the overall quality that combines the user's accuracy and the producer's accuracy.

$$\text{Per-pixel } OQ \text{ for extracted buildings} \;=\; \frac{N_p(\mathbf{C}_b \cap \mathbf{R}_b)}{N_p(\mathbf{C}_b \cap \mathbf{R}_b) + N_p(\mathbf{C}_b - \mathbf{R}_b) + N_p(\mathbf{R}_b - \mathbf{C}_b)}$$

$$=\; \frac{\mathbf{n}_{bb}}{\mathbf{n}_{bb} + (\mathbf{n}_{b+} - \mathbf{n}_{bb}) + (\mathbf{n}_{+b} - \mathbf{n}_{bb})}$$

$$=\; \frac{\mathbf{n}_{bb}}{\mathbf{n}_{b+} + \mathbf{n}_{+b} - \mathbf{n}_{bb}}$$

$$=\; 73.0\ \%$$

$\mathbf{n}_{bb}$ is the number of matching pixels (within the random samples).

$$UA \text{ for extracted buildings} \;=\; \frac{N_p(\mathbf{C}_b \cap \mathbf{R}_b)}{N_p(\mathbf{C}_b \cap \mathbf{R}_b) + N_p(\mathbf{C}_b - \mathbf{R}_b)}$$

$$=\; \frac{\mathbf{n}_{bb}}{\mathbf{n}_{b+}}$$

$$=\; 86.9\ \%$$

$$PA \text{ for extracted buildings} \;=\; \frac{N_p(\mathbf{C}_b \cap \mathbf{R}_b)}{N_p(\mathbf{C}_b \cap \mathbf{R}_b) + N_p(\mathbf{R}_b - \mathbf{C}_b)}$$

$$=\; \frac{\mathbf{n}_{bb}}{\mathbf{n}_{+b}}$$

$$=\; 82.0\ \%$$

The overall quality can also be produced based on *UA* and *PA*:

$$OQ = \frac{1}{\frac{1}{UA} + \frac{1}{PA} - 1}$$

The quality assessment results obtained in this case are considered as quality measures in terms of spatial extent or location since pixels with random locations are used in computing the results. Following the same line of thinking, quality assessment can be applied by counting the number of objects to produce per-object overall quality, correctness and completeness. Per-object quality assessment can be assessed according to various object properties of extracted objects, such as object size and object position. In the remaining part of this section, we will define a number of per-object quality measures based on object properties. Similar measures have also been proposed by Heipke et al. (1997); Wiedemann et al. (1998).

9.3.4 Per-object quality measures based on object properties

Per-object overall quality

The formula remains the same for per-object overall quality but counting the number of objects instead of the number of pixels. The per-object overall accuracy will produce the same figure as when measuring the accuracy of a visual interpretation.

$$\text{Overall quality}_{Object} = \frac{N_o(\mathbf{C}_k \cap \mathbf{R}_k)}{N_o(\mathbf{C}_k \cap \mathbf{R}_k) + N_o(\mathbf{C}_k - \mathbf{R}_k) + N_o(\mathbf{R}_k - \mathbf{C}_k)}$$

$$= \frac{\mathbf{n}_{kk}}{\mathbf{n}_{kk} + (\mathbf{n}_{k+} - \mathbf{n}_{kk}) + (\mathbf{n}_{+k} - \mathbf{n}_{kk})}$$

$$= \frac{\mathbf{n}_{kk}}{\mathbf{n}_{k+} + \mathbf{n}_{+k} - \mathbf{n}_{kk}}$$

$\mathbf{n}_{kk}$ is now the number of matching objects.

Per-object user's accuracy - correctness

The per-object user's accuracy can be regarded as the correctness of obtained results and can be computed by using the same formula as user's accuracy.

$$\text{Correctness} = \frac{N_o(\mathbf{C}_k \cap \mathbf{R}_k)}{N_o(\mathbf{C}_k \cap \mathbf{R}_k) + N_o(\mathbf{C}_k - \mathbf{R}_k)}$$

$$= \frac{\mathbf{n}_{kk}}{\mathbf{n}_{k+}}$$

Per-object producer's accuracy - completeness

The per-object producer's accuracy can be regarded as the completeness of obtained results and can be computed by using the same formula as producer's accuracy.

$$\text{Completeness} \quad = \frac{N_o(\mathbf{C}_k \cap \mathbf{R}_k)}{N_o(\mathbf{C}_k \cap \mathbf{R}_k) + N_o(\mathbf{R}_k - \mathbf{C}_k)}$$

$$= \frac{\mathbf{n}_{kk}}{\mathbf{n}_{+k}}$$

Per-object quality measure defined in terms of similarity in object size (Similarity_{Size})

In many cases, we may wish to find out how good or how similar the extracted objects are in terms of object size as compared with the reference data. A quality measure in terms of object size is proposed that measures similarity between the size of an extracted object and the size of the corresponding object presented in the reference data.

$$\text{Similarity}_{Size}(O_i) = \frac{min(Size_C(O_i), Size_R(O_i))}{max(Size_C(O_i), Size_R(O_i))}$$

$$Mean_{\text{Similarity}_{Size}} = \frac{1}{n} \sum_{i=1}^{n} \text{Similarity}_{Size}(O_i)$$

$$Std_{\text{Similarity}_{Size}} = \frac{1}{n} \sqrt{\sum_{i=1}^{n} (\text{Similarity}_{Size}(O_i) - Mean_{\text{Similarity}_{Size}})^2}$$

where $Size_C(O_i)$ denotes the size of the extracted object O_i and $Size_R(O_i)$ denotes the size of the corresponding object presented in the reference data. The function $min(a, b)$ returns the smaller value of the two arguments. If $min(a, b)$ is a, then $max(a, b)$ is b. $\text{Similarity}_{Size}(O_i)$ measures similarity in terms of object size for object O_i. $Mean_{\text{Similarity}_{Size}}$ is a measure that indicates the average similarity in terms of object size for all extracted objects. $Std_{\text{Similarity}_{Size}}$ is the standard deviation of similarity in terms of object size for all extracted objects.

Per-object quality measure defined in terms of object location ($\text{Quality}_{Location}$)

In many cases, we may wish to find out how good the extracted objects are in terms of object location as compared with the reference data. A quality measure in terms of individual object location is proposed that is measured based on the distance between the centres of mass of an extracted object and the corresponding object in the reference data.

$$\text{Quality}_{Location}(O_i) = \sqrt{(X_C(O_i) - X_R(O_i))^2 + (Y_C(O_i) - Y_R(O_i))^2}$$

$$Mean_{\text{Quality}_{Location}} = \frac{1}{n} \sum_{i=1}^{n} \text{Quality}_{Location}(O_i)$$

$$Std_{\text{Quality}_{Location}} = \frac{1}{n} \sqrt{\sum_{i=1}^{n} (\text{Quality}_{Location}(O_i) - Mean_{\text{Quality}_{Location}})^2}$$

where $X_C(O_i)$ and $Y_C(O_i)$ denote the x and y coordinates of the centre of mass of extracted object O_i. $X_R(O_i)$ and $Y_R(O_i)$ denote the x and y coordinates of the centre of mass of the corresponding object presented in the reference data.

$Mean_{\text{Quality}_{Location}}$ is a measure that indicates an average quality in terms of object location for all extracted objects. $Std_{\text{Quality}_{Location}}$ is the standard deviation of quality measure in terms of object location for all extracted objects.

9.4 Quality assessment of extracted buildings (land-cover objects)

Based on the proposed framework for quality assessment, we obtain quality assessment results for various aspects of extracted buildings from the Amsterdam test site and Ravensburg test site.

9.4.1 Quality assessment of spatial extent of buildings by using randomly generated sample pixels

The per-pixel overall quality computed for extracted buildings from the Amsterdam test site, based on 100,000 sample pixels, is 76.4 % according to the figures shown in Table 9.3. These figures are obtained based on the reference map shown in Figure 6.10. Based on the same figures, the Kappa coefficient, user's accuracy and producer's accuracy are calculated as 85.5 %, 81.8 % and 92.0 % respectively for the Amsterdam test site. The main causes of error in terms of the spatial extent of buildings on the Amsterdam test site are several large parking garages. These have not been extracted because they are directly connected to the nearby road and do not show the desired characteristic of vertical walls. Moreover, a number of metro stations have been extracted as buildings but are not shown on the reference map.

The per-pixel overall quality computed for extracted buildings from the Ravensburg test site, based on 100,000 sample pixels, is 73.0 % according to figures shown in Table 9.2. These figures are obtained based on the reference map shown in Figure 6.15. Based on the same figures, the Kappa coefficient, user's accuracy and producer's accuracy are calculated as 83.7 %, 86.9 % and 82.0 % respectively for the Ravensburg test site. The main causes of error in terms of the spatial extent of buildings on the Ravensburg test site are the existence of many small houses with gable roofs and high trees that are very close to low-rise buildings.

Table 9.3: Error matrix for quality assessment of extracted buildings from the Amsterdam test site, based on 100,000 random samples

		Reference data			
		Building	Not building	Total	User's accuracy
Classified data	Building	9123	2032	11155	81.8 %
	Not building	789	88056	88845	99.1 %
	Total	9912	90088	100000	
	Producer's accuracy	92.0 %	97.7 %		

Overall accuracy: 97.2 %, Kappa: 85.0 %, Overall quality: 76.4 %

9.4.2 Quality assessment by counting numbers of objects

Per-pixel measures can provide information on quality, but they are basically dealing with quality at an 'individual' location for the variables. Quality measures are still needed in an object perspective. Quality assessment by counting the number of objects that are correctly detected, the number of objects that are wrongly detected and the number of objects that are not detected can provide quality information about extracted objects, as shown in Tables 9.4 and 9.5.

Amsterdam test site

The error matrix shown in Table 9.4 is obtained for per-object quality assessment according to figures presented in Table 6.6, which are obtained based on the reference map shown in Figure 6.10. The overall quality of extracted buildings from the Amsterdam test site is calculated as $\frac{683}{683+25+26} = 93.1\%$. Correctness and completeness of the extracted buildings are computed as 96.5 % and 96.3 % respectively.

Ravensburg test site

The error matrix shown in Table 9.5 is obtained for per-object quality assessment according to figures presented in Table 6.7, which are obtained based on the reference map shown in Figure 6.15. The overall quality of extracted buildings from the Ravensburg test site is calculated as $\frac{150}{150+7+23} = 83.3\%$. Correctness and completeness of the extracted buildings are computed as 95.5 % and 86.7 % respectively.

Table 9.4: Error matrix for quality assessment of extracted buildings from the Amsterdam test site in terms of the number of objects

| | | Reference data | | | |
		Building	Others	Total	Correctness
Classified data	Building	683	25	708	96.5 %
	Others	26			
	Total	709			
	Completeness	96.3 %			

Overall quality: 93.1 %

Table 9.5: Error matrix for quality assessment of extracted buildings from the Ravensburg test site in terms of the number of objects

| | | Reference data | | | |
		Building	Others	Total	Correctness
Classified data	Building	150	7	157	95.5 %
	Others	23			
	Total	173			
	Completeness	86.7 %			

Overall quality: 83.3 %
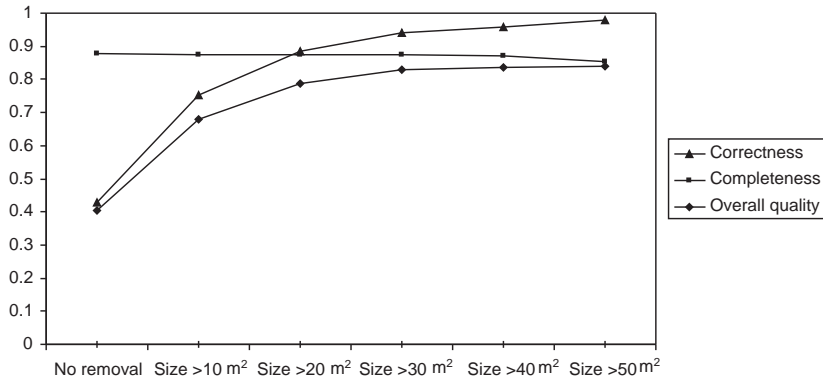
Figure 9.2: Impact of removing small objects on proposed measures.

Relationship between correctness and completeness in the case of the Ravensburg test site

A test was made to see the responses of these measures to the removal of small objects. When all small objects remain, the completeness figure is high, whereas the correctness figure is lowered because of the existence of many small objects. Many small objects are in fact non-building. When small objects are removed to a certain degree, correctness figures climb but completeness figures drop, as shown in Figure 9.2. The overall quality indicates a point of balance between correctness and completeness, which may be useful for general cases such as using these data in planning. In extreme cases, when we provide building data to locate people in these buildings for a rescue operation, completeness is far more important than correctness, since we do not want to miss any building under such circumstances.

9.4.3 Quality assessment in terms of similarity in object size

In many cases, we may wish to find out how good or how similar extracted objects are in terms of object size as compared with reference data. A per-object quality measure in terms of similarity in object size as proposed earlier is obtained for extracted buildings from the Amsterdam test site and the Ravensburg test site.

Amsterdam test site

The mean size similarity value 0.8765 and the standard deviation 0.1272 are calculated based on the classification result and the reference data as shown in Figures 6.7 and 6.8 respectively. The mean size similarity 0.8765 in the scale of 0 to 1 shows high similarity in terms of building size, which means that the major parts of extracted buildings match the corresponding buildings presented in the reference

data. The low standard deviation 0.1272 shows consistent results are obtained for extracted buildings.

Ravensburg test site

The mean size similarity value 0.8574 and the standard deviation 0.1133 are calculated based on the classification result and the reference data as shown in Figure 6.12 and Figure 6.14 respectively. These figures show high similarity in terms of building size; thus the major parts of extracted buildings match the corresponding buildings presented in the reference data. The low standard deviation 0.1133 shows consistent results are obtained for extracted buildings.

9.4.4   Quality assessment in terms of object location

The Euclidean distance between the centres of mass (gravity centres) of an extracted object and the corresponding object presented in the reference data is computed as a measure of the positional quality of the extracted objects.

The mean distance between the centres of mass of corresponding buildings and the standard deviation computed for the Amsterdam test site are 1.9961 m and 4.4102 m. This means that the positions of extracted buildings have shifted about two pixels on average from their positions in the reference data.

The mean distance between the centres of mass of corresponding buildings and the standard deviation computed for the Ravensburg test site are 1.1474 m and 0.9387 m. This means that the positions of extracted buildings have shifted one to two pixels on average from their positions in the reference data.

The results obtained for both test sites show that correctly extracted buildings are very well located.

9.4.5   Quality comparison between buildings extracted from the Amsterdam test site and buildings extracted from the Ravensburg test site

To make a more comprehensive comparison of the building extraction results obtained for the two test sites, we list all quality assessment results obtained based on the proposed quality measures, as shown in Table 9.6. Based on these figures, we can conclude that the extracted result for the Amsterdam test site is generally better than that obtained for the Ravensburg test site according to several overall quality measures such as per-pixel and per-object overall quality and the Kappa coefficient. The main reason is that the Amsterdam test site has many large buildings, which are relatively easier to extract than the many small residential buildings of the Ravensburg test site, especially as these are mixed with trees. Quality measures in terms of building size show that very similar results are obtained from both sites. The quality measures in terms of object location show that the result obtained for the Ravensburg test site is better than the result obtained for the Amsterdam test site. The main reason is likely to be that the reference data for the Amsterdam test site are obtained by digitising the large-scale base map. Many building forecourts

Table 9.6: Comparison of quality assessment results for building extraction obtained from two test sites

Quality measures		Amsterdam	Ravensburg
Per-pixel	Overall quality	76.4 %	73.0 %
	Kappa coefficient	85.0 %	83.7 %
	User's accuracy	81.8 %	86.9 %
	Producer's accuracy	92.0 %	82.0 %
Per-object	Overall quality	93.1 %	83.3 %
	Correctness	96.5 %	95.5 %
	Completeness	96.3 %	86.7 %
	Mean similarity$_{Size}$	0.8765	0.8574
	Std similarity$_{Size}$	0.1272	0.1133
	Mean quality$_{Location}$	1.9961	1.1474
	Std quality$_{Location}$	4.4102	0.9387

were not delineated as parts of buildings in this base map, but they are extracted as parts of buildings in the extracted result. The reference data for the Ravensburg test site are acquired by visual interpretation and screen digitising based on the same image used for building extraction.

9.5 Quality improvement by forming a more compact object shape

A weighted smoothing filter convolution is proposed, and used to obtain a more compact form of an object by applying this filter to overall fuzzy membership values. The proposed weighted smoothing filter as shown in Figure 9.3 has the following properties: $\sum$ p=1 to ensure that the result remains in the range [0, 1]; the central pixel should play a dominant role; the four more closely adjacent pixels should obtain higher weights (i.e. twice as high in this case) than the four corner pixels. The expected role of the smoothing convolution is to improve object quality, obtain a more compact form of an object and eliminate noise and unnecessary links between objects. A test has been made with 100,000 randomly generated samples over the test site to investigate its impact on the overall member function values of pixels as parts of objects in different categories. The test results are shown in Tables 9.7 and 9.8. The mean value (MF_{OA}) of pixels belonging to correctly classified buildings slightly decreased from 0.8048 to 0.7805. The mean value (MF_{OA}) of pixels belonging to wrongly classified buildings significantly decreased from 0.7649 to 0.6679. These figures show that the smoothing filter convolution can decrease the chance of making mistakes in classifying pixels from non-building to building, while no major negative impact is observed for the remaining parts of the figures, as shown in Ta-

0.05	0.1	0.05
0.1	0.4	0.1
0.05	0.1	0.05

Figure 9.3: A weighted smoothing filter for a more compact form of objects.

Table 9.7: Mean and standard deviation (*Std*) of overall fuzzy membership function values (MF_{OA}) before applying the proposed smoothing filter convolution, based on 100,000 random samples

		Reference data			
		Building		Non-building	
		Mean	Std	Mean	Std
Classified data	Building	0.8048	0.0817	0.7649	0.0913
	Non-building	0.1382	0.2250	0.0110	0.0622

bles 9.7 and 9.8. We can also observe by comparing histograms **b** and **b'** in Figure 9.4 that the distribution has been turned to another direction, where higher frequency is found to be close to 1 before applying the smoothing filter and close to 0.5 (0.5 is the threshold for object detection) after applying the smoothing filter for wrongly detected objects. Both observations indicate that the proposed smoothing filter is effective for our objectives mentioned earlier (more compact object shapes can be found in the image).

9.6 Quality and uncertainty assessment of land-use classification (land-use objects)

To assess the land-use classification result, the classified result as shown in Figure 8.22 is compared with the reference data as shown in Figure 8.23. Quantitative assessment is made in terms of object number and with consideration of object size. Uncertainty assessment is made by object-object comparison of the assigned class labels in classified data and the reference data. Quality assessment regarding the spatial extent of land-use units is discussed at the end of this section.

Table 9.8: Mean and standard deviation (*Std*) of overall fuzzy membership function values (MF_{OA}) after applying the proposed smoothing filter convolution, based on 100,000 random samples

		Reference data			
		Building		Non-building	
		Mean	Std	Mean	Std
Classified data	Building	0.7805	0.1034	0.6679	0.1058
	Non-building	0.1230	0.1897	0.0098	0.0538
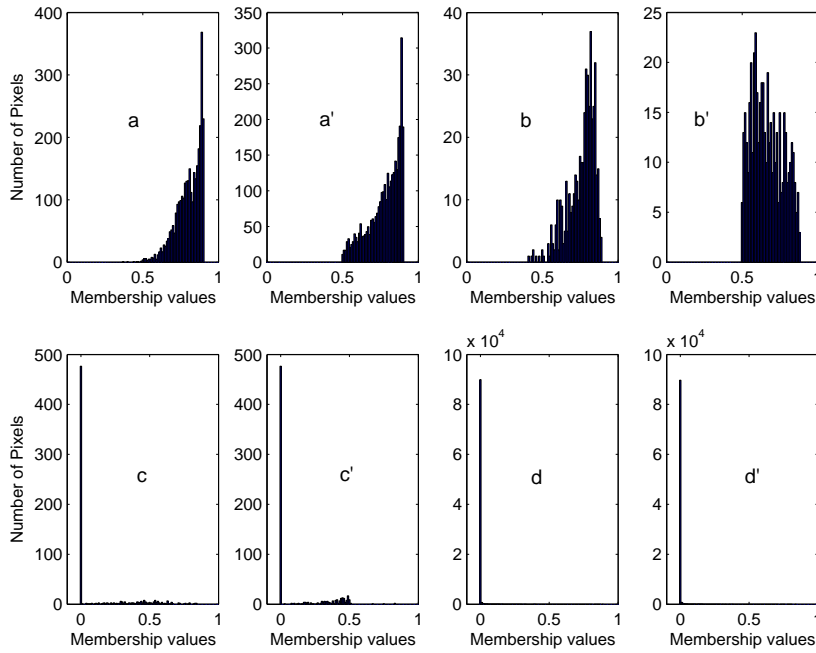
Figure 9.4: Histogram plots for distributions of overall fuzzy membership function values (MF_{OA}) within correct objects (a), mistaken objects (b), undetected objects (c), non-object areas (d) and their counterparts after applying the proposed smoothing filter (a', b', c', d').

Table 9.9: Error matrix of per-object land-use classification obtained from Amsterdam test site

		Reference data				
		Resid.	Comm.	Green	Total	User's accuracy
Classified data	Residential	26	5	0	31	83.9 %
	Commercial	6	21	1	28	75.0 %
	Green space	0	0	43	43	100 %
	Total	32	26	44	102	
	Producer's accuracy	81.3 %	80.8 %	97.7 %		

Overall accuracy: 88.2 %; Kappa coefficient: 82.0 %

9.6.1 Quality assessment by counting numbers of objects

The figures obtained by counting the number of objects that are correct or incorrect when compared with reference data using the per-object approach are presented in Table 9.9. Based on these figures, we can conclude that most land-use objects are correctly classified, with an overall accuracy of 88.2 % and a Kappa coefficient of 82.0 %. Public green space gains both the highest user's accuracy (100 %) and the highest producer's accuracy (97.7 %). Residential area obtains slightly higher user's accuracy (83.9 %) and producer's accuracy (81.3 %) than commercial does (75.0 % and 80.8 %).

9.6.2 Quality assessment by considering object size

The quality assessment figures by considering the object size are acquired by the per-pixel approach, which counts the number of pixels in each land-use unit and checks class labels, as shown in Table 9.10. Based on these figures, we see that most land-use objects are correctly classified taking object size as the criterion. The overall accuracy is 95.3 % and the Kappa coefficient is 92.0 %. Green space gains both the highest user's accuracy (100 %) and the highest producer's accuracy (99.5 %). Residential obtains slightly higher user's accuracy (94.6 %) and producer's accuracy (96.9 %) than commercial does (87.6 % and 81.3 %). The reason for the higher-quality figures obtained by considering object size and using the per-pixel approach (overall accuracy: 95.3 %; Kappa coefficient: 92.0 %), as compared with those obtained by the per-object approach (overall accuracy: 88.2 %; Kappa coefficient: 82.0 %), is mainly because large land-use objects such as large residential areas are more likely to be classified correctly and most misclassified objects are small objects, as discussed earlier in Chapter 8. The relationship between object size, certainty measure and correctness as shown in Figure 9.6 also supports the above judgement.

Table 9.10: Error matrix of per-object land-use classification by considering object size

		Reference data				User's accuracy
		Resid.	Comm.	Green space	Total	
Classified data	Residential	4128568	234057	0	4362625	94.6 %
	Commercial	130240	1015496	13272	1159008	87.6 %
	Green space	0	0	2457512	2457512	100 %
	Total	4258808	1249553	2470784	7979145	
	Producer's accuracy	96.9 %	81.3 %	99.5 %		

Overall accuracy: 95.3 %;   Kappa coefficient: 92.0 %

9.6.3   Uncertainty assessment of classification result

To acquire a 'hard' classification result in fuzzy classification, we normally label the class according to the highest membership value among others. The difference between the highest membership value and the second highest membership value indicates how close these two membership values are for each land-use object. The larger the difference is between the highest membership value and the second highest membership value, the more certain we consider the class assigned to an object. Uncertainty assessment based on this difference is shown in Figure 9.5.

Based on Figure 9.5, we can observe that public green space obtains the highest certainty values, which indicates that public green space is clearly separate from other classes in general. Residential and commercial classes form a considerable mixture.

The average certainty value of 90 correctly classified objects is 0.23; the value for 12 misclassified objects is 0.05. Of 90 correctly classified objects, 26 residential objects get the average certainty value of 0.079, 21 commercial objects get the average certainty value of 0.075, 43 green space objects get the average certainty value of 0.39. These figures indicate that the proposed uncertainty measure based on the difference between the highest and the second highest membership values can reflect the difference and support our observation that green space is easily separated from other classes in general. Residential and commercial are mixed. Based on the certainty values, the users can concentrate their attention on those objects that obtain low certainty values and correct them manually. The relationship between the size of land-use objects, the classification certainty measurement and correctness as shown in Figure 9.6 suggests that the most misclassified objects are small objects and that they received low classification certainty values.
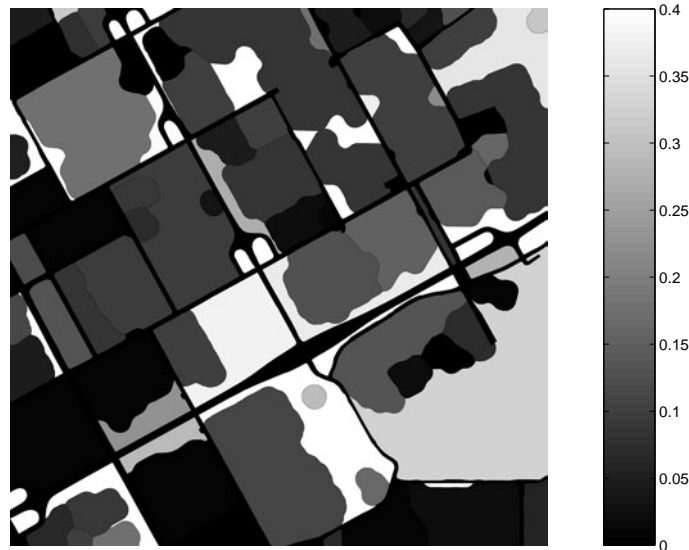
Figure 9.5: Uncertainty assessment measured based on difference between the highest and the second highest fuzzy membership values for each land-use object.

9.6.4 Quality assessment in terms of spatial extent of land-use objects

As we had no authorised land-use map for comparative purposes, no quality assessment has been carried out in terms of the spatial extent of land-use objects. Land-use maps produced by individuals using the visual interpretation method are highly subjective. Although agreement may be reached on assigning a specific class label to a specific area, very often interpreters may not be able to reach agreement on its spatial boundary, especially when no physical feature can be found in the image and in reality to support such delineation. We observe that precision in terms of the spatial extent of land-use units is not considered as important as assigning the right class labels. For instance, in land-use mapping by visual interpretation, people often use straight lines to delineate land-use areas in order to produce more compact and visually attractive results, even in a case where a slightly curved line is suggested by the physical features. In addition, a small area is often ignored and simply delineated as part of a large land-use unit in the surrounding areas.
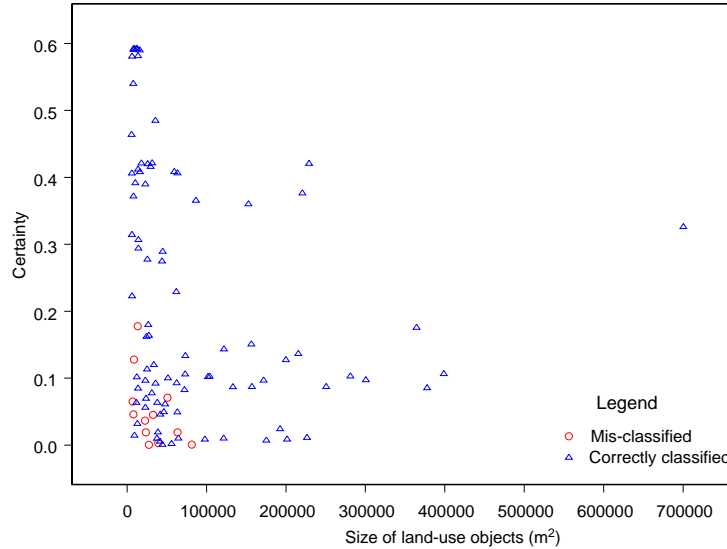
Figure 9.6: Relationship between size of land-use objects, classification certainty and correctness.

9.7 Summary

Quality assessment of objects extracted directly from images is still in its initial stage. We can simply provide users with a fuzzy result, with detailed overall fuzzy membership function values that indicate the possibility of an object being present or absent at a certain location, the spatial extents of objects and the magnitude of uncertainties. Correct interpretation of such results requires solid technical knowledge as well as knowledge about the application fields. Therefore, a crisp classified result and associated quality measures are needed to avoid misunderstanding in the use of such a fuzzy result. Several existing per-pixel-based measures have been tested in the object environment and in the single-class cases. Some of them are no longer valid, while others may be reused, but a different interpretation of these measures is needed in an object environment. Thus a united framework for quality assessment is proposed, utilising per-pixel and per-object quality measures and providing quality measures for the single-class cases. In the united framework, several existing quality measures are explained, and a number of per-object quality assessment measures are proposed. We have explained them for land-cover objects by taking the example of buildings considering the properties, correct classification, size and position. Several per-object quality assessment measures are proposed for land-use objects in terms of 'object number' and considering object size. The difference between the highest membership value and the second highest membership value is

proposed as a measure for uncertainty assessment of land-use classification results. Testing was carried out on results obtained by an object-based approach for building extraction and land-use classification from our test sites. The test results show that the proposed measures can provide useful information from different perspectives concerning uncertainties caused by different sources. These measures provide several choices and allow the users to choose and judge them based on their application requirements. Further investigation are needed, related to application range and the sensitivity to different data sets.

Chapter 10

Conclusions and future research

10.1 Conclusions

1. In this research, several new sensor data have been used and examined for urban land-cover and land-use classification. Triggered by the potential problems of high-resolution data, we have examined the most popular pixel-based classifier, the maximum likelihood classifier, as an example of the behaviour of traditional classifiers toward high-resolution data. A number of problems have been observed and highlighted, such as the existence of sub-clusters in the feature space, which lead to misclassification in the end-member classes. Several remedial measures for the observed problems have been proposed and tested. Land-cover classification accuracy can be improved by modelling the decision surface in the feature space and by selecting samples from both pure pixels and mixed pixels. Spatial partitioning of decision surfaces, which are estimated based on samples of end-member classes, has been the focus of the proposed solutions. The increasing classification accuracy of the experimental results has confirmed the effectiveness of the proposed class integration method using pure and mixed samples. However, other problems such as pixels in shadow areas and relief displacement caused by non-vertical observation remain untouched by the proposed solutions. As pixels in shadow areas have a very similar spectral reflectance to pixels of water surfaces, the spectral responding curve for shadow pixels lies between the spectral responding curves for large water surfaces such as lakes and small water surfaces such as canals as – observed from the IKONOS image of the Amsterdam test site. The shadow problem becomes a real issue when we use high-resolution data (1 m to 4 m) for urban areas, although it can be ignored when using coarse-resolution data or dealing with non-urban sites. This problem is difficult to solve using spectral data alone and per-pixel approaches. This problem has

been largely solved by using spectral data and laser data in our per-object approach.

Although the proposed modifications improve the land-cover classification accuracy of the MLC, we consider the attainable results insufficient for a detailed urban land-use classification because of the complexity of urban environments. Based on our knowledge and experiences obtained from visual image-interpretation, the key features for image-interpretation, such as size, shape, colour, orientation, pattern, association, are directly associated with explicit objects, which are at higher abstraction levels than pixels. These key features should continue to play a key role in image analysis and land-use classification. For instance, we need to check how buildings are spatially distributed in space in order to find out if they belong to the same residential area. We need to know the number of floors of a building to achieve better understanding and classification. We need to know if buildings are similar in size, height or orientation, etc. We need to explore the surrounding features of specific objects. Such information cannot be acquired by per-pixel approaches because they are directly associated with objects, not pixels. Therefore, object-based image processing techniques are considered better for image analysis at higher levels than the pixel level, which provides additional tools and methods for image analysis at higher abstraction levels. Thus an object-based image analysis approach has become the main focus of this research.

2. The combination of high spatial resolution airborne LIDAR data and multi-spectral imagery such as IKONOS, QuickBird and SPOT 5 offers great application opportunities, especially in urban areas. With the remarkable capacities provided by the second generation of airborne scanners that enable both height and spectral information to be acquired simultaneously by laser scanner and multi-spectral scanner, many meaningful features can be derived by combining these two types of data rather than using one type of data alone. With respect to extracting image objects and per-object properties, and identifying explicit topological relations between image objects, several concepts, methods as well as other fundamental issues concerning the object-based approach have been proposed and discussed, such as object hierarchy, image object, the hybrid-raster data model, methods of identifying topological relationships between image objects, and methods of extracting various per-object properties. The hybrid-raster data model has been applied to enable topological relationships between image objects to be explicitly identified. These concepts and methods have been tested and successfully applied in urban land-cover and land-use classification for two test sites. We consider the test results to be promising and to confirm the effectiveness of the proposed per-object approach.

3. A hierarchical object-based approach for urban land-use classification based on high-resolution remote sensing data has been proposed in this research. Three types of hierarchical objects have been outlined: fundamental objects

at pixel level, land-cover objects at the land-cover level and land-use objects at the land-use level. The proposed method consists of three steps: land-cover classification, land-use unit reasoning and delineation, and land-use classification. It incorporates per-pixel-based image processing techniques and per-object-based techniques at different stages. Various techniques have been proposed and tested for object extraction at different aggregation levels. The experimental results produced for various stages, from pixels to land-cover objects and from land-cover objects to land-use objects, show the success of applying hierarchical image objects and structural image analysis techniques in urban land-cover and land-use classification. The experiments show that hierarchically formed image objects are useful tools for image analysis and spatial modelling and are more successful than pixel-based approaches based on aggregation hierarchies. A spatial modelling approach has been proposed for spatial modelling at the sub-pixel level. The test results have confirmed that the proposed sub-pixel approach can be applied to improve classification accuracy, obtain finer class boundaries, and achieve better spatial interpolation.

4. Structural information derived from hierarchical image objects plays an important role in land-use classification of urban areas. Delaunay triangulation has been successfully applied to spatially disjoint objects such as buildings in order to obtain the spatial adjacency relationship and a proximity measurement – useful information for spatial clustering to find spatial land-use units. Several similarity measures have been proposed and tested for similarity measurement based on object properties such as building size and building height. Similarity measurements have proved to be useful information for spatial clustering, in addition to the spatial adjacency relationship and the proximity measurement.

5. Several object properties have been proposed and extracted as object attributes of land-use objects for our two test sites. Fuzzy membership functions have been designed to establish the relationships between extracted land-use object properties and designated land-use classes. A fuzzy classifier has been applied for per-object classification based on extracted land-use units and their object properties. The obtained results show that the proposed per-object land-use classification approach is promising. The extracted land-use object properties are also useful information for urban studies, planning and management.

6. A united framework for quality assessment has been proposed and tested, based on similarity measures between classified data and reference data. This framework utilises per-object and per-pixel measurements, and is also suitable for the quality assessment of single-class cases. The proposed per-object quality measures provide the possibility of obtaining additional quality assessment based on various object properties. The proposed uncertainty measures for extracted land-cover objects and classified land-use objects have been tested, and

proved to be useful information that enables users to concentrate their attention on those uncertain objects.

7. The developed concepts and methods have been implemented by programming in Matlab^{TM}. The implemented system allows different users to express their wishes by specifying characteristics that can be extracted from laser data and spectral data, and related parameters, in order to obtain the desired results. This feature offers planners and other users the opportunity to produce results according to their specific wishes and applications from a detailed data set, and to share the relatively high costs of acquiring high-resolution laser data and spectral data.

10.2   Future research

1. The current image-object concept is defined for representing image regions or area objects in image space. Further research is needed regarding the representation of other types of objects, such as linear objects, in image space. Topological relationships between image objects are defined and extracted based on the hybrid-raster data model for crisp image regions. Additional effort is needed as regarding topological relationships between fuzzy image regions, and linear and point objects.

2. As regards spatial clustering and reasoning for finding spatial land-uses units, the shortest distances between adjacent land-cover objects (buildings), as well as similarity measures in terms of building size and building height, are used as measures in this research. Additional measures may be needed, such as similarity in terms of object shape and orientation, that are considered useful information for obtaining better clustering results. Since similarity measures for shape and orientation are more complicated to formulate and apply, and have not been included in the current research, additional investigation is needed in this respect. It may be worthwhile to derive a more comprehensive measure based on different object properties; additional investigation is required here too.

3. In this research, we have limited ourselves to using laser data and spectral data only, and to extracting as much as possible from these data. It is expected that some more useful products may be produced when additional data such as demographic data, economic data and other GIS data are available. We have demonstrated the successful use of a digitised road map for obtaining refined spatial land-use units. Additional investigation is needed on integrating high-resolution data with other data. Road extraction itself is another interesting research area that is not covered by the current research. We believe our object-based approach can be extended to the extraction of linear features such as roads.

4. Urban areas are complicated. Additional efforts are needed and different settings have to be investigated in order to find out whether additional features are needed and how to specify parameters for situations that may appear in different types of urban areas and different types of cities.

5. Buildings, green spaces, water surfaces and sealed-ground surfaces have been successfully extracted at the land-cover level. Land-use spatial units have been obtained by spatial reasoning based on the extracted land-cover objects. Per-object land-use classification has been made with high accuracy. These achievements have been accomplished using the developed semi-automatic approaches based on high-resolution data. Further research is needed toward fully automatic image interpretation.

References

American Planning Association, 2001. Land-based classification standards (available from http://www.planning.org/lbcs).

Anderson, J. R., Hardy, E. E., Roach, J. T., Witmer, R. E., 1976. A land use and land cover classification system for use with remote sensor data. Tech. rep., U.S. Geological Survey, Professional Paper 964.

Aplin, P., Atkinson, P. M., 2001. Sub-pixel land cover mapping for per-field classification. International Journal of Remote Sensing 22 (14), 2853–2858.

Aplin, P., Atkinson, P. M., Curran, P. J., 1999a. Fine spatial resolution simulated sensor imagery for land cover mapping in the United Kingdom. Remote Sensing of Environment 68, 206–216.

Aplin, P., Atkinson, P. M., Curran, P. J., 1999b. Per-field classification of land use using the forthcoming very fine spatial resolution satellite sensors: problems and potential solutions. In: Atkinson, P. M., Tate, N. J. (Eds.), Advances in Remote Sensing and GIS Analysis. John Wiley and Sons, Chichester etc., pp. 219–239.

Atkinson, P. M., 1997. Mapping sub-pixel boundaries from remotely sensed images. In: Kemp, Z. (Ed.), Innovations in GIS 4. Taylor and Francis, London, pp. 166–180.

Atkinson, P. M., Cutler, M. E. J., Lewis, H. G., 1997. Mapping sub-pixel proportional land cover with AVHRR imagery. International Journal of Remote Sensing 18 (4), 917–935.

Axelsson, P., 1999. Processing of laser scanner data - algorithms and applications. ISPRS Journal of Photogrammetry and Remote Sensing 54 (2-3), 138–147.

Ballard, D. B., Brown, C. M., 1982. Computer Vision. Prentice-Hall, Englewood, Cliffs, NJ.

Barber, C. B., Dobkin, D. P., Huhdanpaa, H. T., 1996. The quickhull algorithm for convex hulls. ACM Transactions on Mathematical Software 22 (4), 469–483.

Barnsley, M. J., Barr, S. L., 1997. Distinguishing urban land-use categories in fine spatial resolution land-cover data using a graph-based, structural pattern recognition system. Computer, Environment and Urban Systems 21 (3/4), 209–225.

Barr, S. L., Barnsley, M. J., 1997. A region-based, graph-theoretic data model for the inference of second-order thematic information from remote-sensed images. International Journal of Geographical Information Science 11 (6), 555–576.

Bastin, L., 1997. Comparison of fuzzy c-means classification, linear mixture modelling and MLC probabilities for unmixing coarse pixels. International Journal of Remote Sensing 18 (17), 3629–3648.

Beaulieu, J.-M., Goldberg, M., 1989. Hierarchy in picture segmentation: a stepwise optimization approach. IEEE Transactions on Pattern Analysis and Machine Intelligence 11 (2), 150–163.

Bian, L., Butler, R., 1999. Comparing effects of aggregation methods on statistical and spatial properties of simulated spatial data. Photogrammetric Engineering and Remote Sensing 65 (1), 73–84.

Bishr, Y., 1997. Semantic Aspects of Interoperable GIS. PhD Dissertation, Wageningen Agricultural University and ITC.

Booch, G., 1993. Object - Oriented Analysis and Design: with Applications, 2nd Edition. Benjamin Cummings, Redwood City etc.

Borges, K. A. V., Davis Jr., C. A., Laender, A. H. F., 2001. OMT-G: an object-oriented data model for geographic applications. GeoInformatica 5 (3), 221–260.

Bradley, P. S., Fayyard, U. M., Reina, C. A., 1998. Scaling EM (Expectation Maximization) clustering to large databases. Microsoft Research Report 98-35, Microsoft.

Brunn, A., Weidner, U., 1997. Extracting buildings from digital surface models. International Archives of Photogrammetry and Remote Sensing 32 (Part 3-4W2), 27–34.

Burrough, P. A., McDonnell, R. A., 1998. Principles of Geographical Information Systems. Oxford University Press, Oxford.

Campbell, B. J., 1981. Spatial correlation effects upon accuracy of supervised classification of land cover. Photogrammetric Engineering and Remote Sensing 47 (3), 355–363.

Campbell, B. J., 1996. Introduction to Remote Sensing, 2nd Edition. Guilford Press, New York.

Campbell, B. J., 2002. Introduction to Remote Sensing, 3rd Edition. Guilford Press, New York.

Chen, B., Song, X., Lin, C., 1989. The study of microcomputer-based urban planning and management information system. In: International Conference on Computers in Urban Planning and Urban Management. Hong Kong: Centre of Urban Studies and Urban Planning, University of Hong Kong, Hong Kong, pp. 119–121.

Coad, P., Yourdon, E., 1990. Object-Oriented Analysis, 2nd Edition. Prentice-Hall, Englewood Cliffs.

Congalton, R. G., Mead, R. A., 1983. A quantitative method to test for consistency and correctness in photo-interpretation. Photogrammetric Engineering and Remote Sensing 49, 69–74.

Cook, R. G., 2001. Hierarchical stimulus processing by pigeons. In: R. G. Cook, (Ed.), Avian Visual Cognition [On-line]. Available: www.pigeon.psy.tufts.edu/avc/cook/.

Couclelis, H., 1992. People maniputate objects (but cultivate fields): beyond the raster-vector debate in GIS. In: Frank, A. U., Campari, I. (Eds.), Theories and Methods of Spatio-Temporal Reasoning in Geographic Space. Lecture Note in Computer Science 639. Springer-Verlag, Berlin etc., pp. 65–77.

Cova, T. J., Goodchild, M., 2002. Extending geographical representation to include fields of spatial objects. International Journal of Geographical Information Science 16 (6), 509–532.

Curran, P. J., 1985. Principles of Remote Sensing. Longman, New York.

Cushine, J. L., 1987. The interactive effects of spatial resolution and degree of internal variability within land-cover types on classification accuracies. International Journal of Remote Sensing 8 (1), 15–29.

de Berg, M., van Kreveld, M., Overmars, M., Schwarzkopf, O., 2000. Computational Geometry: Algorithms and Applications, 2nd Edition. Springer-Verlag, Berlin etc.

Dempster, A. P., Laird, N. M., Rubin, D. B., 1977. Maximum likelihood from incomplete data via the EM algorithm. Journal of the Royal Statistical Society, Series B 39, 1–38.

Dungan, J. L., 2002. Toward a comprehensive view of uncertainty in remote sensing analysis. In: Foody, G. M., Atkinson, P. M. (Eds.), Uncertainty in Remote Sensing and GIS. John Wiley and Sons, Chichester etc., pp. 25–35.

Egenhofer, M. J., 1989. A formal definition of binary topological relationships. In: Litwin, W., Schek, H. J. (Eds.), The Third International Conference on Foundations of Data Organization and Algorithms (DODO), Paris, France. Vol. 367 of Lecture Notes in Computer Science. Springer-Verlag, New York etc., pp. 457–472.

Egenhofer, M. J., 1993. A model for detailed binary topological relationships. Geomatica 47 (3 and 4), 261–273.

Egenhofer, M. J., Franzosa, R. D., 1991. Point-set topological spatial relations. International Journal of Geographical Information Systems 5 (2), 161–174.

Egenhofer, M. J., Herring, J. R., 1991. Categorizing binary topological relations between regions, lines, and points in geographic databases. Tech. rep., Department of Surveying Engineering, University of Maine.

Egenhofer, M. J., Sharma, J., 1993. Topological relations between regions in $\mathbb{R}^2$ and $\mathbb{Z}^2$. In: Abel, D., Ooi, B. C. (Eds.), Advances in Spatial Databases - Third International Symposium on Large Spatial Databases, SSD '93, Singapore. Vol. 692 of Lecture Notes in Computer Science. Springer-Verlag, Berlin etc., pp. 316–336.

Ehlers, M., Schiewe, J., Tufte, L., 2002. Urban remote sensing: new development and challenges. In: Maktav, D., Jürgens, C., Erbek, F. S., Akgün, H. (Eds.), The 3rd International Symposium on Remote Sensing of Urban Areas. Vol. 1. Istanbul Technical University, Istanbul, Turkey, pp. 130–137.

Estivill-Castro, V., Lee, I., 2002. Multi-level clustering and its visualization for exploratory spatial analysis. GeoInformatica 6 (2), 123–152.

Fisher, P., 1997. The pixel: a snare and a delusion. International Journal of Remote Sensing 18 (3), 679–685.

Foody, G. M., 1996. Approaches for the production and evaluation of fuzzy land cover classification from remotely-sensed data. International Journal of Remote Sensing 17 (17), 1317–1340.

Foody, G. M., 1998. Sharpening fuzzy classification output to refine the representation of sub-pixel land cover distribution. International Journal of Remote Sensing 19 (13), 2593–2599.

Foody, G. M., 2000. Accuracy of thematic maps derived from remote sensing. In: Heuvelink, G. B. M., Lemmens, M. J. P. M. (Eds.), Proceedings of 4th International Symposium on Spatial Accuracy Assessment in Natural Resources and Environmental Sciences. Delft University Press, Amsterdam, pp. 217–224.

Foody, G. M., Campbell, N. A., Trodd, N. M., Wood, T. F., 1992. Derivation and applications of probabilistic measures of class membership from the maximum-likelihood classification. Photogrammetric Engineering and Remote Sensing 58 (9), 1335–1341.

Foody, G. M., Lucas, R. M., Curran, P. J., Honzak, M., 1997. Non-linear mixture modelling without end-members using an artificial neural network. International Journal of Remote Sensing 18 (4), 937–953.

Fowler, M., Scott, K., 1999. UML Distilled: A Brief Guide to the Standard Object Modeling Language, 2nd Edition. Addison-Wesley, Boston etc.

Geo-Loket, 2002. AHN: Actual height model of the Netherlands, http://www.minvenw.nl/rws/mdi/geoloket/overig.html.

Gold, C. M., 1991. Problems with handling spatial data - the Voronoi approach. CISM Journal 45, 65–80.

Gold, C. M., 1992. The meaning of 'neighbour'. In: Frank, A. U., Campari, I., Formentini, U. (Eds.), Theories and Methods of Spatio-Temporal Reasoning in Geographic Space. Vol. 39 of Lecture Notes in Computer Science. Springer-Verlag, Berlin etc., pp. 220–235.

Gold, C. M., Edwards, G., 1992. The Voronoi spatial model: two- and three-dimentional applications in image analysis. ITC Journal 1992 (1), 11–19.

Gong, P., Howarth, P. J., 1990. An assessment of some factors influencing multispectral land-cover classification. Photogrammetric Engineering and Remote Sensing 56 (5), 597–603.

Goodchild, M. F., 1992. Geographical data modeling. Computer and Geosciences 18 (4), 401–408.

Goodchild, M. F., 1997. Representing Fields, NCGIA Core Curriculum in GIScience, http://www.ncgia.ucsb.edu/giscc/units/u054/u054.html.

Haala, N., Brenner, C., 1999. Extraction of buildings and trees in urban environments. ISPRS Journal of Photogrammetry and Remote Sensing 54 (2-3), 130–137.

Haala, N., Walter, V., 1999. Automatic classification of urban environment for database revision using LIDAR and color aerial imagery. International Archives of Photogrammetry and Remote Sensing 32 (Part 7-4-3 W6), 76–82.

Hall, P., 2002. Urban and Regional Planning, 4th Edition. Routledge, London etc.

Han, J., Kamber, M., Tung, A. K. H., 2001. Spatial clustering methods in data mining. In: Miller, H. J., Han, J. (Eds.), Geographic Data Mining and Knowledge Discovery. Research Monographs in GIS Series. Taylor and Francis, London etc., pp. 188–217.

Haralick, R. M., Shapiro, L. G., 1985. Survey - image segmentation techniques. Computer Vision, Graphics, and Image Processing 29, 100–132.

Harrison, A., Garland, B., 2001. The national land use database: building new national baseline data of urban and rural land use. In: AGI Conference 2001.

Hartshorn, T. A., 1992. Interpreting the City: An Urban Geography, 2nd Edition. John Wiley and Sons, Chichester etc.

Heipke, C., Mayer, H., Wiedemann, C., Jamet, O., 1997. Evaluation of automatic road extraction. International Archives of Photogrammetry and Remote Sensing 32 (3-2W3), 47–56.

Hoeppner, F., Klawonn, F., Kruse, R., Runkler, T., 1999. Fuzzy Cluster Analysis: Methods for Classification, Data Analysis and Image Recognition. John Wiley and Sons, Chichester etc.

Hug, C., Wehr, A., 1997. Detecting and identifying topographic objects in imaging laser altimeter data. The International Archives of Photogrammetry and Remote Sensing 32 (Part 3-4W2), 19–26.

Janssen, L. L. F., 1994. Methodology for Updating Terrain Object Data from Remote Sensing Data: The Application of Landsat TM Data with Respect to Agricultural Fields. PhD Dissertation. Wageningen Agricultural University and ITC, Enschede, The Netherlands.

Kainz, W., Egenhofer, M. J., Greasley, I., 1993. Modelling spatial relations and operations with partially ordered sets. International Journal of Geographical Information Systems 7 (3), 215–229.

Kaufman, L., Rousseeuw, P. J., 1990. Finding Groups in Data: An Introduction to Cluster Analysis. John Wiley and Sons, New York etc.

Kovalevski, V. A., 1989. Finite topology as applied to image analysis. Computer Vision, Graphics, and Image Processing 46, 141–161.

Le Clercq, F., 1990. Information management within the planning process. In: Scholten, H. J., Stillwell, J. C. H. (Eds.), Geographical Information Systems for Urban and Regional Planning. Kluwer Academic, Dordrecht etc., pp. 59–68.

Lemmens, M. J. P. M., Deijkers, H., Looman, P. A. M., 1997. Building detecting fusing airborne laser-altimeter DEMs and 2D digital maps. The International Archives of Photogrammetry and Remote Sensing 32 (Part 3-4W2), 42–49.

Li, C., Chen, J., Li, Z., 1999. A raster-based method for computing Voronoi diagrams of spatial objects using dynamic distance transformation. International Journal of Geographical Information Science 13 (3), 209–225.

Li, Z., Huang, P., 2002. Quantitative measures for spatial information of maps. International Journal of Geographical Information Science 16 (7), 699–709.

Liu, Y., Molenaar, M., Kraak, M. J., 2002. Semantic similarity evaluation model in categorical database generalization. International Archives of Photogrammetry and Remote Sensing 34, Part 4, 279–285.

Lobovitz, M. L., Masuoka, E. J., 1984. The influence of autocorrelation in signature extraction - an example from a geobotanical investigation of Cotter Basin, Montana. International Journal of Remote Sensing 5 (2), 315–332.

MacQueen, J., 1967. Some methods for classification and analysis of multivariate observation. In: Proceedings of the Fifth Berkeley Symposium on Maths and Statistics Problems. pp. 281–297.

Marble, D. F., Amundason, S. E., 1988. Microcomputer based geographical information systems and their role in urban and regional planning. Environment and Planning B: Planning and Design 15, 305–324.

Marques de Sa, J. P., 2001. Pattern Recognition: Concepts, Methods and Applications. Springer-Verlag, Berlin etc.

Mather, P. M., 1999. Computer Processing of Remotely-Sensed Images: An Introduction, 2nd Edition. John Wiley and Sons, Chichester etc.

McFeeters, S. K., 1996. The use of the Normalized Difference Water Index (NDWI) in the delineation of open water feature. International Journal of Remote Sensing 17 (7), 1425–1432.

McLoughlin, J. B., 1969. Urban and Regional Planning: A System Approach. Faber, London.

Mertikas, P., Zervakis, M. E., 2001. Exemplifying the theory of evidence in remote sensing image classification. International Journal of Remote Sensing 22 (6), 1081–1095.

Molenaar, M., 1998. An Introduction to the Theory of Spatial Object Modelling for GIS. Taylor and Francis, London etc.

Morgan, M. F., 1999. Building extraction from laser scanning data. Msc thesis, ITC.

Okabe, A., Boots, B., Sugihara, K., 1994. Nearest neighbourhood operations with generalized Voronoi diagrams: a review. International Journal of Geographical Information Systems 8 (1), 43–71.

Okabe, A., Boots, B., Sugihara, K., Chiu, S. N., 2000. Spatial Tessellations: Concepts and Applications of Voronoi Diagrams, 2nd Edition. Wiley Series in Probability and Statistics. John Wiley and Sons, Chichester etc.

OMG, 2001. Unified Modeling Language (UML), Specification, version 1.4 [Online]. Available: http://www.omg.org/technology/documents/formal/uml.htm. last update: September 2001.

Osborne, H. R., Bridge, D. G., 1996. A case base similarity framework. In: Smith, I., Faltings, B. (Eds.), Advances in Case-Based Reasoning. Vol. 1168 of Lecture Notes in Artificial Intelligence. Springer-Verlag, Berlin etc., pp. 309–323.

Osborne, H. R., Bridge, D. G., 1997a. Models of similarity for case-based reasoning. In: Proceedings of the Interdisciplinary Workshop on Similarity and Categorisation. pp. 173–179.

Osborne, H. R., Bridge, D. G., 1997b. Similarity metrics: a formal unification of cardinal and non-cardinal similarity measures. In: Leake, D. B., Plaza, E. (Eds.), Case-Based Reasoning Research and Development, Proceedings of the 2nd International Conference on Case-based Reasoning (ICCBR-97). pp. 235–244.

Parker, J. R., 1997. Algorithms for Image Processing and Computer Vision. John Wiley and Sons, New York etc.

Pilouk, M., Tempfli, K., 1992. A digital image processing approach to creating DTMs from digitized contours. The International Archives of Photogrammetry and Remote Sensing 29, Part 4(B), 956–961.

Richards, J. A., Jia, X., 1999. Remote Sensing Digital Image Analysis: An Introduction, 3rd Edition. Springer-Verlag, Berlin etc.

Rodríguez, M. A., Egenhofer, M. J., 2003. Determining semantic similarity among entity classes from different ontologies. IEEE Transactions on Knowledge and Data Engineering 15 (2), 442– 456.

Santini, S., Jain, R., 1995. Similarity matching. In: ACCV. pp. 571–580.

Santini, S., Jain, R., 1999. Similarity measures. IEEE Transactions on Pattern Analysis and Machine Intelligence 21 (9), 871–883.

Schowengerdt, R. A., 1997. Remote Sensing: Model and Methods for Image Processing, 2nd Edition. Academic Press, San Diego.

Shufelt, J., 2000. Geometric Constraints for Object Detection and Delineation. Kluwer Academic, Dordrecht etc.

Skidmore, A. K., 1999. Accuracy assessment of spatial information. In: Stein, A., van der Meer, F., Gorte, B. (Eds.), Spatial Statistics for Remote Sensing. Kluwer, Dordrecht etc., pp. 197–209.

Sonka, M., Hlavac, V., Boyle, R., 1999. Image Processing, Analysis and Machine Vision, 2nd Edition. PWS Publ., an imprint of Brooks/Cole Publ, Pacific Grove, CA [etc.].

Steinwendner, J., 1999. From satellite images to scene description using advanced image processing techniques. In: RSS'99. Remote Sensing Society, Nottingham, pp. 865–872.

Tatem, A. J., Lewis, H. G., Atkinson, P. M., Nixon, M. S., 2001a. Multi-class landcover mapping at the sub-pixel scale using a Hopfield neural network. International Journal of Applied Earth Observation and Geoinformation 3 (2), 184–190.

Tatem, A. J., Lewis, H. G., Atkinson, P. M., Nixon, M. S., 2001b. Super-resolution target identification from remotely sensed images using a Hopfield neural network. IEEE Transactions on Geoscience and Remote Sensing 39, 781–796.

The MathWorks Inc., 2001. Image Processing Toolbox User's Guide. The MathWorks, Inc., Natick, MA, USA.

The Ministry of Construction P. R. China, 2001. National Standard for Urban Land Use Classification, China (GBJ 137-90). Planning Publishing House, Beijing.

TopoSys, 2002. www.toposys.de. Last update: 19 Febuary 2002.

Tso, B., Mather, P. M., 2001. Classification Methods for Remotely Sensed Data. Taylor and Francis, London etc.

Tversky, A., 1977. Features of similarity. Psychological Review 84 (4), 327–352.

UML Bib, 22 October 2002. The UML Bibliography. http://www.db.informatik.uni-bremen.de/umlbib/.

van der Heijden, F., 1994. Image Based Measurement Systems: Object Recognition and Parameter Estimation. John Wiley and Sons, Chichester etc.

Verhoeye, J., De Wulf, R., 2000. Sub-pixel mapping of Sahelian wetlands using multitemporal SPOT VEGETATION images. In: 28th International Symposium on Remote Sensing of Environment, Information for Sustainable Development. CSIR Satellite Applications Centre, Cape Town, pp. category 4, 14–19.

Vincent, L., Soille, P., 1991. Watersheds in digital spaces: an efficient algorithm based on immersion simulations. IEEE Transactions on Pattern Analysis and Machine Intelligence 13 (6), 583–598.

Wiedemann, C., Heipke, C., Mayer, H., Jamet, O., 1998. Empirical evaluation of automatically extracted road axes. In: Bowyer, K., Phillips, P. (Eds.), Empirical Evaluation Techniques in Computer Vision. IEEE Computer Society, Los Alamitos, pp. 172–187.

Winter, S., 1995. Topological relations between discrete regions. In: Egenhofer, M. J., Herring, J. R. (Eds.), Advances in Spatial Databases. Vol. 951 of Lecture Notes in Computer Science. Springer-Verlag, Berlin etc., pp. 310–327.

Winter, S., Frank, A. U., 1999. Functional extensions of a raster representation for topological relations. In: Vckovski, A., Brassel, K. E., Schek, H. J. (Eds.), Interoperating Geographic Information Systems. Vol. 1580 of Lecture Notes in Computer Science. Springer-Verlag, Berlin etc., pp. 293–304.

Winter, S., Frank, A. U., 2000. Topology in raster and vector representation. GeoInformatica 4 (1), 35–65.

Worboys, M. F., 1995. GIS: A Computing Perspective. Taylor and Francis, London.

Worboys, M. F., Hearnshaw, H. M., Maguire, D. J., 1990. Object-oriented data modelling for spatial dadabases. International Journal of Geographical Information Systems 4 (4), 369–383.

Yeh, A. G.-O., 1988. Microcomputer in urban planning: application, constraints, and impacts. Environment and Planning B: Planning and Design 15, 241–254.

Yeh, A. G.-O., 1991. The development and applications of geographical information systems for urban and regional planning in developing countries. International Journal of Geographical Information Systems 5 (1), 5–27.

Yu, D., Chatterjee, S., Sheikholeslami, G., Zhang, A., 1998. Efficiently detecting arbitrary shaped clusters in very large datasets with high dimentions. Computer Science Technical Report 98-08. SUNY Buffalo.

Zadeh, L. A., 1965. Fuzzy set. Information and Control 8, 338–353.

Zhan, Q., Molenaar, M., Gorte, B., 2000. Urban land use classes with fuzzy membership and classification based on integration of remote sensing and GIS. International Archives of Photogrammetry and Remote Sensing 33 (Part B7), 1751–1758.

Zhan, Q., Molenaar, M., Lucieer, A., 2002a. Pixel unmixing at the sub-pixel scale based on land cover classes probabilities: application to urban areas. In: Foody, G. M., Atkinson, P. M. (Eds.), Uncertainty in Remote Sensing and GIS. John Wiley and Sons, Chichester etc., pp. 59–76.

Zhan, Q., Molenaar, M., Tempfli, K., 2002b. Building extraction from laser data by reasoning on image segments in elevation slides. The International Archives of Photogrammetry and Remote Sensing 34, Part 3(B), 305–308.

Zhan, Q., Molenaar, M., Tempfli, K., 2002c. Finding spatial units for land use classification based on hierarchical image objects. The International Archives of Photogrammetry and Remote Sensing 34, Part 4, 263–268.

Zhan, Q., Molenaar, M., Tempfli, K., 2002d. Hierarchical image object-based structural analysis toward urban land use classification using high-resolution imagery and airborne LIDAR data. In: Maktav, D., Jürgens, C., Erbek, F. S., Akgün, H. (Eds.), The 3rd International Symposium on Remote Sensing of Urban Areas. Vol. 1. Istanbul, Turkey, pp. 251–258.

Zhan, Q., Molenaar, M., Tempfli, K., 2003. Uncertainty assessment for geo-spatial objects derived from high-resolution airborne imagery and laser data. In: Shi, W., Goodchild, M. F., Fisher, P. F. (Eds.), Proceedings of the 2nd International Symposium on Spatial Data Quality '03. The Hong Kong Polytechnic University, Hong Kong, pp. 79–88.

Zhan, Q., Molenaar, M., Tempfli, K., under peer review (1). Image-object based approach for building extraction from airborne LIDAR and multi-spectral data. Submitted to ISPRS Journal of Photogrammetry and Remote Sensing .

Zhan, Q., Molenaar, M., Tempfli, K., Shi, W., under peer review (2). Quality assessment for geo-spatial objects derived from high-resolution airborne imagery and laser data. Submitted to International Journal of Remote Sensing .

Zhan, Q., Molenaar, M., Xiao, Y., 2001. Hierarchical object-based image analysis of high-resolution imagery for urban land use classification. In: IEEE - ISPRS Joint Workshop on Remote Sensing and Data Fusion over Urban Areas. IEEE, Rome, Italy, pp. 35–39.

Zobel, J., Moffat, A., 1998. Exploring the similarity space. SIGIR Forum 32 (1), 18–34.

Author's Bibliography

1. Zhan, Q., Molenaar, M., and Gorte, B., 2000. Urban land use classes with fuzzy membership and classification based on integration of remote sensing and GIS. In: *The International Archives of Photogrammetry and Remote Sensing*, Vol. 33, Part B7, Amsterdam, The Netherlands, pp. 1751-1758.

2. Zhan, Q., Molenaar, M., and Xiao, Y., 2001. Hierarchical object-based image analysis of high-resolution imagery for urban land use classification. In: *IEEE - ISPRS Joint Workshop on Remote Sensing and Data Fusion over Urban Areas*, Rome, Italy, pp. 35-39.

3. Zhan, Q., Molenaar, M., and Lucieer, 2002. Pixel unmixing at the sub-pixel scale based on land cover class probabilities: application to urban areas. In: G. Foody and P. Atkinson (eds), *Uncertainty in Remote Sensing and GIS*, pp. 59-76. Chichester etc., John Wiley and Sons.

4. Zhan, Q., Molenaar, M., and Tempfli, K., 2002. Finding spatial units for land use classification based on hierarchical image objects. In: *The International Archives of Photogrammetry and Remote Sensing*, Vol. 34, Part 4, Ottawa, Canada, pp. 263-268.

5. Zhan, Q., Molenaar, M., and Tempfli, K., 2002. Hierarchical image-object based structural analysis toward urban land use classification using high-resolution imagery and airborne LIDAR data. In: *Proceedings of the 3rd International Symposium on Remote Sensing of Urban Areas*, Istanbul, Turkey, pp. 251-158.

6. Zhan, Q., Molenaar, M., and Tempfli, K., 2002. Building extraction from laser data by reasoning on image segments in elevation slices. In: *The International Archives of Photogrammetry and Remote Sensing*, Vol. 34, Part 3 (B), Graz, Austria, pp. 305-308.

7. Zhan, Q., Molenaar, M., and Tempfli, K., 2003. Uncertainty assessment for geo-spatial objects derived from high-resolution airborne imagery and laser data. In: *Proceedings of the 2nd International Symposium on Spatial Data Quality '03*, Hong Kong, China, pp. 79-88.

8. Zhan, Q., Molenaar, M., and Tempfli K., Under peer review. Image-object based approach for building extraction from airborne LIDAR and multi-spectral data, Submitted to *ISPRS Journal of Photogrammetry and Remote Sensing*.

9. Zhan, Q., Molenaar, M., Tempfli, K., and Shi, W., Under peer review. Quality assessment for geo-spatial objects derived from high-resolution airborne imagery and laser data, Submitted to *International Journal of Remote Sensing*.

Appendix A: Land-Use Classification Systems

Table A.2: General land use classification system (Anderson et al., 1976)

1 Urban or Built-up Land	**5 Water**
11 Residential	51 Streams and Canals
12 Commercial and Services	52 Lakes
13 Industrial	53 Reservoirs
14 Transportation, Communications, and Utilities	54 Bays and Estuaries
15 Industrial and Commercial Complexes	**6 Wetland**
16 Mixed Urban or Built-up Land	61 Forested Wetland
17 Other Urban or Built-up Land	62 Nonforested Wetland
	7 Barren Land
2 Agricultural Land	71 Dry Salt Flats
21 Cropland and Pasture	72 Beaches
22 Orchards, Groves, Vineyards, Nurseries, and Ornamental Horticultural Areas	73 Sandy Areas other than Beaches
	74 Bare Exposed Rock
23 Confined Feeding Operations	75 Strip Mines, Quarries, and Gravel Pits
24 Other Agricultural Land	76 Transitional Areas
	77 Mixed Barren Land
3 Rangeland	
31 Herbaceous Rangeland	**8 Tundra**
32 Shrub and Brush Rangeland	81 Shrub and Brush Tundra
33 Mixed Rangeland	82 Herbaceous Tundra
	83 Bare Ground Tundra
	84 Wet Tundra
4 Forest Land	85 Mixed Tundra
41 Deciduous Forest Land	
42 Evergreen Forest Land	**9 Perennial Snow or Ice**
43 Mixed Forest Land	91 Perennial Snowfields
	92 Glaciers

Table A.3: Land-based classification standards: function dimension, US

1000	**Residential or accommodation functions**		**6000**	**Education, public admin., health care, and other inst.**
1100	Private household		6100	Educational services
1200	Housing services for elderly		6200	Public administration
1300	Hotels, motels, or accommodation services		6300	Other government functions
			6400	Public safety
2000	**General sales or services**		6500	Health and human services
2100	Retail sales or service		6600	Religious institutions
2200	Finance and insurance		6700	Death care services
2300	Real estate, and rental and leasing		6800	Associations, nonprofit organizations, etc.
2400	Business, professional, scientific, and technical services		**7000**	**Construction-related businesses**
2500	Food services		7100	Building, developing, and general contracting
2600	Personal services			
2700	Pet and animal sales or service (except veterinary)		7200	Machinery-related
			7300	Special trade contractor
3000	**Manufacturing and wholesale trade**		7400	Heavy construction
			8000	**Mining and extraction establishments**
3100	Food, textile, and related products			
3200	Wood, paper, and printing products		8100	Oil and natural gas
			8200	Metals (iron, copper, etc.)
3300	Chemicals, and metals, machinery, and electronics manufacturing		8300	Coal
			8400	Nonmetallic mining
3400	Miscellaneous manufacturing		8500	Quarrying and stone cutting establishment
3500	Wholesale trade establishment			
3600	Warehouse and storage services		**9000**	**Agriculture, forestry, fishing and hunting**
4000	**Transportation, communication, information, and utilities**			
			9100	Crop production
4100	Transportation services		9200	Support functions for agriculture
4200	Communications and information		9300	Animal production including slaughter
4300	Utilities and utility services			
5000	**Arts, entertainment, and recreation**		9400	Forestry and logging
			9500	Fishing, hunting and trapping, game preserves
5100	Performing arts or supporting establishment		9900	Unclassifiable function
5200	Museums and other special purpose recreational institutions			
5300	Amusement, sports, or recreation establishment			
5400	Camps, camping, and related establishments			
5500	Natural and other recreational parks			

Source: American Planning Association, http://www.planning.org/lbcs

Table A.4: NLUD land-use classification v3.3, UK (Harrison and Garland, 2001)

1	**Agricultural**	**7**	**Recreation**
1.1	Field crops	7.1	Leisure and recreational buildings
1.2	Fallow land	7.2	Outdoor recreation
1.3	Horticulture and orchards	7.3	Allotments
1.4	Improved pasture		
1.5	Field margins	**8**	**Transport**
		8.1	Roads
2	**Woodland**	8.2	Public car parks
2.1	Conifer woodland	8.3	Railways
2.2	Mixed woodland	8.4	Airports
2.3	Broadleaved woodland	8.5	Docks
2.4	Undifferentiated young woodland		
2.5	Scrub	**9**	**Residential**
2.6	Felled woodland	9.1	Residential
2.7	Land cultivated for afforestation	9.2	Institutional and communal accommodation
3	**Unimproved Grassland and Heathland**	**10**	**Community Buildings**
3.1	Unimproved grassland	10.1	Institutional buildings
3.2	Heathland	10.2	Educational buildings
3.3	Bracken	10.3	Religious buildings
3.4	Upland mosaics		
		11	**Industrial and Commercial**
4	**Water and Wetland**	11.1	Industry
4.1	Sea/estuary	11.2	Offices
4.2	Standing water	11.3	Retailing
4.3	Running water	11.4	Storage and warehousing
4.4	Freshwater marsh	11.5	Utilities
4.5	Salt marsh	11.6	Agricultural buildings
4.6	Bog		
		12	**Vacant Land and Buildings**
5	**Rock and Coastal Land**	12.1	Vacant land previously developed
5.1	Inland rock	12.2	Vacant buildings
5.2	Coastal rocks and cliffs	12.3	Derelict land and buildings
5.3	Inter-tidal sand and mud		
5.4	Dunes	**13**	**Defence Land and Buildings**
6	**Minerals and Landfill**		
6.1	Mineral workings and quarries		
6.2	Landfill waste disposal		

Source: http://www.nlud.org.uk/

Table A.5: National standard for urban land-use classification, China

R	**Residential**	**S**	**Road and Street**
R1	Good infrastructure and environment with low-rise building ≤ 3 floors	S1	Street
		S2	Square
R2	Good infrastructure and environment with buildings > 3 floors	S3	Parking space
R3	Medium infrastructure and poor environment	**U**	**Municipal Utilities**
R4	Poor infrastructure, poor environment and poor-quality buildings	U1	Supplying facility
		U2	Traffic facility
		U3	Posts and telecommunications
		U4	Environmental facility
C	**Commercial and Public Facilities**	U5	Construction and maintenance
C1	Government offices	U6	Funeral facility
C2	Commercial and financial services	U9	Others
C3	Cultural and recreational use		
C4	Sport facilities	**G**	**Green Space**
C5	Medical treatment and health centre	G1	Public green space
C6	Educational and research use	G2	Productive or protective land
C7	Culture heritage and historic sites		
C9	Others	**D**	**Specially Designated**
		D1	Military
M	**Industrial, Manufacturing**	D2	Diplomatic (embassy or consulate)
M1	Non-pollution industrial	D3	Prison and security place
M2	Light pollution industrial		
M3	Heavy pollution industrial	**E**	**Water Area and Others**
		E1	Water bodies
W	**Warehouse**	E2	Agricultural land
W1	Warehouse for normal material	E3	Garden plot
W2	Warehouse for dangerous material	E4	Forest land
W3	Open storage site	E5	Rangeland
		E6	Villages and small towns
T	**Transportation**	E7	Bare land
T1	Railway	E8	Mining land
T2	Road		
T3	Pipeline (e.g. water, oil, gas) or conveyor belt		
T4	Harbour		
T5	Airport		

Source: The Ministry of Construction, P. R. China, 1991

Table A.6: Urban land use classification system used in this research

Land Use Classes	Main Features
R **Residential**	Uniformed building clusters
R1 Detached and semi-detached houses	Small size (50 - 200 m^2) and low rise (1-3 story)
R2 Low-rise apartment buildings	Medium size (50 - 200 m^2) and low rise (1-3 story)
R3 Medium- and high-rise apartment buildings	Medium size, medium and high rise ($\geq$ 4 stories)
R4 Mixed settlements	Vary in size or irregular distribution
C **Commercial and Services**	
C1 Commercial or service district	Buildings vary in size and height, high proportion of concrete surfaces
C2 Educational and research complex	Buildings vary in size and height
C3 Mixed	
M **Industrial and Warehouse**	Large buildings
S **Road, Street and Square**	Linear concrete surfaces inside urban areas
T **Transportation**	
T1 Road	Linear concrete surfaces outside urban areas
T2 Railway	Linear concrete surfaces outside urban areas
T3 Harbour	High proportion of concrete surfaces along coast or river bank
T4 Airport	Very large concrete surfaces
G **Green Space and Recreational**	
U **Utilities**	
E **Non-urban (Water Area and Others)**	
E1 Water bodies	Low tone in all bands
E2 Agricultural land	Vegetation
E3 Rangeland	Vegetation in hilly areas
E4 Forest land	High vegetation
E5 Others	

List of ITC PhD Theses

1. **Akinyede,** 1990, Highway cost modelling and route selection using a geotechnical information system

2. **Pan, He Ping,** 1990, 90-9003757-8, Spatial structure theory in machine vision and applications to structural and textural analysis of remotely sensed images

3. **Bocco Verdinelli, G.,** 1990, Gully erosion analysis using remote sensing and geographic information systems: a case study in central Mexico

4. **Sharifi, M.,** 1991, Composite sampling optimization for DTM in the context of GIS

5. **Drummond, J.,** 1991, Determining and processing quality parameters in geographic information systems

6. **Groten, S.,** 1991, Satellite monitoring of agro-ecosystems in the Sahel

7. **Sharifi, A.,** 1991, 90-6164-074-1, Development of an appropriate resource information system to support agricultural management at farm enterprise level

8. **Zee, D. van der,** 1991, 90-6164-075-X, Recreation studied from above: air photo interpretation as input into land evaluation for recreation

9. **Mannaerts, C.,** 1991, 90-6164-085-7, Assessment of the transferability of laboratory rainfall-runoff and rainfall - soil loss relationships to field and catchment scales: a study in the Cape Verde Islands

10. **Wang, Ze Shen,** 1991, 90-393-0333-9, An expert system for cartographic symbol design

11. **Zhou, Yunxian,** 1991, 90-6164-081-4, Application of Radon transforms to the processing of airborne geophysical data

12. **Zuviria, M. de,** 1992, 90-6164-077-6, Mapping agro-topoclimates by integrating topographic, meteorological and land ecological data in a geographic information system: a case study of the Lom Sak area, north central Thailand

13. **Westen, C. van,** 1993, 90-6164-078-4, Application of geographic information systems to landslide hazard zonation

14. **Shi, Wenzhong,** 1994, 90-6164-099-7, Modelling positional and thematic uncertainties in integration of remote sensing and geographic information systems

15.   **Javelosa, R.,** 1994, 90-6164-086-5, Active Quaternary environments in the Philippine mobile belt

16.   **Lo King-Chang,** 1994, 90-9006526-1, High quality automatic DEM, digital elevation model generation from multiple imagery

17.   **Wokabi, S.,** 1994, 90-6164-102-0, Quantified land evaluation for maize yield gap analysis at three sites on the eastern slope of Mt. Kenya

18.   **Rodriguez, O.,** 1995, Land use conflicts and planning strategies in urban fringes: a case study of western Caracas, Venezuela

19.   **Meer, F. van der,** 1995, 90-5485-385-9, Imaging spectrometry and the Ronda peridotites

20.   **Kufoniyi, O.,** 1995, 90-6164-105-5, Spatial coincidence: automated database updating and data consistency in vector GIS

21.   **Zambezi, P.,** 1995, Geochemistry of the Nkombwa Hill carbonatite complex of Isoka district, northeast Zambia, with special emphasis on economic minerals

22.   **Woldai, T.,** 1995, The application of remote sensing to the study of the geology and structure of the Carboniferous in the Calañas area, pyrite belt, SW Spain

23.   **Verweij, P.,** 1995, 90-6164-109-8, Spatial and temporal modelling of vegetation patterns: burning and grazing in the paramo of Los Nevados National Park, Colombia

24.   **Pohl, C.,** 1996, 90-6164-121-7, Geometric aspects of multisensor image fusion for topographic map updating in the humid tropics

25.   **Jiang, Bin,** 1996, 90-6266-128-9, Fuzzy overlay analysis and visualization in GIS

26.   **Metternicht, G.,** 1996, 90-6164-118-7, Detecting and monitoring land degradation features and processes in the Cochabamba Valleys, Bolivia: a synergistic approach

27.   **Hoanh Chu Thai,** 1996, 90-6164-120-9, Development of a computerized aid to integrated land use planning (CAILUP) at regional level in irrigated areas: a case study for the Quan Lo Phung Hiep region in the Mekong Delta, Vietnam

28.   **Roshannejad, A.,** 1996, 90-9009284-6, The management of spatio-temporal data in a national geographic information system

29.   **Terlien, M.,** 1996, 90-6164-115-2, Modelling spatial and temporal variations in rainfall-triggered Landslides: the integration of hydrologic models, slope stability models and GIS for the hazard zonation of rainfall-triggered landslides with examples from Manizales, Colombia

30.   **Mahavir, J.,** 1996, 90-6164-117-9, Modelling settlement patterns for metropolitan regions: inputs from remote sensing

31.   **Al-Amir, S.,** 1996, 90-6164-116-0, Modern spatial planning practice as supported by the multi-applicable tools of remote sensing and GIS: the Syrian case

32.   **Pilouk, M.,** 1996, 90-6164-122-5, Integrated modelling for 3D GIS

33. **Duan, Zengshan,** 1996, 90-6164-123-3, Optimization modelling of a river-aquifer system with technical interventions: a case study for the Huangshui river and the coastal aquifer, Shandong, China

34. **Man, W.H. de,** 1996, 90-9009-775-9, Surveys: informatie als norm: een verkenning van de institutionalisering van dorp - surveys in Thailand en op de Filippijnen

35. **Vekerdy, Z.,** 1996, 90-6164-119-5, GIS-based hydrological modelling of alluvial regions: using the example of the Kisaföld, Hungary

36. **Pereira, Luisa,** 1996, 90-407-1385-5, A robust and adaptive matching procedure for automatic modelling of terrain relief

37. **Fandino Lozano, M.,** 1996, 90-6164-129-2, A framework of ecological evaluation oriented at the establishment and management of protected areas: a case study of the Santuario de Iguaque, Colombia

38. **Toxopeus, B.,** 1996, 90-6164-126-8, ISM: an interactive spatial and temporal modelling system as a tool in ecosystem management: with two case studies : Cibodas biosphere reserve, West Java Indonesia: Amboseli biosphere reserve, Kajiado district, central southern Kenya

39. **Wang, Yiman,** 1997, 90-6164-131-4, Satellite SAR imagery for topographic mapping of tidal flat areas in the Dutch Wadden Sea

40. **Asun Saldana-Lopez,** 1997, 90-6164-133-0, Complexity of soils and soilscape patterns on the southern slopes of the Ayllon Range, central Spain: a GIS assisted modelling approach

41. **Ceccarelli, T.,** 1997, 90-6164-135-7, Towards a planning support system for communal areas in the Zambezi valley, Zimbabwe: a multi-criteria evaluation linking farm household analysis, land evaluation and geographic information systems

42. **Peng, Wanning,** 1997, 90-6164-134-9, Automated generalization in GIS

43. **Lawas, C.,** 1997, 90-6164-137-3, The resource users' knowledge, the neglected input in land resource management: the case of the Kankanaey farmers in Benguet, Philippines

44. **Bijker, W.,** 1997, 90-6164-139-X, Radar for rain forest: a monitoring system for land cover change in the Colombian Amazon

45. **Farshad, A.,** 1997, 90-6164-142-X, Analysis of integrated land and water management practices within different agricultural systems under semi-arid conditions of Iran and evaluation of their sustainability

46. **Orlic, B.,** 1997, 90-6164-140-3, Predicting subsurface conditions for geotechnical modelling

47. **Bishr, Y.,** 1997, 90-6164-141-1, Semantic aspects of interoperable GIS

48. **Zhang, Xiangmin,** 1998, 90-6164-144-6, Coal fires in northwest China: detection, monitoring and prediction using remote sensing data

49. **Gens, R.,** 1998, 90-6164-155-1, Quality assessment of SAR interferometric data

50. **Turkstra, J.,** 1998, 90-6164-147-0, Urban development and geographical information: spatial and temporal patterns of urban development and land values using integrated geo-data, Villaviciencia, Colombia

51. **Cassells, C.,** 1998, Thermal modelling of underground coal fires in northern China

52. **Naseri, M.,** 1998, 90-6164-195-0, Characterization of salt-affected soils for modelling sustainable land management in semi-arid environment: a case study in the Gorgan region, northeast, Iran

53. **Gorte, B.G.H.,** 1998, 90-6164-157-8, Probabilistic segmentation of remotely sensed images

54. **Tenalem Ayenew,** 1998, 90-6164-158-6, The hydrological system of the lake district basin, central main Ethiopian rift

55. **Wang, Donggen,** 1998, 90-6864-551-7, Conjoint approaches to developing activity-based models

56. **Bastidas de Calderon, M.,** 1998, 90-6164-193-4, Environmental fragility and vulnerability of Amazonian landscapes and ecosystems in the middle Orinoco river basin, Venezuela

57. **Moameni, A.,** 1999, Soil quality changes under long-term wheat cultivation in the Marvdasht plain, south-central Iran

58. **Groenigen, J.W. van,** 1999, 90-6164-156-X, Constrained optimisation of spatial sampling: a geostatistical approach

59. **Cheng, Tao,** 1999, 90-6164-164-0, A process-oriented data model for fuzzy spatial objects

60. **Wolski, Piotr,** 1999, 90-6164-165-9, Application of reservoir modelling to hydrotopes identified by remote sensing

61. **Acharya, B.,** 1999, 90-6164-168-3, Forest biodiversity assessment: a spatial analysis of tree species diversity in Nepal

62. **Akbar Abkar, Ali,** 1999, 90-6164-169-1, Likelihood-based segmentation and classification of remotely sensed images

63. **Yanuariadi, T.,** 1999, 90-5808-082-X, Sustainable land allocation: GIS-based decision support for industrial forest plantation development in Indonesia

64. **Abu Bakr, Mohamed,** 1999, 90-6164-170-5, An integrated agro-economic and agro-ecological framework for land use planning and policy analysis

65. **Eleveld, M.,** 1999, 90-6461-166-7, Exploring coastal morphodynamics of Ameland (the Netherlands) with remote sensing monitoring techniques and dynamic modelling in GIS

66. **Yang Hong,** 1999, 90-6164-172-1, Imaging spectrometry for hydrocarbon Microseepage

67. **Mainam, Félix,** 1999, 90-6164-179-9, Modelling soil erodibility in the semiarid zone of Cameroon

68. **Bakr, Mahmould,** 2000, 90-6164-176-4, A stochastic inverse-management approach to groundwater quality

69. **Zlatanova, Z.,** 2000, 90-6164-178-0, 3D GIS for urban development

70. **Ottichilo, Wilber K.,** 2000, 90-5808-197-4, Wildlife dynamics: an analysis of change in the Masai Mara ecosystem

71. **Kaymakci, Nuri,** 2000, 90-6164-181-0, Tectono-stratigraphical evolution of the Cankori Basin (central Anatolia, Turkey)

72. **Gonzalez, Rhodora,** 2000, 90-5808-246-6, Platforms and terraces: bridging participation and GIS in joint-learning for watershed management with the Ifugaos of the Philippines

73. **Schetselaar, Ernst,** 2000, 90-6164-180-2, Integrated analyses of granite-gneiss terrain from field and multisource remotely sensed data: a case study from the Canadian Shield

74.   **Mesgari, Saadi,** 2000, 90-3651-511-4, Topological cell-tuple structure for three-dimensional spatial data

75. **Bie, Cees A.J.M. de,** 2000, 90-5808-253-9, Comparative performance analysis of agro-ecosystems

76. **Khaemba, Wilson M.,** 2000, 90-5808-280-6, Spatial statistics for natural resource management

77.   **Shrestha, Dhruba,** 2000, 90-6164-189-6, Aspects of erosion and sedimentation in the Nepalese Himalaya: highland-lowland relations

78. **Asadi Haroni, Hooshang,** 2000, 90-6164-185-3, The Zarshuran gold deposit model applied in a mineral exploration GIS in Iran

79. **Raza, Ale,** 2001, 90-3651-540-8, Object-oriented temporal GIS for urban applications

80. **Farah, Hussein,** 2001, 90-5808-331-4, Estimation of regional evaporation under different weather conditions from satellite and meteorological data: a case study in the Naivasha Basin, Kenya

81. **Zheng, Ding,** 2001, 90-6164-190-X, A neuro-fuzzy approach to linguistic knowledge acquisition and assessment in spatial decision making

82. **Sahu, B. K.,** 2001, Aeromagnetics of continental areas flanking the Indian Ocean: with implications for geological correlation and Gondwana reassembly

83. **Alfestawi, Y.,** 2001, 90-6164-198-5, The structural, paleogeographical and hydrocarbon systems analysis of the Ghadamis and Murzuq Basins, west Libya, with emphasis on their relation to the intervening Al Qarqaf Arch

84. **Liu, Xuehua,** 2001, 90-5808-496-5, Mapping and modelling the habitat of giant pandas in Foping Nature Reserve, China

85. **Oindo, Boniface Oluoch,** 2001, 90-5808-495-7, Spatial patterns of species diversity in Kenya

86. **Carranza, Emmanuel John,** 2002, 90-6164-203-5, Geologically-constrained mineral potential mapping

87 **Rugege, Dennis,** 2002, 90-5808-584-8, Regional analysis of maize-based land use systems for early warning applications

88 **Liu, Yaolin,** 2002, 90-5808-648-8, Categorical database generalization in GIS

89 **Ogao, Patrick Job,** 2002, 90-6164-206-X, Exploratory visualization of temporal geospatial data using animation

90. **Abadi, Abdulbaset M.,** 2002, 906164-205-1, Tectonics of the Sirt Basin: – interferences from tectonic subsidence analysis, stress inversion and gravity modeling

91. **Geneletti, Davide,** 2002, ISSN 0169-4839, Ecological evaluation for environmental impact assessment

92. **Sedogo, Laurent G.,** 2002, ISBN 90-5808-751-4, Integration of participatory local and regional planning for resources management using remote sensing and GIS

93. **Montoya, Ana Lorena,** 2002, ISBN 90-6164-2086, Urban disaster management: a case study of earthquake risk assessment in Cartago, Costa Rica

94. **Ahmad, Mobin-ud-Din,** 2002, ISBN 90-5808-761-1, Estimation of net groundwater use in irrigated river basins using geo-information techniques: a case study in Rechna Doab, Pakistan

95. **Said, Mohammed Yahya** 2003, ISBN 90-5808-794-8, Multiscale perspective of species richness in East Africa

96. **Schmidt, Karin S.** 2003, ISBN 90-5808-830-8, Hyperspectral remote sensing of vegetation species distribution in a saltmarsh

97. **López Binnqüist, Citlalli** 2003, ISBN 9036519004, The endurance of mexican amate paper: exploring additional dimensions to the sustainable development concept

98. **Huang, Zhengdong,** 2003, ISBN 90-6164-211-6, Data integration for urban transport planning

99. **Cheng, Jianquan,** 2003, ISBN 90-6164-212-4, Modelling spatial and temporal urban growth

100. **Campos dos Santos, Jose Laurindo,** 2003, 90-6164-214-0, A biodiversity information system in an open data/metadatabase architecture

101. **Hengl, Tomislav,** 2003, 90-5808-896-0, Pedometric mapping, bridging the gaps between conventional and pedometric approaches

102. **Barrera Bassols, Narciso,** 2003, 90-6164-217-5, Symbolism, knowledge and management of soil and land Resources in indigenous communities: ethnopedology at global, regional and local scales

Samenvatting

Zhan, Q., 2003. Een hiërarchische, objectgeoriënteerde aanpak voor stedelijke landgebruik classificatie van remote sensing data. PhD dissertatie

Informatie over landbedekking en landgebruik is essentieel voor stedelijke planning en management. Traditionele landgebruik kartering door middel van visuele interpretatie is duur, tijdrovend en vaak subjectief. Onderzoekers zoeken sinds lange tijd naar automatische en semi-automatische methodes. De combinatie van vliegtuig LIDAR data met hoge ruimtelijke resolutie en multi-spectrale beelden zoals IKONOS, QuickBird en SPOT 5, biedt uitstekende kansen voor toepassingen in stedelijke gebieden. Uit de combinatie van dergelijke beelden kunnen vele relevante objecten geëxtraheerd worden. De algemene doelstelling van dit onderzoek is de ontwikkeling van automatische of semi-automatische methoden voor de classificatie van landbedekking en landgebruik, gebaseerd op laser hoogte data en multi-spectrale beelden en de ontwikkeling van methoden voor de consistente aggregatie van elementaire objecten tot samengestelde objecten op hoge abstractie niveaus.

In dit onderzoek zijn verschillende moderne typen sensor data gebruikt voor de classificatie van stedelijke landbedekking en landgebruik. We hebben de meest populaire pixel gebaseerde classificatie, de traditionele maximum likelihood classifier (MLC), toegepast op hoge resolutie data. Een aantal problemen konden geïdentificeerd worden. Verscheidene oplossingen zijn voorgesteld en getest. De nauwkeurigheid van landbedekking classificatie kan verbeterd worden door de specificatie van een beslissingsvlak in de attribuutruimte en door het selecteren van zowel pure als gemengde trainingspixels. De voorgestelde oplossingen richten zich op de ruimtelijke partitionering van beslissingsvlakken, aan de hand van referentie klassen. Experimentele resultaten tonen de effectiviteit van de voorgestelde integratie methode aan. Deze methode gebruikt zowel pure als gemengde monsters.

Ondanks deze verbetering van de landbedekking classificatie met het MLC algoritme, beschouwen we de resultaten als onvoldoende voor een gedetailleerde landgebruik classificatie. De belangrijkste kenmer-

ken voor beeldinterpretatie, zoals grootte, vorm, kleur, oriëntatie, patroon en associatie worden gezien als kenmerken van geaggregeerde objecten op een hoger abstractie niveau. Deze kenmerken hebben een belangrijke rol in beeldanalyse en landgebruik classificatie. Objectgeoriënteerde beeldanalyse technieken betreffen een abstractie niveau hoger dan het pixel niveau. Daarom is een objectgeoriënteerde aanpak het belangrijkste aandachtspunt van dit onderzoek.

Dit onderzoek ontwikkelt een hiërarchische objectgeoriënteerde aanpak voor stedelijk landgebruik classificatie gebaseerd op hoge resolutie remote sensing. De methode bestaat uit drie stappen: de classificatie van landbedekking, de definitie en identificatie van landgebruik eenheden en de classificatie van landgebruik. Deze methode combineert pixel en object gebaseerde technieken in verschillende fasen. Verschillende technieken worden voorgesteld en getest voor object identificatie op verschillende aggregatie niveaus.

Verschillende concepten en methoden worden voorgesteld and bediscussieerd om objecten en object eigenschappen uit beelden af te leiden en om expliciete topologische relaties tussen objecten te identificeren. Een hybride raster model wordt toegepast om topologische relaties tussen objecten te identificeren. Deze concepten en methoden zijn getest en succesvol toegepast op stedelijke landbedekking en landgebruik classificatie in twee test gebieden. We beschouwen de test resultaten als veelbelovend. De resultaten bevestigen de effectiviteit van de voorgestelde objectgeoriënteerde aanpak.

Structurele informatie afgeleid uit hiërarchisch geordende beeld objecten speelt een belangrijke rol in landgebruik classificatie van stedelijke gebieden. Delaunay triangulatie toegepast op ruimtelijk verspreide objecten levert een goede methode om topologische relaties en een nabijheidsmaat te definiëren. Verscheidene maten worden voorgesteld om similariteit van gebouwen te testen. Deze similariteitsmaten in combinatie met de topologische relaties en nabijheidsmaten triangulatie geven essentiële informatie voor de ruimtelijke clustering van objecten die landgebruikseenheden vormen.

Verschillende landgebruik object eigenschappen worden voorgesteld voor onze twee testgebieden. Fuzzy membership functies zijn ontworpen om de relaties tussen deze eigenschappen en landgebruik klassen vast te stellen. Een fuzzy object classificatie wordt toegepast, gebaseerd op landgebruik eenheden en hun object eigenschappen. De verkregen resultaten laten zien dat de voorgestelde objectgeoriënteerde landgebruik classificatie veelbelovend is. De verkregen landgebruik object eigenschappen bieden nuttige informatie voor stedelijke studies, planning en management.

Verder wordt voorgesteld en getest voor kwaliteitsbeoordeling, gebaseerd op gelijkheidsmaten tussen geclassificeerde en referentie data. Deze beoordeling maakt gebruik van object en pixel gebaseerde maten. De voorgestelde kwaliteitsmaat geeft mogelijkheden om een additionele

kwaliteitsbeoordeling te verkrijgen gebaseerd op verscheidene object eigenschappen. De voorgestelde onzekerheidsmaten voor de verkregen landbedekking en landgebruik objecten zijn getest. Deze lijken bruikbaar voor het controleren van het classificatieproces.

De ontwikkelde concepten en methoden zijn geïmplementeerd in Matlab. Het systeem staat verschillende gebruikers toe om karakteristieken te specificeren voor het verkrijgen van de gewenste resultaten. Dit biedt planners en andere gebruikers de mogelijkheid om uit een gedetailleerde data set die informatie te halen die voldoet aan hun specifieke wensen en eisen. Dit maakt meervoudig gebruik mogelijk van de relatief dure hoge resolute laser data en spectrale data. De experimentele resultaten laten de mogelijkheden zien van hiërarchische objectmodellering in combinatie met structurele beeld analyse technieken voor stedelijke landbedekking en landgebruik classificatie.

Dit onderzoek laat zien dat gebouwen, groene ruimte, water lichamen en dichte oppervlaktes succesvol geïdentificeerd kunnen worden op het niveau van landbedekking. Landgebruik eenheden zijn verkregen door combinatie en aggregatie van landbedekking objecten. De hoge kwaliteit van de objectgeoriënteerde landgebruik classificatie werd vastgesteld door vergelijking met resultaten van een visuele interpretatie.

Trefwoorden

Remote Sensing, GIS, Beeldbewerking, beeldclassificatie, kenmerk identificatie, object modellering, landgebruik, stedelijke planning, LIDAR.

Curriculum Vitae

Qingming Zhan was born in Fujian province, China, on 28 September 1964. He received his BSc and MSc in photogrammetry and remote sensing from Wuhan Technical University of Surveying and Mapping (WTUSM), Wuhan, China, in 1985 and 1995 respectively. He had followed the postgraduate course in urban survey and human settlement analysis at ITC from 1986 to 1987. He became a staff member of the remote sensing section in the department of photogrammetry and remote sensing, WTUSM, in 1985. He joined the Educational Centre for Urban/Rural Survey, Planning and Management (EURSPAM), a joint educational institute of WTUSM and ITC, in 1987. He became a staff member of the School of Urban Studies (SUS), WTUSM, which was formed by merging EURSPAM and the department of civil engineering, in 1993. He has been intensively involved in the cooperation projects between WTUSM (now Wuhan University) and ITC for more than 15 years.

Qingming Zhan has been involved in more than 20 scientific research projects projects since 1985. He has undertaken scientific research projects with the Natural Science Foundation of China, the State Bureau of Surveying and Mapping of China, World Bank, UNDP, and the Dutch development aid fund. His research experience covers GIS, remote sensing and applications in urban planning and management. His current research interests include automatic feature extraction and image understanding based on remote sensing images, object-based analysis of remote sensing images, and land-cover and land-use classification methods. He has supervised 20 MSc students and more than 50 BSc students. He is the author and co-author of over 30 scientific articles. He is the author and co-author of four books, including this one.