Multispectral and Digital Elevation Model Information Fusion for Classification and Change Detection in Earth Observation

Isabel Rodes Arnau

Université de Bourgogne (France)
Universitat de Girona (Spain)
Heriot-Watt University (United Kingdom)

Supervisor: Dr. H. Chaabouni-Chouayakh
Deutsches Zentrum für Luft- und Raumfahrt (DLR)
Weßling, München, Germany

A Thesis Submitted for the Degree of MSc Erasmus Mundus in Vision and Robotics (VIBOT)

· 2011 ·

Abstract

A review of the state-of-the-art in land cover classification with satellite imagery of the Earth is provided. Information fusion of different typologies (in terms of processing, operation levels and information sources) for classification and change detection purposes is also surveyed. Three different approaches for automatic unsupervised classification of land cover have been implemented to exploit the fusion of multispectral and digital elevation model information for subsequent change detection processes. In broad terms, the first method is based on feature extraction and clustering, the second makes use of a hierarchical decision tree, and the third one constitutes a hybrid, incorporating further techniques. Different types of features have been extracted, at several levels of the process, and merged also at different fusion levels. The algorithms have been tested on two case scenarios, defined by three datasets. A comprehensive set of evaluation measures has been used for the assessment of the classifications, in the three cases scoring high, and especially with the third method, with completeness and correctness values of 97.1877% and 97.3453%. The potential of subsequently applying post-classification change detection techniques with the obtained results is demonstrated, allowing not only for the detection of changes, but for their identification. Implementation of the algorithms and the software system and graphical interface generated has been undertaken in Matlab v.R2010a and IDL v.6.2.

You need to choose a dream, and go for it . . .

Contents

Acknowledgments **viii**

1 Introduction **1**

2 Problem Statement **5**

2.1 Motivation 5

2.2 Problem statement 6

3 State of the Art **10**

3.1 Earth Observation satellites and sensors 10

3.2 IKONOS imagery 11

3.3 Classification of satellite imagery 13

3.3.1 Height information-based classification 15

3.3.2 Spectral information-based classification 16

3.4 Fusion 18

3.4.1 Fusion for classification 19

3.4.2 Fusion for change detection 20

3.5 Remarks 22

4 Methodology **23**

4.1 Dataset 24

4.2 Feature analysis . . . 26
4.2.1 Spectral properties . . . 27
4.2.2 Shape properties . . . 34
4.2.3 Textural properties . . . 37
4.2.4 Gabor features . . . 38
4.3 Feature fusion for classification . . . 39
4.3.1 Method 1 . . . 39
4.3.2 Method 2 . . . 45
4.3.3 Method 3 . . . 46
4.4 Change detection . . . 49

5 Results **51**
5.1 Evaluation measures . . . 51
5.2 Classification results . . . 53
5.2.1 Method 1 . . . 54
5.2.2 Method 2 . . . 56
5.2.3 Method 3 . . . 58
5.3 Change detection results . . . 61

6 Conclusions **63**
6.1 Evaluation . . . 63
6.2 Future work . . . 64

Bibliography **77**

List of Figures

1.1 Rondonia, Brazil. (a) ASTER false colour image (2nd January 2001). (b) IKONOS true colour image (16th January 2001). From NASA's Visible Earth [2]. 2

1.2 Discovery of a new mass burial site in Izbica, Kosovo. Aerial photographs (April 1999). From GlobalSecurity [36]. 3

1.3 Japan tsunami affected area. Rapid Eye images. (a) 5th September 2010. (b) 12th March 2011. From Deutsches Zentrum für Luft- und Raumfahrt (DLR). [31]. 3

2.1 Destruction in Mbare Township, Zimbabwe. (a) QuickBird image (16th April 2005). (b) IKONOS image (27th June 2005). From UNOSAT (2011). 6

2.2 Reconstruction in Banda Aceh, Indonesia. (a) QuickBird image (2005). (b) GeoEye image (2009). From [63]. 6

2.3 München, Germany. Examples of the semantic class *building*. IKONOS image (2005). 7

2.4 Spectral curves for the examples of class *building* in Fig. 2.3, presenting dissimilarities both value and curve-wise. 7

2.5 München, Germany. Examples of the semantic classes *building* and *water*. IKONOS image (2005). 8

2.6 Spectral curves for the examples of classes *building* and *water* in Fig. 2.5 present similarities. 8

2.7 München, Germany. IKONOS true colour image (year) with close-up of critical area, and the corresponding digital elevation model. 9
2.8 Mesh of the digital elevation model for München, generated from the stereo pair, and used for orthorectification. From Deutsches Zentrum für Luft- und Raumfahrt (DLR). 9

3.1 (a) IKONOS spectral response curves. From Pagnutti et al. [85] (b) IKONOS average revisit times. From Dial et al. [24]. 13
3.2 AVIRIS and Landsat spectral channels. From Simmonds [37]. 14
3.3 Levels of fusion: low or pixel-level, intermediate or feature-level and high or decision-level fusion. Adapted from [114]. 19

4.1 Overview of the processing flow. The different methods here presented implement fusion for classification and change detection at different levels 23
4.2 IKONOS generated digital elevation models for München (2005, see Table 4.1 for parameters) and Yongbyon (2006, see Table 4.2). 26
4.3 Reflectance spectra for wheat, dry and wet soil. From Jackson and Huete [50]. . 27
4.4 Tasseled Cap Transformation and PVI. From Jackson and Huete [50]. 28
4.5 Red and near-infrared RVI, NDVI, and PVI isolines of equal vegetation amounts. From Jackson and Huete [50]. 30
4.6 München and Yongbyon true colour IKONOS images. 40
4.7 NDVI values histogram. Gaussian fitting is used to identify vegetation and water areas. 40
4.8 Two-dimensional principal component transformation, here used in higher dimensional spaces. From [42]. 42
4.9 München dataset. Calculated Principal Components. 44
4.10 Method 1 classification results. (a) München (2005) (b) Yongbyon (2006). 45
4.11 München dataset. Segmented vegetation areas and water bodies. 46
4.12 Method 2 classification results. (a) München (2005) (b) Yongbyon (2006). 47

4.13 München dataset. Excentricity and solidity values for building areas. 48
4.14 Normalised dem above zero values and derived building map. 48
4.15 Method 3 classification results for München (2005) and Yongbyon (2006). 49
4.16 (a) Yongbyon dataset (2010) (b) Detection results for positive changes in the central area of image (a). 50

5.1 Unsupervised clustering of (a) multispectral information (b) multispectral and height information 54
5.2 München dataset. Classification detail. (a) Method 1 (b) Method 2. 60
5.3 München dataset. Classification detail. (a) Method 2 (b) Method 3. 60
5.4 Graphical user interface of the implemented software. 61

List of Tables

3.1 IKONOS specifications. From Dial et al. [24] and Grodecki and Dial [38] 12

3.2 IKONOS multispectral bands. From Dial et al. [24] and Grodecki and Dial [38] . 12

4.1 München dataset acquisition parameters 24

4.2 Yongbyon dataset acquisition parameters for 2006 25

4.3 Yongbyon dataset acquisition parameters for 2010 25

4.4 München dataset. Calculation of variances and cumulative variance percentages for the first six PCs 43

5.1 Confusion matrix for results of Method 1 (in pixels). NB: Water bodies are not detected, and only central areas of regions are considered, due to the difficulty of manually delimiting border regions for the creation of ground truth. 54

5.2 Confusion matrix for results of Method 1 (in percentages). 55

5.3 Comission and omission measures for results of Method 1 (in percentages). . . . 55

5.4 Producer accuracy and user accuracy measures for results of Method 1 (in percentages). 55

5.5 Pixel-based results evaluation for Method 1 in percentages: *completeness* C_mp, *correctness* C_rp, *branching factor* B_f, *miss factor* M_f, *quality* Q_p, *accuracy* A and *Helden's index* H_m 55

5.6 Confusion matrix for results of Method 2 (in pixels). Only central areas of regions are considered. 57

5.7 Confusion matrix for results of Method 2 (in percentages). Only central areas of regions are considered. 57
5.8 Comission and omission measures for results of Method 2 (in percentages). Only central areas of regions are considered. 57
5.9 Producer accuracy and user accuracy measures for results of Method 2 (in percentages).Only central areas of regions are considered. 57
5.10 Pixel-based results evaluation for Method 2 in percentages: *completeness* C_mp, *correctness* C_rp, *branching factor* B_f, *miss factor* M_f, *quality* Q_p, *accuracy* A and *Helden's index* H_m 58
5.11 Confusion matrix for results of Method 3 (in pixels). Only central areas of regions are considered. 58
5.12 Confusion matrix for results of Method 3 (in percentages). Only central areas of regions are considered. 59
5.13 Comission and omission measures for results of Method 3 (in percentages). Only central areas of regions are considered. 59
5.14 Producer accuracy and user accuracy measures for results of Method 3 (in percentages).Only central areas of regions are considered. 59
5.15 Pixel-based results evaluation for Method 3 in percentages: *completeness* C_mp, *correctness* C_rp, *branching factor* B_f, *miss factor* M_f, *quality* Q_p, *accuracy* A and *Helden's index* H_m 60
5.16 Pixel-based change detection evaluation results in percentages. 61
5.17 Object-based change detection evaluation results in percentages. 61

Acknowledgments

I would like to express my gratitude to my supervisor Dr. H. Chaabouni-Chouayakh, for having accepted to supervise this work in difficult times, and for all her time, advice and support; to Prof. P. Reinartz, for having given me the opportunity of joining his department at the Deutsches Zentrum für Luft- und Raumfahrt; to Dr. A. Makarau, for all his valuable advice; and in general to all the special people I have had the opportunity of meeting at the DLR.

I would also like to dedicate this thesis to my parents and my brother, for all their love and support over the years. In particular, I would like to thank my partner and *coéquipier* Alexandru for all his help and patience, and for all the happy times we have shared, not only during the writing of our theses at the DLR, but during the whole MSc VIBOT period. I would not have got here without you.

I am grateful to the coordinators, lecturers and administration personnel of the MSc VIBOT course, for having given me the opportunity of joining such a programme, of enjoying the teaching quality of the participating universities, and of living this international experience. Particularly, I wish to say thank you to Prof. D. Fofi and Prof. Y. Fougerolle from Université de Bourgogne and Prof. J. Salvi from Universitat de Girona for their availability and willingness to provide help.

Finally, I would like to express my good wishes to all the colleagues and friends of the course. It has been a great academic and life experience, and I wish you all the best!

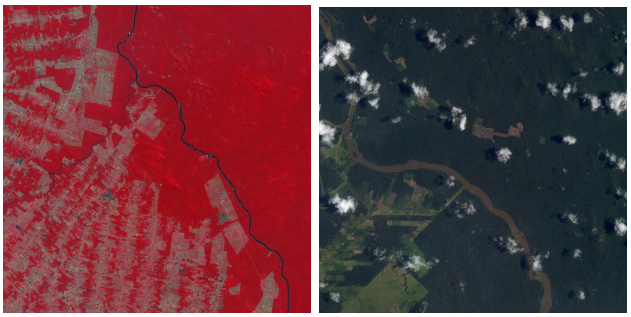
Chapter 1

Introduction

Earth observation satellites provide comprehensive and continually updated information for the monitoring and investigation of our planet's atmosphere, oceans and coasts, land surface, and polar regions, among others, for purposes like disaster relief, climate research, and environmental and security monitoring. Accurate and timely change detection of the Earth's surface features and changes is important in these areas in order to aid policy making and emergency action.

Many activities illustrate the importance of remote sensing satellite imaging. In the polar regions, the melting of ice areas due to an increase of average temperature in the last decades, and the subsequent rise of sea levels, are some of the aspects modelled and monitored by use of remote sensing data. In the marine environment, changing conditions, storm propagation and distribution of temperature and wind are analysed. In the atmosphere, water and humidity fluxes, aerosol-cloud interactions, and pollutant and greenhouse gases distributions are remotely sensed and studied in an attempt to further understand climate change and complex processes in the atmosphere [35].

Remote sensing of the Earth's surface is of particular relevance, as it allows the study of changes in the planet's landscapes. Desertification in arid and semiarid regions is expanding, fresh water resources are declining in availability, and fertile soils are being exhausted with the world's expanding population needs and the present-day unprecedented climatic and environmental changes. Resource scarcity and environmental degradation affect human security in different dimensions, as they also play a role in generating and exacerbating violent conflict [120]. Fig. 1.1 shows ASTER (Advance Spaceborne Thermal Emission and Reflection Radiometer) and IKONOS imaging of a region in Rondonia, Brazil. In the ASTER false colour composite (combining near-infrared, red, and green bands), tropical rainforest corresponds to bright red, brown highlights deforested land, and black and grey indicate recently burnt areas.



(a) (b)

Figure 1.1: Rondonia, Brazil. (a) ASTER false colour image (2nd January 2001). (b) IKONOS true colour image (16th January 2001). From NASA's Visible Earth [2].

The IKONOS true colour image corresponds to a fraction of the same area.

Different change forces, both climate-related and human-induced, can be identified, such as natural disasters (earthquakes, floods, volcanic eruptions, fires, droughts), land and ocean changes (desertification, rising water levels, shifting vegetation areas), overexploitation and exhaustion (fisheries, arable land) and ecosystem interference (forests, coral reefs). Scarcity and environment degradation result in population stress, economic productivity decrease, social difference deepening, and migration, which in turn can result in ethnic clashes and conflicts over decreasing supplies, such as water [119].

The key role of remote sensing for land monitoring is thus made clear, as it provides varied and rich information for strategy formulation on a variety of aspects, ranging from crop and settlement monitoring (mapping of crop illnesses, identification of physical and social vulnerabilities of human settlements) to nuclear and arms facilities building observation for peace keeping, treaty supervision (existence of illegal diamond mines, drug cultivation fields, toxic waste dumping), and in situations of humanitarian crisis (monitoring of population movements, refugee camps, landmine scatter, mass grave proliferation in relation to ethnic cleansing) [97]. Fig. 1.2 shows, during the Kosovo conflict, one of the proofs presented by NATO indicating ethnic cleansing activities [36].

Furthermore, the application of remote sensing in natural disaster and crisis relief action is of the outmost importance, as it supplies timely and precise disaster information for prompt assistance in crisis management and provision of humanitarian aid. Scenarios varying from fire and landslide emergencies to floods, volcanic eruptions and earthquake situations, the rapid assessment of location and extent of damaged areas, affected population, needs and transport infrastructure for the relief teams is a critical issue [33]. Fig. 1.3 shows imaging of the Japanese coast before and after the tsunami on the 11th of March 2011, acquired in the frame of the

Figure 1.2: Discovery of a new mass burial site in Izbica, Kosovo. Aerial photographs (April 1999). From GlobalSecurity [36].

(a) (b)

Figure 1.3: Japan tsunami affected area. Rapid Eye images. (a) 5th September 2010. (b) 12th March 2011. From Deutsches Zentrum für Luft- und Raumfahrt (DLR). [31].

International Charter on Space and Major Disasters [1], providing for the supply of satellite data to relief groups in the event of major disasters.

The need for the development of automatic methods and algorithms allowing the efficient and fast use of satellite image data for classification and change detection of land cover is as a consequence determinant, both in emergency cases and in day-to-day monitoring of the environment, as explained above. In this sense, the availability of new Earth observation satellites yielding very high resolution (VHR) imaging of the Earth's surface, such as IKONOS or the more recent WorldView, has paved the way towards new possibilities for the classification of land cover and change detection. These newly attained resolutions have determined the emergence of a new breed of data acquisition techniques, complementing and progressively replacing previously existent methods, and currently being researched and expanded for greater effectiveness and efficiency.

In this context, the present study aims at offering a new approach for the automatic and unsupervised classification and change detection analysis of very high resolution IKONOS satellite imagery of the Earth's surface, non-dependent on user input and not requiring additional

geographical or 3D information coming from other sources (such as LIDAR or SAR imagery), and based solely on the information gathered and supplied by the satellite's stereo pairs. In particular, the present method relies on information fusion techniques, integrating IKONOS multispectral images and the generated digital elevation model data derived from the supplied imagery for classification and change detection over a particular region.

Thus, the main objectives of this study are the following:

1. To provide an overview of the existing approaches for land cover classification and their application to change detection with remotely sensed high resolution satellite data, as well as to provide insight on the associated domain-specific knowledge.

2. To implement an effective, fully automatic and user-independent multi-spectral IKONOS *Geo* image classification system by usage of digital elevation models generated from the satellite's stereo pairs, and to apply this classification for change detection purposes.

3. To analyse and implement different approaches for the fusion of information (pixel and object-oriented; operating at decision, feature and pixel levels; and integrating information from different sources), as well as the impact on classification of the usage of spectral, height, contour, region, and texture features.

4. To develop a comprehensive software system exploiting the strengths of multispectral images and digital elevation models in order to achieve an accurate classification and change detection tool.

Implementation has been undertaken in Matlab v.R2010a and IDL (Interactive Data Language) v.6.2, using the processing and analysis software ENVI v.4.7, and making use of the Deutsches Zentrum für Luft- und Raumfahrt (DLR, German Aerospace Centre) in-house processing software XDibias. The DLR is the German national research centre for aerospace, energy, transportation, and the exploration of the Earth and the solar system. IKONOS satellite high resolution images have been supplied by Space Imaging and GeoEye.

This work is organised in six chapters; the introduction to the topic, relevance of the study, and description of objectives being given here. Chapter 2 offers information on the topic basics, and a description of the main problematic areas to be overcome. Chapter 3 gives a literature review on a variety of approaches for classification, fusion and change detection with remote sensing imagery. In Chapter 4, the methodology applied for the development of this study is explained. Chapter 5 presents both classification and change detection results obtained with the application of the chosen methods, as well as discussion and comparison based on the analysis against ground truth by usage of a range of evaluation statistic measures. Finally, in Chapter 6, discussion on results is given, general conclusions are drawn, and future directions of research are identified.

Chapter 2

Problem Statement

2.1 Motivation

As explained in the previous section, the development of methods and algorithms for classification and change detection of land cover based on satellite image data is of great value, for example for emergency response in crisis situations, for monitoring of human migrations and vulnerable settlements, and for the supervision of humanitarian agreements and safety treaties. The new Earth observation satellites such as IKONOS or QuickBird, yielding very high resolution imaging of the planet's surface, provide greatly valuable multi-temporal information that can be exploited with these aims. Fig. 2.1 shows imagery of these two satellites covering a location in Zimbabwe affected by a clean-up operation undertaken by the government in 2005, a consequence of which an estimated 700,000 people lost their homes, their source of income, or both [55]. The two images, separated by a time span of 2 months, show the extent of the building destruction undertaken.

The need for the implementation of classification and change detection methods for the analysis of Earth observation satellite imagery can also be understood in terms of generation of land cover maps, which are relevant both in ecological and land management aspects, for example in terms of urban planning, water management, and agricultural statistics. Furthermore, classification and change detection are valuable for the assessment of damage and building destruction, as in the previous example, but also for the monitoring of building construction and reconstruction after disasters, or to supervise compliance with arms non-proliferation treaties in risk designated areas. Fig. 2.2 shows imagery of a location in Sumatra, affected by a tsunami in 2004. The images were captured by QuickBird in 2005 and GeoEye in 2009, showing the extent to which the region was reconstructed in a period of four years.

(a) (b)

Figure 2.1: Destruction in Mbare Township, Zimbabwe. (a) QuickBird image (16th April 2005). (b) IKONOS image (27th June 2005). From UNOSAT (2011).

Figure 2.2: Reconstruction in Banda Aceh, Indonesia. (a) QuickBird image (2005). (b) GeoEye image (2009). From [63].

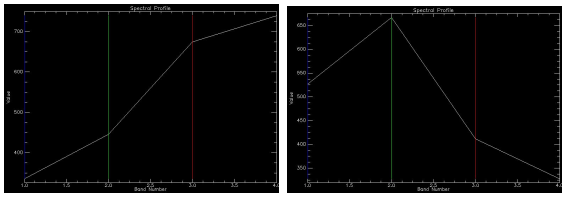
2.2 Problem statement

Many methods for classification and change detection exist, with varying degrees of user and parameter dependence. In remote sensing, classification and change detection from multispectral imagery require complex algorithms and techniques, due to the variability of the data in relation to semantic categories. Furthermore, most approaches in the literature make use of supervised approaches, requiring the user to input data, set parameters, or train method classifiers. This work presents a comparison between different algorithms for totally automatic, unsupervised classification of IKONOS imagery, not requiring any information input by the user, for subsequent change detection tasks.

Automatic unsupervised classification of multispectral satellite imagery is challenging in a number of ways. Firstly, as explained, a high degree of inhomogeneity is present; the semantic class *buildings*, for example, presents varying spectral characteristics due to the differences in materials to be found in rooftops, such as tiles, concrete or metal. The training of a classifier with such varying properties is hard when a high accuracy is aimed for [70]. Secondly, the lack of input by the user for the learning of the classes makes the classification more difficult, as no

(a) (b)

Figure 2.3: München, Germany. Examples of the semantic class *building*. IKONOS image (2005).



(a) (b)

Figure 2.4: Spectral curves for the examples of class *building* in Fig. 2.3, presenting dissimilarities both value and curve-wise.

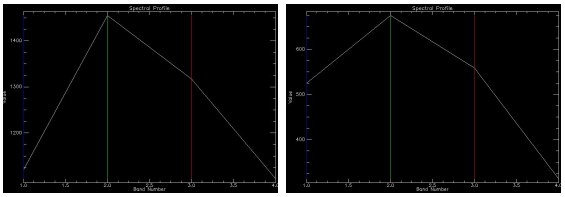
specification of the desired properties of a class is given. Fig. 2.4 shows the dissimilarity to be encountered between two spectral curves, both corresponding to buildings, as shown in Fig. 2.3.

Thirdly, the opposite case occurs frequently, with two different classes presenting similar spectral curves. Fig. 2.6 shows the similarity that can be found between certain types of roof and water bodies in the images in Fig. 2.5. In the fourth place, shape and neighbourhood information are also relevant for the classification, a property which makes the implementation of a classical two-phase feature extraction-classification approach difficult to implement [49]. Finally, as the implemented classification algorithm is to be subsequently applied for change detection, pixel accuracy is demanded, which excludes the possibility of applying level sets methods or line fitting algorithms, as any misclassification in the classification phase results necessarily in greater error rates in the following change detection step.

Images obtained by IKONOS allow for the identification of surface objects such as vegetation, buildings, water, shadows and ground with a precision higher than that previously attainable. However, the associated resolution limit and the presence of shadows are still drawbacks of this imagery, and require appropriate treatment of the data. For this reason, most of

(a) (b)

Figure 2.5: München, Germany. Examples of the semantic classes *building* and *water*. IKONOS image (2005).



(a) (b)

Figure 2.6: Spectral curves for the examples of classes *building* and *water* in Fig. 2.5 present similarities.

the current literature uses additional data sets such as LIDAR (Light Detection and Ranging), SAR (Synthetic Aperture Radar) or GIS (Geographic Information System) information in order to complement the extraction of information from IKONOS images.

This work presents a classification based on the following general land cover categories: *buildings*, *shadows*, *water areas*, *low vegetation*, *ground*, and *high vegetation*. Logically, no spectral information can supply a satisfactory partitioning between low and high vegetation, or between buildings and ground, due to the similarity of spectral properties between them. For this purpose, the digital elevation model derived from the IKONOS stereo pairs is fused with the existent multispectral image. This poses new problems for the classification and change detection steps, as the quality of the generated height information is affected by a number of factors, such as the accuracy of the source data, the available ground control data, and the degree of complexity of the area. In addition to this, the interpolation of points needed to fill the digital elevation model also generates erroneous data that is added to the final model. Fig. 2.7 exemplifies the height model quality obtained with one of the used IKONOS stereo pairs.

Finally, DEM inaccuracies affect the orthorectification process that remote sensing images undergo before integration into a geographic information system, data fusion, or subsequent

(a) (b)

Figure 2.7: München, Germany. IKONOS true colour image (year) with close-up of critical area, and the corresponding digital elevation model.

processing for classification, for example. Orthorectification refers to to the treatment by which geometric distortions contained in raw images are rectified, so that in a final unit of terrain-geocoded data, all distortions including the relief are corrected [108]. As further explained in section 4, the digital elevation model is applied on the input images to obtain this correction [110], which translates in any error contained in the DEM being further propagated in the product images that eventually undergo subsequent processing.

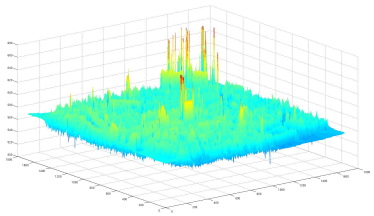


Figure 2.8: Mesh of the digital elevation model for München, generated from the stereo pair, and used for orthorectification. From Deutsches Zentrum für Luft- und Raumfahrt (DLR).

Thus, the aim of this thesis is to overcome the stated difficulties and to improve IKONOS image classification by fusion of multispectral information with digital elevation model data, as well as to apply this classification to perform change detection in multi-temporal images. Different totally automatic and unsupervised image processing approaches have been implemented and compared, as well as the impact on performance of the use of spectral, contour, regional, and texture features. A fully working software system demonstrating the obtained results has been implemented.

Chapter 3

State of the Art

A general introduction to Earth observation satellites with a focus on GeoEye's IKONOS and its imagery is provided in this section, as well as survey and evaluation of the state-of-the-art in classification and change detection based on the extraction of information from satellite multispectral images and digital elevation models, among others, and on the fusion of these data.

3.1 Earth Observation satellites and sensors

Artificial satellites can be categorized according to the variety of purposes with which they were deployed. Telecommunication satellites (Inmarsat, Intelsat, Iridium) provide data transmission, telephone calls and television broadcasting over great distances; navigation satellites enable localization on the Earth (Global Positioning System, GPS); meteorological satellites observe the formation and evolution of hurricanes and clouds (NOAA, Meteosat); military satellites survey conflict and strategic areas; and scientific satellites explore changes in vegetation, glaciers, and mineral deposits, and perform astronomical observations [82].

The launch in 1957 of the first Earth-orbiting artificial satellite, the Russian Sputnik-1, marked the start of what is known as the Space Age. In 1959, the American Explorer 6 took the first pictures of the Earth, and in 1960 with the also American Discoverer 14 the first recovery of the planet's pictures on film was possible [3]. In just half a century since then, a wealth of other Earth observation satellites has been launched into orbit in order to watch our planet, with varied objectives, capabilities, and orbit characteristics.

With aims ranging from weather and oceanography research to military, atmospheric pollution and ozone monitoring, more than 150 Earth-observation are currently in orbit, with sensors

measuring different regions of the electromagnetic spectrum, and maintaining either a geostationary orbit (orbiting in synchrony with the Earth, imaging a particular location), or a polar orbit (sun-synchronous, giving global coverage, revisiting the same spot at regular intervals, and going over the poles) [107].

Apart from IKONOS, other representative Earth Observation satellites allowing for photogrammetric elevation data generation are, among many others, SPOT (*Satellite pour l'Observation de la Terre*), ALOS (Advanced Land Observing Satellite) and GeoEye. These offer different spectral and radiometric resolutions: 20 m for multispectral and 2.5-10 m for panchromatic with SPOT, 1.64 m for multispectral and 0.41 m for panchromatic with GeoEye, and 2.5 m panchromatic with ALOS. The last of the latter series, ALOS2, has been recently reported lost by the Japanese Aerospace Exploration Agency, due to a sudden loss of power on 22nd April 2011, after 5 years of operation [51].

Other techniques for obtaining elevation data are LIDAR laser altimetry and SAR interferometry, generating digital elevation models of high accuracy (30-15 cm planimetric, and 15-10 cm altimetric) by, respectively, active light pulse transmission and radar, the latter offering the advantage of penetrating clouds and operating in conditions where other optical sensors cannot function, for example during the night. Both SAR and LIDAR have been widely used in the literature to complement multispectral imaging for classification and change detection.

3.2 IKONOS imagery

Initially named Commercial Remote Sensing System (CRSS), and then renamed to IKONOS (a variant of the Greek *eikon*, 'image'), this satellite was originally a project of Space Imaging, and is now operated by GeoEye. With a body size of 1.83x1.57 m and a mass of 728 kg, the first IKONOS (IKONOS-1) was launched in April 1999, failing to reach a stable orbit around the Earth, and finally reentering the Earth's atmosphere over the Pacific ocean [60]. The second IKONOS (IKONOS-2), identical to the first version, and in fact built in parallel, was successfully launched in September 1999 and is still in operation.

IKONOS constitutes a remarkable development in land observation, as it became the supplier at the time of the best spatial resolution publicly available from space: 1-meter panchromatic and 4-meter multispectral with blue (B), green (G), red (R) and near-infrared (NIR) bands. Together with DigitalGlobe's QuickBird, launched in 2001, it initiated a new management model for Earth observation based on privately funded satellites, at a time when most systems had been government-developed and funded [127]. Table 3.1 summarizes its main characteristics [24], [38], including temporal, spatial, spectral and radiometric resolutions, and Table 3.2 introduces its multispectral bandpass ranges.

Specification	Value
Orbit height	681 km
Orbit inclination	98.1, sun synchronous
Descending node time	10:30 a.m., local solar time
Field of regard	Up to 45 off nadir
Revisit time at mid-latitude	3 days at 60 elevation
	11 days at 72 elevation
	141 days at 89 elevation
Image sensors	Multispectral and panchromatic
Field of view	11 km at nadir
Radiometric resolution	11 bits
Panchromatic ground sample distance (GSD)	0.82 m at nadir
	1 m at 30 off nadir
Multispectral GSD	3.28 m at nadir
	4 m at 30 off nadir

Table 3.1: IKONOS specifications. From Dial et al. [24] and Grodecki and Dial [38]
.

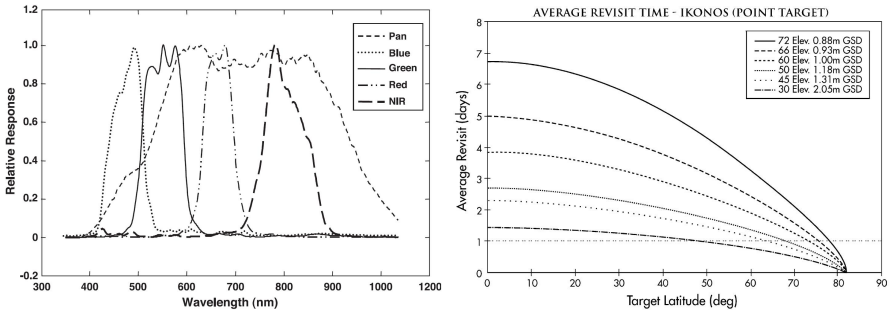
Specification	Value
B bandpass	445-516 nm
G bandpass	506-595 nm
R bandpass	632-698 nm
NIR bandpass	757-853 nm

Table 3.2: IKONOS multispectral bands. From Dial et al. [24] and Grodecki and Dial [38]
.

IKONOS orbit altitude, although stated as 681 km, varies effectively between 681 and 709 km. It circles the Earth with a period of 98 min and a repeat cycle of maximum 14 days in a sun synchronous orbit. IKONOS four multispectral bands are similar to the first four LANDSAT TM bands, with the NIR being modified to minimise atmospheric water absorption, and with varying GSD for multispectral and panchromatic information depending on the type of imaging, nadir or off nadir (3.28 to 4 m and 0.82 to 1 m, respectively), as shown in Table 3.1.

Unlike other satellites like SPOT (rolling side-to-side) and LANDSAT (imaging only at nadir), IKONOS can be rotated to different angles to image to the side, forward or aft of its position [38]). Regarding radiometric resolution, the satellite's dynamic range is 11 bits, data being compressed to 2.6 bits/pixel for transmission and then restored on the ground with minimal impact [85]. Fig. 3.1 (a) shows the IKONOS spectral response curves for panchromatic and multispectral bands.

Although newer satellites such as QuickBird, WorldView and GeoEye obtain higher resolutions than IKONOS in different respects, the latter represents a good compromise regarding



(a) (b)

Figure 3.1: (a) IKONOS spectral response curves. From Pagnutti et al. [85] (b) IKONOS average revisit times. From Dial et al. [24].

the trade-off between spatial and spectral resolution, offering high spatial resolution (0.6-4 m, like SPOT-5, CARTOSAT-1, ALOS, FORMOSAT-2, and the three aforementioned) with medium to high temporal resolution (4-16 days to <3 days), and medium spectral resolution (3-15 bands). In relation to temporal resolution, Fig. 3.1 (b) shows the average revisit times of the satellite depending on target latitude, elevation angles and GSD [24].

With its good compromise between spatial, spectral and temporal resolutions, IKONOS manages features that other satellites cannot resolve [85], such as ASTER, LANDSAT-7 and CBERS-2, with lower spatial resolutions, or others with lower dynamic ranges. As a consequence, information obtained from this satellite has been chosen as the only source of information used for classification and change detection of land cover, without resource to fusion with any additional data coming from other types of sensors, as it is common in the literature.

However, IKONOS imagery is still poorer than aerial images in terms of spatial resolution, and often presents weak reflectance around building boundaries, with a subsequent loss of geometric cues [103]. Data fusion can be used as a means of compensating for the weaknesses of the different data elements. In this work, multispectral and panchromatic IKONOS information have been fused with the derived digital elevation data obtained by stereo pair matching, with an aim to provide a classification of surface objects in urban and semi-urban areas in Germany and Asia, and to perform subsequent change detection analysis by incorporating the previously obtained classification.

3.3 Classification of satellite imagery

A multiplicity of methods and techniques have been developed since the mid-1990 for classification of satellite imagery, of which comprehensive categorization is of great difficulty, due

to the heterogeneity encompassed by the field. For example, in relation to the complexity of data, differences in terms of resolution and content are to be found, with satellites offering different radiometric, spatial, and temporal resolutions as discussed in the previous section. ASTER gives coverage of 14 spectral bands, with spatial resolutions between 15 m and 90 m; SPOT provides 4 spectral bands, with resolutions between 2.5 m and 20 m; GeoEye1 yields 4 multispectral bands and one panchromatic, achieving 1.64 m and 0.41 m spatial resolution; Orbview4, which failed to orbit in 2001, was equipped with over 200 bands, with the same spatial resolution as IKONOS [66].

Sensors may be of different nature, such as microwave and radar, thermal, multispectral or hyperspectral. Successful classification is thus necessarily tailored to the specific characteristics of the data to be exploited (as possibilities are different for example for multispectral and hyperspectral imagery), as well as to the determinant needs in terms of informational classes. The rich spectral information provided by narrow band hyperspectral imagery, for example, allows for the use of spectral libraries that can identify material types based on their spectral signatures. Fig. 3.2 illustrates the different data complexity and volume imaged by the 224 channels of AVIRIS in comparison to the 6 bands of Landsat TM. In addition to data complexity, other factors posing difficulties to the categorization of classification methods are the complexity of the model, which can be of varied nature, with different geometric and radiometric approaches, as well as sensor, object and scene models, and representation formalisms.

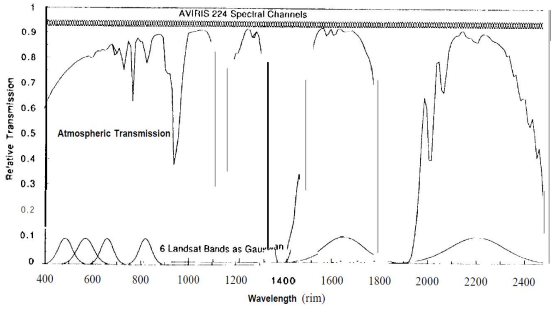


Figure 3.2: AVIRIS and Landsat spectral channels. From Simmonds [37].

Finally, the complexity of the strategy can present variations in the type of control, search, grouping and fusion of the data, in an attempt to offer a manageable categorization of building extraction methods. The taxonomy of classification methods offered here, however, uses another approach, based on usage of the sources of information relevant for the problem to be solved: spectral data, elevation data, or their fusion with each other or with other sources of information. In general terms, in time, a tendency towards mature holistic or general techniques is observable,

as well as some steps towards semi-automation, as well a growing use of combined cues. The next subsections present an overview of different approaches, based on height or spectral information, or their combined application.

3.3.1 Height information-based classification

Height information extraction has been much researched in the literature in its varied sources: photogrammetry, using stereo imagery, interferometry, based on radar imagery, and altimetry, from radar imagery. The obtained digital elevation model or DEM is the representation of surface heights above a datum [66] in a selected area, including vegetation canopies and buildings, unlike the digital surface model or DSM, where only ground surface elevation values are registered.

Photogrammetrically obtained DEMs are generated stereoscopically, namely, calculated form measurements from two or more images acquired from different positions, and with sufficient overlaps. Some satellites allowing for derivation of height data are ASTER, IKONOS, SPOT, EROS, Orbview, WorldView, ALOS and GeoEye [66]. On the other hand, altimetrically produced DEMs are obtained with LIDAR systems, based on return travel times of laser beams, and with high accuracies. Both methodologies for height information generation have been exploited in data fusion for classification [58] by use of techniques that are relevant for the objectives of this work, and thus studied here.

In land cover classification, research has mostly been focused on laser scanning, and applied for building extraction and reconstruction. LIDAR data has been studied for this purpose since the mid-1990s, with many of the approaches including a segmentation step, followed by building hypothesis generation, and subsequent checking against building models [30]. In 1995, for example, Weidner and Forstner [117] applied thresholding on normalized DEMs for building extraction and reconstruction with geometric models, parametric for rectangular-shaped targets, and prismatic, for more complex blocks. They differentiated buildings from vegetation by analysing the variance of the surface normals, and applied a minimum distance principle to a library of models for posterior reconstruction.

In 1999, Haala and Brenner [40] used ground plans for building area detection and roof slope hypothesis formulation by segmentation of the DEM and roof matching with RANSAC [29] for posterior reconstruction. More recently, Forlani et al. [30] also used LIDAR information in a rule-based scheme for classification and 3D reconstruction of buildings. The authors applied region growing for the identification of areas determined by steep edges, and performed rule-based classification based on building outlines and roof slopes. In 2007, Zingaretti et al. [84] used a tree-structured classification algorithm, Adaboost, to detect ground, vegetation and buildings from raw LIDAR data. Finally, in 2011, Dorninger and Nothegger [26] present a

method based on cloud points obtained from airborne laser scanning (ALS), in which a high density of points per square meter (over 20) allows for the identification of planar faces and the calculation of their intersection for building reconstruction.

Land cover classification and building extraction based on height information using solely on radiometric images has a number of remarkable difficulties, and indeed approaches in the literature often rely on additional data to compensate for the source derived pitfalls. Even when presently sophisticate techniques are used for DEM generation from stereo pairs, results are very much affected by the presence of shadows and occlusions. In this sense, LIDAR information is superior, as height data provides information on geometry and shape without being affected by illumination and texture [30]. Although resolution with optical images is increasing and thus allowing for better DEM generation, techniques used on LIDAR cloud points are still not directly applicable on radiometrically derived DEMs.

3.3.2 Spectral information-based classification

Different mathematical pattern recognition techniques, learning strategies, or decision tree-based methodologies have been applied in the literature for classification of land objects and land cover types using spectral information, as well as both supervised and unsupervised methods. Unsupervised approaches, undertaking grouping without foreknowledge of the properties of the classes, has frequently been performed with clustering methods such as k-means, ISODATA, or hierarchical clustering. These techniques, when not combined with other methods, present difficulties in the labelling of informational classes.

Many attempts in the literature based solely on spectral information make use of supervised approaches, relying on previous knowledge on the classes learned from user input, or suggesting methods that require setting of parameters, either by the user or tailored to the targeted areas to be processed. For example, in [92], a recursive user-based partitioning method is proposed, using spectral and geometric properties, and a comparison between pixel and object-based classification is given for the integration of results in the GIS of Rio de Janeiro in Brazil.

In the same line, in [64] a method for classification of typical land cover is given by use of a set of threshold values based on vegetation and water indexes and textural properties, in an object-oriented decision-rule approach, for the identification of areas such as forests, dry lands, rivers and shades from the four bands of SPOT5-XS images. In [88], an evaluation of different supervised classification algorithms for multispectral images is provided. Finally, in [112] a weighted pixel statistics method for archaeological land use analysis based on mean and variance values is given; the accuracy of the method is however not quantitatively evaluated, and it requires manual setting of weights by the user.

Another identifiable trend in multispectral information classification focuses on edge detec-

tion for the selection of boundaries that can identify objects in a scene. For example, [116] propose an automatic method for the extraction of buildings in panchromatic Quickbird images by applying unsupervised clustering with histogram peak selection, verification of building existence with corresponding shadows, and edge detection by Hough transform to build the building polygon descriptions. The authors report underdetection of buildings due to initial under and oversegmentations. In [96] a method in three phases is presented: edge detection with a multispectral gradient filter, unsupervised ISODATA clustering using the initial centroids from the preceding step, and merging of regions using mean and covariance calculations.

Texture analysis has been many times incorporated for the purpose of classification. In [115], application of both spectral and texture features is shown to increase classification accuracy with QuickBird imagery, using both a region-based and a pixel-based approach. The authors use mean-shift segmentation, a weighted minimum distance classifier, and multi-scale and multi-direction Gabor filter banks for textures. Accuracies are reported to range between 64.16% and 96.46% for the first approach, and between 20.81% and 98.81% for the second, depending on the classes.

In [126] a variety of texture features are calculated, and their performance is evaluated by using a finite mixture model expectation-maximization algorithm, with accuracies ranging between 82.16% and 98.66%. However, only one spectral and one texture feature are finally applied for results estimation. Similarly, texture features are incorporated in [4] for the monitoring of greenhouse areas with QuickBird and IKONOS imagery and applying Maximum Likelihood classification. The authors report no improvement of results between the use of only spectral properties and of both spectral and textural features, and expectedly overall superior results for the higher resolution Quickbird image-based tests.

Finally, two other major currents in supervised classification are the application of neural networks and, now being a trend, support vector machines (SVMs). A comparative study of different neural networks for multispectral images was given by [79], using multi-layer perceptron, binary diamond and radial basis function algorithms. In [111], a multi-layer perceptron is used with input data and/or k-means clustered input data for land cover classification, obtaining high accuracies with 2, 3 and 4 hidden neurons. Only three classes were used: vegetation, water and soil. Much research on SVMs in remote sensing application has been published this year; a very comprehensive review is given in [77], and a comparison of one-class versus v-SVM for the classification of roof materials with hyperspectral data is given in [11], with similar results for both methods.

3.4 Fusion

Fusion refers to the merging of different related data, such as different imaging or mapping of a certain area, or information coming from different sensors. It aims at obtaining a product improving those originated from single sensor measurements, and it is applied not only in remote sensing, but in a wide variety of fields. It can fuse multi-temporal, multi-sensor, and multi-resolution data, which are closely related to remote sensing applications. In this field, fusion is performed for example to obtain more accurate spatial resolutions, for example when multispectral satellite bands are improved by use of panchromatic bands. It is also common for the inclusion of imaging data to a variety of datasets in GIS.

Fusion can occur at different conceptual levels, and indeed three main categories of image data fusion are generally identified in the literature [114] for our purpose, namely, pixel-level, feature-level, and decision level fusion. When analysing practical cases, it is made clear that mostly a combination of these is normally applied, rather methods relying on only a single identifiable approach. Fusion can merge data from different sources; for example, research has been developed for the fusion of multispectral or hyperspectral images with LIDAR data and, in some few cases, with DEM data. Pansharpening of multispectral images is based on the fusion of multispectral and panchromatic images acquired by the same unit, and DEM generation itself is based on stereo pair information fusion.

Among the different approaches to fusion, pixel-level algorithms, working at the lowest level, offer the highest potential in terms of detection while also the highest computational costs [114]. Feature-level methods, on the other hand, require independent extraction of features and their subsequent combination in a decision space. At this level, model-driven algorithms have been developed, alternative to the traditional statistical data-driven models. Finally, decision-level methods are based on initial independent detection for each unit, and their subsequent combination by different decision types, either hard (boolean, by weighting of units, or confirming agreement), or soft (using Bayes or Dempster's rule, or fuzzy logic) [114]. Fig. 3.3 illustrates the different approaches to fusion.

In the present field of interest, pixel-level treatment often refers to the combination of input data with the application of feature reduction techniques; feature-level has to do with the extraction of texture or edge parameters that can be later further exploited. Finally, decision-level works at a higher level, combining the decisions from different units. As explained, however, mostly combinations of the previous methods are to be found in the literature, rather than pure approaches. The suitability of each one is necessarily to be evaluated depending on available input information, and needs in terms of output [118]. A variety of hybrid and multilevel approaches have been developed, of which an overview is provided in the the next section.

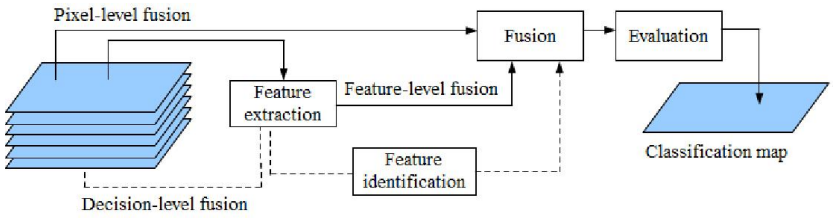


Figure 3.3: Levels of fusion: low or pixel-level, intermediate or feature-level and high or decision-level fusion. Adapted from [114].

3.4.1 Fusion for classification

In the last years, more efforts have been directed towards the possibilities for landscape classification offered by the fusion of multispectral and height information, rather than focusing on the properties of only one of these sources. This fusion is semantically classifiable in two abstract levels: in relation to the variety of employed sources, and also in terms of the classification introduced above, that is to say, with respect to the undertaken processing approach (pixel, feature, or decision-oriented, or a hybrid of these). As with classification based only on either multispectral or height information, most research has focused on building detection and extraction, due to its variety of applications, such as urban growth analysis and map updating. Most approaches in the literature, however, use only LIDAR derived height information, which is not world-wide available, with only a minority making use of stereo-pair based DEMs.

Fusion was applied in 1999 by Haala and Brenner [40], combining multispectral imagery and laser altimeter data in order to extract buildings, trees, and grass-covered areas with an unsupervised classification algorithm, and also laser data and 2D plan information to generate 3D reconstructions of buildings by use of polyhedral primitives. It required, though, user parameter setting, and the existence of ground maps for the area. In 2002, Kim and Muller [59] integrated LIDAR with multispectral IKONOS imagery. Later, Rottensteiner et al. [94] presented a hierarchical approach, making use of thresholding for cues like height, size, NDVI, and surface roughness for an iterative application of morphological operators with of user-specified size, and thus having parameter and user-dependence. The same authors achieved pixel-level fusion of the same sources later [93] by application of the Dempster-Shafer theory, using as features height variance, colour and surface roughness.

More recently, Sohn and Dowman [103] proposed a method for automatic extraction of buildings by data fusion of IKONOS imagery with LIDAR data, in an approach combining normalized difference vegetation values (NDVI) and height information for detection, followed by convex polygon partition of the detected building blobs, and both data- and model-driven

line generation for building outline detection. Reported results were of 88.2% completeness and 90.1% correctness for building detection, with the drawbacks of assuming only straight lines for building outlines, and computation time. Similarly, Lee et al. [61] use line segment matching on colour region segmentation results from initial LIDAR derived building regions.

Another recent geometric approach is given by Awrangjeb et al. [6], based on the identification of potential building boundary areas from LIDAR and hue information, and line segment search in the area, with reported average completeness of 78.27% and correctness of 88.90% for pixel-based evaluation in the test areas. A comprehensive performance evaluation, with different indexes for pixel and object-based accuracy estimation is provided by the authors. Similarly, Khoshelham et al. [58] present a comparative analysis of different methods for building detection with aerial imagery and laser data, pixel and object-based, with the following supervised classification methods: normalised DEM thresholding, AdaBoost, Dempster-Shafer, minimum distance, and maximum likelihood classification. The authors use morphological operations for the removal of small areas from the resulting classifications, which enhances the presentation of results, but also affects morphologically all areas and might hinder subsequent processing of the data, such as change detection procedures.

Other methods in the literature apply fusion of multispectral and height information not only for building detection tasks, but also to the purpose of general classification of land cover. These are generally supervised and user dependent, although some attempts at automation and user independence exist. For instance, Kim and Muller [59] propose an approach for classification of IKONOS multispectral images by fusion with DEM information, both from LIDAR and from stereo pairs, for the detection of buildings and trees. The method is hierarchical, using thresholding and splitting by NDVI calculation, and thus fusion takes place at decision level.

Gercek [34] proposes a supervised approach incorporating topographic data generated from the corresponding DEM to multispectral LANDSAT imaging. A 10% improvement in accuracy is obtained with the inclusion of elevation, slope and aspect for maximum likelihood classification into *agricultural land*, *shrub*, *herb* and *forest*. In [10], also different types of vegetation are automatically classified by decision-level fusion of multispectral maximum likelihood classification results and height information decision-rule classification results. It is to be pointed out that thresholds used in the decision tree are tailored to the specific characteristics of the rangeland vegetation in the study area.

3.4.2 Fusion for change detection

Change detection is a major remote sensing research pursuit [47], with four major objectives: detecting changes in remotely sensed images, estimating the geographic location of these changes, identifying their type, and finally quantifying their relevance. Many algorithms have been de-

veloped in the literature, of which a brief typology is given here; common drawbacks include the fact that most of them are supervised and require interaction with a user, with the associated costs, and that in many cases the type of change is not estimated, as they are based on change enhancement. The approach presented in this work is automatic and identifies different types of change, as described in section 4.

Selection of significant changes, such as construction and destruction of objects, motion of elements, or changes in shape of units is an inherent difficulty; it is not a trivial task to discern them from other types of change, such as sensor noise or illumination variations [91]. Furthermore, the classification of changes by semantic type is of further difficulty.

In general terms, change detection algorithms can be classified into four main types: algebraic, transformation-based, classification-based, and other hybrid approaches. The first group encompasses the most commonly used method, image differencing, based on the pixel-by-pixel subtraction of the images; a disadvantage is generally the need for thresholds [68]. Other algebraic approaches are image regression calculation and subtraction, image ratioing, and change vector analysis. Transformation-based approaches include those making use of Principal Component Analysis (PCA) to emphasize difference information, as recently in [28]. Finally, classification-based methods include unsupervised clustering and posterior detection of classes in the second image, or spectral-temporal analysis, and the technique here undertaken, post-classification comparison. This method has the advantage over others that it provides a complete matrix of change information, with the disadvantage that the quality of the change detection mask is highly dependent on the quality of the previous classification. For a complete review of methods for change detection, see [68] and [91].

Many methods, as explained, make use of supervised approaches in a diversity of ways, some of them undertaking data fusion. Indeed, some of the methods here mentioned make use of data fusion inherently, such as for example those applying PCA, or are based on previous fusion stages for classification. In [113], change detection based on a previous supervised maximum likelihood classification is undertaken, by inclusion of information from a geographical information system database. Another common approach for change detection is the use of neural networks with backpropagation to train the model, or of support vector machines, as in [80].

Recent approaches show a tendency towards the integration of extracted features with previous established methods. For example, Liang et al. [63] propose an image differencing method (IR-MAD or iteratively re-weighted multivariate alteration detection) combined with textural information; Dianat and Kasaei [25] incorporate spatial information to a conventional polynomial regression method; and Pagot and Pesaresi [86] propose the integration of the common use of SVMs for change detection with different levels of information; best results are obtained with the use of only panchromatic and pansharpened images, rather than with any combination using morphologically-derived structural information or the IR-MAD difference image approach.

Finally, the most recent research incorporates information fusion to change detection by applying multispectral and digital elevation information for the probability estimation of detected changes, by use of an SVM classifier and shape features, as in [14] and [13].

3.5 Remarks

The analysis of the literature for classification and change detection with satellite imagery that has been given in the preceding subsections reveals several facts. In spite of the fact that the application of several techniques shows promising results, still a breakthrough in terms of a generally applicable and fully automatic, unsupervised classification and change detection procedure for the monitoring of satellite imaged areas is needed.

Many drawbacks can be identified in the approaches presented in the literature. For example, the possible complexity of the scene is sometimes underestimated when simple building models are used, which only cover for a limited percentage of real scenarios. Also, when these building models are applied recursively for the recognition of more complex buildings, computational needs become an issue. Dependence on parameter settings or user input, for example with supervised methods, can also be considered as a drawback in this context, due to the associated costs when dealing with great volumes of data.

The type of sensor data is also an aspect to be considered, as not always a variety of sources is available, especially in cases of for example treaty monitoring in conflict areas, or nuclear facility surveillance. Laser scans, a preferred source in the literature for the extraction of height information, are generally not available in such cases, and also yet not economical. The alternative of computational stereo, as seen, is however problematic due to the existence of occlusions and shadows, which require a comprehensive method able to tackle these difficulties.

In spite of the complexity of the problem, and the drawbacks identified in the literature, this thesis aims at contributing to research in the pursuit of a generally applicable and fully automatic, unsupervised classification and change detection procedure for the monitoring of satellite imaged areas. Different totally automatic procedures, applicable on IKONOS imaging for the land cover analysis of buildings, grass, trees, ground, water bodies, and shadows have been designed, implemented and compared, and applied for change detection purposes. The benefits of fusion of spectral and height information are exploited in the different approaches. User and parameter dependence have been eliminated, as well as the use of simplified building models, and required sources are solely the widely available commercial IKONOS satellite stereo pairs, without any need for additional data.

Chapter 4

Methodology

This chapter describes the methodology and techniques that have been applied for the purposes stated in chapter 2. Firstly, the flow chart of the project is given in order to provide an overview of the process. Secondly, different aspects are commented, such as preprocessing steps, feature analysis and extraction employed approaches, and feature fusion implementation. This fusion has been attained by using three proposed methods, of different nature, and with different variants. Finally, results for two case scenarios are given in chapter 5.

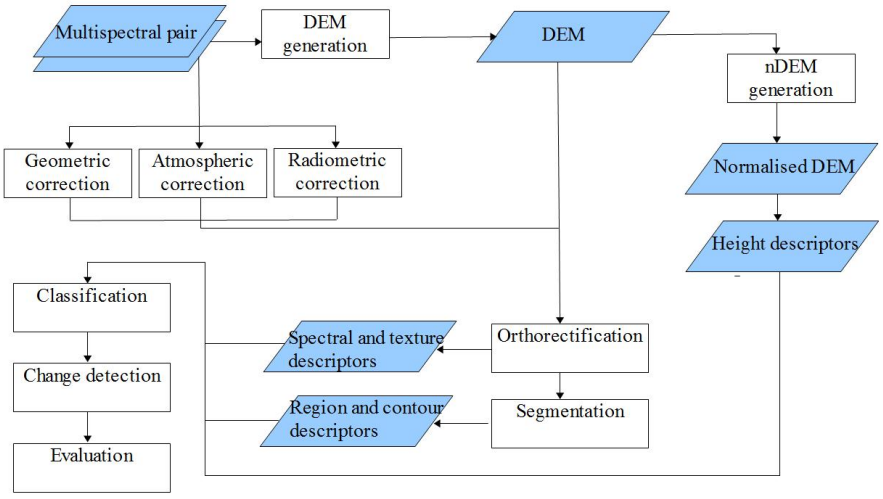


Figure 4.1: Overview of the processing flow. The different methods here presented implement fusion for classification and change detection at different levels

In general terms, the process can be described as follows. Firstly, from the raw stereo pairs

provided by IKONOS, normalised digital elevation model (nDEM) generation is undertaken. This is achieved through the generation of the epipolar images, and the computation of the disparity map by use of semi-global matching. The digital elevation model (DEM) is filled by interpolation, and from this DEM a normalised nDEM is generated, where ground height is referenced to zero, therefore providing information on building heights independent from the terrain [101]. From the nDEM, height descriptors can be obtained.

Secondly, orthorectification of the multispectral image is necessary, as explained in section 2, in order to georeference the data for further processing. Geometric, atmospheric and radiometric corrections are applied, and height information is used for the orthorectification. Pansharpening is applied, and from the product images, spectral and textural descriptors can be obtained. Then, segmentation into regions allows for the generation of region and contour descriptors. Finally, the processing chain includes the steps of classification (with the different methods here presented, implementing different levels of fusion), change detection, and final evaluation.

4.1 Dataset

The methods here introduced have been tested on two study areas. The first one comprises the city centre of München, Germany, imaged by IKONOS on 15th July 2005; the second area covers a region of Yongbyon in North Korea, captured by the satellite on 23rd February 2006 and 12th January 2010, respectively. The first set of stereo pair images for the city of München have been used to investigate the performance of the different classification methods here suggested; the other two sets of stereo pairs have been used to apply the best performing approach for classification in order to undertake change detection and analyse results. Thus, multi-temporal and multi-season analysis is here performed. Table 4.1 summarizes the parameters of the first dataset, for the first case scenario. Tables 4.2 and 4.3 give the parameters of the second and third datasets, for the second case scenario.

Specification	Image 1	Image 2
Scan Azimuth	180.06 degrees	0.06 degrees
Scan Direction	Reverse	Forward
Nominal Collection Azimuth	354.9351 degrees	263.2657 degrees
Nominal Collection Elevation	80.75830 degrees	86.55718 degrees
Sun Angle Azimuth	153.9184 degrees	154.0430 degrees
Sun Angle Elevation	61.50598 degrees	61.52575 degrees
Acquisition Date/Time	2005-07-15 10:28 GMT	2005-07-15 10:28 GMT
Percent Cloud Cover	0	0

Table 4.1: München dataset acquisition parameters

.

Specification	Image 1	Image 2
Scan Azimuth	180.0411263902 degrees	180.0411263902 degrees
Scan Direction	Reverse	Reverse
Nominal Collection Azimuth	7.3730 degrees	217.6230 degrees
Nominal Collection Elevation	65.84380 degrees	83.23550 degrees
Sun Angle Azimuth	157.8630 degrees	158.1090 degrees
Sun Angle Elevation	37.88420 degrees	37.94370 degrees
Acquisition Date/Time	2006-02-23 02:40 GMT	2006-02-23 02:41 GMT
Percent Cloud Cover	0	0

Table 4.2: Yongbyon dataset acquisition parameters for 2006

.

Specification	Image 1	Image 2
Scan Azimuth	359.9579269685 degrees	359.9579269685 degrees
Scan Direction	Forward	Forward
Nominal Collection Azimuth	32.2735 degrees	122.7043 degrees
Nominal Collection Elevation	61.48764 degrees	79.07821 degrees
Sun Angle Azimuth	160.1576 degrees	160.3700 degrees
Sun Angle Elevation	26.19710 degrees	26.25266 degrees
Acquisition Date/Time	2010-01-12 02:29 GMT	2010-01-12 02:30 GMT
Percent Cloud Cover	0	0

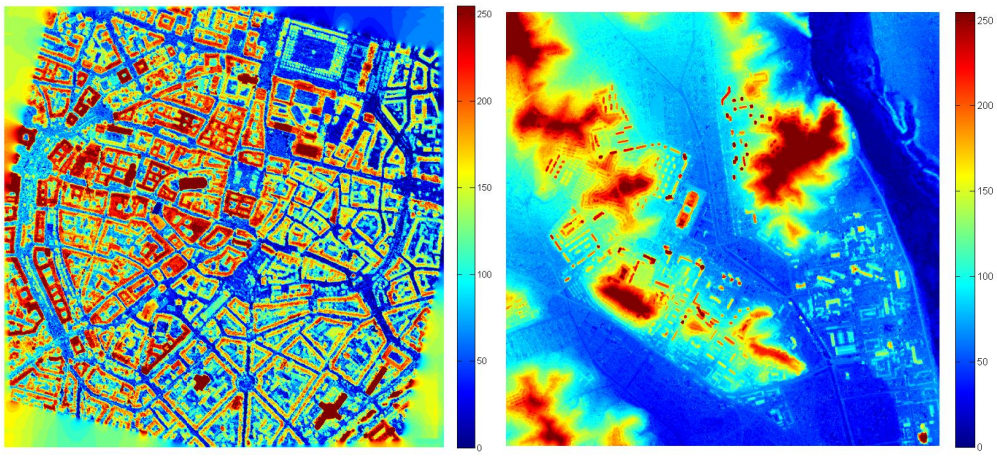
Table 4.3: Yongbyon dataset acquisition parameters for 2010

.

The pairs of IKONOS stereo images used here were provided by the satellite's managing company at the time of purchase, the Germany-based European Space Imaging (©EUSI under the EC/ESA GSC-DA) for the München case, and GeoEye for the Asian scenes, and they were obtained radiometrically corrected and not orthorectified. Six levels of products for IKONOS were available (and they are also presently with the current company GeoEye), with positional accuracies ranging from 15 m exclusive of terrain effects (for the non-orthorectified version) and from 50 m to 2 m in the CE90 standard (circular error at 90% confidence) in relation to the different options *Geo*, *Standard Ortho*, *Reference*, *Pro*, *Precision* and *Precision Plus*, respectively [48].

Users seeking metre-level accuracy need to acquire images in the range of *Precision* or *Precision Plus*, with five to ten times the cost of the *Geo* images, of lower positional reliability. Here, however, high accuracy results are sought by making use of *Geo* images, by joint application of Deutsches Luft- und Raumfahrt Zentrum developed preprocessing approaches and the presented classification and change detection methods.

Thus, the original *Geo* stereo pair images, acquired separately, as the product does not give a stereo option, have been used to generate a digital elevation model using the in-house developed



(a) (b)

Figure 4.2: IKONOS generated digital elevation models for München (2005, see Table 4.1 for parameters) and Yongbyon (2006, see Table 4.2).

semi-global matching algorithm (SGM) [43] [21] with a delta fill technique for interpolating data in hole areas generated by matching failures. The generation of the normalised DEM is done using a technique based on hierarchical filtering of the digital elevation model also developed at DLR by Arefi [5], after which orthorectification is undertaken. Extraction of height, spectral, texture, region and contour features is applied at different levels of the former process, as described in Fig. 4.1.

Figs. 4.2 (a) and (b) correspond to the elevation models for both test sites, with histogram DEM values of [510,674[and [43,116[, respectively, coherent with the altitudes of the areas, with a quantisation of [0,255]. As it can be observed, although both scenarios present a height range of over 100 m, the nature of the areas is very different, with little ground variation and high edification in the Munich case, and great ground variation and lower construction heights in the Asian area.

4.2 Feature analysis

Feature analysis of the input data has been focused on different properties: spectral characteristics, shape (both based on contours and on regions), and textural features. This section introduces the extraction that has been undertaken, and section 4.3 gives the details on how the fusion making use of these descriptors has been attained at different levels. Firstly, in relation to spectral characteristics, implemented vegetation indices, water body detection approaches and

shadow analysis methods are described. Secondly, shape has been analysed with region-based moment invariants and contour-based descriptors, such as solidity or compactness. Finally, textural characteristics have been inspected by use of the gray level coocurrence matrix and a set of derived features.

4.2.1 Spectral properties

Spectral responses of different types of land cover change are not static, and a one-to-one correspondence between semantic categories of objects and spectral signatures does not exist. As seen in section 2, different types of roof are characterised by different spectral curves that difficult their classification into a common class *buildings*. Both solar irradiance and the surface reflectance properties of the different land cover materials are variable, and thus the corresponding reflected amounts of light. Soil background, solar and view angle, atmosphere, moisture condition, material degradation and non-uniformity are some of the factors conditioning this variability; to reduce its impact rationing from different bands has been applied. Furthermore, spectra change timely (agricultural fields, for example, have growth cycles) and spatially (forests have changing density, for instance). However, some patterns can be identified for some classes of interest (water regions, shadows, and vegetation), and they have been exploited in this work.

Vegetation

Vegetation presents a characteristic spectral response that differentiates it from other types of land cover. While in the visible wavelengths most of the light is absorbed, in the near-infrared wavelengths it is transmitted and reflected. In Fig. 4.3, the spectra for wheat, dry and wet soil are compared, using LANDSAT's red and infrared band widths. Corresponding expected spectral values for each class are indicated with A-D.

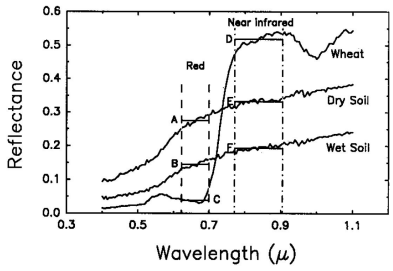


Figure 4.3: Reflectance spectra for wheat, dry and wet soil. From Jackson and Huete [50].

Three different approaches for vegetation enhancement have been implemented in this work:

the Normalised Difference Vegetation Index (NDVI) by Deering [23], the Soil Adjusted Vegetation Index by Huete [46], and the Vmap for IKONOS imagery described in the work of Cheng et al. [16]. The first two methods have then been extended to allow for automatic classification into vegetated and non-vegetated areas automatically and without user intervention.

- **The Normalised Difference Vegetation Index (NDVI)** enhances vegetation by exploiting the spectral behaviour of vegetation in the red and infrared bands. It is an improvent over the previously used ratio vegetation index $RVI = NIR/R$, which was problematic with red reflectances approaching zero. Here, values range from one to around zero, with small negative values being possible. The ratio is given in Eq. 4.1 and it is also formulated in relation to RVI. Mathematical proof of their informational equivalence is given in [87]; analysis of their different sensitivities is provided in [50].

$$NDVI = \frac{NIR - R}{NIR + R} = \frac{RVI - 1}{RVI + 1} \tag{4.1}$$

- **The Tasseled Cap transformation** was formulated by Kauth and Thomas [57], an extended version of which, tailored for IKONOS images, is used here. The former describes the variation in terms of the relationship between red and infrared band values of vegetated areas with different levels of cover. In Fig. 4.4, points A, B, C and D correspond, respectively, to dry soil, wet soil, fully vegetated and partially vegetated points, with the Perpendicular Vegetation Index (PVI) or distance from a point to the line A-B indicating vegetative growth.

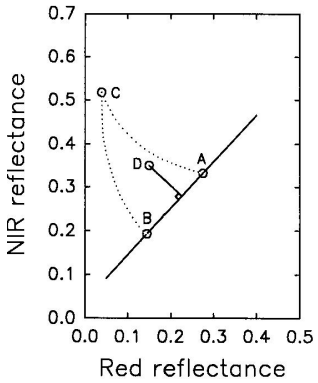


Figure 4.4: Tasseled Cap Transformation and PVI. From Jackson and Huete [50].

Cheng et al. [16] propose an approach to generate enhanced vegetation maps for IKONOS imagery, which has been implemented in this work. Based on the IKONOS Tasseled Cap

coefficients by Horne [44], four components (TCs) are calculated from the G,B,R and NIR bands, as shown in Eq. 4.2-4.5.

$$TC_1 = 0.326B + 0.509G + 0.560R + 0.567NIR \tag{4.2}$$

$$TC_2 = -0.311B - 0.356G - 0.325R + 0.819NIR \tag{4.3}$$

$$TC_3 = -0.612B - 0.312G + 0.722R - 0.081NIR \tag{4.4}$$

$$TC_4 = -0.650B + 0.719G - 0.243R - 0.031NIR \tag{4.5}$$

An enhanced vegetation index $VI_{TC} = aTC_2 - bTC_1 - cTC_3$ is proposed, calculated with a pseudo-Karhaunen-Loeve transform, from which $a = -1/4$, $b = 1/2$, $c = -1/4$. The resulting vegetation map suggested by the authors is based on thresholding above and below $\theta = 0$, as shown in Eq. 4.6.

$$V_{map} = \begin{cases} VI_{TC} & \text{where } VI_{TC} \geq \theta \\ 0 & \text{where } VI_{TC} < \theta \end{cases} \tag{4.6}$$

- **The Soil Adjusted Vegetation Index** was formulated by Huete [46] to minimize the impact of soil on the estimation of vegetation, which is present at medium and low densities. Thus, soil influence is greater in arid regions and early phases of growth [50], and also with yellow and red soils. Illustration of the impact of the soil on NDVI and RVI is observable in Fig. 4.5. Given A, a vegetated pixel in dry soil, and D, the same pixel after the soil is moistured, two other points are observable: C, the expected PVI index for equal amount of vegetation, and B, the expected NDVI. The solution provided by the author is to include in the ratio $L = l_1 + l_2$, the addition of the coordinate results of the extension of the line $A - D$ until its intersection with the soil line. The ratio is given in Eq. 4.7, L having been estimated by the author as optimal at $L = 0.5$, and using a multiplicative factor to maintain the NDVI bounds.

$$SAVI = \frac{NIR - R}{NIR + R + L}(1 + L) \tag{4.7}$$

In general terms, NDVI shows less sensitivity than RVI to view and solar angle changes; however, atmospheric turbidity decreases the indexes in general, and canopy architecture has an impact on their values as well due to the reflection direction of the light depending on types of leaves [50]. Further recent research on the topic includes [105], a comparative analysis of vegetation indexes with IKONOS, SPOT and ETM+ data, [52], and [53], a

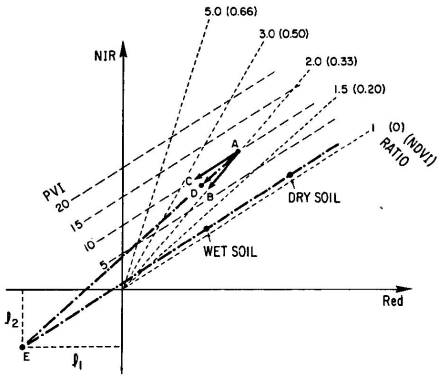


Figure 4.5: Red and near-infrared RVI, NDVI, and PVI isolines of equal vegetation amounts. From Jackson and Huete [50].

new vegetation index tailored for MODIS.

Water bodies

Water body extraction from remotely sensed imagery has been researched in the literature, and attempts to provide automatic methods have been proposed. However, most of these methods make use of middle-infrared bands available for example in Landsat ETM+ imagery or SPOT, not available with IKONOS, and are also based on water enhancement techniques that require subsequent manual thresholding, or supervised training and classification [99].

Different indices have been proposed for the enhancement of water in satellite images. NDVI, formulated as explained in the previous subsection for the highlighting of vegetation, can be also applied for the detection of water. While vegetation yields positive values tending to 1 because of its high reflectance in the near-infrared portion of the spectrum, water bodies are represented with negative values owing to the absorption of water [99]. Mcfeeters proposed the Normalized Difference Water Index (NDWI) [75] in order to exploit the low reflectance of water in the near-infrared and the behaviour of vegetation and soil in the green band. Thus, the former presents positive values, while soil and vegetation range from null to negative values. This index is given in Eq. 4.8.

$$NDWI = \frac{G - NIR}{G + NIR} \tag{4.8}$$

Based on the spectral relationship between ETM+ bands, the Water Reflection Index was formulated as $WRI = (G + R)/(NIR + MIR)$, incorporating the middle-infrared band. In the same line, Xu [123] developed the Modified $MNDVI = (G - MIR)/(G + MIR)$ proving

good results for the extraction of water in urban areas. These indexes were later applied by Shen [99] in a supervised Adaboost approach, and by Liu et al. [65] for supervised neural network classification. Other approaches also make use of the specular reflectance in the microwave wavelength of water that results in dark tones and backscattering in SAR images [65].

Few approaches exist in the literature for automatic extraction of water bodies based on a limited number of bands; in [27] comparison between supervised and unsupervised approaches using a hierarchical decision tree based on MIR and fixed-valued thresholded DEM information for water body extraction is reported, results being successful only for the former. In this work, a method has been devised to automatically derive water areas, based on enhancement methods, and without requiring any input from the user or manual thresholding, and without making use of any additional information aside from the four-band spectral information of IKONOS imagery.

Shadow areas

Shadow is one of the most problematic factors in optical satellite imagery processing, affecting treatment, in particular, in urban areas. Different shadow detection methods have been devised in the literature, based on statistical models, physical or geometrical properties of light propagation, and color space transformations [71]. Most of the statistical approaches make use of training for the learning of a classifier and require thresholding, which makes generalisation difficult and user dependency unavoidable. Tappen et al. [106], for example, use classifier training to identify patterns in an image and local evidence to discard unsuitable hypothesis. Wu and Tang [121] apply user input and Bayesian classification.

Most geometrical methods, on the other hand, are based on the geometric properties of shadows, and the fact that chromaticity information is generally not affected by changes in illumination [18]. Thus, shadowed areas can be identified by local analysis of neighbouring pixels with similar chromaticity information and confirmation with geometrical cues. These approaches are difficult to implement with satellite imagery, due to the attainable resolutions and the data volume for processing, specially in urban areas, and have been not applied in this work.

Regarding colour-space transformation methods, several approaches are to be mentioned. Polidorio et al. [90] detected shadows by focusing on two properties, low luminance and highly saturated blue-violet wavelength, by transforming the RGB composite into the HSI model and segmenting in the last two bands of the latter. Huang et al. [45] focused on other properties, large hue, low blue, and small difference between green and blue values, and suggested the application of different experimental threshold values.

Tsai [109] used transformations into HSI (hue, saturation, intensity), HSV (hue, satura-

tion, value), YC_bC_r (luma, blue-difference chroma, red-difference chroma), HCV (hue, chroma, value), and YIQ (luminance, hue, saturation) colour models, and obtained best results by thresholding the map generated by ratioing hue over intensity. The method by Tsai was further developed by Chung et al. [18] by using a successive thresholding scheme instead of global thresholding. As best results with Tsai's method were accomplished by the HSI colour model, it has been applied in this work as explained in section 4.3.

Another approach for shadow detection implemented here is an extension of the method proposed by Marchant and Onyago [72] [73], further developed by Makarau [71], and finally extended here for automatic and user independent extraction. As initially formulated by Marchant and Onyago [72], the physical process of light reflection and interpretation by a camera is given in Eq. 4.9, with C_I being the colour channel output, S_I the channel spectral sensitivity, ρ the surface reflectance, E the illumination, and λ the wavelength, with G_I being a gain factor defined by the camera parameters (integration time, aperture, electronics) and positioning angles between illumination, camera and surface.

$$C_I = G_I \int S_I(\lambda)\rho(\lambda)E(\lambda)d\lambda \tag{4.9}$$

An approximation to Eq. 4.9 is given in Eq. 4.10, assuming narrow passbands of the camera filter, so that they can be represented with impulse functions, with λ_{cI} the centre of each each channel I, and g_I dependent on G_I and the shape of the filter. Also, the relationship between a colour measured at an illumination m and at an illumination reference ref can be expressed as in Eq. 4.11.

$$C_I = g_I\rho(\lambda_{cI})E(\lambda_{cI}) \tag{4.10}$$

$$C_{Iref} = \frac{C_{Im}E_{ref}(\lambda_{cI})}{E_m(\lambda_{cI})} \tag{4.11}$$

Planck's spectral radiant exitance of a blackbody at temperature T per wavelength interval given in Eq. 4.12 can also be approximated using Wien's method as shown in Eq. 4.13, with λ being the wavelength and c_1 and c_2 constants.

$$M_{e\lambda} = c_1\lambda^{-5}[exp(c_2/T\lambda) - 1]^{-1} \tag{4.12}$$

$$M_{e\lambda} = c_1\lambda^{-5}exp(-c_2/T\lambda) \tag{4.13}$$

Using the explained colour scaling and exitance formulations, together with the use of band ratios defined by Barnard et al. [7] $r = C_R/C_B$ and $g = C_G/C_B$, the basis is set for the following

development. Applying scaling to the previous ratios, Eq. 4.14 is obtained.

$$r_{ref} = r_m/s_r \quad g_{ref} = g_m/s_g \tag{4.14}$$

where

$$s_r = \frac{E_{ref}(\lambda_{cB} E_m(\lambda_{cR}))}{E_{ref}(\lambda_{cR} E_m(\lambda_{cB}))} \quad s_g = \frac{E_{ref}(\lambda_{cB} E_m(\lambda_{cG}))}{E_{ref}(\lambda_{cG} E_m(\lambda_{cB}))} \tag{4.15}$$

Then, substituting $M_{e\lambda}$ for illuminant E from Eq. 4.13 into Eq. 4.15 we obtain Eq. 4.16.

$$\begin{aligned} s_r &= exp\left[c_2\left(\frac{1}{T_{ref}} - \frac{1}{T_m}\right)\left(\frac{1}{\lambda_{cR}} - \frac{1}{\lambda_{cB}}\right)\right] \\ s_g &= exp\left[c_2\left(\frac{1}{T_{ref}} - \frac{1}{T_m}\right)\left(\frac{1}{\lambda_{cG}} - \frac{1}{\lambda_{cB}}\right)\right] \end{aligned} \tag{4.16}$$

where

$$s_r = s_g^A \quad \text{and } A = \frac{1/\lambda_{cR} - 1/\lambda_{cB}}{1/\lambda_{cG} - 1/\lambda_{cB}} \tag{4.17}$$

Finally, from Eq. 4.14 and Eq. 4.17, Marchant and Onyago's initial thesis is derived, as shown in Eq. 4.18.

$$r_m = F g_m^A \quad \text{where } F = r_{ref}/g_{ref}^A \tag{4.18}$$

Marchant and Onyago extended their method in [73] by increasing the number of channels and defining the band ratios for any wavelength channel. Here, instead of Wien's approximation to Planck's formula, a new representation for the CIE daylight family is used. The reason for this is that, although the initial method gives good results for conventional cameras with a range 440-610 nm, there are divergences with the CIE daylight standard for high CCTs in the blue and infrared ranges. The chosen relationship is given in Eq. 4.19, of which Wien's approximation is in fact a case with $f = -c_2/T$, $u = 1/\lambda$, and $h = c_1\lambda^{-5}$.

$$E(\lambda, T) = h(\lambda) exp[u(\lambda) f(T)] \tag{4.19}$$

Given the illumination model expressed in Eq. 4.19, the authors derive an F function independent of T, that is to say, independent of any illumination changes coming from the same family, as follows. From Eq. 4.19, the exponent A is given by Eq. 4.20, being λ_c the camera filters centre frequencies.

$$A = \frac{u(\lambda_{cR}) - u(\lambda_{cB})}{u(\lambda_{cG}) - u(\lambda_{cB})} \tag{4.20}$$

Taking now band ratios of any wavelength, we have $y_\lambda = (C_\lambda)/(C_{\lambda_n})$, where λ_n is a normalizing wavelength, similarly to the blue band in the initial method. As g is cancelled out, the ratio is now independent of any camera or positioning factors, as expressed in Eq. 4.21. From Eq. 4.18 and Eq. 4.20, we obtain for any two band ratios λ_1 and λ_2 Eqs. 4.22 and 4.23.

$$y_\lambda = \frac{\rho(\lambda)E(\lambda, T)}{\rho(\lambda_n)E(\lambda_n, T)} \tag{4.21}$$

$$F_{12} = (y_\lambda)/y_{\lambda 2}^{A_{12}} \tag{4.22}$$

$$A_{12} = \frac{u(\lambda_1) - u(\lambda_n)}{u(\lambda_2) - u(\lambda_n)} \tag{4.23}$$

From Eqs. 4.19, 4.21, 4.22 and 4.23, Eqs. 4.24 and 4.25 are derived.

$$F_{12} = y_{\lambda_1}/y_{\lambda_2}^{A_{12}} \tag{4.24}$$

where

$$a_1 = \frac{\rho(\lambda_1)h(\lambda_1)}{\rho(\lambda_n)h(\lambda_n)} \quad \text{and } a_2 = \frac{\rho(\lambda_2)h(\lambda_2)}{\rho(\lambda_n)h(\lambda_n)} \tag{4.25}$$

Finally, fixing λ_2 and allowing λ_1 to have any values, the final formulation $F_\lambda = (y_\lambda)/y_{\lambda_2}^{A_\lambda}$ is achieved. Makarau [71] further extended the method by Marchant and Onyago for calculating invariant spectra of light reflected from surfaces explained until here to the field or remote sensing by applying their results in satellite image processing for the purpose of shadow enhancement. Following suggestions by the author, the method has been further expanded here for automatic detection of shadows.

4.2.2 Shape properties

Like spectral properties, shape is a fundamental unit of perception for object recognition, and thus a variety of shape analysis techniques exist. In broad terms, these can be classified into contour-oriented (representing shapes based on sampling of the contour of the studied region, or on the extraction of landmark points) or region-based (considering all the interior and boundary pixels of this region). Contour-based descriptors include features such as compactness, eccentricity and solidity, while region-based descriptors are generally based on moments, such as Hu's seven moment invariants. These descriptors have been applied in this work in order to achieve description of regions regardless of size, position and orientation, which makes them

invariant against transformations such as scale change, translation and rotation.

Hu's seven moment invariants

Moment-based invariants are the most commonly used region-based descriptors [15]. From the definition of regular moments, central and normalised central moments can be derived. Based on the second and third order moments, Hu formulated the seven invariants that have been applied here. Given the moment of order $(p+q)$ of the continuous image function $f(x,y)$ m_{pq}, the central moment of $f(x,y)$ is μ_{pq}. Their definitions are given in Eqs. 4.26 and 4.27, with the centroid of the region being $\bar{x} = m_{10}/m_{00}$ and $\bar{y} = m_{01} - m_{00}$. The normalised central moments are expressed in Eq. 4.28.

$$m_{pq} = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} x^p y^q f(x,y)\, dx\, dy \quad p,q = 0,1,2,... \tag{4.26}$$

$$\mu_{pq} = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (x-\bar{x})^p (y-\bar{y})^q f(x,y)\, dx\, dy \tag{4.27}$$

$$\eta_{pq} = \frac{\mu_{pq}}{\mu_{00}^{\gamma}} \quad \text{with} \quad \gamma = (p+q+2)/2, \quad p+q = 2,3,... \tag{4.28}$$

Hu's seven moment invariants are given in Eqs. 4.29-4.35. They are invariant to scaling, translation, and rotation, as explained. However, their non-orthogonal basis results in information redundancy, and the higher order of moments is also sensitive to noise. The use of powers results also in wide dynamic range values, which requires normalisation.

$$\phi_1 = \eta_{20} + \eta_{02} \tag{4.29}$$

$$\phi_2 = (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2 \tag{4.30}$$

$$\phi_3 = (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2 \tag{4.31}$$

$$\phi_4 = (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2 \tag{4.32}$$

$$\begin{aligned}\phi_5 = {} & (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} - \eta_{03})^2] \\ & + (3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2]\end{aligned} \tag{4.33}$$

$$\begin{aligned}\phi_6 = {} & (\eta_{20} - \eta_{02})[(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \\ & + 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03})\end{aligned} \tag{4.34}$$

$$\begin{aligned}\phi_7 = {} & (3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] \\ & + (3\eta_{21} - \eta_{30})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2]\end{aligned} \tag{4.35}$$

Contour-based descriptors

Given a set of vertices v_0, v_1, ..., v_{N-1} of a region represented polygonally by (x_0, y_0), (x_1, y_1), ..., (x_{N-1}, y_{N-1}), different shape features can be calculated, as in [122]:

- **Compactness** is defined in Eq. 4.36, A and P being the area and perimeter of the region, its values ranging from one (circular shape) to null (maximally elongated shape).

$$C = \frac{4\pi A}{P^2} \tag{4.36}$$

- **Eccentricity** represents the ratio of the distance between the foci of the region best fitting ellipse and its major axis lengths, with values varying from one (maximum elongation of the ellipse) to null (circular shape). The feature is defined in Eq. 4.37, with $(\bar{x}, \bar{y})$ the centroid of the object, and $\mu_{pq} = \Sigma\Sigma(x - \bar{x})^p(y - \bar{y})^q$ the (p, q) order central moment of the shape.

$$E = \frac{\mu_{20} + \mu_{02} - \sqrt{(\mu_{20} - \mu_{02})^2 + 4\mu_{11}^2}}{\mu_{20} + \mu_{02} + \sqrt{(\mu_{20} - \mu_{02})^2 + 4\mu_{11}^2}} \tag{4.37}$$

- **Solidity** stands for the ratio of the area of the polygon and the convex hull area of the polygon approximating the shape, with values ranging from one (convex shape) to null.

- **Asymmetry** is calculated as suggested by Huan et al. [122], where m and n are the major and the minor axis of the ellipse best fitting the area, as shown in Eq. 4.38, its values approaching one with the asymmetry of the region.

$$As = 1 - \frac{n}{m} \tag{4.38}$$

- **Rectangular fit** refers to the fit of the region area A to its bounding box area A_0, as defined in Eq. 4.39, with values approaching zero for a perfect fit [122].

$$As = 1 - \frac{A_0}{A} \tag{4.39}$$

- **Length to width ratio** is calculated using Eq. 4.40, where a and b are the bounding box length and width, A is the region area, and A_0 its bounding box area [122].

$$\gamma = \frac{a^2 + ((1 - f)b)^2}{A} \quad \text{where} \quad f = \frac{A}{A_0} \tag{4.40}$$

Other shape features include the Euler number, a topological indicator defined as the number of objects in the region minus the number of holes in those objects, and depending on the classification, the area of the shape. None of these two features have been used in this work, due to their particular variability, for example in terms of building shapes in rural and urban areas, as this work aims at providing a comprehensive and general method for classification and change detection.

4.2.3 Textural properties

As it is the case with spectral and shape features, textural properties can aid classification, as they constitute an important local descriptor. Two types of texture extraction have been applied here: the Gray Level Coocurrence matrix approach, allowing for the derivation of a set of established textural measures, and Gabor filtering, which has been applied for texture discrimination. Both approaches are described in the next subsections.

Gray Level Coocurrence Matrix

Among the variety of existing methods for texture analysis, one of the most well known statistical approaches for the extraction of information is the Gray Level Coocurrence Matrix (GLCM) proposed by Haralick [41]. This matrix mathematically represents spatial dependence in an image based on the orientation and distance between pixels. Briefly, the coocurrence matrix $C(i,j)$ is calculated by using a displacement vector $d_{xy} = (\delta x, \delta y)$ maximizing the statistical measure in each particular case, and studying all pairs separated a distance d_{xy} with gray levels i and j. After normalisation, from the GLCM matrix a variety of second-order statistics can be computed. Haralick originally suggested a set of 14 features [41]; the ones applied in this work, in addition to mean and dissimilarity measures, and following the author's original notation (with μ and σ for mean and standard deviation, respectively), are given in Eqs. 4.41-4.46.

$$\text{Angular Second Moment} = \sum_i \sum_j \{p(i,j)\}^2 \tag{4.41}$$

$$\text{Contrast} = \sum_{n=0}^{N_{g-1}} n^2 \{ \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p(i,j) \} \quad |i-j| = n \tag{4.42}$$

$$\text{Correlation} = \frac{\sum_i \sum_j (ij) p(i,j) - \mu_x \mu_y}{\sigma_x \sigma_y} \tag{4.43}$$

$$\text{Variance} = \sum_i \sum_j (i-\mu)^2 p(i,j) \tag{4.44}$$

$$\text{Inverse Difference Moment or Homogeneity} = \sum_i \sum_j \frac{p(i,j)}{1+(i-j)^2} \tag{4.45}$$

$$\text{Entropy} = -\sum_i \sum_j p(i,j) \log\{p(i,j)\} \tag{4.46}$$

Angular Second Moment is a measure of homogeneity, giving few gray tone transitions in homogeneous images; contrast measures the amount of local variation that is present by assigning weights that increase as the distance from the GLCM diagonal widens. Entropy is a measure of complexity, complex textures tending to have high entropy. Correlation analyses the dependency of neighbour pixels, and tends to be high when the scale of the texture is larger than the distance used. The Inverse Difference Moment represents the inverse of the contrast of the GLCM, as it is a measure of the amount of local uniformity. Finally, entropy is a measure of complexity, complex textures tending to have high entropy.

It is to be noted that the proposed texture features are functions of a distance and an angle and thus, as suggested by Haralick, the obtained angularly-dependent features are not used directly here. Instead, a rotation invariant average function is calculated for the features obtained with different angles. Also, a multi-scale approach has been applied, and the impact of the use of textural features in the particular case of IKONOS satellite imagery has been analysed.

4.2.4 Gabor features

Another common approach for texture extraction is the application of Gabor filters, which have an orientation and frequency that resembles that of the human vision system, and that have proven to be effective for texture discrimination. Indeed, cells in the visual cortex of mammalians can be modelled by Gabor functions, as proven by Daugman [22]. Mathematically, Gabor filters can be formulated by the multiplication of a harmonic function by a Gaussian, and they have two components, a real and an imaginary one, representing orthogonal directions. The specification of the complex, real and imaginary forms is given in Eqs. 4.47, 4.48, and 4.49.

$$g(x,y;\lambda,\theta,\psi,\sigma,\gamma) = \exp\left(-\frac{x'^2+\gamma^2 y'^2}{2\sigma^2}\right)\exp\left(i\left(2\pi\frac{x'}{\lambda}+\psi\right)\right) \tag{4.47}$$

$$g(x,y;\lambda,\theta,\psi,\sigma,\gamma) = \exp\left(-\frac{x'^2+\gamma^2 y'^2}{2\sigma^2}\right)\cos\left(2\pi\frac{x'}{\lambda}+\psi\right) \tag{4.48}$$

$$g(x,y;\lambda,\theta,\psi,\sigma,\gamma) = \exp\left(-\frac{x'^2+\gamma^2 y'^2}{2\sigma^2}\right)\sin\left(2\pi\frac{x'}{\lambda}+\psi\right) \tag{4.49}$$

$$\text{where} \quad x' = x\cos\theta + y\sin\theta \quad y' = -x\sin\theta + y\cos\theta \tag{4.50}$$

Notation is as follows: γ is the aspect ratio defining ellipticity, σ the sigma of the Gaussian, ψ the phase offset, θ the orientation of the normal to the Gabor function ridges, and λ the wavelength of the harmonic function. A filter bank consisting of Gabor filters with different orientations and periods of the sine component has been created for convolution and generation of the Gabor space for the given images, as it is further explained in section 4.3.

4.3 Feature fusion for classification

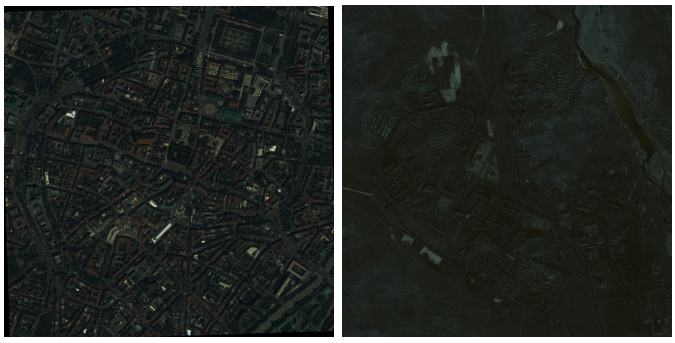
Three different methods have been implemented to obtain a good automatic, general and user independent classification with IKONOS images. The first method is based on the analysis of multispectral, height and other properties, as described in section 4.2, and on the extraction of relevant features for posterior coordinate transformation and unsupervised classification. The second method also exploits height and multispectral features, in a rule-based hierarchical classification, for the extraction of the desired classes *low vegetation*, *buildings*, *shadows*, *high vegetation*, *ground*, and *water bodies*. Finally, a hybrid model is presented, applying of a combination of rule-based hierarchical classification and coordinate transformation with clustering, and more complex conceptually, as it incorporates to the previous multispectral and height information also shape features and processing of the digital elevation model for an improved building identification, in a multilevel fusion approach, and using both object and pixel-oriented analysis.

4.3.1 Method 1

Method 1 is based on the pixel-based analysis of several properties of the image, and the extraction of spectral, textural and Gabor features. Relevant features are selected for the creation of a multidimensional composite that undergoes feature reduction by application of principal components analysis; cumulative variances are calculated to this effect. Finally, unsupervised clustering is used to obtain the final classification.

Feature extraction

Spectral features are obtained as explained in section 4.2.1. Firstly, Deering's Normalised Vegetation Index (NDVI) [23] values are calculated by exploiting the relative responses of the nearinfrared and red bands of the spectrum, as given in Eq. 4.1. This index is a vegetation enhancing method, and thus in this work this property has been further exploited by analysing the histogram of values obtained, and applying histogram thresholding to isolate the peak corresponding to vegetation, assuming Gaussian distributions. Figs. 4.6 (a) and (b) show the true colour images of the two datasets on which the method is demonstrated.



(a) (b)

Figure 4.6: München and Yongbyon true colour IKONOS images.

Histogram thresholding of the extracted NDVI values is obtained by moving average filter smoothing and Gaussian curve fitting. This has been attained in two ways: by localising the first Gaussians, corresponding to the targeted area, and conversely by subtracting the main Gaussian for the enhancement of the rest of distributions. Once a peak is found, a height equivalent to half of the peak is searched for on both sides. Then, the Gaussian is calculated according to Eqs. 4.51 and 4.52, where both minimum or maximum values can be chosen for the fit, $x(a)$ and $x(b)$ being the searched positions, and $x(h)$ the peak location. Figure 4.7 shows the plot of the NDVI values histogram on which calculations are undertaken.

$$\sigma = min||max(x(b) - x(h), x(h) - x(a)) * 2/2.35; \tag{4.51}$$

$$Gauss = h * \exp\left(\frac{-(x - x(h))^2}{2 * \sigma^2}\right) \tag{4.52}$$

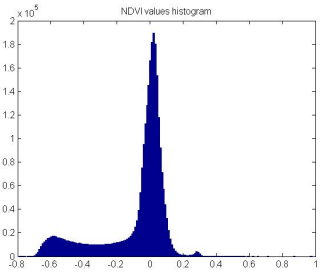


Figure 4.7: NDVI values histogram. Gaussian fitting is used to identify vegetation and water areas.

The NDVI values have also been used for the extraction of water bodies, as water enhancement also occurs, although in the opposite extreme of the histogram. In addition to the NDVI, the Normalized Difference Water Index (NDWI) has been used as formulated in 4.8, based on the relative responses of the green and the nearinfrared bands of the spectrum. For vegetation, also the Tasseled Cap Transformation variant by Cheng et al. [16] to generate enhanced vegetation maps for IKONOS images has been applied; please see section 4.2.1 for details. Finally, the Soil Adjusted Vegetation Index by Huete [46] has been calculated, as given in Eq. 4.7

A bank of Gabor wavelets has been created, for four periods of the sine component of the filter ($\pi/4, \pi/2, 3\pi/4$ and π) and six orientations ($0, \pi/6, \pi/3, \pi/2, 2\pi/3, 5\pi/6$) with $\sigma = 1$, giving thus 24 filters that have been applied on one of the spectral bands, generating 24 features per pixel. In the same way, textures have been analysed as given in Eqs. 4.41 to 4.46, giving 8 additional features per pixel. As suggested by Haralick [41], four angle-dependent features with steps $[1, 1]$, $[1, -1]$, $[-1, -1]$ and $[-1, 1]$ have been calculated in a multiscale fashion for two sizes of window (3 and 5, given the resolution) and averaged. Both Gabor and GLCM-based texture features have not generated positive results in the classification, as it is further discussed in section 5, and thus have been discarded.

Shadow areas have been studied by calculation of the F index by Marchant and Onyago [72] [73], and its extension by Makarau [71]. Detection of shadowed areas has been undertaken by histogram thresholding after moving average smoothing in order to optimally isolate the Gaussian representing the shadow areas for each image. Finally, the original multispectral bands, detected shadow, vegetation and water areas, digital elevation model, high elevation band, NDVI, F, SAVI, NDWI and VITC values have been used to generate a multidimensional composite of 13 bands, with redundant information, for subsequent processing and fusion of this information.

Dimensionality reduction

Principal Components Analysis (PCA) is a technique for analysis of correlated multivariable datasets based on statistics and algebraic matrix operations. Using coordinate rotation, it concentrates the information of correlated spectral bands or information composite into uncorrelated principal components, reducing thus the size of the initial dataset. In terms of multispectral or hyperspectral images, it is observable that correlation increases with narrower bands, making the storage of information inefficient.

PCA has been applied to the composite of multispectral bands, digital elevation model and extracted multispectral and height features explained above in order to produce uncorrelated output bands and to reduce the dimensionality of the dataset. Although used in this work for higher dimensional spaces, an illustration of a simplification for two dimensions is given in Fig.

4.8. Given an m-band composite as an m-dimensional dataset in an m-dimensional coordinate system, oblique ellipsoid clusters indicate correlation; their axes formulate a new orthogonal coordinate system representing the data with $n \leq m$ independent principal components (PCs).

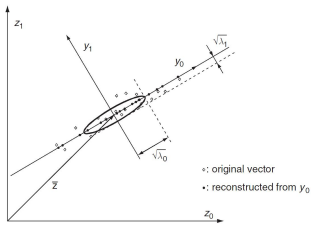


Figure 4.8: Two-dimensional principal component transformation, here used in higher dimensional spaces. From [42].

Mathematically, PCA is a G transformation diagonalizing the covariance matrix Σ_x of the m-band composite X to produce another n-PC object Y of reduced dimensionality $n \leq m$ with a diagonal covariance matrix Σ_y. As developed in [66], given an n-dimensional pixel vector $x_j = (x_{j1}, x_{j2}, ..., x_{jm})^T$, where $x_j \in x$, $j = 1, 2, ..., N$, m_x the mean vector of X, and N the total number of pixels in X, the covariance matrix of X is given in Eqs. 4.53 and 4.54.

$$\Sigma_x = \varepsilon\{(x - m_x)(x - m_x)^T\} \approx \frac{1}{N-1}\sum_{j=1}^{N}(x_j - m_x)(x_j - m_x)^T \tag{4.53}$$

$$m_x = \varepsilon x = \frac{1}{N-1}\sum_{j=1}^{N} x_j \tag{4.54}$$

Similarly, given the transformation G that diagonalizes Σ_x, $y = Gx$, y ($y_j \in y$, $j = 1, 2, ..., N$), m_y the mean vector of Y, the covariance matrix of Y is given in Eq. 4.55.

$$\Sigma_y = \varepsilon\{(y - m_y)(y - m_y)^T\} \tag{4.55}$$

Thus, $m_y = \varepsilon y = \varepsilon Gx = G\varepsilon x = Gm_x$, from which $\Sigma_y = \varepsilon(Gx - Gm_x)(Gx - Gm_x)^T = G\varepsilon\{(x - m_x)(x - m_x)^T\}G^T = G\Sigma_x G^T$. Then, G is the nxm transposed matrix of the eigenvectors of Σ_x, and Σ_y is a diagonal matrix with the eigenvalues of Σ_x, as shown in Eq. 4.56 [66].

$$G = \begin{pmatrix} g_{11} & g_{12} & \cdots & g_{1m} \\ g_{21} & g_{22} & \cdots & g_{2m} \\ \vdots & \vdots & \vdots & \vdots \\ g_{n1} & g_{n2} & \cdots & g_{nm} \end{pmatrix} = \begin{pmatrix} g_1^T \\ g_2^T \\ \vdots \\ g_n^T \end{pmatrix} \quad \text{and} \quad \Sigma_y = \begin{pmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix} \tag{4.56}$$

The eigenvalues λ_i indicate the variance of the PC_i object, and thus its informational content, determined by $\lambda_1 > \lambda_2 > ... > \lambda_n$ [66]. Given an identity matrix I of dimension m, the eigenvalues of Σ_x from with the eigenvector matrix G is calculated can be obtained with $|\Sigma_x - \lambda I|$, the eigenvectors of Σ_x being g ($g \in G$) satisfying $(\Sigma_x - \lambda I)g = 0$ or $\Sigma_x g = \lambda g$. Finally, each obtained PC constitutes a linear combination of the original composite layers, as formulated in Eq. 4.57:

$$PC_i = g_i^T X = \sum_{k=1}^{m} g_{ik} Band_k \tag{4.57}$$

Eigenvalues, variance and eigenvectors have been calculated for each obtained principal component for each of the studied composites. Table 4.4 shows, for the München dataset, total variance and cumulative variance explained by these PCs, indicating a remarkable 43.3410% for the first component, and an account by the first 5 PCs of 98.5779% of the total variance. These 5 PCs have been thus selected for processing in the next steps of the methods, thus achieving feature reduction, with the consequent computation time limitation, and feature fusion. Fig. 4.9 shows the calculated PC layers.

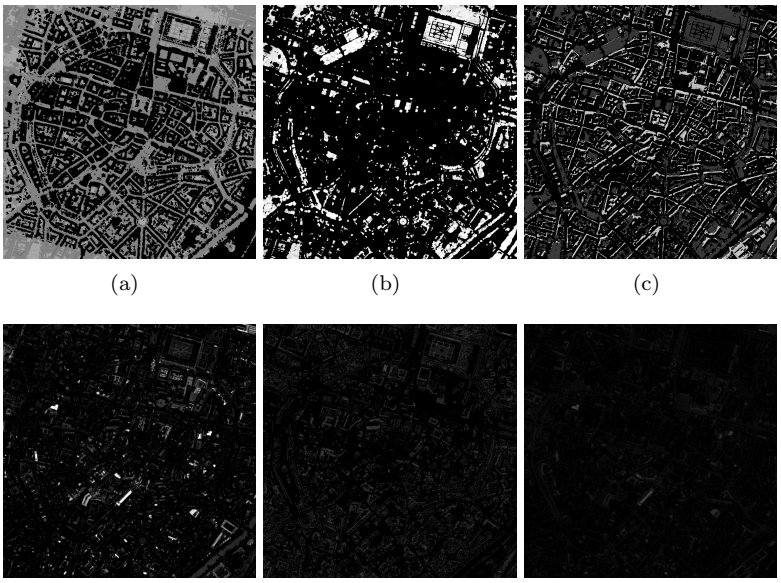
PCs	Variance %	Cumulative
PC1	43.3410	43.3410
PC2	32.7771	76.1181
PC3	17.8294	93.9475
PC4	3.1337	97.0812
PC5	1.4967	98.5779
PC6	0.9239	99.5017

Table 4.4: München dataset. Calculation of variances and cumulative variance percentages for the first six PCs

.

Unsupervised clustering

Once the principal components covering a high cumulative variance have been selected, unsupervised clustering is applied. To this aim, the k-means method has been chosen, with $k = 6$



(a) (b) (c)

(d) (e) (f)

Figure 4.9: München dataset. Calculated Principal Components.

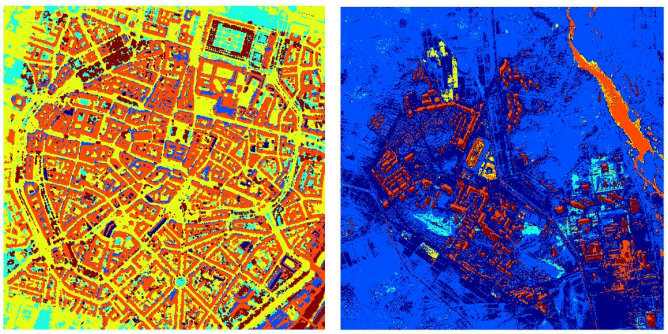
for the 6 targeted semantic categories. Briefly, this technique has as an objective the partition of n observations into k clusters in which each of these observations is assigned to the cluster with the nearest mean. Mathematically, given $(x_1, x_2, ..., x_n)$, where x is a vector of dimension d, the targeted partition into k sets ($k \leq n$) is that minimising the internal sum of squares for each cluster $S = S_1, S_2, ..., S_k$, with μ_i being the mean of S_i, as stated in Eq. 4.58.

$$\arg\min_{\mathbf{S}} \sum_{i=1}^{k} \sum_{\mathbf{x}_j \in S_i} \|\mathbf{x}_j - \boldsymbol{\mu}_i\|^2 \tag{4.58}$$

Squared Euclidean distances and random initialisation have been chosen, with an iterative refinement technique that leads the algorithm to convergence when the assignments no longer change. Two phases are alternated: an assignment phase, assigning each observation to the closest cluster mean, and an update phase, recalculating the new cluster centroids after assignation. This is given in Eqs. 4.59 and 4.60.

$$S_i^{(t)} = \left\{ \mathbf{x}_j : \left\|\mathbf{x}_j - \mathbf{m}_i^{(t)}\right\| \leq \left\|\mathbf{x}_j - \mathbf{m}_{i^*}^{(t)}\right\| \text{ for all } i^* = 1, \ldots, k \right\} \tag{4.59}$$

$$\mathbf{m}_i^{(t+1)} = \frac{1}{|S_i^{(t)}|} \sum_{\mathbf{x}_j \in S_i^{(t)}} \mathbf{x}_j \tag{4.60}$$



(a) (b)

Figure 4.10: Method 1 classification results. (a) München (2005) (b) Yongbyon (2006).

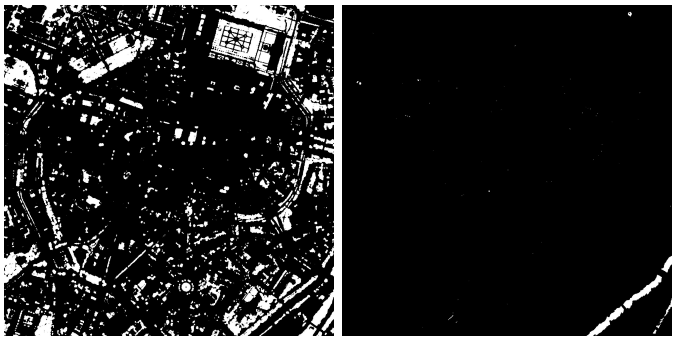
Finally, the obtained classifications are shown in Figs. 4.10 (a) and (b). For the München dataset, low vegetation is shown in green, and high vegetation in dark red; water and ground are confused, and thus labelled with the same colour, yellow; shadows appear in blue; finally, buildings are split into two classes, shown in dark blue and orange, for high reflectivity and normal reflectivity surfaces, respectively. Areas in the four corners correspond to undefined values of the digital elevation model, and are not considered part of the obtained classification. Results are quantified and discussed in chapter 5.

For the Yongbyon dataset, results are similar. Vegetation is shown in cyan, buildings in red, and water areas in yellow; the river, of different characteristics, is merged with shadows and, as the clustering looks for six classes, ground is then split in two, both represented in darker shades of blue.

4.3.2 Method 2

Method 2 makes use of the techniques presented in Method 1 by integration in a hierarchical decision tree. Both height and multispectral features are also exploited here for the extraction of the desired classes, in a rule-based hierarchical classification. To this aim, classes are identified in order, according to separability and index segmentation overlapping, as follows. Firstly, Deering's Normalised Vegetation Index (NDVI) [23] values are calculated as described in Method 1, and automatic histogram thresholding is used to isolate the distributions for vegetation and water bodies. Fig. 4.11 shows the segmented vegetation and water areas. The former group is later fused in the decision tree with height information, in order to distinguish between high vegetation (trees) and low vegetation (grass and bushes).

Then, shadow is segmented by applying the technique presented in Method 1, F value calculation as formulated by Marchant and Onyago [72] [73] and Makarau [71]. Again, the



(a) (b)

Figure 4.11: München dataset. Segmented vegetation areas and water bodies.

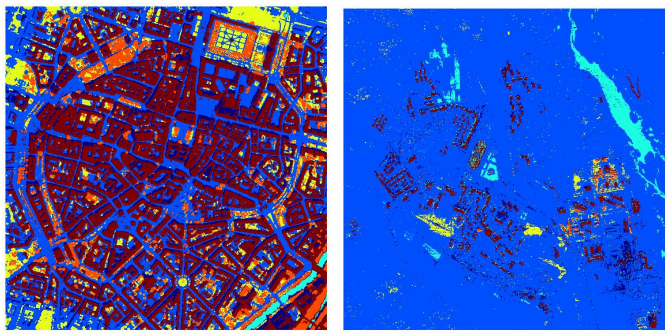
histogram of values is plotted, smoothed with a moving average filter, and thresholded by identifying the Gaussian curve corresponding to shadowed areas. Otsu's method [83], an image threshold technique minimizing the intraclass variance of the resulting binary partition, is sometimes used in the literature, but proves not to be adequate when dealing with urban area IKONOS images due to the presence of semishadowed regions. The decision-based approach can handle the fact that F values enhance not only shadows, but also water and vegetation, by initially assigning labels to the latter two and then using the F results in the non-classified areas. Finally, a knowledge-based decision rule is introduced for the detection of buildings, based on the thresholding of the normalised digital elevation model above a certain height.

The obtained classifications are shown in Figs. 4.12 (a) and (b). For the München dataset, low vegetation is shown in yellow, and high vegetation in orange. An improvement in relation to the previous method is attained, as water and ground are not confused, and shown in pale blue and blue; shadows appear in dark blue; finally, buildings are correctly identified in dark red. As before, areas in the four corners correspond to undefined values of the digital elevation model, and are not considered part of the obtained classification. Results are discussed in chapter 5.

For the Yongbyon dataset, results are similar, and again better than those obtained with Method 1. Classes are shown with the same code as that of München, and water areas, shadows, high and low vegetation, ground and buildings are correctly identified, with only some small confusion of some pixels of a highly reflective building roof with water.

4.3.3 Method 3

Having analysed the advantages and disadvantages of the previous methods, and their faults, a hybrid model is presented. It applies a combination of Method 1 and Method 2, with some additional improvements, integrating rule-based hierarchical classification with principal com-



(a) (b)

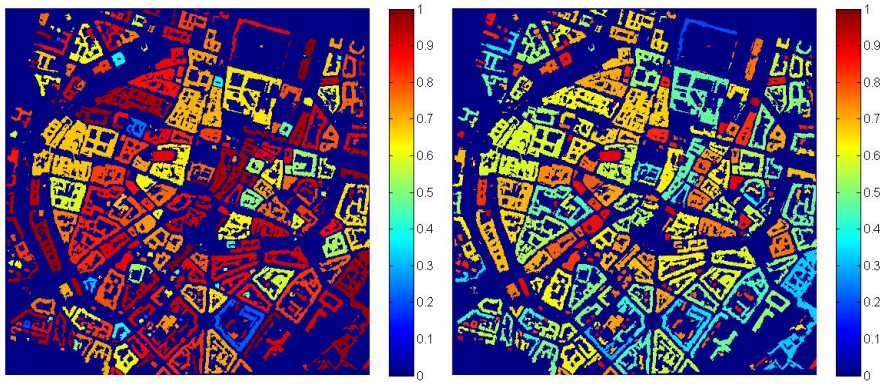
Figure 4.12: Method 2 classification results. (a) München (2005) (b) Yongbyon (2006).

ponents analysis feature reduction and subsequent unsupervised clustering. It is also more complex than the previous approaches, as it incorporates not only multispectral and height information but also shape features, as well as further processing of the digital elevation model for an improved building identification, applying a multilevel fusion approach using both object and pixel-oriented analysis.

The method uses the techniques explained in Method 2 for the segmentation of vegetation, water and shadows. Then, a composite is selected with spectral, height, and shape features, without the inclusion of multispectral features such as NDVI or F, to avoid redundancy. In particular, the R,G,B and NIR bands, together with the Hue band of the HSV transform, the digital elevation model, and the shape features described in Eqs. 4.36 to 4.40 have been included. In relation to region features, all Hu moments have been calculated, but only the first has been applied due to separability reasons. Regarding contour features, their effective application here proves their usefulness for building detection and classification. Figs. 4.13 (a) and (b) show two of the contour-based calculated values.

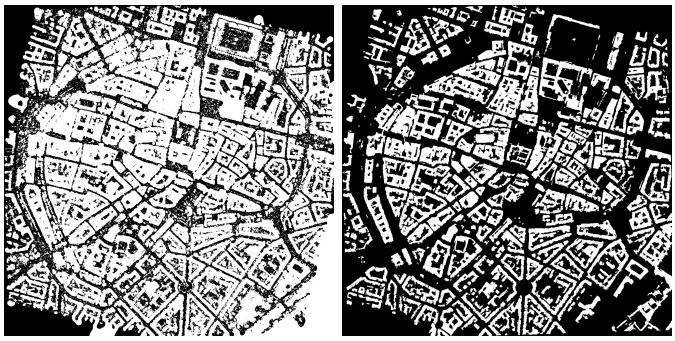
For the processing of the normalised DEM, mean-shift segmentation has been applied on the München RGB bands by using the system by Georgescu and Christoudias [76] [17] [20]. A segmentation with a high number of uniform areas has been aimed for. Then, every coloured area overlapping with the blobs of the nDEM over a considerable percentage (here set to 60%) has been incorporated to the a building map. With this, the need of thresholding the normalised digital elevation model disappears, as now every point in the height model presenting some elevation can be considered.

The growth of the building areas is controlled by the coloured image; to prevent cases of overgrowth, a small dilation with a disk or radius 1 of the original nDEM blobs has been applied as a limiting factor. Areas having exceeded the boundaries defined by the dilation are



(a) (b)

Figure 4.13: München dataset. Excentricity and solidity values for building areas.

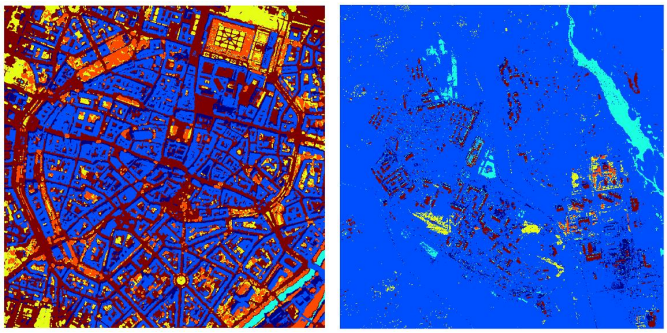


(a) (b)

Figure 4.14: Normalised dem above zero values and derived building map.

eliminated. An object-based approach is applied to calculate the previously explained region areas with the obtained map, while the main approach is still pixel-based. Equally, fusion is thus used at decision level, with the initial segmentation phase, at feature level and at pixel level, with the composite configuration merging feature and pixel values, the following principal components analysis, and the final clustering. Fig. 4.14 shows the obtained building map and the initial nDEM from which it has been generated.

The obtained classifications are given in Figs. 4.15 (a) and (b). For the München dataset, low vegetation is shown in yellow; high vegetation in orange; water in pale blue; ground in brown; and finally shadows in dark blue. An improvement in relation to the previous method is attained, as buildings are correctly identified, and the impact of the imprecise borders of the digital elevation model is minimised. Furthermore, no knowledge-based threshold has to



(a) (b)

Figure 4.15: Method 3 classification results for München (2005) and Yongbyon (2006).

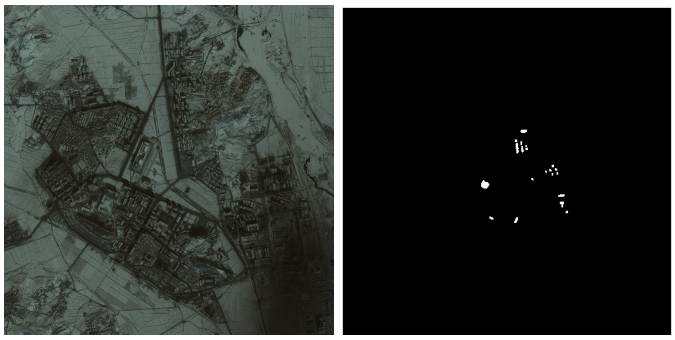
be set, and low buildings are better identified. As before, areas in the four corners correspond to undefined values of the digital elevation model, and are not considered part of the obtained classification. Results are quantified and discussed in chapter 5.

For the Yongbyon dataset, results are again similar. Classes are shown with the same code as that of München, and water areas, shadows, high and low vegetation, ground and buildings are correctly identified, with only some small confusion, as with Method 2, of some pixels of a highly reflective building roof with water.

4.4 Change detection

As explained in section 1, this work tries to demonstrate how the fusion of multispectral and digital elevation model information applied for the generation of an accurate land cover classification can be further exploited to undertake change detection. Some of the common drawbacks of traditional change detection approaches, as described in 3.4.2, include their supervised nature, the requirement of interaction with a user, and the fact that the type of change is generally not estimated, as most techniques are based on change enhancement. Here, the best performing classification approach Method 3 is applied to automatically perform change detection.

Advantages of applying the previously obtained classification are multiple. With common image differencing change detection, many artefacts are to be found, with different causes. For example, virtual detected changes are generated by precision errors in the height computation, varying levels of vegetation growth, and by erroneous interpolation over shadowed areas. Different approaches have been suggested in the literature, such as the elimination of vegetation areas, the elimination of shadows or the subtraction of interpolated points in the normalised digital elevation model.



(a) (b)

Figure 4.16: (a) Yongbyon dataset (2010) (b) Detection results for positive changes in the central area of image (a).

In the approach presented here, the proper selection of buildings in the two datasets that are compared is implicit in the previous classification, and thus the information included in the comparison is the result of the fusion with multispectral information, instead of that traditionally used and only based on height information. Thus, positive and negative changes in buildings have been investigated by subtraction of the two building maps obtained via the application of classification Method 3. Refinement of the borders of buildings has been undertaken by applying contour masking. This approach is thus fast, automatic, and provides promising detection of changes. Fig. 4.16 (a) shows the true colour image of the Yongbyon dataset for 2010, and 4.16 (b) presents the results of the explained approach for the central area of image (a). Results are quantified and evaluated in chapter 5.

Chapter 5

Results

Although a variety of methods have been proposed, there is currently an absence of standards and lack of uniform evaluation systems for performance assessment of classification results [95]. Building detection published methods, for example, often lack quantitative result evaluation or present only a pair of evaluation indices [6].

5.1 Evaluation measures

In order to evaluate classification results, twelve pixel-based evaluation indices have been chosen. Other evaluation metrics, such as object-based and geometrical approaches, have been discarded due to the nature of the objectives here pursued and that of the imagery being used. Firstly, objects such as buildings can vary in shape from clearly detached rectangular units to intricate asymmetrical urban conglomerates, blurring the notion of object to be applied. The same applies for vegetation areas, where it is difficult to decide which assemblies of trees can be considered as an object. Secondly, available ground truth, manually delineated, would bias the accuracies defined by geometrical indices, and thus their use has also been discarded. For the evaluation of change detection results, with a focus on buildings, two object-based measures have also been used.

Pixel-based performance evaluation indices here applied are objective metrics based on confusion matrices obtained by comparison between the obtained classifications and generated ground truth. Each confusion matrix CM_{ij} indicates the number of elements in class i labelled by each classification method as j. Ideally, a close to perfect classification would concentrate all elements in the diagonal of the matrix, with the rest of entries being close to zero. Confusion matrices obtained for each of the scenarios analysed, for each of the classification methods, are given in Section 5.2.

Thus, chosen evaluation metrics are the following: *completeness* or *matched overlay* C_{mp}, *correctness* C_{rp}, *branching factor* B_f, *miss factor* M_f, *quality* Q, *area omission error* A_o, and *area commission error* A_c, as suggested in [6]. Computation has been undertaken as formulated in Eqs. 5.1-5.5, as in [103], and in Eqs. 5.5 and 5.6, as in [104]. The terms TP, FP, FN and FP refer to true positives, false positives, false negatives and false positives for a two-class comparison; as more than two classes are considered in this work, an approach one-versus-all has been applied here.

$$C_{mp} = 100 * TP/(TP + FN) \tag{5.1}$$

$$C_{rp} = 100 * TP/(TP + FP) \tag{5.2}$$

$$B_f = FP/TP \tag{5.3}$$

$$M_f = FN/TP \tag{5.4}$$

$$Q_p = TP/(TP + FN + FP) A_o = 100 * FN/(TP + FN) \tag{5.5}$$

$$A_c = 100 * FP/(TP + FP) \tag{5.6}$$

Generally known *sensitivity* (*true positive rate* or *recall*) corresponds to C_{mp}, (Eq. 5.1), *precision* (*positive predictive value*) refers to C_{rp} (Eq. 5.5), and finally *false discovery rate* corresponds to A_c (Eq. 5.6). As above, a one-versus-all approach has been applied for the estimation. Finally, the well known Kappa statistic has been estimated, as well as the traditional *accuracy* A and Helden's *mean accuracy* H_m, as in Eqs. 5.7 and 5.8. [74].

$$A = (TP + TN)/(P + N) \tag{5.7}$$

$$H_m = 2 * TP/(TP + FN + P) \tag{5.8}$$

The better the classification method, the higher the C_{mp} and C_{rp} values, and the lower the A_c and A_o. Q gives a compromise between C_{mp} and C_{rp} [74]. Also, H_m should be high, as well as the rest of traditional statistics. The use of the proposed indexes aims at offering a comprehensive tool for the analysis of classification results, with an extensive evaluation of the process results, and not allowing for the biased assessment that using only few indexes might permit [104]. The combination of indexes here provided aims at providing insight into inaccuracies from several perspectives, for a better understanding and analysis of results.

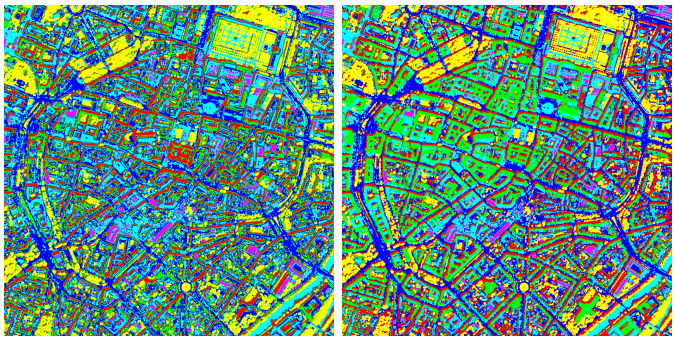
5.2 Classification results

In this section, classification results for each of the three implemented methods described in chapter 4 are evaluated with the assessment statistics given in section 5.1. For each of them, in turn, overall accuracy and kappa coefficient are given, followed by the confusion matrices containing all classes (one in pixels, the other in percentages), and the calculated evaluation measures, distributed in three tables: one for omission and commission errors (in pixels and in percentages), another for user's and producer's accuracies (in pixels and percentages), and finally another for the rest of statistics (C_mp, C_rp, B_f, M_f, Q_p and H_m, in percentages). As it is shown, the three implemented methods obtain good assessment figures, each of them improving the previous one, and thus the highest scores being obtained with the proposed hybrid Method 3.

Classification statistics have been computed for the München dataset, and change detection statistics have been estimated for the Yongbyon images, as two sets are available, making multi-temporal analysis possible. Results attained with the classification algorithms are positively high; however, it must be pointed out, only central areas of the regions have been considered in the ground truth. The reason for this is that reference ground truth can only be generated manually, and edges of buildings, vegetation, shadows and water appear blurry in the input images. This is a consequence of all that has been explained in chapter 2: the resolution limits of the images, multiple-cause errors introduced in the stereo-generated digital elevation model, and further propagation of these errors in subsequent processing stages.

Thus, in order not to introduce error by manually erroneously deciding on unclear edges, only central and clearly classifiable areas have been chosen. Finally, it must be pointed out, the clustering method that has been used can converge slightly differently in different runs of the algorithm, due to the random initialisation of cluster centres. However, results obtained have shown sufficient uniformity due to the chosen input bands, and thus statistic measures have been given here for one of these runs, a representative one, instead of for mean values of several runs.

For comparison purposes, results obtained with unsupervised clustering of the provided multispectral bands are given here, as well as those obtained with the application of the same clustering method on the composite of these multispectral bands with the normalised digital elevation model. Figs. 5.1 (a) and (b) give an indication of the low accuracy classifications obtained, especially when only multispectral bands are considered, making the classes unidentifiable. Height information improves the outcome, although not reaching an acceptable classification: water bodies, buildings and ground are merged and split in conglomerate classes, high and low vegetation are indistinguishable, and shadow areas include building regions.



(a) (b)

Figure 5.1: Unsupervised clustering of (a) multispectral information (b) multispectral and height information

5.2.1 Method 1

A ground truth comprising an average of just above 8500 pixels per class (51030 pixels in total) has been created for evaluation purposes. Results obtained with Method 1 have scored an overall accuracy of 82.2911% (34991/42521 pixels) and a Kappa coefficient of 0.7786. It must be pointed out that water areas have not been considered in the evaluation, as they are not identified by the algorithm, and thus for Method 1, a total of 42521 pixels is used.

Five evaluation tables are provided. Tables 5.1 and 5.2 give the confusion matrices for the different classes, in pixels and percentages. Table 5.3 gives the omission and commission errors (in pixels and in percentages), and Table 5.4 the user's and producer's accuracies (in pixels and percentages). Finally, Table 5.5 provides the rest of statistics (C_mp, C_rp, B_f, M_f, Q_p and H_m, in percentages), as described in section 5.1.

Class/GT (pixels)	High veg.	Low veg.	Shadow	Buildings	Ground	Total
High veg.	8410	146	3166	485	2477	14684
Low veg.	30	8289	3	0	0	8322
Shadow areas	2	0	5056	2	13	5073
Buildings	53	0	81	7566	346	8046
Ground	10	71	194	451	5670	6396
Total	8505	8506	8500	8504	8506	42521

Table 5.1: Confusion matrix for results of Method 1 (in pixels). NB: Water bodies are not detected, and only central areas of regions are considered, due to the difficulty of manually delimiting border regions for the creation of ground truth.

From the information contained in the given tables, it can be concluded that Method 1 provides an overall good classification, as it can be seen in Fig. 4.10. For the München dataset,

Class/GT(%)	High veg.	Low veg.	Shadow	Buildings	Ground	Total
High veg.	98.88	1.72	37.25	5.70	29.12	34.53
Low veg.	0.35	97.45	0.04	0.00	0.00	19.57
Shadow areas	0.02	0.00	59.48	0.02	0.15	11.93
Buildings	0.62	0.00	0.95	88.97	4.07	18.92
Ground	0.12	0.83	2.28	5.30	66.66	15.04
Total	100.00	100.00	100.00	100.00	100.00	100.00

Table 5.2: Confusion matrix for results of Method 1 (in percentages).

Class	Commission (%)	Omission (%)	Commission (Pixels)	Omission (Pixels)
High veg.	41.57	1.12	6274/14684	95/8505
Low veg.	0.40	2.55	33/8322	217/8506
Shadow areas	0.34	40.52	17/5073	3444/8500
Buildings	5.97	11.03	480/8046	938/8504
Ground	11.35	33.34	726/6396	2836/8506

Table 5.3: Comission and omission measures for results of Method 1 (in percentages).

Class	Prod. Acc.	User Acc.	Prod. Acc.	User Acc.
High veg.	98.88	57.27	8410/8505	8410/14684
Low veg.	97.45	99.60	8289/8506	8289/8322
Shadow areas	59.48	99.66	5056/8500	5056/5073
Buildings	88.97	94.03	7566/8504	7566/8046
Ground	66.66	88.65	5670/8506	5670/6396

Table 5.4: Producer accuracy and user accuracy measures for results of Method 1 (in percentages).

Class	C_mp	C_rp	B_f	M_f	Q_p	H_m
High veg.	98.8830	57.2732	0.7460	0.0113	56.9051	98.8830
Low veg.	97.4489	99.6035	0.0040	0.0262	97.0723	97.4489
Shadow areas	59.4824	99.6649	0.0034	0.6812	59.3636	59.4824
Buildings	88.9699	94.0343	0.0634	0.1240	84.2164	88.9699
Ground	66.6588	88.6492	0.1280	0.5002	61.4168	66.6588
Average	82.2886	87.8450	0.1890	0.2686	71.7948	82.2886

Table 5.5: Pixel-based results evaluation for Method 1 in percentages: *completeness* C_mp, *correctness* C_rp, *branching factor* B_f, *miss factor* M_f, *quality* Q_p, *accuracy* A and *Helden's index* H_m

used for classification evaluation, shadows, buildings and both low and high vegetation are on the whole well identified. Water bodies, however, are not distinguished from ground, and thus, as the clustering looks for six classes, a further subtype of building of high roof reflectivity is identified. Accuracies for high and low vegetation are high (98.88% and 97.45%), with shadow

areas presenting the lowest values (59.48%) due to mostly error of omission. User accuracies are high for all classes (88.65-99.60%) except high vegetation, with mostly errors of commission (41.57%). Average values show an overall good classification with 82.2886% completeness and 87.8450% correctness.

Following Makarau et al. [69] and Longbotham et al. [67] and attempt to include Gabor features was undertaken, as explained in chapter 4. However, although the authors report an improvement of results with such an approach, worse scores were attained here. Reasons might be found in the fact that both approaches use WorldView images, of higher resolution than those of IKONOS and thus presenting less difficulties in edges, and also in the different objectives of their research, as the number of classes targeted by the authors is of higher count and they encompass for example differences in roof composition. In this case, the results are more fragmented than it is desired.

Following Zhang [128] and Makarau et al. [70], also the extraction of textures based on the GLCM matrix has been implemented, as explained in chapter 4. Again, both authors report an improvement in building detection and classification, respectively. However, again the objectives are different, with a much higher number of targeted classes in the former, and a higher number in the latter, with the resulting fragmentation when a smaller number of semantic classes is aimed for.

5.2.2 Method 2

The same ground truth used for the evaluation of Method 1 has been applied here, with 51030 pixels in total. Results with Method 2 have an overall accuracy of 95.8946% (48935/51030 pixels) and a Kappa coefficient of 0.9507. Again, five evaluation tables are provided. Tables 5.6 and 5.7 give the confusion matrices for the different classes, in pixels and percentages. Table 5.8 gives the omission and commission errors (in pixels and in percentages), and Table 5.9 the user's and producer's accuracies (in pixels and percentages). Finally, Table 5.10 provides the rest of statistics (C_mp, C_rp, B_f, M_f, Q_p and H_m, in percentages), as described in section 5.1.

It can be concluded that Method 2 provides an overall better classification than Method 1, with a good classification for all classes, as it can be seen in Fig. 4.12. For the München dataset, all classes are well identified, including water bodies. Accuracies for high and low vegetation are high (98.86% and 96.64%), with buildings giving the lowest value (88.97%) due to a combination of errors of commission and especially omission. User accuracies are high for all classes (89.45-99.99%). Average values show an overall good classification with 95.8943% completeness and 95.9744% correctness.

Class/GT (pixels)	High veg.	Water	Low veg.	Shadow	Buildings	Ground	Total
High veg.	8408	215	0	0	0	0	8623
Low veg.	30	8220	0	3	0	0	8253
Water	0	0	8373	1	0	0	8374
Shadow areas	4	0	0	8221	286	13	8524
Buildings	49	0	125	62	7566	346	8148
Ground	14	71	11	213	652	8147	9108
Total	8505	8506	8509	8500	8504	8506	51030

Table 5.6: Confusion matrix for results of Method 2 (in pixels). Only central areas of regions are considered.

Class/GT(%)	High veg.	Low veg.	Water	Shadow	Buildings	Ground	Total
High veg.	98.86	2.53	0.00	0.00	0.00	0.00	16.90
Low veg.	0.35	96.64	0.00	0.04	0.00	0.00	16.17
Water	0.00	0.00	98.40	0.01	0.00	0.00	16.41
Shadow areas	0.05	0.00	0.00	96.72	3.36	0.15	16.70
Buildings	0.58	0.00	1.47	0.73	88.97	4.07	15.97
Ground	0.16	0.83	0.13	2.51	7.67	95.78	17.85
Total	100.00	100.00	100.00	100.00	100.00	100.00	100.00

Table 5.7: Confusion matrix for results of Method 2 (in percentages). Only central areas of regions are considered.

Class	Commission (%)	Omission (%)	Commission (Pixels)	Omission (Pixels)
High veg.	2.49	1.14	215/8623	97/8505
Low veg.	0.40	3.36	33/8253	286/8506
Water bodies	0.01	1.60	1/8374	136/8509
Shadow areas	3.55	3.28	303/8524	279/8500
Buildings	7.14	11.03	582/8148	938/8504
Ground	10.55	4.22	961/9108	359/8506

Table 5.8: Comission and omission measures for results of Method 2 (in percentages). Only central areas of regions are considered.

Class	Prod. Acc.	User Acc.	Prod. Acc.	User Acc.
High veg.	98.86	97.51	8408/8505	8408/8623
Low veg.	96.64	99.60	8220/8506	8220/8253
Water bodies	98.40	99.99	8373/8509	8373/8374
Shadow areas	96.72	96.45	8221/8500	8221/8524
Buildings	88.97	92.86	7566/8504	7566/8148
Ground	95.78	89.45	8147/8506	8147/9108

Table 5.9: Producer accuracy and user accuracy measures for results of Method 2 (in percentages).Only central areas of regions are considered.

Class	C_mp	C_rp	B_f	M_f	Q_p	H_m
High veg.	98.8595	97.5067	0.0256	0.0115	96.4220	98.8595
Low veg.	96.6377	99.6001	0.0040	0.0348	96.2642	96.6377
Water bodies	98.4017	99.9881	0.0001	0.0162	98.3901	98.4017
Shadow areas	96.7176	96.4453	0.0369	0.0339	93.3886	96.7176
Buildings	88.9699	92.8571	0.0769	0.1240	83.2710	88.9699
Ground	95.7794	89.4488	0.1180	0.0441	86.0568	95.7794
Average	95.8943	95.9744	0.0436	0.0441	92.2988	95.8943

Table 5.10: Pixel-based results evaluation for Method 2 in percentages: *completeness* C_mp, *correctness* C_rp, *branching factor* B_f, *miss factor* M_f, *quality* Q_p, *accuracy* A and *Helden's index* H_m

5.2.3 Method 3

Ground truth used for the evaluation of Method 1 and Method 2 has also been applied here, with 51030 pixels in total. Results with Method 3 have an overall accuracy of 97.1879% (49595/51030 pixels) and a Kappa coefficient of 0.9663. Again, five evaluation tables are provided. Tables 5.11 and 5.12 give the confusion matrices for the different classes, in pixels and percentages. Table 5.13 gives the omission and commission errors (in pixels and in percentages), and Table 5.14 the user's and producer's accuracies (in pixels and percentages). Finally, Table 5.15 provides the rest of statistics (C_mp, C_rp, B_f, M_f, Q_p and H_m, in percentages), as described in section 5.1.

Class/GT (pixels)	High veg.	Water	Low veg.	Shadow	Buildings	Ground	Total
High veg.	8408	215	0	0	0	0	8623
Low veg.	30	8220	0	3	0	0	8253
Water	0	0	8373	1	0	0	8374
Shadow areas	4	0	0	8221	286	0	8511
Buildings	1	4	0	13	7875	8	7901
Ground	62	67	136	262	343	8498	9368
Total	8505	8506	8509	8500	8504	8506	51030

Table 5.11: Confusion matrix for results of Method 3 (in pixels). Only central areas of regions are considered.

Finally, it can be concluded that Method 3 provides an overall highly accurate classification, better than those provided by Method 1 and Method 2, with very good results for all classes, as it can be seen in Fig. 4.15. For the München dataset, all classes are well identified, and an improvement is visible in terms of a lower impact of the digital elevation model inaccuracies, and in terms of building recognition, as even low buildings in construction are identified by the method. Accuracies for all classes are markedly high, ranging from 92.60% to 99.91%. Errors of

Class/GT(%)	High veg.	Low veg.	Water	Shadow	Buildings	Ground	Total
High veg.	98.86	2.53	0.00	0.00	0.00	0.00	16.90
Low veg.	0.35	96.64	0.00	0.04	0.00	0.00	16.17
Water	0.00	0.00	98.40	0.01	0.00	0.00	16.41
Shadow areas	0.05	0.00	0.00	96.72	3.36	0.00	16.68
Buildings	0.01	0.05	0.00	0.15	92.60	0.09	15.48
Ground	0.73	0.79	1.60	3.08	4.03	99.91	18.36
Total	100.00	100.00	100.00	100.00	100.00	100.00	100.00

Table 5.12: Confusion matrix for results of Method 3 (in percentages). Only central areas of regions are considered.

Class	Commission (%)	Omission (%)	Commission (Pixels)	Omission (Pixels)
High veg.	2.49	1.14	215/8623	97/8505
Low veg.	0.40	3.36	33/8253	286/8506
Water bodies	0.01	1.60	1/8374	136/8509
Shadow areas	3.41	3.28	290/8511	279/8500
Buildings	0.33	7.40	26/7901	629/8504
Ground	9.29	0.09	870/9368	8/8506

Table 5.13: Comission and omission measures for results of Method 3 (in percentages). Only central areas of regions are considered.

Class	Prod. Acc.	User Acc.	Prod. Acc.	User Acc.
High veg.	98.86	97.51	8408/8505	8408/8623
Low veg.	96.64	99.60	8220/8506	8220/8253
Water bodies	98.40	99.99	8373/8509	8373/8374
Shadow areas	96.72	96.59	8221/8500	8221/8511
Buildings	92.60	99.67	7875/8504	7875/7901
Ground	99.91	90.71	8498/8506	8498/9368

Table 5.14: Producer accuracy and user accuracy measures for results of Method 3 (in percentages).Only central areas of regions are considered.

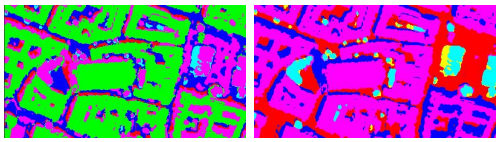
commission are more frequent with the ground class, while errors of omission are more present with buildings. Producer's and user's accuracies are very high for all classes, above 92.60% for the former and above 90.71 for the latter. Average values show the best performing classification with 97.1877% completeness and 97.3453% correctness.

Improvements reported by each method can be observed in the following examples. In Fig. 5.2 (a) a fragment of the classification for München given by Method 1 is shown. Ground is classified as two different classes, while in Fig. 5.2 (b) the correct classification obtained by Method 2 is shown, with good detection of both ground and shadow. Similarly, in Fig. 5.3 (a) the classification generated by Method 2 for a low building in the lower central area can

Class	C_mp	C_rp	B_f	M_f	Q_p	H_m
High veg.	98.8595	97.5067	0.0256	0.0115	96.4220	98.8595
Low veg.	96.6377	99.6001	0.0040	0.0348	96.2642	96.6377
Water bodies	98.4017	99.9881	0.0001	0.0162	98.3901	98.4017
Shadow areas	96.7176	96.5926	0.0353	0.0339	93.5267	96.7176
Buildings	92.6035	99.6709	0.0033	0.0799	92.3212	92.6035
Ground	99.9059	90.7131	0.1024	0.0009	90.6357	99.9059
Average	97.1877	97.3453	0.0284	0.0296	94.5933	97.1877

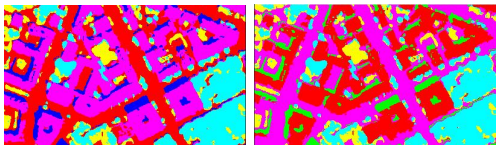
Table 5.15: Pixel-based results evaluation for Method 3 in percentages: *completeness* C_mp, *correctness* C_rp, *branching factor* B_f, *miss factor* M_f, *quality* Q_p, *accuracy* A and *Helden's index* H_m

be observed. In In Fig. 5.3 (b), the low building is better detected. Somewhat more ragged contours, however, can also be appreciated, but with the advantage of no height threshold being necessary.



(a) (b)

Figure 5.2: München dataset. Classification detail. (a) Method 1 (b) Method 2.



(a) (b)

Figure 5.3: München dataset. Classification detail. (a) Method 2 (b) Method 3.

The graphical user interface created for the integration of the algorithms in a usable system is presented here. Attention has been paid to functionality issues and a help panel has been incorporated to provide usage advice to the user. Images can be browsed and selected for loading, the desired classes and method is chosen, and classification results are displayed in the panel, showing the classes selected for visualization. Further possibilities have been included of enlarging the image, and saving it by selecting the appropriate storage folder. Change detection can be performed by loading to sets of images, classifying them, and then proceeding with the change detection panel. Fig. 5.4 shows a screenshot of the created GUI.

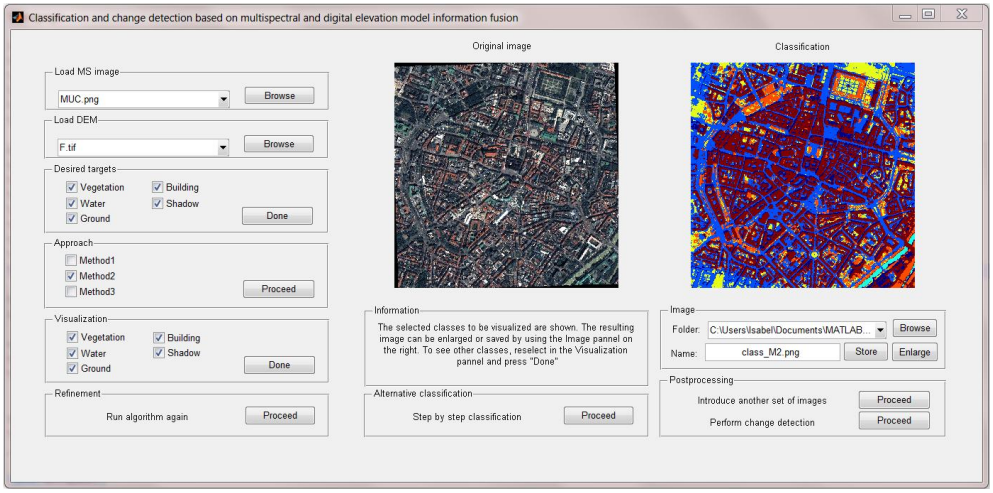


Figure 5.4: Graphical user interface of the implemented software.

5.3 Change detection results

A ground truth comprising an area of 800x800 pixels, approximately just over 10% of the dataset number of pixels, has been manually created for the evaluation of results. This ground truth includes both positive and negative changes, against which the previously explained assessment measures have been computed. As explained, evaluation includes both pixel and object-based measurements. Table 5.16 gives the former group, C_mp, C_rp, B_f, M_f, Q_p and H_m, in percentages, and as described in 5.1. Table 5.17 provides the latter group, by applying *completeness* and *correctness* in an object-based fashion, as suggested by [104].

Type	C_mp	C_rp	B_f	M_f	Q_p	H_m
Positive changes	59.512	82.715	0.20896	0.68035	52.929	59.512
Negative changes	49.767	81.633	0.225	1.0094	44.755	49.767

Table 5.16: Pixel-based change detection evaluation results in percentages.

Type	C_mo	C_ro
Positive changes	66.667	90.909
Negative changes	50.000	100.000

Table 5.17: Object-based change detection evaluation results in percentages.

Although the evaluation measures do not score high, and still further improvements need to be achieved for the change detection step, results are promising. Object-based measures have

been determined by the small amount of buildings evaluated; a greater area needs to be considered for the creation of ground truth, so that these scores are more representative, in particular for negative changes. For example, the measures for the latter group are only calculated over two ground truth buildings, and thus a 100% *correctness* is achieved. For positive changes, a 90.909% *correctness* is obtained. Regarding pixel-based evaluation, *correctness* scores are relatively high for both positives and negatives (82.715 and 81.633%, respectively), with low completeness (59.512% and 49.767%). Although results need to be improved, the approach is still fast, automatic, simple, and it incorporates implicit fusion of multispectral and height information.

Chapter 6

Conclusions

6.1 Evaluation

Objectives defined in the introduction to this work have been achieved. The state-of-the-art in land cover classification with satellite imagery of the Earth has been reviewed and analysed, as well as the opportunities offered by information fusion of different typologies (in terms of processing, operation levels and information sources, as explained in section 3.4. Different types of features have been extracted, and their impact analysed. Finally, a fully automatic and user-independent system has been implemented to exploit the fusion of multispectral and digital elevation model information to provide an effective classification and change detection tool.

Three different automatic methods have been implemented and integrated in a software system providing a graphical user interface. Analysis of the advantages and disadvantages of the first two methods (of very different nature, one based on clustering, the other on rule-based extraction) has lead to the creation of a hybrid model, and the integration of further improvements, resulting in high accuracy results. Indeed, scores obtained with this system attain average completeness and correctness of 97.1877% and 97.3453%, and a Helden's quality measure of 97.1877%, which can be considered very satisfactory, and outperforming many reviewed methods in its particularities. Also, the effectiveness of subsequently applying post-classification change detection techniques with the obtained results is demonstrated, allowing not only for the detection of changes, but for their identification.

The significance of the work is to be found in the lack of general automatic methods for the classification and posterior change detection analysis of remotely sensed imagery. As stated, most literature provides only supervised approaches requiring user expertise, or are automatic with a limited application to determined region characteristics. In this thesis, an attempt to

provide a general, automatic, and user input independent method for the classification and change detection of IKONOS *Geo* images has been undertaken. Even with this product, of the lowest accuracy provided by GeoEye, high accuracies have been obtained. New techniques for the extraction of information on *low vegetation*, *ground*, *high vegetation*, *buildings* and *water areas* from IKONOS satellite images are the original contribution of this work.

Furthermore, it is also demonstrated how the fusion of multispectral and digital elevation model information for classification can be further exploited to undertake change detection, as the simple comparison of the building maps resulting from the first step implicitly includes both height and multispectral information, instead of only the former, as it is the case with traditional digital elevation model differencing approaches.

Drawbacks also need to be considered. Testing of the algorithms has been undertaken only on two case scenarios, of very different nature, but still not representative of all the cases of land cover to be encountered. Thus, further testing and development in this sense is required. In addition to this, post classification methods present the difficulty of further extending errors included in the classification. Overall, however, results are promising, although further improvements can still be obtained. Future areas of work are identified in section 6.2.

6.2 Future work

Different aspects have been identified, offering opportunities and challenges for the expansion of the study here presented. Future work includes the extraction of surface parameters to include slope angle and orientation in the classification process, dependent on the quality of the DEM, and the introduction of mathematical morphological transformation operations for the generation of structural information, as suggested by Benediktsson in his line of research [9] in the plenary session of the Joint Remote Sensing Event (JURSE) in April 2011 [8], and also applied by several authors [86] [78] [89].

Also, the treatment of sparse vegetation deserves to be studied in more detail, and denoising of the IKONOS initial data will be implemented, as suggested by Fraser and Baltsavias [32], by use of anisotropic diffusion for noise reduction and edge sharpening, and Wallis filtering for contrast enhancement. Furthermore, target areas for classification can also be expanded with ice areas and burned land, which have determinate spectral responses that have been studied in the literature [12], [56]. For example, one of the next steps in the continuation of this work is the analysis of the behaviour and integration of the Burned Area Index by Martin [19].

Road detection can also be incorporated to the approach, in order to differentiate bare soil from paved roads. Multiple methods have been proposed in the literature [124], [39], [98], [54], [125], a review of which is provided by [100]. Finally, known azimuth and elevation angles of the

satellite at the time of acquisition will be exploited for an estimation of expected shadow area percentages, as in [81] [62], in order to complement and make more robust the shadow threshold-based segmentation that has been applied in this work. Also, the obtained segmented buildings will be applied for DEM enhancement, in a conceptually similar approach to that of [102] [101], for surface regularisation purposes.

Bibliography

[1] International Charter for Space and Major Disasters, 2011.

[2] National Aeronautics and Space Administration (NASA). Visible earth, 2011.

[3] European Space Agency. History of europe in space. 50 years of earth observation, July 2009.

[4] F. Aguera, F. J. Aguilar, and M. A. Aguilar. Using texture analysis to improve per-pixel classification of very high resolution images for mapping plastic greenhouses. *ISPRS Journal of Photogrammetry and Remote Sensing*, 63(6):635–646, November 2008.

[5] H. Arefi and M. Hahn. A morphological reconstruction algorithm for separating off-terrain points from terrain points in laser scanning data. *International Society for Photogrammetry and Remote Sensing Workshop on Laser Scanning*, pages 120–125, 2005.

[6] M. Awrangjeb, M. Ravanbakhsh, and C. S. Fraser. Automatic detection of residential buildings using LIDAR data and multispectral imagery. *ISPRS Journal of Photogrammetry and Remote Sensing*, 65(5):457–467, September 2010.

[7] K. Barnard, G. Finlayson, and B. Funt. Colour constancy for scenes with varying illumination. *Computer Vision and Image Understanding*, 65:311–321, 1997.

[8] J. A. Benediktsson. Plenary session. Classification of urban remote sensing imagery based on mathematical morphology. In *IEEE GRSS and ISPRS Joint Urban Remote Sensing Event*, 2011.

[9] J. A. Benediktsson, M. Pesaresi, and K. Arnason. Classification and feature extraction for remote sensing images from urban areas based on morphological transformations. *IEEE Transactions on Geoscience and Remote Sensing*, 41(9):1940–1949, 2003.

[10] E. W. Bork and J. G. Su. Integrating LIDAR data and multispectral imagery for enhanced classification of rangeland vegetation: A meta analysis. *Remote Sensing of Environment*, 111(1):11–24, November 2007.

[11] A.C. Braun, W. Uwe, and S. Hinz. Classifying roof materials using data fusion through kernel composition - comparing nv-SVM and one-class SVM. In U. Stilla, P. Gamba, C. Juergens, and D. Maktav, editors, *IEEE GRSS and ISPRS Joint Urban Remote Sensing Event*, München, April 2011.

[12] A. Brenning. Benchmarking classifiers to optimally integrate terrain analysis and multispectral remote sensing in automatic rock glacier detection. *Remote Sensing of Environment*, 113(1):239–247, January 2009.

[13] H. Chaabouni-Chouayakh, P. d'Angelo, T. Krauss, and P. Reinartz. Automatic urban area monitoring using digital surface models and shape features. In *IEEE GRSS and ISPRS Joint Urban Remote Sensing Event*, Muenchen, 2011.

[14] H. Chaabouni-Chouayakh and P. Reinartz. Towards automatic 3D change detection inside urban areas by fusing height and shape features. 2011.

[15] Q. Chen, E. Petriu, and X. Yang. A comparative study of fourier descriptors and hu's seven moment invariants for image recognition. In *Canadian Conference on Electrical and Computer Engineering 2004 (IEEE Cat. No.04CH37513)*, pages 103–106, Niagara Falls, Ont., Canada, 2004.

[16] W. Cheng, J. Chang, C. Chang, Y. Su, and T. Tu. A Fixed-Threshold approach to generate High-Resolution vegetation maps for IKONOS imagery. *Sensors*, 8(7):4308–4317, 2008.

[17] C. Christoudias, B. B. Georgescu, and P. Meer. Synergism in low-level vision. *16th International Conference on Pattern Recognition*, 4:150–155, 2002.

[18] K. Chung, Y. Lin, and Y. Huang. Efficient shadow detection of color aerial images based on successive thresholding scheme. *Geoscience and Remote Sensing, IEEE Transactions on*, 47(2):671–682, 2009.

[19] E. Chuvieco, M. P. MartIn, and A. Palacios. Assessment of different spectral indices in the red-near-infrared spectral domain for burned land discrimination. *International Journal of Remote Sensing*, 23(23):5103, 2002.

[20] D. Comanicu and P. Meer. Mean shift: a robust approach toward feature space analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24:603–619, 2002.

[21] P. d'Angelo, M. Lehner, T. Krauss, D. Hoja, and P. Reinartz. Towards automated dem generation from high resolution stereo satellite images. In *International Society for Photogrammetry and Remote Sensing*, pages 1137–1342, 2008.

[22] J. Daugman. Complete discrete 2d gabor transforms by neural networks for image analysis and compression. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 36(7):1169–1179, 1988.

[23] D. W. Deering. *Rangeland reflectance characteristics measured by aircraft and spacecraft sensors*. PhD thesis, Texas A&M University, 1978.

[24] G. Dial, H. Bowen, F. Gerlach, J. Grodecki, and R. Oleszczuk. IKONOS satellite, imagery, and products. *Remote Sensing of Environment*, 88(1-2):23–36, November 2003.

[25] R. Dianat and S. Kasaei. Change detection in optical remote sensing images using Difference-Based methods and spatial information. *Geoscience and Remote Sensing Letters, IEEE*, 7(1):215–219, 2010.

[26] P. Dorninger and C. Nothegger. 3D segmentation of unstructured point clouds for building modelling. In U. Stilla et al., editor, *PIA07. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, volume 36, pages 191–196., 2001.

[27] J. Du, X. Feng, Z. Wang, Y. Huang, and E. Ramadan. The methods of extracting water information from spot image. *Chinese Geographical Science*, 12:68–72, 2002.

[28] D. Ehrlich and C. Bielski. Texture based change detection of built-up on spot panchromatic imagery using pca. In *Joint Urban Remote Sensing Event*, 2011.

[29] M. A. Fischler and R. C. Bolles. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM*, 24(6):381–395, 1981.

[30] G. Forlani, C. Nardinocchi, M. Scaioni, and P. Zingaretti. Complete classification of raw LIDAR data and 3D reconstruction of buildings. *Pattern Analysis and Applications*, 8(4):357–374, 2006.

[31] Deutsches Zentrum fr Luft-und Raumfahrt e.V. Erdbeobachtung, 2011.

[32] C. S. Fraser, E. Baltsavias, and A. Gruen. Processing of ikonos imagery for submetre 3d positioning and building extraction. *ISPRS Journal of Photogrammetry and Remote Sensing*, 56:177–194, 2002.

[33] P. Gamba, P. Du, C. Juergens, and D. Maktav. Foreword to the special issue on Human settlements: A global remote sensing challenge? *Selected Topics in Applied Earth Observations and Remote Sensing, IEEE Journal of*, 4(1):5–7, 2011.

[34] D. Gercek and U. Zeydanl. Object-based classification of landscape into land management units (lmus). In *GEOBIA10*, 2010.

[35] P. Poete. German Aerospace Center (DLR). EOS Network. Integrated Earth Observation System, 2011.

[36] GlobalSecurity. Kosovo operation allied force imagery, 2005.

[37] R. O. Green and J. J. Simmonds. A role for aviris in the landsat and advanced land remote sensing system program. *Aviris Workshop Bibliographies. NASA Jet Propulsion Laboratory*, 1993.

[38] J. Grodecki and G. Dial. Ikonos geometric accuracy. In *Joint Workshop of ISPRS Working Groups I/2, I/5 and IV/7 on High Resolution Mapping from Space*, 2001.

[39] A. Grote, M. Butenuth, M. Gerke, and c. Heipke. Segmentation based on normalized cuts for the detection of suburban roads in aerial imagery. In *4th IEEE/GRSS/ISPRS Joint Workshop*, Paris, 2007.

[40] N. Haala and C. Brenner. Extraction of buildings and trees in urban environments. *ISPRS Journal of Photogrammetry and Remote Sensing*, 54:130—137, 1999.

[41] R. M. Haralick, K. Shanmugam, and I. Dinstein. Textural features for image classification. *IEEE Transactions on Systems, Man and Cybernetics*, 6:610–621, 1973.

[42] F. Heijden, R. P. W. Duin, D. Ridder, and D. M. J. Tax. *Classification, Parameter Estimation, and State Estimation. An Engineering Approach using Matlab.* John Wiley & Sons, 2004.

[43] H. Hirschmuller. Stereo processing by semiglobal matching and mutual information. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 30(2):1137–1342, 2008.

[44] J. H. Horne. A tasseled cap transformation for ikonos-2 images. *ASPRS03 Annual Conference Proceedings*, pages 60–70, 2003.

[45] J. Huang, W. Xie, and L. Tang. Detection of and compensation for shadows in colored urban aerial images. *Proceedings of the 5th World Congress in Intelligence, Control and Automation*, pages 3098–3100, 2004.

[46] A.R Huete. A soil-adjusted vegetation index (SAVI). *Remote Sensing of Environment*, 25(3):295–309, August 1988.

[47] J. Im and J. R. Jensen. A change detection model based on neighborhood correlation image analysis and decision tree classification. *Remote Sensing of Environment*, 99(3):326–340, November 2005.

[48] Space Imaging. Space imaging's ikonos imagery, 2003.

[49] J. Inglada. Automatic recognition of man-made objects in high resolution optical remote sensing images by SVM classification of geometric image features. *ISPRS Journal of Photogrammetry and Remote Sensing*, 62(3):236–248, August 2007.

[50] R. D. Jackson and A. R. Huete. Interpreting vegetation indices. *Preventive Veterinary Medicine*, 11(3-4):185–200, December 1991.

[51] Japanese Aerospace Exploration Agency (JAXA). Advanced Land Observation Satellite (ALOS), 2011.

[52] Z. Jiang, A. R. Huete, J. Chen, Y. Chen, J. Li, G. Yan, and X. Zhang. Analysis of NDVI and scaled difference vegetation index retrievals of vegetation fraction. *Remote Sensing of Environment*, 101(3):366–378, April 2006.

[53] Z. Jiang, A. R. Huete, K. Didan, and T. Miura. Development of a two-band enhanced vegetation index without a blue band. *Remote Sensing of Environment*, 112(10):3833–3845, October 2008.

[54] X. Jin and Ch. Davis. Automatic road extraction from high-resolution multi-spectral IKONOS imagery. *Proceedings of International Geosciences and Remote Sensing Symposium*, page 17301732, 2003.

[55] A. Kajumulo. Report of the fact-finding mission to Zimbabwe to assess the scope and impact of operation Murambatsvina by the UN Special Envoy on Human Settlements Issues in Zimbabwe. Technical report, United Nations, 2005.

[56] J. S. Kargel, M. J. Abrams, M. P. Bishop, A. Bush, G. Hamilton, H. Jiskoot, H. H. Kieffer, E. M. Lee, F. Paul, F. Rau, B. Raup, J. F. Shroder, D. Soltesz, D. Stainforth, L. Stearns, and R. Wessels. Multispectral imaging contributions to global land ice measurements from space. *Remote Sensing of Environment*, 99(1-2):187–219, November 2005.

[57] R. J. Kauth and G. S. Thomas. The tasseled cap. A graphic description of the spectral-temporal development of agricultural crops as seen by Landsat. In *Symposium on Machine Processing of Remotely Sensed Data*, 1976.

[58] K. Khoshelham, C. Nardinocchi, E. Frontoni, A. Mancini, and P. Zingaretti. Performance evaluation of automated approaches to building detection in multi-source aerial data. *ISPRS Journal of Photogrammetry and Remote Sensing*, 65(1):123–133, January 2010.

[59] J. R. Kim and J. P. Muller. 3D reconstruction from very high resolution satellite stereo and it's application to object identification. In *ISPRS*, Ottawa, Canada, 2002. July 8-12.

[60] H. J. Kramer. *Observation of the earth and its environment: survey of missions and sensors*. Springer, 2002.

[61] D. Lee, K. Lee, and S. Lee. Fusion of LiDAR and imagery for reliable building extraction. *Photogrammetric Engineering and Remote Sensing, Journal of the American Society for Photogrammetry and Remote Sensing*, 74(2):215–225, February 2008.

[62] T-Y. Lee and T. Kim. Reconstruction of 3D building structures from IKONOS images through monoscopic line and shadow analysis. In *Asian Conference on Remote Sensing (ACRS)*, 2011.

[63] W. Liang, D. Hoja, M. Schmitt, and U. Stilla. Comparative study of change detection for reconstruction monitoring based on very high resolution optical data. In U. Stilla, P. Gamba, C. Juergens, and D. Maktav, editors, *IEEE GRSS and ISPRS Joint Urban Remote Sensing Event*, 2011.

[64] W. Lin, Q. Wang, S. Zha, and J. Li. Construction and application of characteristic bands of typical land cover based on spectrum-photometric method. In *Geoinformatics, 2010 18th International Conference on*, pages 1–6, 2010.

[65] J. Liu, N. Currit, and X. Meng. Extraction of water bodies from remotely sensed images. *International Symposium on Intelligent Signal Processing and Communication Systems*, 2010.

[66] J. G. Liu and P. J. Mason. *Image Processing and GIS for Remote Sensing*. Wiley-Blackwell, 2009.

[67] N. Longbotham, C. Bleiler, C. Chaapel, C. Padwick, W. Emery, and F. Pacifici. Spatial classification of worldview-2 multi-angle sequence. In *IEEE GRSS and ISPRS Joint Urban Remote Sensing Event*, 2011.

[68] D. Lu, P. Mausel, E. Brondazio, and E. Moran. Change detection techniques. *International Journal of Remote Sensing*, 25(12):2365, 2004.

[69] A. Makarau, G. Palubinskas, and P. Reinartz. Classification accuracy increase using multisensor data fusion. In *ISPRS Hannover Workshop*, June 2011.

[70] A. Makarau, G. Palubinskas, and P. Reinartz. Multisensor data fusion for urban area classification. In U. Stilla, P. Gamba, C. Juergens, and D. Maktav, editors, *IEEE GRSS and ISPRS Joint Urban Remote Sensing Event*, 2011.

[71] A. Makarau, R. Richter, R. Muller, and P. Reinartz. Adaptive shadow detection using a blackbody radiator model. *IEEE Transactions on Geoscience and Remote Sensing*, 2011.

[72] J. A. Marchant and C. M. Onyango. Shadow-invariant classification for scenes illuminated by daylight. *Journal of the Optical Society of America. A, Optics, Image Science, and Vision*, 17(11):1952–1961, November 2000.

[73] J. A. Marchant and C. M. Onyango. Spectral invariance under daylight illumination changes. *Journal of the Optical Society of America. A, Optics, Image Science, and Vision*, 19(5):840–848, May 2002.

[74] L. Matikainen, J. Hyypp, E. Ahokas, L. Markelin, and H. Kaartinen. Automatic detection of buildings and changes in buildings for updating of maps. *Remote Sensing*, 2(5):1217–1248, 2010.

[75] S. K. Mcfeeters. The use of the normalized difference water index NDWI in the delineation of open water features. *International Journal of Remote Sensing*, 17:1425–1432, 1996.

[76] P. Meer and B. Georgescu. Edge detection with embedded confidence. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23:1351–1365, 2001.

[77] G. Mountrakis, J. Im, and C. Ogole. Support vector machines in remote sensing: A review. *ISPRS Journal of Photogrammetry and Remote Sensing*, 66(3):247–259, May 2011.

[78] M. Dalla Mura, A. Villa, J.A. Benediktsson, J. Chanussot, and L. Bruzzone. Classification of hyperspectral images by using extended morphological attribute profiles and independent component analysis. *IEEE Geoscience and Remote Sensing Letters*, 8:(In Press), 2011.

[79] S. D. Murnion. Comparison of back propagation and binary diamond neural networks in the classification of a landsat tm image. *Computers & Geosciences. Special issue: neural network applications in the geosciences*, 22(9):995–1001, 1996.

[80] H. Nemmour and Y. Chibani. Multiple support vector machines for land cover change detection: An application for mapping urban extensions. *ISPRS Journal of Photogrammetry and Remote Sensing*, 61(2):125–133, November 2006.

[81] J. Nossin. Slope visibility and shadows in side-look spot imagery. *ISPRS Journal of Photogrammetry and Remote Sensing*, 46:133–146, 1991.

[82] Stefanik Observatory of Prague (Czech Republic). Uses of artificial satellites, 2011.

[83] N. Otsu. A threshold selection method from gray-level histograms. *IEEE Transactions on Systems, Man, and Cybernetics*, 9:62–66, 1979.

[84] E. Frontoni P. Zingaretti, G. Forlani, and C. Nardinocchi. Automatic extraction of lidar data classification rules. *14th International Conference on Image Analysis and Processing ICIAP- IEEE*, pages 273–278, 2007.

[85] M. Pagnutti, R. E. Ryan, M. Kelly, K. Holekamp, V. Zanoni, K. Thome, and S. Schiller. Radiometric characterization of IKONOS multispectral imagery. *Remote Sensing of Environment*, 88(1-2):53–68, November 2003.

[86] E. Pagot and M. Pesaresi. Systematic study of the urban postconflict change classification performance using spectral and structural features in a support vector machine. *Selected Topics in Applied Earth Observations and Remote Sensing, IEEE Journal of*, 1(2):120–128, 2008.

[87] C. R. Perry and L. F. Lautenschlager. Functional equivalence of spectral vegetation indices. *Remote Sensing of the Environment*, 14:169–182, 1984.

[88] K. Perumal and R. Bhaskaran. Supervised classification performance of multispectral images. *Journal of Computing*, 2(2), February 2010.

[89] A. Plaza, J. A. Benediktsson, J. W. Boardman, J. Brazile, L. Bruzzone, G. Camps-Valls, J. Chanussot, M. Fauvel, P. Gamba, A. Gualtieri, M. Marconcini, J. C. Tilton, and G. Trianni. Recent advances in techniques for hyperspectral image processing. *Remote Sensing of Environment*, 113(Supplement 1):S110–S122, September 2009.

[90] A. M. Polidorio, F. C. Flores, N. N. Imai, A. M. G. Tomamaselli, and C. Franco. Automatic shadow segmentation in aerial color images. *Proceedings of the Brazilian Symposium on Computer Graphics and Image Processing*, pages 270–277, 2003.

[91] R. Radke, S. Andra, O. Al-kofahi, and B. Roysam. Image change detection algorithms: A systematic survey. *IEEE Transactions on Image Processing*, 14:294—307, 2005.

[92] L.F.G. Rego and B. Koch. Automatic classification of land cover with high resolution data of the rio de janeiro city brazil. In *Remote Sensing and Data Fusion over Urban Areas, 2003. 2nd GRSS/ISPRS Joint Workshop on*, pages 172–176, 2003.

[93] F. Rottensteiner. Building change detection from digital surface models and multi-spectral images. In *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, volume XXXVI 3/W49B, 2007.

[94] F. Rottensteiner, J. Trinder, S. Clode, and K. Kubik. Building detection using LIDAR data and multispectral images. In *APRS*, volume 2, pages 673–682, Sydney, 2003.

[95] M. Rutzinger, F. Rottensteiner, and N. Pfeifer. A comparison of evaluation techniques for building extraction from airborne laser scanning. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2(1):11–20, 2009.

[96] A. Rydberg and G. Borgefors. Integrated method for boundary delineation of agricultural fields in multispectral satellite images. *Geoscience and Remote Sensing, IEEE Transactions on*, 39(11):2514–2520, 2001.

[97] S. Schneiderbauer. Monitoring multilateral humanitarian agreements. *Brief 27. Monitoring Environment and Security. Integrating concepts and enhancing methodologies. Global Monitoring for Security and Stability (GMOSS). Bonn International Center for Conversion (BICC)*, 2011.

[98] A. Shackelford and C.H. Davis. Fully automated road network extraction from high-resolution satellite multispectral imagery. In *Geoscience and Remote Sensing Symposium, 2003. IGARSS '03. Proceedings. 2003 IEEE International*, volume 1, pages 461–463 vol.1, 2003.

[99] L. Shen and C. Li. Water body extraction from landsat ETM+ imagery using Adaboost algorithm. In *Geoinformatics, 2010 18th International Conference on*, pages 1–4, 2010.

[100] B. Sirmacek. *Object Detection in Aerial and Satellite Images.* PhD thesis, Yeditepe University, 2009.

[101] B. Sirmacek, P. d'Angelo, T. Krauss, and P. Reinartz. Enhancing urban digital elevation models using automated computer vision techniques. In *ISPRS Commission VII Symposium*, Vienna, Austria, July 2010.

[102] B. Sirmacek, P. d'Angelo, and P. Reinartz. Detecting complex building shapes in panchromatic satellite images for digital elevation model enhancement. In *ISPRS*, Istanbul, Turkey, October 2010.

[103] G. Sohn and I. Dowman. Data fusion of high-resolution satellite imagery and LiDAR data for automatic building extraction. *ISPRS Journal of Photogrammetry and Remote Sensing*, 62(1):43–63, May 2007.

[104] W. Song and T. Haithcoat. Development of comprehensive accuracy assessment indexes for building footprint extraction. *IEEE Transactions on Geoscience and Remote Sensing*, 43(2):402–404, 2005.

[105] K. Soudani, C. Francois, G. le Maire, V. Le Dantec, and E. Dufrne. Comparative analysis of IKONOS, SPOT, and ETM+ data for leaf area index estimation in temperate coniferous and deciduous forest stands. *Remote Sensing of Environment*, 102(1-2):161–175, May 2006.

[106] M. F. Tappen, W. T. Freeman, and E. H. Adelson. Recovering intrinsic images from a single image. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27(9):1459–1472, 2005.

[107] A. J. Tatem, S. J. Goetz, and S. I. Hay. Fifty years of earth observation satellites. *American Scientist*, 96(5):390–398, September 2008.

[108] T. Toutin. Geometric processing of remote sensing images: models, algorithms and methods. *International Journal of Remote Sensing*, 25(10):1893–1924, 2004.

[109] V. J. D. Tsai. A comparative study on shadow compensation of colour aerial images in invariant color models. *IEEE Transactions on Geoscience and Remote Sensing*, 44:1661–1671, 2006.

[110] S. Vassilopoulou, L. Hurni, V. Dietrich, E. Baltsavias, M. Pateraki, E. Lagios, and I. Parcharidis. Orthophoto generation using IKONOS imagery and high-resolution DEM: a case study on volcanic hazard monitoring of nisyros island Greece. *ISPRS Journal of Photogrammetry and Remote Sensing*, 57:24–38, 2002.

[111] Y. V. Venkatesh and S. K. Raja. On the classification of multispectral satellite images using the multilayer perceptron. *Pattern Recognition. Kernel and Subspace Methods for Computer Vision*, 36(9):2161–2175, 2003.

[112] I.E. Villalon-Turrubiates and M.J. Llovera-Torres. Multispectral classification of remote sensing imagery for archaeological land use analysis: Prospective study. In *Geoscience and Remote Sensing Symposium (IGARSS), 2010 IEEE International*, pages 323–326, 2010.

[113] V. Walter. Object-based classification of remote sensing data for change detection. *ISPRS Journal of Photogrammetry and Remote Sensing*, 58(3-4):225–238, January 2004.

[114] E. Waltz. The principles and practice of image and spatial data fusion. In *Multisensor Data Fusion*. CRC Press, D. Hall and J. Llinas edition, 2001.

[115] L. Wang, Q. Dai, C. Zheng, and C. Wang. A comparative study of pixel level and region level classification of land use types using QuickBird imagery. In *Geoscience and Remote Sensing (IITA-GRS), 2010 Second IITA International Conference on*, volume 1, pages 219–222, 2010.

[116] Y. Wei, Z. Zhao, and J. Song. Urban building extraction from high-resolution satellite panchromatic image using clustering and edge detection. In *Geoscience and Remote Sensing Symposium, 2004. IGARSS '04. Proceedings. 2004 IEEE International*, volume 3, pages 2008–2010 vol.3, 2004.

[117] U. Weidner and W. Forstner. Towards automatic building extraction from high-resolution digital elevation models. *ISPRS Journal of Photogrammetry and Remote Sensing*, 50(4):38–49, August 1995.

[118] W. Wenbo, Y. Jing, and K. K. Tingjun. Study of remote sensing image fusion and its application in image classification. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 37:1141–1146, 2008.

[119] L. Wirkus and J. Bogardi. The global water crisis. are water-related violent conflicts becoming more likely? *Brief 27. Monitoring Environment and Security. Integrating concepts and enhancing methodologies. Global Monitoring for Security and Stability (GMOSS). Bonn International Center for Conversion (BICC)*, 2011.

[120] L. Wirkus and J. Schure. Environmental change, natural resources and violent conflict. *Brief 27. Monitoring Environment and Security. Integrating concepts and enhancing methodologies. Global Monitoring for Security and Stability (GMOSS). Bonn International Center for Conversion (BICC)*, 2011.

[121] T. P. Wu and C. K. Tang. A bayesian approach for shadow extraction from a single image. *Proceedings of the IEEE International Conference on Computer Vision*, 1:480–487, 2005.

[122] H. Xie, C. Heipke, P. Lohmann, U. Soergel, X. Tong, and W. Shi. Ein neuer binarcodierungs-algorithmus zur gemeinsamen regionenbasierten klassifikation von hyperspektraldaten und digitalen oberflachenmodellen. *Photogrammetrie - Fernerkundung - Geoinformation*, 2011(1):17–33, February 2011.

[123] H. Xu. Modification of normalised difference water index NDWI to enhance open water features in remotely sensed imagery. *International Journal of Remote Sensing*, 27:3025–3033, 2006.

[124] D. Yan and Z. Zhao. Road detection from quickbird fused image using IHS transform and morphology. In *Geoscience and Remote Sensing Symposium, 2003. IGARSS '03. Proceedings. 2003 IEEE International*, volume 6, pages 3967–3969 vol.6, 2003.

[125] K. Yang, M. Li, Y. Liu, and C. Jiang. Multi-points fast marching: A novel method for road extraction. In *Geoinformatics, 2010 18th International Conference on*, pages 1–5, 2010.

[126] Q. Yin and P. Guo. Multispectral remote sensing image classification with multiple features. In *Machine Learning and Cybernetics, 2007 International Conference on*, volume 1, pages 360–365, 2007.

[127] V. M. Zanoni and S. N. Goward. A new direction in earth observations from space: IKONOS. *Remote Sensing of Environment*, 88(1-2):1–2, November 2003.

[128] Y. Zhang. Optimisation of building detection in satellite images by combining multispectral classification and texture filtering. *ISPRS Journal of Photogrammetry and Remote Sensing*, 54(1):50–60, February 1999.