

# **3D change detection from high and very high resolution satellite stereo imagery**

**Dissertation**

**zur Erlangung  
des Doktorgrades der Naturwissenschaften (Dr. rer. nat.)  
des Fachbereichs Mathematik / Informatik  
der Universität Osnabrück**

**vorgelegt**

**von**

**Jiaojiao Tian**

**aus**

**Hebei/China**

**Osnabrück, 2013**

# **3D change detection from high and very high resolution satellite stereo imagery**

**Dissertation**

**zur Erlangung  
des Doktorgrades der Naturwissenschaften (Dr. rer. nat.)  
des Fachbereichs Mathematik / Informatik  
der Universität Osnabrück**

**vorgelegt**

**von**

**Jiaojiao Tian**

**Betreuer:**

**Hon. - Prof. Dr. - Ing. Peter Reinartz**

**Zweitgutachter:**

**Prof. Dr. - Ing. Manfred Ehlers**

Tag der mündlichen Prüfung: 29. Oktober 2013

*Dedicated to my beloved family*



## **Abstract**

Change detection is one of the most essential processing steps for monitoring urban and forest areas using remote sensing data. Even though 2D data obtained from satellite images from different dates can already provide plenty of useful information, it is usually insufficient when dealing with changes in the vertical direction. Moreover, if only one class of changes, such as buildings or forest, is of interest, it is often difficult to distinguish between relevant and irrelevant changes. In such cases, the information provided by Digital Surface Models (DSMs) is crucial, as it provides additional height information, which can be indispensable when analyzing changes. This dissertation addresses the challenge of using DSMs generated by satellite stereo images for 3D change detection.

DSM generation techniques based on stereo imagery from space have been improved continuously in recent years, enhancing the quality of the generated DSMs considerably. Nevertheless, up to now these DSMs have not been widely adopted for change detection methods. Available 3D change detection approaches prefer LiDAR data, which are more accurate but have the drawback of being more expensive and exhibit a comparatively low temporal repetition rate. The characteristic and quality of DSMs based on satellite stereo imagery have so far hardly been considered within 3D change detection procedures. Therefore, more in-depth investigations concerning the adoption of these DSMs for 3D change detection should be performed.

In this dissertation, DSMs based on stereo imagery have been visually and numerically evaluated and subsequently analyzed for various land cover types. Taking into account the quality of DSMs generated with the described methods, three DSM-assisted change detection approaches are developed. The first method, called “DSM-assisted change localization”, describes a robust change difference map generation method followed by DSM denoising. The generated change map is refined using vegetation and shadow masks and finally shape feature are used to consolidate the results through distinguishing relevant from irrelevant objects.

Concerning fusion-based change detection, two methods, feature fusion and decision fusion, are proposed. The proposed feature fusion methods make use of the fact that panchromatic images feature much sharper contours than DSMs. To alleviate the shortcomings of the DSMs, the designed region-based change detection framework extracts the original regions from the ortho-

images. For this approach, a new robust region-merging strategy is proposed to combine segmentation maps from two dates. Regarding the uncertain information contained in the DSMs and spectral images, a decision fusion method is proposed as the second fusion-based change detection method. The extracted features are classified as change indicators and no-change indicators, while two steps of the Dempster-Shafer fusion model are implemented for the final change detection.

Post-classification is a common DSM-assisted change detection method, since the DSM can be very helpful for building extraction. In this third approach, the changed building's location is obtained by comparing the new building mask with existing (often outdated) building footprint information, e.g. from GIS databases. To extract the boundaries of newly built buildings, a robust building extraction method has been developed by also considering the quality of the DSM.

These three approaches are evaluated experimentally using four representative data sets. Quantitative and qualitative experimental results obtained from each data set are analyzed in detail. The experimental results show that all of the proposed approaches are able to determine the change status of the objects of interest. The results achieved vary according to the DSM quality, the object of interest and the test area. Furthermore, it is shown experimentally that, by making proper use of DSMs in complex decision frameworks, both the efficiency and the accuracy of the change detection are improved in comparison to 2D change detection approaches. In addition, the developed approaches enable the rapid localization of changes concerning objects of interest, such as buildings and forest, which is valuable for many applications such as fast response systems after earthquake or other disasters.

## Zusammenfassung

Veränderungsanalyse ist einer der wichtigsten Verarbeitungsschritte bei der Überwachung von Stadt- und Waldgebieten mit Hilfe von Fernerkundungsdaten. Auch wenn 2D-Satellitenbilder von verschiedenen Zeitpunkten hierfür bereits viele nützliche Informationen liefern können, reichen diese Daten für gewöhnlich nicht aus um Veränderungen in vertikaler Richtung zu untersuchen. Wenn nur eine Veränderungsklasse untersucht wird, z.B. Gebäude oder Wald, ist es oft schwierig zwischen den relevanten und irrelevanten Veränderungen zu unterscheiden. In solchen Fällen kann es notwendig sein Informationen aus digitalen Oberflächenmodellen (DOMs) zu verwenden. Diese Daten enthalten ergänzende Höheninformationen, die entscheidend zur Bestimmung von Veränderungen sein können. Ziel dieser Dissertation ist es aufzuzeigen in wieweit aus satellitenbasierten Stereobildern erzeugte DOMs verwendet werden können, um verlässliche 3D-Veränderungsanalysen durchzuführen.

In den letzten Jahren wurden die Methoden zur Erstellung von DOMs basierend auf satellitengestützten Stereobildern kontinuierlich verbessert, wodurch auch die Qualität der erstellten DOMs deutlich zugenommen hat. Bisherige Verfahren zur 3D Veränderungsanalyse verwenden LiDAR-Daten, die zwar genauer, aber teurer und für gewöhnlich auch nur mit höherem zeitlichem Abstand verfügbar sind. Bis jetzt wurden die Beschaffenheit und Qualität von DOMs aus Stereobildern, in Bezug auf 3D Veränderungsanalyse kaum untersucht. Deshalb ist es zunächst von Interesse detaillierte Untersuchungen zum Einsatz von DOMs basierend auf satellitengestützten Stereobildern für die 3D-Veränderungsanalyse durchzuführen.

In dieser Dissertation werden DOMs basierend auf Stereobildern visuell und numerisch untersucht und anschließend für einige Objektklassen detailliert verglichen. Unter Berücksichtigung der Qualität der DOMs, werden auf Grund der unterschiedlichen Qualitäten drei DOM-gestützte Verfahren zur Veränderungsdetektion entwickelt. In der ersten Methode, hier als „DOM-gestützte Veränderungsdetektion“ bezeichnet, wird eine Methode zur robusten Erzeugung einer Veränderungskarte vorgeschlagen, gefolgt von einer DOM-Fehlerbeseitigung. Diese Veränderungskarte wird anschließend unter Zuhilfenahme von Vegetations- und Schattenmasken verbessert. Schließlich werden Formmerkmale verwendet um nicht relevante von relevanten Veränderungen zu unterscheiden.

Bei der fusionsbasierten Veränderungsdetektion werden zwei verschiedene Methoden entwickelt: die Merkmals- und Entscheidungsfusion. Die Merkmalsfusion basiert darauf, dass Konturen (z.B. von Gebäuden) in panchromatischen Bildern schärfer abgebildet werden als in DOMs, diese werden daher aus den Orthobildern extrahiert. Für dieses Verfahren wird eine neue robuste Strategie zur Zusammenführung von Regionen entwickelt, mit der die Segmentierungen von zwei verschiedenen Zeitpunkten zusammengeführt werden. Um mit ungenauen Informationen in den DOMs und Spektralbildern richtig umzugehen, wird außerdem eine neue Methode zur Entscheidungsfusion entwickelt. Die extrahierten Merkmale werden dabei in die Klassen „verändert“ und „unverändert“ eingeteilt. Für die endgültige Veränderungsdetektion werden zwei Prozessschritte des Dempster-Shafer Fusionsmodells implementiert.

Die Post-Klassifizierung ist eine gebräuchliche Methode zur DOM-basierten Veränderungsanalyse, da das DOM zur Gebäudeextraktion sehr gut verwendet werden kann. Bei diesem dritten Verfahren erhält man die neuen oder veränderten Gebäudepositionen indem die neue Gebäudemasken mit existierenden (oft auch überholten) Informationen zu den Gebäudeumrissen verglichen wird. Um die Grundrisse der neugebauten Gebäude zu erhalten, wird unter Berücksichtigung der Qualität des DOMs eine robuste Methode entwickelt.

Alle drei Verfahren werden an vier beispielhaften Datensätzen experimentell untersucht. Die quantitativen und qualitativen Resultate dieser Experimente werden detailliert analysiert. Die experimentellen Ergebnisse zeigen, dass alle vorgestellten Verfahren in der Lage sind den Veränderungszustand untersuchter Objekte festzustellen. Die erzielten Resultate variieren jedoch je nach Qualität des DOMs und der vorhandenen Objekte in den Testgebieten. Außerdem wird experimentell gezeigt, dass sowohl die Effizienz als auch die Genauigkeit der Veränderungsdetektion bei Verwendung von satellitengestützten Stereobildern gegenüber den reinen 2D-Verfahren verbessert werden. Zusätzlich ermöglichen die vorgestellten Methoden auch eine schnelle Lokalisierung der Veränderung bestimmter Objekte, z.B. Gebäude oder Wald. Das ist für viele Anwendungen sehr wichtig, wie z.B. bei der Verwendung in der Katastrophenhilfe bei Erdbeben- oder anderen Großschadenslagen.



## Acknowledgments

I would like to sincerely thank Prof. Dr. Peter Reinartz, department head of the department Photogrammetry and Image Analysis at the German Aerospace Center (DLR), Remote Sensing Technology Institute (IMF), for giving me the opportunity to work in his research group and to perform my research on this interesting topic. He has always listened to me and has given suggestions concerning my ideas. Being a mentor, he has provided support for my questions during the creation of all of my papers and also this dissertation. All of his efforts have been a great help for my PhD study. I would further like to thank Prof. Dr. Manfred Ehlers of the Institute for Geoinformatics and Remote Sensing at Osnabrück University, for invaluable advice concerning my dissertation.

I would also like to thank the German Academic Exchange Service (DAAD) for financial support of my PhD study and European Space Imaging (EUSI) and Euromap GmbH for providing the satellite data for this research.

I would like to thank all my colleagues in the Photogrammetry and Image Analysis department, IMF, DLR, especially Dr. Pablo d'Angelo, who has generated all of the DSMs required for this thesis and Shiyong Cui, who has provided valuable programming support for this research as well as Dr. Aliaksei Makarau, who has generated the shadow maps used for this work. Then I also want to express my thanks to Mr. Rolf Stätter for carefully proof reading my thesis. I am also grateful to our team leader Thomas Krauß, and all the colleagues in our group, Dr. Danielle Hoja, Georg Kusch, Ke Zhu, Dr. Hossein Arefi and Mathias Schneider for their help and suggestions during my daily work.

I would also like to express my gratitude towards Dr. Emmanuel Baltsavias and Prof. Konrad Schindler for their guidance and suggestions during my stay at ETH Zürich. I also need to thank all the colleagues in the Photogrammetry and Remote Sensing group in the Institute of Geodesy and Photogrammetry. I would like to thank Dr. Franz Rottensteiner, Institute of Photogrammetry and Geoinformation, University of Hannover, for providing his suggestions on using Dempster-Shafer fusion theory.

My special thanks go to Prof. Qin Yan from the Chinese Academy of Surveying and Mapping, who has been my guide, not only in research but also in life. She has supported me during my studies and academic career and has always believed in me and encouraged me. I also want to thank my previous supervisor during my master's study, Prof. Xinming Tang from the Satellite Surveying and Mapping Application Center, China, who has treated all of his students like his own children, and given me valuable suggestions even during my PhD study.

Last but not least, I would like to thank my family and all my friends for their love and support during my stay in Germany. Finally, I want to thank my husband Peter, who has driven away my homesickness and given me a home in Germany and who has also provided many valuable suggestions concerning my work.

# Table of Contents

ABSTRACT .....	I
ZUSAMMENFASSUNG .....	III
ACKNOWLEDGMENTS .....	V
1. INTRODUCTION.....	1
1.1 MOTIVATION AND PROBLEM STATEMENT.....	1
1.2 GOALS AND SCIENTIFIC CONTRIBUTION .....	3
1.3 ORGANIZATION OF THE DISSERTATION.....	4
2. DSM GENERATION AND QUALITY ASSESSMENT .....	7
2.1 STEREO SENSORS.....	8
2.2 INTRODUCTION TO DSM GENERATION .....	10
2.2.1 Rational Polynomial Coefficient (RPC) model correction.....	11
2.2.2 Semi-Global Matching (SGM).....	12
2.2.3 Image orientation and DSM generation .....	13
2.2.4 DSM Refinement .....	13
2.3 3D CO-REGISTRATION .....	13
2.4 DSM QUALITY ASSESSMENT .....	15
2.4.1 Visualization of the quality .....	17
2.4.2 Statistical comparison .....	21
2.5 DISCUSSION.....	22
2.6 SUMMARY .....	24
3. REVIEW OF CHANGE DETECTION TECHNIQUES .....	25
3.1 INTRODUCTION.....	25
3.2 2D CHANGE DETECTION .....	25
3.2.1 Pre-processing.....	25
3.2.2 Pixel-based change detection .....	27
3.2.3 Region-based change detection.....	30
3.2.4 Binary change map generation.....	31
3.2.5 Shortcomings of 2D change detection methodology .....	33
3.3 DSM-ASSISTED CHANGE DETECTION .....	33
3.3.1 Change candidates .....	34
3.3.2 Fusion-based change detection .....	35
3.3.3 Post-classification .....	36
3.4 SUMMARY .....	38
4. DSM-ASSISTED CHANGE LOCALIZATION .....	41

4.1	INTRODUCTION.....	41
4.2	ROBUST DSM DIFFERENCING.....	43
4.3	CHANGE MAP DENOISING .....	46
4.3.1	Morphological filtering of the change map.....	46
4.3.2	Shape feature based filtering.....	47
4.4	CHANGE MAP REFINEMENT USING LAND COVER CLASSIFICATION .....	49
4.4.1	Fusion with vegetation covers.....	49
4.4.2	Fusion with shadow covers .....	50
4.5	SUMMARY .....	53
5.	FUSION-BASED CHANGE DETECTION .....	55
5.1	INTRODUCTION.....	55
5.2	FEATURE FUSION-BASED CHANGE DETECTION .....	57
5.2.1	Feature fusion .....	57
5.2.2	Mean-shift segmentation.....	58
5.2.3	Greedy strategy for small region merging .....	60
5.2.4	Region-based change detection.....	62
5.3	DECISION FUSION-BASED CHANGE DETECTION.....	63
5.3.1	Decision fusion .....	63
5.3.2	Dempster-Shafer fusion .....	64
5.3.3	Change and no-change indicators .....	67
5.3.4	Two-step DS fusion model.....	69
5.4	SUMMARY .....	71
6.	POST-CLASSIFICATION CHANGE DETECTION USING GIS DATA .....	73
6.1	INTRODUCTION.....	73
6.2	CHANGE CLASSIFICATION.....	73
6.2.1	nDSM generation .....	73
6.2.2	Destroyed building detection .....	74
6.2.3	New-built building location detection.....	75
6.3	BUILDING RECONSTRUCTION.....	77
6.3.1	Hough Line-based segmentation.....	79
6.3.2	Automatic training data selection.....	80
6.3.3	Random forest-based segments classification.....	82
6.4	SUMMARY .....	83
7.	EXPERIMENTS .....	85
7.1	EXPERIMENTAL SET-UP AND QUALITY ASSESSMENT METHODS .....	86
7.1.1	Experimental set-up .....	86
7.1.2	Evaluation method .....	87
7.2	DONG-AN TEST AREA, NORTH KOREA.....	89
7.2.1	Data introduction .....	89
7.2.2	Results.....	91

7.2.3	Evaluation of results .....	94
7.3	MUNICH TEST AREA, GERMANY .....	96
7.3.1	Data introduction .....	96
7.3.2	Results.....	98
7.3.3	Evaluation of results .....	102
7.4	ISTANBUL TEST AREA, TURKEY.....	105
7.4.1	Data introduction .....	105
7.4.2	Results.....	107
7.4.3	Evaluation of results .....	110
7.5	OBERAMMERGAU TEST AREA, GERMANY.....	112
7.5.1	Data introduction .....	112
7.5.2	Results.....	114
7.5.3	Evaluation of results .....	117
7.6	DISCUSSION AND SUMMARY.....	119
8.	CONCLUSIONS AND OUTLOOK.....	123
8.1	SUMMARY OF ACHIEVEMENTS.....	123
8.2	CONCLUSION.....	124
8.3	OUTLOOK.....	125
	BIBLIOGRAPHY.....	127



## List of Abbreviations

2D	2 Dimensional
3D	3 Dimensional
AUC	Area Under ROC Curve
BCV	Belonging-Certainly Value
CVA	Change-Vector Analysis
DS	Dempster-Shafer
DOMs	Digitalen Oberflächenmodellen
DSM	Digital Surface Model
DTM	Digital Terrain Model
FD	False Detected rate
FDN	False Detected Number
FN	False Negative
FP	False Positive
GCP	Ground Control Points
GIS	Geographic Information System
GSD	Ground Sample Distance
HR	High Resolution
ISODATA	Iterative Self-Organizing Data Analysis Technique
KIA	Kappa Index of Agreement
KL	Kullback-Leibler
LiDAR	Light Detection and Ranging
LiDAR-DSM	DSM generated from LiDAR
MAD	Multivariate Alteration Detection
MS	Multispectral channels
NDVI	Normalized Difference Vegetation Index
NIR	Near-Infrared
NMAD	Normalized Median Absolute Deviation
nDSM	normalized DSM
OA	Overall Accuracy
OBCD	Object-Based Change Detection
PDF	Probability Density Function

PCA	Principal Components Analysis
PRISM	Panchromatic Remote-Sensing Instrument for Stereo Mapping
ROC	Receiver Operating Characteristics (ROC)
RE	Reprojection Error
RPC	Rational Polynomial Coefficients
RVI	Ratio Vegetation Index
SAR	Synthetic Aperture Radar
SGM	Semi-Global Matching
Stereo-DSM	DSM generated from stereo imagery
SRTM	Shuttle Radar Topography Mission
TD	True Detected rate
TDN	True Detected Number
TN	True Negative
TP	True Positive
VHR	Very High Resolution



## List of Figures

Fig. 1-1. Example of change detection at different seasons. ....	2
Fig. 2-1. DSM generation flow chart. ....	10
Fig. 2-2. 3D co-registration results.....	15
Fig. 2-3. Examples of different land covers. ....	16
Fig. 2-4. Building and building boundary. ....	16
Fig. 2-5. Profile comparison for an urban area.....	17
Fig. 2-6. Profile comparison for a forest and grass area.....	18
Fig. 2-7. Histograms of the height errors in urban areas from three satellites. ....	19
Fig. 2-8. Histograms of the height errors in and close to forest areas. ....	20
Fig. 2-9. Stereoscopic image acquisition in low-density forest areas. ....	24
Fig. 3-1. Minimum Error thresholding.....	33
Fig. 4-1. Comparison of DSMs generated from LiDAR and stereo matching.....	42
Fig. 4-2. DSM subtraction results of a building site.....	44
Fig. 4-3. Different image filtering results.....	44
Fig. 4-4. Robust difference results with positive changes.....	46
Fig. 4-5. Morphological based filtering.....	47
Fig. 4-6. Vertical change value evaluation.....	48
Fig. 4-7. Convex Hull calculation. ....	48
Fig. 4-8. WorldView-2 images and classified changes. ....	50
Fig. 4-9. Shadow observed in stereo data.....	51
Fig. 4-10. DSM quality in shadow area.....	51
Fig. 4-11. Detected shadows based on the scene shown in Fig. 4-10. ....	53
Fig. 4-12. Changed building mask and shadow mask. ....	53
Fig. 4-13. Flow chart of the DSM-assisted change location method proposed.....	54
Fig. 5-1. Processing levels of image fusion.....	55
Fig. 5-2. Boundary sharpness comparison between panchromatic image.....	57
Fig. 5-3. Flow chart of the feature fusion approach. ....	58
Fig. 5-4. Segmentation intersection procedure.....	59
Fig. 5-5. Small regions at the borders resulting from imperfect registration. ....	60
Fig. 5-6. Merging problem statement.....	60
Fig. 5-7. The probability mass function. ....	66

Fig. 5-8. Comparison of the change map before and after sigmoid curve stretching.....	66
Fig. 6-1. DSM vs nDSM. nDSM (b) is generated from DSM (a). .....	74
Fig. 6-2. Footprint-assisted destroyed building extraction.....	75
Fig. 6-3. Building location detection procedure.....	76
Fig. 6-4. Flowchart of the proposed method .....	78
Fig. 6-5. Line representation. ....	79
Fig. 6-6. Building boundary detection procedure.....	80
Fig. 6-7. Generated training data (with five classes).....	82
Fig. 6-8. Building footprint extraction procedure .....	83
Fig. 7-1. Locations of the four test areas.....	85
Fig. 7-2. Example of selected regions used for 3D-shift determination.....	87
Fig. 7-3. Dataset in the Dong-an test area. ....	90
Fig. 7-4. Change map and masks generated in the Dong-an Test area with Method I.....	92
Fig. 7-5. Change map and masks generated in the Dong-an test area with Method II.....	93
Fig. 7-6. Change map and masks generated in the Dong-an test area with Method III .....	94
Fig. 7-7. Dataset in the Munich test area.....	97
Fig. 7-8. Change map and masks generated in the Munich test area with Method I.....	99
Fig. 7-9. Change map and masks generated in the Munich test area with Method II .....	100
Fig. 7-10. Change map and masks generated in the Munich test area with Method III.....	101
Fig. 7-11. Change masks generated in the Munich test area with Method IV .....	102
Fig. 7-12. An example of non-positive building change.....	104
Fig. 7-13. Three extracted changed buildings. ....	104
Fig. 7-14. Dataset in the Istanbul test area. ....	106
Fig. 7-15. Change map and masks generated in the Istanbul test area with Method I .....	108
Fig. 7-16. Change map and masks generated in the Istanbul test area with Method II.....	109
Fig. 7-17. Change map and masks generated in the Istanbul test area with Method III.....	110
Fig. 7-18. Dataset in the Oberammergau forest test area. ....	113
Fig. 7-19. Change map and masks generated in the Oberammergau test area with Method I .....	115
Fig. 7-20. Change map and masks generated in the Oberammergau test area with Method II .....	116
Fig. 7-21. Change map and masks generated in the Oberammergau test area with Method III..	117
Fig. 7-22. Example of a mis-detected building change.....	119
Fig. 7-23. Example of a reconstructed building. ....	120

## List of Tables

Table 2-1. Summary of the stereo sensor properties. ....	9
Table 2-2. Accuracy measures for DSM quality assessment. ....	21
Table 2-3. Quality assessment for seven land covers in three sensors. ....	21
Table 4-1. Shape difference of possible buildings and artifact objects. ....	48
Table 5-1. Decision model for the six classes. ....	70
Table 6-1. Knowledge-based training data selection rules. ....	82
Table 7-1. AUC comparison in the Dong-an test area. ....	95
Table 7-2. Pixel-based change masks evaluation in the Dong-an test area. ....	96
Table 7-3. Region-based change masks evaluation in the Dong-an test area. ....	96
Table 7-4. AUC comparison in the Munich test area. ....	103
Table 7-5. Pixel-based change masks evaluation in the Munich test area. ....	103
Table 7-6. Object-based change masks evaluation in the Munich test area. ....	104
Table 7-7. AUC comparison in the Istanbul test area. ....	111
Table 7-8. Pixel-based change masks evaluation in the Istanbul test area. ....	111
Table 7-9. Object-based change masks evaluation in the Istanbul test area. ....	111
Table 7-10. AUC comparison in the Oberammergau test area. ....	118
Table 7-11. Pixel-based change masks evaluation in the Oberammergau test area. ....	118



# 1. Introduction

Observing the earth with remote sensing sensors is becoming increasingly important nowadays. Along with image quality improvements, the requirements for image process efficiency and the accuracy of results are also becoming much higher, especially for change detection applications. When handling emergency problems, such as earthquakes or floods, more automatic and real-time processing systems are preferable to manual image interpretation. In urban area change detection, the focus of interest is shifting from overall classification to monitoring of single objects like buildings, and even to detailed changes in building boundaries. In forest change detection, it is important to determine changes in both area and volume. Height information derived from digital surface models (DSM) can be very helpful for these tasks.

## 1.1 Motivation and problem statement

Fast-paced urban and rural development (e.g., urban growing, deforestation) as well as more frequent natural disasters (e.g., earthquakes, hurricanes, tsunamis) have increased the demand for efficient city monitoring and disaster assessment. These topics are related to change detection and have led to the establishment of a fundamental research field in remote sensing. Many works focus on comparing multi-temporal images. However, there are several challenges associated with change detection using only 2D information extracted from satellite images. Due to the nature of the sensors, only changes related to reflectance values and / or local textural changes can be detected. High resolution images from sensors like WorldView-2 and IKONOS allow the derivation of detailed information; however, automatic change detection methodologies applied to these very high resolution (VHR) data often lead to the detection of irrelevant changes between images from two dates. Differences in sensor object geometry, sun angle, soil moisture, atmospheric conditions, vegetation phenology and other factors influence the radiometric and geometric information from satellite images, thus limiting the effectiveness of change detection techniques based on these data.

On the other hand, two images of an actually changed area or object can be very similar. For example, two panchromatic images from the IKONOS sensor are shown in Fig. 1-1. One is from the year 2006 (early summer) and the other is from the year 2010 (winter). In the center of Fig. 1-1b, a new building can be recognized, which was built on the concrete area visible in Fig. 1-1a.

## 1 Introduction

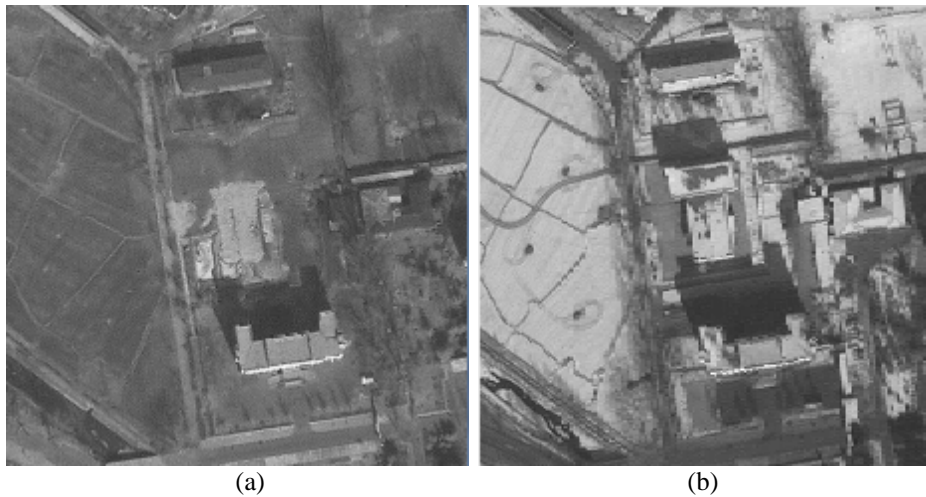


Fig. 1-1. Example of change detection at different seasons. Panchromatic image from June 2006 (a) and from December 2010 (b).

It is possible that the building's roof and the previous concrete floor consist of similar material: thus, they will share similar radiometric characteristics and even similar textures. Without the height information from a DSM, these changes in vertical direction, such as building height or forest growth, are easily overlooked. Such information can play an important role in various applications, such as disaster assessment (e.g. after an earthquake) and urban area construction and/or destruction monitoring.

Fig. 1-1 also shows that due to seasonal differences, most of the unsealed area in Fig. 1-1b is snow-covered, while in Fig. 1-1a, fields and bare soil are visible, which would lead to large areas of change. A similar problem may also occur in and near forest areas. Some trees have completely different radiometric properties in different seasons or from different viewing angles. These are just a few examples in which comparing only the spectral images would lead to false positives in the change detection process. In these cases, height is a very important feature in highlighting the object-specific changes. A considerable body of literature exists for 2D change detection but very little research is available that also incorporates height information (referred to hereafter as "3D change detection"), especially with regard to satellite data (see also Chapter 3).

Many applications related to 3D change detection make use of DSM from Light Detection and Ranging (LiDAR) which unfortunately is not ideal for frequent and large-scale change monitoring because of the high acquisition costs and low time resolution. Matching-based DSM generation from satellite stereo data and its application has become a very important task, and an increasing interest in both remote sensing and computer vision research has recently become evident (Zhang,

2005; Hirschmüller, 2008; d'Angelo et al., 2008). However, DSMs generated from stereo matching (Stereo-DSM) have not yet been widely used for change detection. Most 3D change detection methods rely on a simple subtraction of two DSMs. This may be suitable if DSMs from very high quality data, such as DSMs from LiDAR (LiDAR-DSMs), are available. However, due to the restrictions to the quality and resolution of DSMs generated from spaceborne data, directly subtracting two Stereo-DSMs does not generate higher change detection accuracy than using just images. Therefore, a systematic analysis of the quality of the stereo-generated DSMs is necessary and, for improved change detection, methods are developed within this thesis to combine the height information with features from original images to achieve higher accuracies in automatic as well as semi-automatic procedures.

## **1.2 Goals and scientific contribution**

This thesis is focused on change analysis for urban and forest areas, since the height component plays a major role in these domains. Prior to introducing DSM data from stereo matching for automatic or semi-automatic change detection, the analysis of their quality and suitability is essential. The main questions to be answered are the following:

- How good is the quality of the generated DSM and how can it be measured?
- How can DSMs of different quality be used for change detection procedures?
- How can the information from panchromatic and multispectral data be used?
- Can a DSM of lower quality also be used to improve the change detection accuracy and what are the limitations?

Concerning DSM quality assessment methods for urban and forest areas, a lot of fundamental research still has to be done, especially for spaceborne data. Both a visual and a numerical evaluation strategy will be used in this work to exhaustively analyze and compare the quality of DSMs from several spaceborne sensors with different resolutions. Assessments are performed on regions with various land covers. The results from the DSM quality assessment lead to a better understanding of the characteristics of these DSMs and provide decision guidance on their potential usability for change detection.

A major advantage of working with Stereo-DSM data is that the corresponding spectral information for the same time and area is always available. After orthorectification of these images by using the generated DSM, the ortho-image and Stereo-DSM are co-registered well and can be used together. As well as providing possible change features, the ortho-images can also help to

evaluate the quality of the DSM, and to alleviate some errors of the DSMs. Due to the varying DSM quality (e.g., from high or low resolution stereo imagery and acquisitions with different viewing and convergence angles), variations in the number of channels (e.g., only panchromatic images or multispectral channels (MS)) and differences in the investigated objects (e.g., buildings or trees), it is not possible to develop one uniform method which works optimally for all of these scenarios. Therefore, three different methodologies for DSM-assisted change detection are developed in this thesis. Each of them is focused on solving different parts of the abovementioned problems. The approaches differ in how and when the DSM information is integrated into the change detection workflow. The first approach focuses on refining a change map generated using height subtraction. Subtracting two DSMs is a straightforward way to locate change regions quickly. However, a change map generated in this way might contain false alarms. Two aspects should be considered for this approach: how to improve the quality of the initial height change map, and how to detect the real changes within the initial candidates. Not only the spectral information but also the shape features should be considered. The second approach operates through direct fusion of the Stereo-DSMs and the original images. As well as height change, radiometric changes should be also considered when obtaining the initial change map. Thus, a proper fusion model has to be developed. When an outdated building footprint is available, post-classification methodology is a better choice. Therefore, the third proposed approach uses post-classification and focuses on developing a robust building extraction method. The Stereo-DSM, panchromatic image and multi-spectral channel are fused in a more logical way for automatic building extraction.

To study the behavior of the proposed methods, they are tested on different types of test site. The data from the test sites differ in their DSM quality and the types of object of interest. The extracted change maps and change masks are visually and numerically compared with the change reference data provided. Thus the suitability of the proposed approaches to various datasets can be analyzed. The correct and false detections are further studied. Thus, the requirements and potential of using DSMs for urban and forest areas are evaluated in detail.

### **1.3 Organization of the dissertation**

Chapter 2 illustrates the automatic DSM generation procedure and analyzes the quality of DSM from different sensors for various land covers.



Chapter 3 reviews the different methods that are used for image preprocessing and change detection and points out previous work in which 3D information is discussed and used. Shortcomings are shown and ideas for further developments presented.

Chapters 4 to 6 present a detailed description of the three successful 3D change detection methodologies developed within this dissertation. Chapter 4 describes a DSM-assisted localization approach. Although the quality of DSMs from spaceborne stereo imagery is restricted, subtracting two DSMs can provide candidate locations of changed objects. Real changes can be selected from these candidates. Chapter 5 presents the fusion-based approach. Two fusion methods have been developed for this approach. One is feature fusion-based, focusing on fusing height change features from DSM and sharp boundaries from panchromatic images. For this approach, a novel region merging approach is proposed. The second method is designed based on the Dempster-Shafer decision fusion theory. It combines the building change features and no-building change features to highlight the building changes. Chapter 6 proposes a post-classification building change detection approach. A fast change location strategy is introduced. To extract accurate boundaries for the newly built buildings, a novel building extraction method is proposed, which fuses the DSM, panchromatic and multispectral image. The basic algorithms and proposed methods were implemented using the image processing toolbox in MATLAB. The code is attached on CD-ROM.

The developed methods are evaluated in Chapter 7 with four datasets, including DSMs featuring different resolutions and quality. The datasets and experimental results are described in this chapter in detail. A discussion is included at the end of this chapter.

Chapter 8 closes this thesis with conclusions obtained from the present work, along with an outlook on further work.



## **2. DSM generation and quality assessment**

A DSM represents the height values of the first reflective or visible surface of the earth, while in contrast, the Digital Terrain Model (DTM) refers to the bare ground surface, without natural or manmade object like buildings, trees or bridges (Zhang, 2005).

Height information from DSM is especially important for change detection of objects exhibiting height values above ground, such as buildings or trees. Several technologies are available for DSM generation: LiDAR, Interferometric Synthetic Aperture Radar and photogrammetry (Zhang, 2005). This work focuses on DSMs generated using photogrammetry, especially from satellite stereo imagery, since these data have special benefits, especially the repeatability in short time frames. The advantages of these data include:

- Low cost: The images from optical satellites are not as expensive as airborne SAR or LiDAR data, since obtaining a pair of satellite stereo images is much cheaper than hiring a plane and acquiring airborne stereo imagery, especially for medium size regions.
- Easy availability with high quality and temporal repetition rate. The satellite camera system is able to obtain more frequent images covering the same area, which is very important for timely change detection.
- As the DSMs are generated from optical satellite images, the spectral channels can also be used for change detection in a fusion process. The data are acquired without any time difference.
- The increasing quality of both the image resolution for spaceborne stereo data and the improved DSM generation technologies from computer vision in recent years.

Since the generation of DSM is not the focus of this thesis, only a brief introduction to spaceborne stereo sensors and the generation of DSM is given. More focus is placed upon the 3D co-registration and the DSM quality measures, since these procedures and their analysis are prerequisites for the further processing and analysis of 3D change detection, which is the main topic of the thesis.

## 2.1 Stereo sensors

Extracting height information from stereo imagery is a well-known technique. Zhang (2005) has listed most of the existing airborne stereo systems in detail. In recent years, the possibilities for obtaining stereo imagery from spaceborne platforms have increased substantially, and stereo data quality and image resolution are improving continuously. There are two ways to obtain stereo data: one is along-track stereoscopy from the same orbit using two or more views, either using multiple cameras or by adjusting the viewing angles; the other is across-track stereoscopy from two or more adjacent orbits. For cross-track stereoscopy, the quality of the generated DSM is impaired by changes of the surface reflectance during the two acquisition dates from both orbits. Therefore, along-track stereoscopy technology is employed by most new satellites, for instance SPOT-5, IKONOS, GeoEye-1, WorldView-2, QuickBird, ALOS-PRISM, Cartosat-1 and Ziyuan-3. Basic characteristics of these stereo sensors are summarized in Table 2-1.

In the following, these satellites, which represent the most important satellites with stereo acquisition capabilities, are briefly described:

- WorldView-2 (US)

WorldView-2 is a commercial Earth observation satellite that provides VHR space borne images. Being Digital Globe's third satellite, it was launched on October 8<sup>th</sup>, 2009. It provides 0.5 meter panchromatic images and 2 meter resolution 8 band multi-spectral images, four standard colors (red, green, blue, and near-infrared 1) and four new bands (coastal, yellow, red edge and near infrared 2). Using a pushbroom sensor, WorldView-2 is able to collect large areas of stereo-view or multi-view imagery in a single pass by adjusting the view angle of the camera. The average revisit time of the satellite is 1.1 days using off-nadir viewing (Satellite Imaging Cooperation, 2012g).

- GeoEye-1 (US)

The American satellite GeoEye-1 was launched on September 6<sup>th</sup>, 2008. It provides 0.5 meter (original 0.41 meter) resolution panchromatic and 2 meter (original 1.54 meter) multispectral images in 15.2 km swaths to the public. The multispectral images consist of four standard color channels (blue, green, red, and near-infrared). GeoEye-1 can acquire imagery at an off-nadir angle of up to 60° with a revisit time of less than 3 days (Satellite Imaging Cooperation, 2012c).

Table 2-1. Summary of the stereo sensor properties.

Country Sensor	Ground sampling distance (m)	Multi-spectral	Rev. int. (days)	Stereo mode
U.S. WorldView-2	0.5	8 bands	1.1	Along-track
U. S. GeoEye-1	0.41	4 bands	3	Along-track /Cross track
U. S. IKONOS	1	4 bands	3	Along-track /Cross-track
U. S. Quickbird	0.61	4 bands	1-3.5	Along-track /Cross-track
France SPOT	10	4 bands	2-3	Along-track /Cross-track
India Cartosat-1	2.5	pan only	5	Along-track
China Ziyuan-3	2.1 /3.6	4 bands	5	Along-track
Japan PRISM	2.5	4 bands	2	Along-track /Cross track

- IKONOS (US)

IKONOS supplies commercial customers with 1 meter resolution panchromatic images and 4 meter resolution multi-spectral images, with the four standard colors. It has a revisit time of about 3 days at 40° latitude. The off-nadir view can be adjusted up to 60° in order to obtain stereo imagery (Satellite Imaging Cooperation, 2012d).

- Quickbird (US)

Quickbird is another commercial satellite operated by DigitalGlobe. It was launched on October 18<sup>th</sup>, 2001. The revisit time of Quickbird is 1 to 3.5 days depending on the latitude. The off-nadir view is up to 30°. It has a swath width of 16.5 km at nadir view. The panchromatic image has a resolution of 0.61 meter in nadir view and 0.72 meter in 25° off-nadir view. The multispectral images have a resolution of 2.44 meter in nadir view and 2.88 meter in 25° off-nadir view (Satellite Imaging Cooperation, 2012e).

- Cartosat-1 (India)

Cartosat-1 was launched on May 5<sup>th</sup>, 2005 by the Indian Space Research Organization. It carries two panchromatic cameras that are able to obtain black-and-white stereo imagery, including a forward view of +26° (Band F) and a nadir view of -5° (Band A). The panchromatic images

## 2 DSM generation and quality assessment

obtained have a spatial resolution of 2.5 meter and cover a swath of 30 km. The satellite has a maximum revisit time of 5 days (Satellite Imaging Cooperation, 2012b).

- PRISM (Japan)

The Panchromatic Remote-Sensing Instrument for Stereo Mapping (PRISM) is one of the three instruments on board the Japanese satellite ALOS (Advanced Land Observation satellite). ALOS was launched on January 24<sup>th</sup>, 2006 and stopped working on April 21<sup>st</sup>, 2011. PRISM can provide 2.5 meter resolution stereo imagery with up to 70 km swath width. It has a revisit time of 2 days (Osawa, 2004; Satellite Imaging Cooperation, 2012a).

- SPOT-5 (France)

The French satellite SPOT 5 was launched on May 4<sup>th</sup>, 2002 by an Ariane 4 rocket. The newly designed HR stereoscopic instrument contains two fixed cameras, with a  $\pm 20^\circ$  off-nadir viewing capability. The stereo imagery has 10 meter resolution and covers a swath of 120 km (Vadon, 2003). The revisit time is about 2 to 3 days depending on the latitude (Satellite Imaging Cooperation, 2012f).

- Ziyuan-3 (China)

Ziyuan-3 is a new Chinese Satellite, which was launched on January 9<sup>th</sup>, 2012. The panchromatic imagery has 2.1 meter resolution at nadir view, and 3.6 meter resolution for the forward and backward view (Tang and Xie, 2012). It has a revisit time of about 5 days.

## 2.2 Introduction to DSM generation

Extracting height automatically from stereo imagery is a task of great importance for many applications. The main workflow of DSM generation using Rational Polynomial Coefficients (RPC) and dense matching methodology is shown in Fig. 2-1.

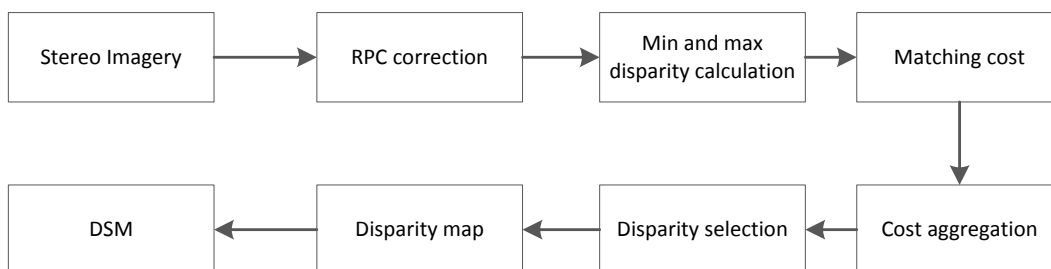


Fig. 2-1. DSM generation flow chart.

Since it is the basis of all further processing, the major difficulty of this work is accurate stereo matching. The matching points' density and accuracy directly influence the DSM quality.

### 2.2.1 Rational Polynomial Coefficient (RPC) model correction

Terrain extraction from satellite stereo imagery requires an accurate RPC sensor model. This model is used to transform the three-dimensional object space coordinates into two-dimensional image space coordinates. The initial RPC can be derived from orbit and attitude information. However, the original RPC usually has a lower absolute accuracy than the ground resolution. Thus numerous methods have been used in attempts to refine the RPC model using bundle adjustment. Toutin (2003) analysed the bundle adjustment accuracy for IKONOS in-track images. Grodecki and Dial (2003) have improved the bundle adjustment directly using the physical camera model by introducing the geometric properties of the camera model parameters into the adjustment parameter. That made the bundle adjustment more stable than before and generally applicable to all photogrammetric cameras.

Correcting the RPC for high quality geolocation requires sub-pixel accurate Ground Control Points (GCP). In most literature (Toutin, 2003; Grodecki and Dial, 2003; Xiong and Zhang, 2011), GCPs of this quality are usually obtained via ground surveys or high resolution orthorectified images and DSMs. Unfortunately such resources are not always available, which is why image-based GCP extraction and correction methods are used in contemporary research and applications (Müller et al., 2007). For the higher resolution images, like IKONOS and WorldView-2, this RPC model can be adjusted by incorporating a priori constraints into the adjustment model (Grodecki et al., 2004). If the image resolution is not too high (e.g. for Cartosat-1 and ALOS-PRISM), the Shuttle Radar Topography Mission (SRTM: Rodriguez et al., 2005) DSM can be used as a geolocation reference (d'Angelo et al., 2008) for RPC correction. However, if the quality of the RPC is good enough, RPC correction is not necessary.

The bundle adjustment assumes that a number of  $N$  ( $N = m \times n$ ) 3D points are obtained from different views. After projecting one pixel  $i$  onto the image  $j$ , it is located in  $X_{ij}$ . To denote whether  $X_{ij}$  is located outside of the image  $j$ , the variable  $v_{ij}$  is either set to 1 (inside) or 0 (outside). As illustrated in Eq. (2-1), the bundle adjustment is used to minimize the reprojection error ( $RE$ ) between the image locations of observed and predicted image points.

$$RE = \min_{a_j, b_i} \sum_{i=1}^n \sum_{j=1}^m v_{ij} d(Q(a_j, b_i), x_{ij})^2 \quad (2-1)$$

where  $Q(a_j, b_i)$  is the predicted projection of point  $i$  on image  $j$  and  $d(x, y)$  denotes the Euclidean distance between the image points represented by vectors  $x$  and  $y$ .

### 2.2.2 Semi-Global Matching (SGM)

The SGM-based DSM generation procedure – one of the most powerful methods from the dense matching category in computer vision – can be divided into three main tasks: epipolar stereo pair generation, matching and disparity map generation. First, high quality tie points between the stereo pairs are established using pyramid-based local least squares matching. Quasi-epipolar stereo pairs are generated by aligning the columns of the two stereo images. SGM is performed based on the generated epipolar stereo pairs. This matching procedure, which avoids using matching windows, is thus able to reconstruct even sharp edges like building walls. Instead of strong local assumption on the local surface shape, the matching step is cast into an energy minimization problem. It performs a semi-global optimization by aggregating costs from sixteen directions, which finds an approximate solution to the global energy function  $E$ :

$$E(D) = \sum_p \left\{ C(p, D_p) + \sum_{q \in N_p} P_1 T[|D_p - D_q| = 1] + \sum_{q \in N_p} P_2 T[|D_p - D_q| > 1] \right\} \quad (2-2)$$

Function  $C$  is the pixel-wise matching cost between the image and the disparities  $D$ . A combined use of the Mutual Information and the Census cost function is employed, as described by Hirschmüller (2008) and d’Angelo et al. (2008). The second and third terms of the energy function  $E$  calculate penalties for disparity changes in the neighborhood  $N_p$  for all pixels  $p$ .  $T$  is a conditional function: it is equal to 1 if the argument is true, and to 0 otherwise. In the case of small disparity changes, a constant penalty  $P_1$  is added to the energy function. When a higher penalty disparity change occurs, a larger constant penalty  $P_2$  will be added. This energy function consists of a data term, measuring the similarity of possibly corresponding pixels in the two images, and a regularization term, which favors similar disparities for neighboring pixels, but also allows large jumps at discontinuities.

Finally, the disparity map is recomputed by choosing the disparities with the smallest  $E$ . Sub-pixel disparities are computed by fitting a parabola to the aggregated cost values next to the disparity with minimum aggregated cost. To detect errors and occlusions, SGM is performed from the first to the second and from the second to the first image, and only consistent disparities are kept. Additionally, small isolated disparity regions are rejected as outliers, as described in d’Angelo (2010). The core SGM algorithm is described in Hirschmüller (2008).



### **2.2.3 Image orientation and DSM generation**

A preliminary point cloud consisting of approximately one million points is generated from a thinned SGM disparity map by spatial intersection. This point cloud is aligned with the SRTM DSM using a 3D affine transformation. This aligned point cloud and the known image locations are used in the image resection, yielding affine RPC correction parameters. The complete process is described in d'Angelo et al. (2008).

The DSM is generated by spatial intersection of the complete disparity map using the previously estimated RPC correction parameters and subsequent reprojection into a local coordinate system.

### **2.2.4 DSM Refinement**

Due to matching errors in regions with weak contrast or occlusions, the generated DSMs possibly contain a small number of outliers. To eliminate large outliers, in this paper a check against the SRTM DSM is performed. All points whose height deviates more than 200 m from the unfilled SRTM DSM are removed. Removing outliers will produce holes in the DSMs. Moreover, unmatched pixels also appear as holes. These holes are filled with SRTM data using the delta surface fill method proposed by Grohman et al. (2006). The delta surface fill algorithm effectively interpolates small holes and seamlessly fills large holes, such as those created by clouds or water areas, with SRTM data. Remaining holes in areas where both the generated DSM and the SRTM DSM fail to provide data are filled using inverse distance weighted interpolation.

The DSMs used in this dissertation are generated using the DLR in-house software XDibias. All of the approaches and algorithms mentioned in this section are available as modules of XDibias.

## **2.3 3D co-registration**

When more than one image is used for a designed task like change detection, the co-registration among these images is essential. It is necessary to make sure that all corresponding pixels represent the same geo-location. Only then can the information from these data be combined or compared correctly. When information about the third dimension (height) is used, it should also be considered in the co-registration procedure. In this work, 3D co-registration is not only important for change detection, but also necessary for DSM quality assessment. The DSMs to be evaluated should be co-registered to each other or to a reference DSM before starting the comparison.

## 2 DSM generation and quality assessment

In the co-registration procedure, one DSM is used as the reference DSM ( $D_0$ ), and register the other DSM ( $D_1$ ) to this DSM. The 3D coordinate is defined by X, Y, Z. All of the pixels in these two DSMs can be represented as:

$$A_0 = \{P_{01}, P_{02}, \dots, P_{0i}, \dots, P_{0n}\} \quad (2-3)$$

$$A_1 = \{P_{11}, P_{12}, \dots, P_{1i}, \dots, P_{1n}\} \quad (2-4)$$

In 3D least square matching (Gruen and Akca, 2005), not only the distance shifts in three dimensions but also the rotation shifts in three dimensions are considered. In this scenario, all of the DSMs are based on stable satellite stereo images, and in the DSM generation procedure, the SRTM DSM data are used for RPC correction: therefore, the shift in rotation is generally very low.

That is why a simple version of least square matching, linear shifting, is used for 3D co-registration to reduce the computation burden and because it is stable and sufficiently accurate. The shift distances in three dimensions ( $X_s, Y_s, Z_s$ ) are estimated via iterative 3D shifts adjustment based on the minimization of Eq. (2-5):

$$\sum_{x=1}^m \sum_{y=1}^n (Z_{ref}(x, y) - (Z_{2nd}(x + X_s, y + Y_s) + Z_s))^2 \quad (2-5)$$

In which  $m$  and  $n$  are the column and rows of the DSMs,  $(x, y)$  are the plane coordinates, and  $Z_{ref}$  and  $Z_{2nd}$  represent the height values from the two DSMs respectively, one DSM is chose as reference DSM (*ref*) in this step. The shift-values  $X_s, Y_s$ , and  $Z_s$  are adjusted iteratively, leading to a final shift result. In order to avoid large changes between two DSMs, some parameters are added to control the iterative procedure. One is the outlier rate, which can be adjusted according to the respective situation: for instance, when both DSMs exhibit a large difference in spatial resolution or time distance, which can result in large differences or changes between them. In such a case, a larger outlier rate is chosen. The pixels classified as outliers are not included in the shift distance calculation procedure.

Fig. 2-2 illustrates the effects of the 3D co-registration. Fig. 2-2a shows the displacement between the DSMs before the registration. The generated DSMs from stereo matching and the LiDAR-DSM are displayed in a checkerboard pattern with  $50 \text{ m} \times 50 \text{ m}$  size interval squares. As can be seen, the two DSMs are displaced in all three dimensions. After 3D co-registration, the situation has been improved significantly in all three dimensions (shown in Fig. 2-2b).

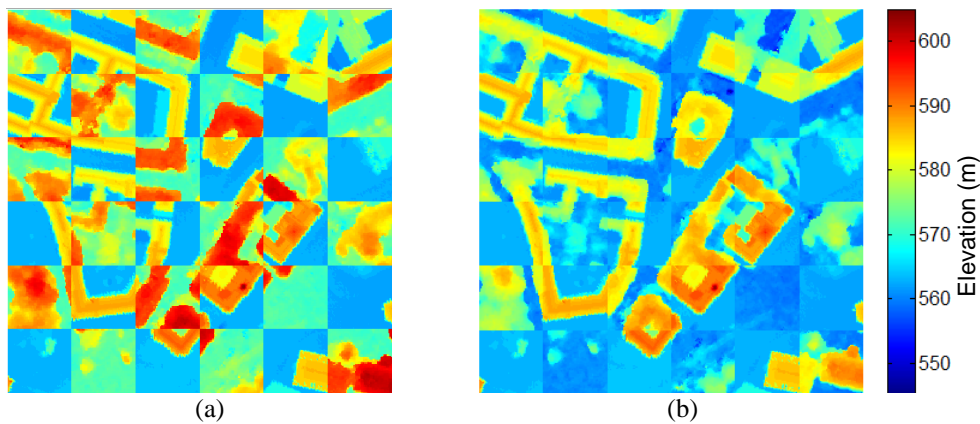


Fig. 2-2. 3D co-registration results: (a) Before 3D co-registration; (b) After 3D co-registration procedure.

## 2.4 DSM quality assessment

In the context of this section, the quality of the DSM means the accuracy of the height value it provides. All of the DSMs used here are provided as raster images, and have been derived from the point clouds generated by matching and interpolation as described in section 2.2.3. In principle, the evaluation of DSM quality requires real survey height information from the selected test area at the same time. Unfortunately these data are not available for our research. That is why, since the DSMs from photogrammetry methods have relatively lower quality in comparison to LiDAR-DSMs from dense LiDAR point clouds, the LiDAR-DSM is used as reference data here.

DSMs quality can be very different for distinct land covers. Firstly, each land cover has its unique spectral or texture information, which might influence the matching procedure. Secondly, the height derived from different land covers might have different influence from the stereo view angle (e.g., forest areas are more sensitive to stereo view angles). Moreover, two Stereo-DSMs for the same land cover can have different levels of accuracy if they are based on different sensors. Therefore, it is necessary to specify and assess the accuracy for different land covers and different stereo sensors.

- Land covers

The following land cover classes are considered in this work. Three land cover classes are from urban areas (Fig. 2-3a). They are buildings, building boundaries and ground (roads). Buildings and building boundaries are considered separately. Buildings and building boundaries are defined in Fig. 2-4, the background of this figure is a LiDAR-DSM, and the red polygon represents the inner part of a building. Pixels inside this red polygon are used to evaluate roof height. The blue polygon is a little larger than the real building. Pixels located between these two polygons are building boundary. It is defined as a separate land cover to evaluate the sharpness of the stereo-DSM.



Fig. 2-3. Examples of different land covers, three urban regions (a) and four forest regions (b).

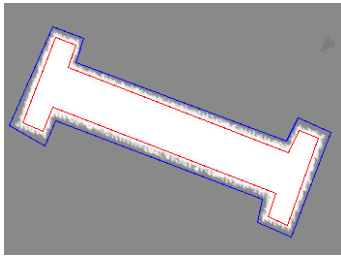


Fig. 2-4. Building and building boundary. Minimum building area (red), maximum building area (blue), the region between them is the building boundary.

The other four land covers are from forest areas and their surroundings. They have been marked in Fig. 2-3b with rectangles in different colors and are named: high-density forest region (Fig. 2-3b, marked with a yellow rectangle), low-density forest region (marked with a blue rectangle), built-up region (with low height and high density houses, Fig. 2-3b, marked with a red rectangle) and grassland (Fig. 2-3b, marked with a white rectangle).

- Stereo Sensors

The following stereo sensors are considered in this work: VHR images (WorldView-2, IKONOS) and high resolution (HR) images (Cartosat-1). In the urban area, all three datasets have been used. Since the IKONOS stereo imagery is not available in the test area, only DSMs from WorldView-2 and Cartosat-1 are evaluated. It is necessary to confirm that the reference DSM has a higher accuracy than the DSM to be evaluated. The DSM quality is evaluated in two steps. The first is the visualization of quality and the second is the statistical comparison. Both of these steps are discussed in the following sections.

### 2.4.1 Visualization of the quality

The quality of the DSM can be demonstrated graphically by comparing it with the reference DSM. The visualization methods include error distribution, absolute horizontal errors, results of filtering on the top of the topographic map, profiles of the DSM heights etc., as summarized by Höhle and Potuckova (2011). Here the profiles of DSM elevations along a given line and height error histogram analysis are used to visualize the quality.

Profile comparison can provide a direct overview of the DSM quality. Fig. 2-5 shows an example of the profile comparison in an urban area. All of the three relevant land cover types in urban areas (buildings, building boundaries and roads) are included. Three of the DSMs shown are based on satellite stereo sensors (WorldView-2, IKONOS, Cartosat-1), and are named as WorldView-DSM, IKONOS-DSM and Cartosat-DSM respectively. Another was produced using a LiDAR point cloud. The LiDAR-DSM is used as reference data. It can be seen that WorldView-DSM and IKONOS-DSM match better with the LiDAR-DSM. Although Cartosat-DSM presents more outliers than other DSMs, in most cases building areas still obtain higher height values than ground areas. Fig. 2-6 shows a profile generated in a forest area, including high and low density forest as well as grassland.

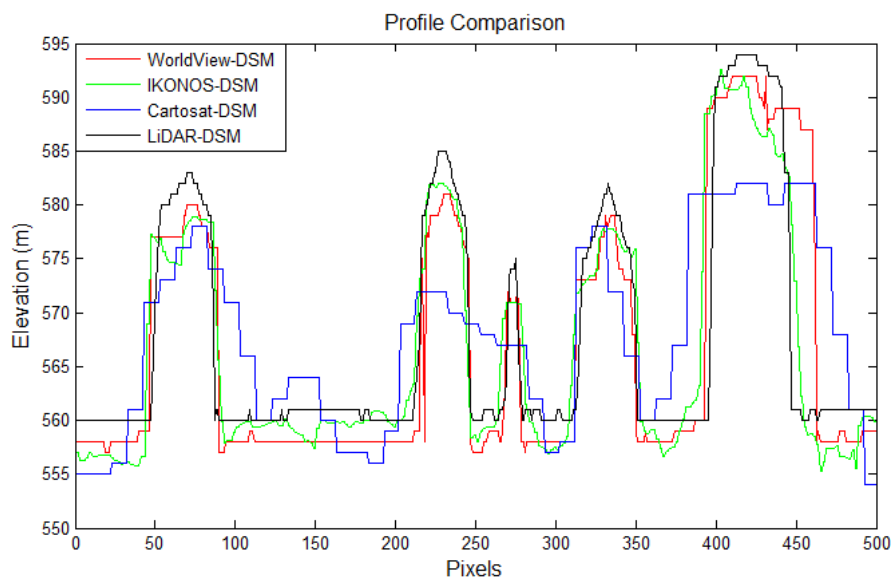


Fig. 2-5. Profile comparison for an urban area.

## 2 DSM generation and quality assessment

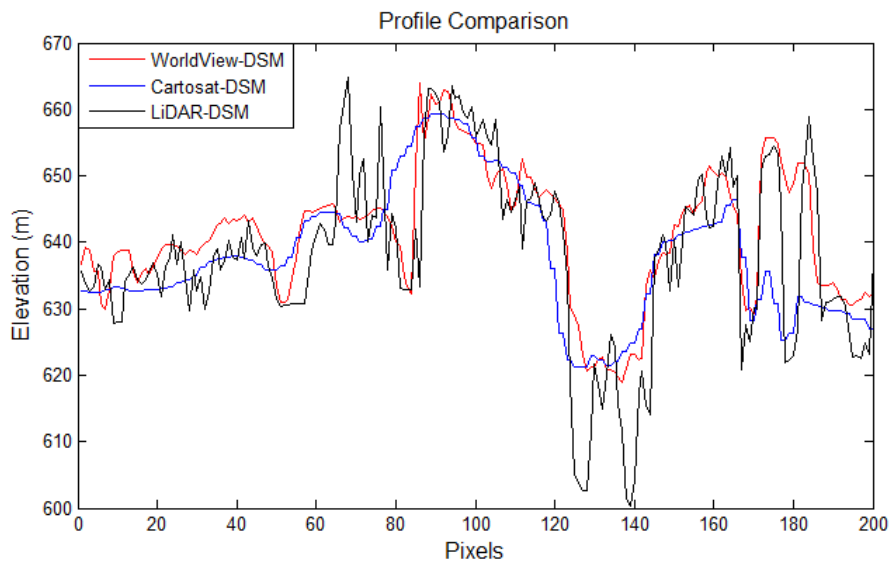


Fig. 2-6. Profile comparison for a forest and grass area.

A further view on the quality of the DSMs can be visualized using a relative frequency histogram of the height errors. If there were only regular (non-specific) differences between the generated DSM and the reference DSM, the error distribution should follow a normal distribution. The error distributions of all seven investigated land covers are shown below in Fig. 2-7. As is shown, the  $x$ -axis represents the error values (m), while the  $y$ -axis shows the relative frequency at which these errors occur.

As displayed in the height error histograms, the quality of WorldView-DSM is slightly better than that of IKONOS-DSM. In Fig. 2-5, the building roof areas are much better represented in comparison to the building boundary areas in all of the DSMs. This character is also well presented in Fig. 2-7. In the WorldView-DSM, the simulated normal distribution for the building roof region is much slimmer than the one based on the building boundary area, and most of the errors are centred between -5m and 5m for building roofs and range between -10 m to 20 m in building boundary regions. The Cartosat-1 DSM in the boundary areas shows higher noise and variations leading to very blurred representation of building walls, which can also be seen indirectly through widely spread histogram values in Fig. 2-7f, which follow approximately a normal distribution range from -40 m to 40 m. This is a clear indicator that Cartosat-1 is not suitable for urban building monitoring.

Fig. 2-8 displays the height errors in forest areas. The DSMs from four kinds of land cover, namely high-density forest regions, low-density forest regions, grassland regions and built-up regions, are analysed. As the figure shows, DSMs in high-density forest regions show higher accuracies than those in low-density forest regions.

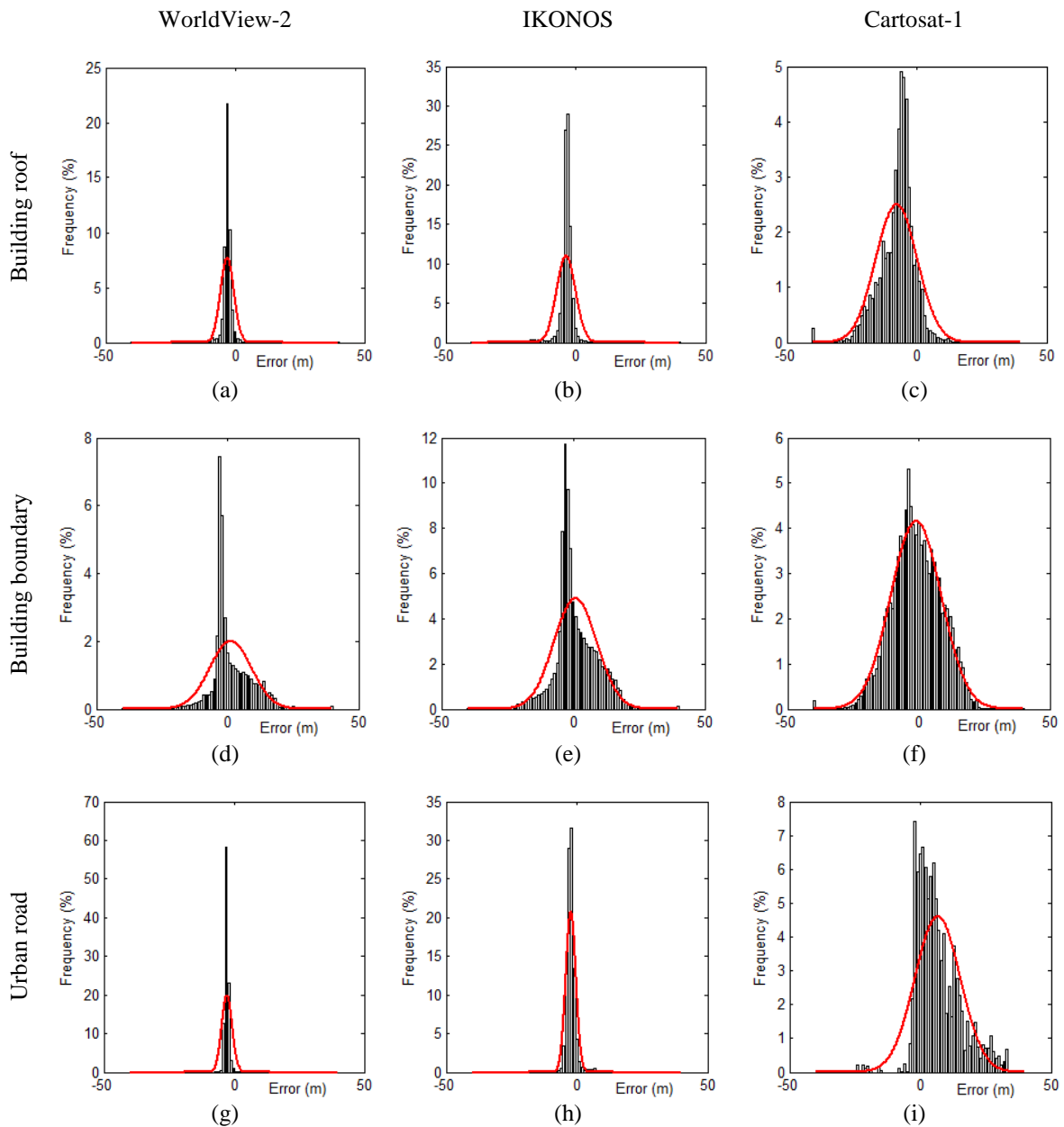


Fig. 2-7. Histograms of the height errors in urban areas from three satellites. Figs. a, b, c: Building roof regions; Figs. d, e, f: Building boundary regions; Figs. g, h, i: Urban roads.

In Figs. 2-8a and 2-8b, the small peak on the right can be explained by Fig. 2-6. Normally Stereo-DSMs feature higher height values in between trees than LiDAR-DSMs. It must be noted that the quality difference of the DSM from VHR WorldView-2 and HR Cartosat-1 in forest areas is not as obvious in urban areas, especially in the high-density forest regions (as shown in Fig. 2-8a and b). But in the built-up regions, the height value from WorldView-DSM is more accurate than from Cartosat-DSM.

## 2 DSM generation and quality assessment

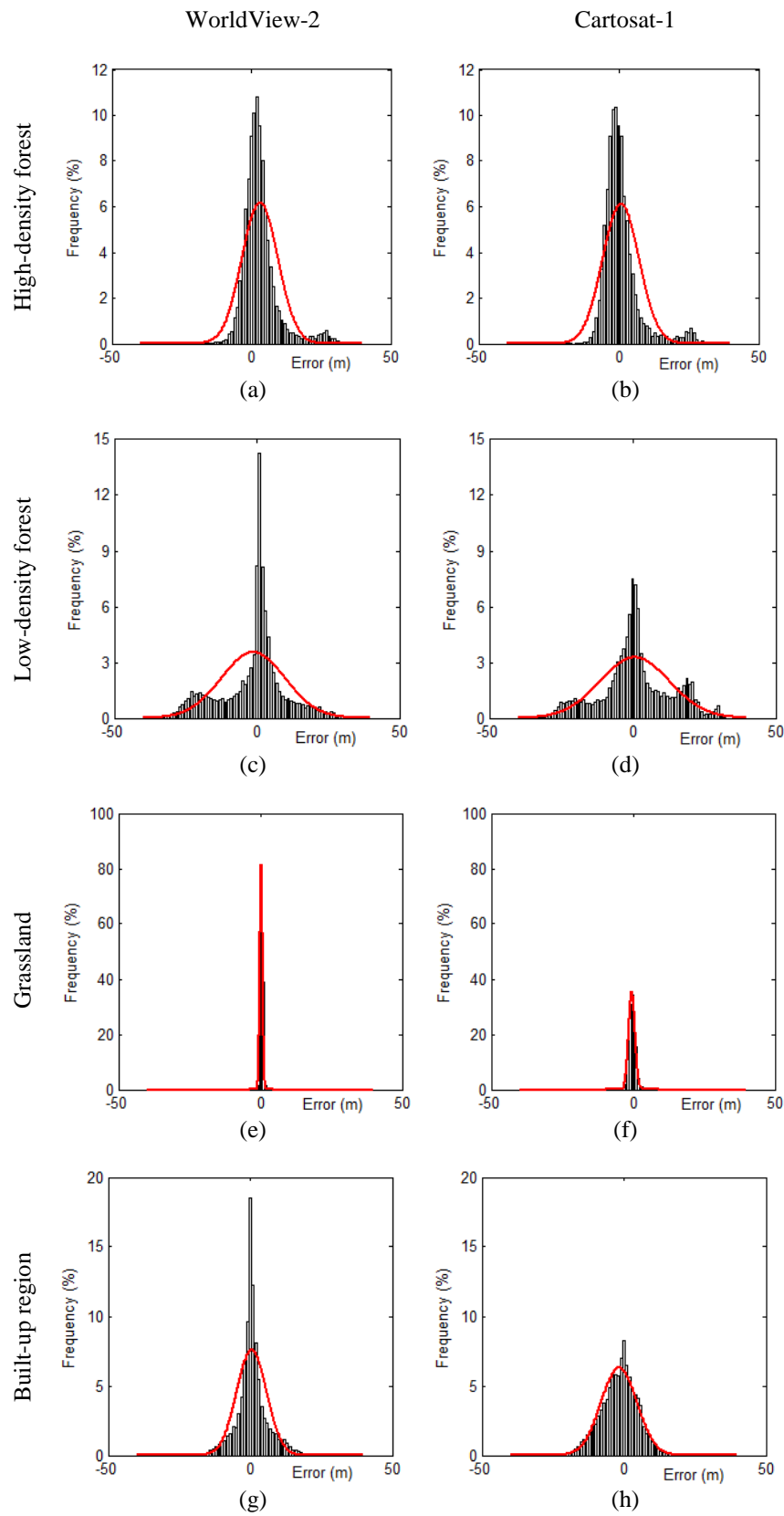


Fig. 2-8. Histograms of the height errors in and close to forest areas. Figs. a, b: High-density forest regions; Figs. c, d: Low-density forest regions; Figs. e, f: Grassland regions; Figs. g, h: Built-up regions.



Table 2-2. Accuracy measures for DSM quality assessment.

Number of checkpoints (N)	n
Vertical errors (Error)	$\Delta h =  h_{DSM} - h_{ref\_Height} $
Mean errors (MEAN)	$\hat{\mu} = \frac{1}{n} \sum_{i=1}^n \Delta h_i$
Standard Deviation (STD)	$\hat{\sigma} = \sqrt{\frac{1}{(n-1)} \sum_{i=1}^n (\Delta h_i - \hat{\mu})^2}$
Normalized Median Absolute Deviation (NMAD)	$NMAD = 1.4826 \times median( \Delta h_i - median(\Delta h) )$

Table 2-3. Quality assessment for seven land covers in three sensors.

Land cover type	Satellite	N	MEAN	STD	NMAD
Buildings	WorldView-2	216035	3.19	2.59	1.48
	IKONOS	216035	3.88	3.62	1.26
	Cartosat-1	216035	8.24	7.97	4.45
Build boundary	WorldView-2	125021	5.60	7.95	4.45
	IKONOS	125021	5.99	8.14	5.45
	Cartosat-1	125021	7.57	9.58	8.90
Ground (roads)	WorldView-2	49708	2.94	1.99	0
	IKONOS	49708	2.56	1.92	1.13
	Cartosat-1	49708	7.84	8.66	7.41
High-density forest	WorldView-2	12475	4.65	6.46	3.95
	Cartosat-1	12475	4.21	6.52	3.98
Low-density forest	WorldView-2	16997	7.91	11.26	6.01
	Cartosat-1	16997	9.02	12.21	8.82
Grassland	WorldView-2	25007	0.49	0.49	0.44
	Cartosat-1	25007	0.92	1.13	1.05
Built-up region	WorldView-2	11207	3.61	5.27	3.14
	Cartosat-1	11207	5.06	6.32	5.97

### 2.4.2 Statistical comparison

The accuracy assessments are measured using the difference between the Stereo-DSMs and the reference LiDAR-DSM ( $\Delta h$ ). The mean error ( $\hat{\mu}$ ) represents the mean absolute shift value for all pixels. The standard deviation ( $\hat{\sigma}$ ) is derived from the difference of two DEMs without the systematic error between them, which also indicates the surface smoothness of the generated DSM. Another robust statistical based accuracy measurement is the Normalized Median Absolute Deviation (NMAD) (Höhle and Höhle, 2009). This model can be used to evaluate the quality of

DSMs in the presence of outliers and non-normal distributions. The NMAD is thus proportional to the median of the absolute differences between errors and the median error. It can be considered as an estimate for the standard deviation that is more resilient to outliers in the dataset (Höhle and Höhle, 2009). Table 2-2 displays the formulas of these accuracy measures.

Calculations of accuracy measures have been carried out for the seven predefined land covers. The results are displayed in Table 2-3. Among them, the first three land covers (building, building boundary and road) are evaluated in the city area (Munich dataset as shown in Chapter 7). All DSMs in the city areas are resampled to 0.5 meter resolution. The other four land covers (high-density and low-density forest regions, built-up regions and grassland regions) are evaluated in countryside, near rural/forest area.

### 2.5 Discussion

The interpretation of qualitative and quantitative results leads to the following statements: in general, results of DSMs from stereo imagery are better when the resolution is higher, especially in urban areas. The quality differences for different sensors for building roofs are not as obvious as for building boundaries. The Cartosat-1 data show low quality for building boundary areas. As shown in Fig. 2-7f, the height errors are nearly normally distributed with a maximum at approximately 5%. In the profile comparison results (Fig. 2-5), even though the buildings in the Cartosat-DSM show higher elevation values for building areas relative to the ground areas, as can be seen in the profile comparison results, in areas with high building density, it is very difficult to separate single buildings (Fig. 2-5 at pixel 250). Thus, Cartosat-1 stereo data can hardly be used directly in building change detection in dense urban areas such as in the city of Munich. It can only be used when building boundaries or building footprints can be provided as reference.

WorldView-2 data are more suitable to generate DSMs in city areas, as they are a good match with the LiDAR-DSM, especially in areas featuring roads. As presented in Fig. 2-7, the height of building roofs and road matched LiDAR-DSM well. Although some false alarms existed in the building boundaries area, the situation is still much better than Cartosat-DSM. This demonstrates that the resolution of the original stereo imagery can directly influence the DSM quality in urban areas. The road error distribution map for Cartosat-1 data has more positive than negative values (Fig. 2-7i). This can be explained by the fuzzy building boundaries, which affect the road elevation. If the stereo matching failed in areas featuring roads, the interpretation of the unfilled DSM by using the neighborhood pixels would give a higher height value. Traffic on the road, e.g. buses or trucks, may also influence the elevation for all data sets. But in forest areas, the resolution of the satellite image does not influence DSMs' quality as obviously as in urban areas. As

displayed in Table 2-3, the mean, standard deviation and NMAD obtained from WorldView-DSM and IKONOS-DSM are quite similar in most of the regions near forest areas, except the built-up regions. Fig. 2-8 also shows a similar situation. This indicates that Cartosat-1 stereo imagery might be suitable for forest change monitoring.

In the profile comparison of forest area displayed in Fig. 2-6, the LiDAR-DSM is able to measure terrain height and canopy height for forest areas simultaneously, but the Stereo-DSMs only show the top-of-canopy height. However, because first pulse data also show part of the terrain height, which is not usually measured by Stereo-DSMs, there are some discrepancies between the DSMs from stereo matching and the LiDAR-DSM in low-density areas. As shown in Fig. 2-9, the nadir and forward view may observe different parts of the tree crowns, which leads to inaccurate matching results. This situation is even worse for low-density forest areas under the same stereo view angle and leads to discrepancies between the Stereo-DSMs and the LiDAR-DSM in low-density forest areas. However, the DSMs from both WorldView-2 and Cartosat-1 show the main trend of forest distribution with accurate canopy height, which can be very helpful in forest change detection or even timber volume change estimation. The mean forest canopy height is correct even in low-density regions: the main errors stem from the pixels between the trees or forest boundary area.

Without influence from buildings or other land cover, the grassland regions display a higher accuracy with mean errors less than one meter in Cartosat-DSM. This can be interpreted as underlying noise and will not introduce large false alarms in forest change detection. Similar to the assessment results in urban regions, the DSM of built-up areas from WorldView-2 is of better quality than from Cartosat-1, due to the lower image resolution of the latter. Therefore WorldView-2 data is more suitable in this case. In cooperation with the Bavarian State Institute of Forestry, a more detailed quality evaluation study has been conducted by comparing the DSMs with inventory points (Straub et al., 2013). In that study, DSMs were evaluated for both height and timber volume. According to the evaluation result, Cartosat-DSM is only slightly worse than WorldView-DSM. However, the differences among Cartosat-DSM, WorldView-DSM and LiDAR-DSMs are quite limited.

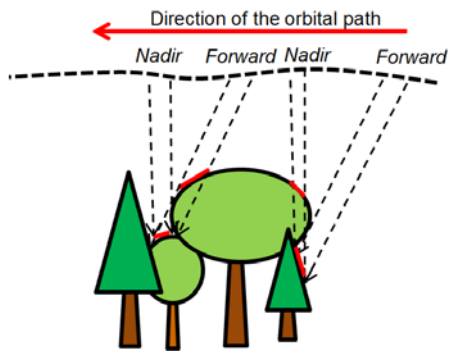


Fig. 2-9. Stereoscopic image acquisition in low-density forest areas.

## 2.6 Summary

This chapter provides an introduction to the available stereo imagery and how the DSMs can be generated. In this dissertation, all of the DSMs have been generated using SGM. Thus, the SGM-based DSM generation and refinement procedure has been described in detail. In order to better understand the accuracy of Stereo-DSMs from difference sensors, they are compared with LiDAR-DSMs. As it is a necessary pre-process of the comparison work, a 3D co-registration approach has been described in this chapter. This 3D co-registration guarantees that all of the compared pixels indicate the same location. In the evaluation, both visual and numerical quality evaluation methods have been applied. Firstly, profiles of all Stereo-DSMs have been compared with LiDAR-DSM in both urban and forest regions. Secondly, the height difference between these Stereo-DSMs and LiDAR-DSM, named as errors, has been analyzed using the histogram and statistical parameters: mean errors, standard deviation and NMAD. It has been analyzed for various land covers respectively.

## **3. Review of change detection techniques**

### **3.1 Introduction**

The purpose of land cover change detection is to identify changes between two or more images of the same location acquired at different times. This has always been one of the most heavily investigated topics in satellite image applications. Numerous detection methods and many kinds of image are used according to the different applications and user requirements (Singh, 1989; Bruzzone and Serpico, 1997; Lu et al., 2004). Among these applications, the change of interest can be of any possible class or only pertaining to specific objects, for example buildings or trees. The standard automatic or semi-automatic change detection methods can be broadly divided into pixel- and region-based methods. Depending on the data sources used, change detection can be classified into 2D change detection and 3D change detection.

Hereafter, the most important change detection methods are reviewed. Later, the height information involved in 3D change detection, called DSM-assisted change detection, which is developed in this work, is discussed.

### **3.2 2D change detection**

Change detection algorithms are based on the features which can be extracted from two data sets. These features can be grey values or height values, obtained directly from the image and digital surface model. They can also be generated with statistical methods by using information from multi-spectral or hyperspectral channels. Numerous approaches have been developed for change detection using only the 2D satellite images. Many articles reviewing the existing change detection methodologies can be found, such as, Singh (1989), Coppin and Bauer (1996), Macleod and Congalton (1998), Mas (1999), Lu et al. (2004), and Radke et al. (2005). Recent developments in change detection can be divided into pixel-based and region-based methods. Both require that the multi-temporal remote sensing images are properly co-registered.

#### **3.2.1 Pre-processing**

The quality of the images directly influences the performance of any feature detection, and thus change detection. Multi-temporal images which are used for change detection are often acquired

### *3 Review of change detection techniques*

by different sensors under different atmospheric conditions. Thus, co-registration in radiometry and geometry is necessary to make these images comparable. Pre-processing involves (1) atmospheric calibration and/or radiometric normalization between multi-temporal images; and (2) geometric co-registration.

- **Radiometric correction**

Many change alarms are produced when multi-temporal and multi-sensor data are employed, and if the strength or position of the light source (sun) changes (Radke et al, 2005). In this case, appropriate radiometric correction is required for successful change detection (Paolini et al., 2006). Two types of radiometric correction are usually employed: absolute and relative correction.

Absolute radiometric correction requires various calibration techniques, including modelling algorithms to describe the ground reflectance and atmospheric absorption etc. (Bowen, 2002; Schroeder et al., 2006). Unfortunately, for many change detection applications, absolute radiometric correction is expensive and impractical. Song et al. (2001) demonstrated that absolute radiometric correction can improve the accuracy of change detection, but more complicated algorithms did not necessarily lead to higher accuracy.

The relative correction takes one image as a radiometric reference image, and adjusts the radiometric properties (including the grey value range and grey value distribution) of the other image according to this reference image. After the correction, the same land cover in the two images exhibits similar grey values. The real reflectance value on the ground is not necessarily considered (Radke et al., 2005; Bavolo, 2006). The relative radiometric correction methods include dark object subtraction (Chavez Jr, 1988) by selecting pseudo-invariant features (Du et al., 2002). The selection of the radiometric correction method depends on the requirements of the application.

- **Geometric co-registration**

In the change detection procedure, in order to compare images from two dates, the co-registration between the two images is also very important. The accuracy of the geometric co-registration directly influences the co-responding features, especially when pixel-based change detection methods are adopted. In the literature, there are two main approaches for 2D image co-registration: a) semi-automatic and b) automatic co-registration. For the semi-automatic method, the matching control points are manually extracted from the two corresponding images so that the parameters for the transformation model can be calculated based on these points (Ton and Jain, 1989). A detailed review of automatic registration has been provided by Brown (1992), and later by Wyawahare et al. (2009). Automatic co-registration is based on automatic feature detection,

followed by feature matching methods such as least square matching or SIFT matching. Using the matched points, the parameters for the transformation model can be estimated.

A common agreement has been achieved that to perform reliable change detection on multi-temporal images, an accurate (in the order of at least 0.3 pixel) spatial co-registration is required (Singh, 1989). Detailed research on how co-registration accuracy can influence change detection results has been performed by Dai and Khorram (1998) and Roy (2000). Dai and Khorram (1998) used satellite images as an example to illustrate that higher registration accuracy was required to achieve reasonable change detection results. Roy (2000) demonstrated that when multi-temporal images are used, mis-registration errors can significantly influence the thematic change detection results (e.g. comparison of vegetation mask).

#### 3.2.2 Pixel-based change detection

Pixel-based change detection methods are based on features extracted by combining images at pixel level, e.g. pixel-based subtraction, ratioing and image regression. When multi-spectral information is available, a feature vector can be obtained for each pixel. Beside the pixel spectrum, more features, such as features after color transformation or texture features, can be extracted. To reduce data redundancy between bands and to emphasize the difference between objects and background, image transformations can be performed, such as color space transformation or Principal Components Analysis (PCA) (Lillesand and Kiefer, 1987), Gramm-Schmidt, and Chi-square transformation. Fung (1990) compared image differencing, principal components analysis and tasseled-cap transformation based on Landsat TM data. These combinations can improve the mathematical combination result, and the original images can be filtered or transformed to achieve a better change map.

In the literature, various techniques have been proposed for performing change detection. For the 2D method, the main methods are binary change detection, which aims to separate changed from unchanged areas, and 'from-to change detection', which provides a matrix of change directions with results including the nature of change and the spatial pattern of changes.

*Image differencing*, which works by simply subtracting the pixel grey values of the images from two dates, is one of the most basic change detection method (Hayes and Sader, 2001). This technique has been widely used and is documented in the early studies of remote sensing (Singh, 1986; Muchoney and Haack, 1994; Green et al., 1994; Coppin and Bauer, 1996; Macleod and Congalton, 1998). Except for direct satellite image subtraction, image differencing has also been developed for change feature subtraction. For instance, Normalized Difference Vegetation Index

### 3 Review of change detection techniques

(NDVI) subtraction is used to monitor vegetation changes (Mas, 1999), while the normalized difference of near infrared bands and red bands can highlight vegetation (Singh, 1986). Hayes and Sader (2001) clustered the NDVI images from each date to several levels, and used the level differences to distinguish change / no-change. Post-classification is also a kind of image differencing, but it uses the class cover instead of images. Similar to image differencing, pixel-based image ratio is another widely applied change detection method (Singh, 1989).

*Change vector analysis (CVA)* was proposed by Malila (1980). A change vector can be described by an angle of change (vector direction) and a magnitude of change from date 1 to date 2 (Jensen, 1996). Originally only the multi-spectral information from remote sensing images are used for the features; then the  $N$  dimensional Euclidean distance shows the change magnitude. A larger magnitude indicates a higher possibility of change. The changes can be identified by manual or automatic thresholding. The direction of these vectors can be used to classify different kinds of change (Chen et al., 2003). Bovolo and Bruzzone (2007) have tried to improve the accuracy of change detection by introducing CVA to the polar coordinate domain instead of Cartesian coordinates.

*Multivariate alteration detection (MAD)* was introduced to the field of change detection by Nielsen et al. (1998). It is a well-established linear change detection method (Nielsen et al., 1998; Marpu et al., 2011). The main aim of this method is to establish a better background of 'no change', which consequently enables the identification of real changes. The most important advantage of MAD is that it is an unsupervised change detection method. It considers all of the feature channels generated from the images from the two dates  $F = (F_1, F_2, F_3, \dots, F_k)^T$  and  $G = (G_1, G_2, G_3, \dots, G_k)^T$ . The changes can be expressed by the linear combination of the features from two dates (as shown in Eq. (3-1)).  $a_i$  and  $b_i$  are the coefficients calculated by applying canonical correlation analysis.

$$\begin{bmatrix} F \\ G \end{bmatrix} \Rightarrow \begin{bmatrix} a_1^T F - b_1^T G \\ \vdots \\ a_k^T F - b_k^T G \end{bmatrix} \quad (3-1)$$

Iteratively reweighted-MAD (Canty and Nielsen, 2008) is an iterative scheme to put high weights on areas with little change. These weights are included in the calculation of the mean, variance and covariance. The iteration stops when the largest absolute change in the canonical correlations reaches a preset value, such as  $10^{-6}$ .



*Statistical based similarity measures* are able to highlight changes by computing information's theoretical similarity. Many change indicators, such as difference ratios or neighborhood correlation (Im and Jensen, 2005), are less meaningful when multi-sensor images are adopted. For such cases, similarity measures have been proposed to extract change indicators. Information similarity measures, such as Mutual Information, are primarily used for image co-registration (Suri and Reinartz, 2010; Reinartz et al., 2011) but they have also been introduced to change detection (Inglada and Mercier, 2007; Bovolo et al., 2008; Alberga, 2009; Gueguen et al., 2011).

*Kernel-based change detection.* Introducing machine learning algorithms to remote sensing image processing has led to considerable improvements in this field, especially with regard to the kernel method and support vector machines. Initially these methods were mainly used for classification. It has been proven that nonlinear decision boundaries are more flexible and effective than a linear classifier (Camps-Valls and Bruzzone, 2009). Recently kernel-based methods have also been adopted in change detection (Volpi et al., 2012). Kernel PCA (Nielsen and Canty, 2008) and kernel minimum noise fraction (Nielsen, 2011) have been proposed to improve the accuracy of change detection.

*Post-classification comparison.* This approach is based on analysis of two or more independently produced classification maps, or extracted object maps, such as building change detection based on the comparison of two building footprints. As discussed in Singh (1989), since the classification of each map is performed independently, the multi-sensor problem can be minimized. In this case, precise co-registration is less important than for other change detection methods. However, the classification result directly influences the change detection result (Singh, 1989; Mas, 1999). Compared to two-time comparison, multi-temporal comparisons are less sensitive in their classification accuracy. Liu and Zhou (2004) proposed a rule-based rationality analysis to improve the post-classification change detection results based on multi-temporal images. This approach has also been widely used to update databases, for instance for road map updates (Mena, 2003) and building footprint map revision (Champion et al., 2008). The existing database can be used not only as the classification result of one date, but also as training data (Matikanien et al., 2003). The information extracted from existing maps can enhance the image interpretation procedure, which can reduce the change search space and minimize false alarms (Bentabet et al., 2003; Bouziani et al., 2010).

### 3.2.3 Region-based change detection

The increasingly high resolution of satellite images allows the extraction of more detailed changed objects with higher accuracy, but also introduces some false alarm that are not related to the changes of interest. The main drawback of pixel-based change detection, called the “salt and pepper” effect, is a result of these false alarms. Therefore, region-based change detection methods have gained more interest in recent years (Blaschke, 2010). Instead of analyzing pixels independently, region-based approaches take the pixels inside one meaningful homogeneous region. Some papers refer to this method as ‘object-based change detection (OBCD)’ (Hall and Hay, 2003; Blaschke, 2005), as one feature of OBCD is to extract meaningful objects from images. Initially, these meaningful objects were obtained from Geographic Information System (GIS) databases (Coppin and Bauer, 1996; Walter, 2004; Durieux et al., 2008). However, in more recent work, these objects are often extracted using segmentation: thus, the obtained regions might not always be a whole object, such as a complete building or bridge, but very probably only one part of it. Therefore, referring to these methods as ‘region-based’ is more suitable.

Region-based change detection evolved from ‘region-based image analysis’, which combines spectral images with segmentation or GIS data for image understanding or land cover classification. The GIS database can also be used as training data, since it includes not only the boundary but also the attributes of each region (Walter, 2004). As the use of GIS data is limited to the available data sources, object-based change detection using image data is usually preferred, along with the development of automatic segmentation techniques (Comaniciu and Meer, 2002; Ning et al., 2010). After segmentation, the images can be divided into a number of homogeneous regions. The research in this direction is thus focused on obtaining an appropriate segmentation level for all land covers of interest. Bruzzone and Prieto (2000) stated that the original segments that can be used for change detection should be elementary ‘homogeneous regions’. However, it is difficult to provide an ideal definition of ‘homogeneous’ when various objects are of interest: for example, the cars on the road, the cars in car parks etc. Multi-level segments are preferred for detecting changes to various classes of objects (Bovolo, 2009).

Hall and Hay (2003) proposed a multi-scale change detection framework for forest changes. In their framework, the multi-scale segmentation was derived using the watershed methodology. After segmentation, the mean grey value change of each segment was used as a change feature for change detection. In their experiments, the finest segmentation resulted in the best change detection accuracy. Instead of using only panchromatic images, Desclée et al. (2006) successfully monitored forest changes based on multispectral SPOT images. In this paper, a multi-data segmentation approach was proposed for the first time. The mean and the standard deviation of

each region were calculated based on the pixel-wise difference images. Multivariate iterative trimming was adopted to separate the forest change from no-change regions.

Besides segmentation, the initial region boundary can be directly obtained from existing GIS data. Walter (2004) used multi-spectral channels of airborne data (blue, green, red and near-infrared) for classification. Besides the mean value and variance of these channels in each GIS region, the gray-level co-occurrence matrices textures from the blue channel, the vegetation index and the pixel-based classification results were used as input for the classification. The classification results were compared with existing GIS data for change detection.

A post-classification method in object range has been proposed by Dronova et al. (2011). In the region-based classification procedure, NDVI and normalized difference water index and brightness obtained from multi-spectral channels were adopted to separate vegetation, water and sand classes.

A significant body of research has been performed with a focus on the comparison of region-based and pixel-based change detection methods for remote sensing data. Walter (2004) used region-based classification results from two dates to generate a land cover change map. The original regions were obtained from existing GIS data. Region-based features can thus be extracted, which can be used in classification. Desclée et al. (2006) used the region-based approach for forest change detection, where it exhibited a much higher Kappa Index of Agreement (KIA = 60%) than the pixel-based method (KIA = 49%) using the same features from multi-spectral satellite images. Im et al. (2008) compared region- and pixel-based change detection methods. The results have shown that the region-based change detection method can reach a higher KIA (about 90%) than the pixel-based technique (KIA = 80% ~ 85%). Duveiller et al. (2008) applied the forest / non-forest classification method from Desclée et al. (2006) in region-based deforestation detection from 571 image pairs, where an initial image selection was performed to exclude all poor quality images and image pairs with no-forest change. Aguirre-Gutiérrez et al. (2012) compared pixel- and region-based methods and their combination in land cover classification. A higher accuracy was achieved based on a combined classification method along with a region-based change detection analysis. Considering the above comparisons between the pixel- and region-based change detection methods, it can be concluded that in general, the region-based method performs better than the pixel-based method if a proper segmentation level can be achieved.

#### 3.2.4 Binary change map generation

The change map obtained from both pixel- and object-based change detection methods must be processed in an appropriate way to obtain the final change mask. In a useful change map, the

### 3 Review of change detection techniques

changed areas are highlighted by a considerably higher value, so that these areas can be separated from the background using a threshold value  $T$ . However, in many cases, automatically obtaining this threshold value is not an easy task. Several thresholding methods have been proposed for this purpose (Melgani et al., 2002).

Methods applicable for choosing a proper  $T$  were surveyed and analysed by Rosin (2002; Rosin and Hervás, 2005). Three of the most popular thresholding methods are Otsu's algorithm (Otsu, 1979), the Iterative Self-Organizing Data Analysis Technique (ISODATA) (Ridler and Calvard, 1978), and Minimum Error thresholding (Kittler and Illingrowth, 1986).

Otsu's algorithm (Otsu, 1979) defines that an image is composed of objects and background. A discriminant analysis is performed by minimizing the intra-class variance ( $\sigma_{\omega}^2(t), t = 0, 1, \dots, n$ ).  $n$  is the maximum value in the image.

$$\sigma_{\omega}^2(t) = \omega_1 \sigma_1^2(t) + \omega_2 \sigma_2^2(t) \quad (3-2)$$

Weights  $\omega_i$  are the probabilities obtained from the image histogram that are separated by a threshold  $t$ .  $\sigma_i^2$  are the variances of the two classes.

ISODATA (Ridler and Calvard, 1978) takes an initial threshold value to separate the image ( $I$ ) into object and background. Then the average pixel values of classified objects and the background are calculated. An updated threshold value is the average of those two values. This process is repeated until the threshold value does not change any more.

The detail calculation procedure can be represented as:

$$T_{n+1} = \frac{\mu(I(i,j), I(i,j) < T_n) + \mu(I(i,j), I(i,j) \geq T_n)}{2} \quad (3-3)$$

Background:  $I(i,j), I(i,j) < T_n$

Object:  $I(i,j), I(i,j) \geq T_n$

$I(i,j)$  is the pixel value in the image  $I$  at the position  $(i,j)$ .

The final  $T$  is chosen when  $|T_{n+1} - T_n| < \varepsilon$ ,  $\varepsilon$  is a predefined small value.

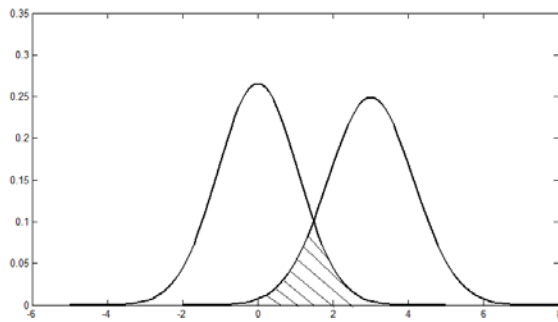


Fig. 3-1. Minimum Error thresholding.

Minimum Error thresholding (Kittler and Illingworth, 1986) also obtains the threshold through an iterative procedure. It assumes that the image histograms follow a Gaussian mixture model, and that each class (object and background) follows a normal Gaussian distribution. As shown in Fig. 3-1, the threshold value ( $T$ ) is located at the intersection points of these two normal distributions. After iteration the final value  $T$  which leads to the smallest size of the shadowed area is selected.

### 3.2.5 Shortcomings of 2D change detection methodology

Although numerous change detection methods exist, many methods are tested using ideal pairs of images. One of the main problems related to the 2D change detection based on satellite images lies in the lack of height information, meaning that only changes that influence reflectance values or local texture changes can be detected, while changes in the vertical direction – such as height changes of buildings - are not available and are therefore ignored. Furthermore, in some cases, the changes detected from spectral images do not represent real changes, but just the image difference caused by changes in satellite position, sensor properties or seasonal change (Heller et al., 2001). Moreover, when only certain specific objects of change are of interest, for example buildings, it is very difficult to separate those changes from other changes without height information. Many irrelevant changes will be mixed with building changes, especially when the data are acquired from different sensors or under different imaging geometries (e.g. viewing directions). Thus, recently DSM-assisted 3D change detection is attracting more interest, particularly as more 3D data become available.

## 3.3 DSM-assisted change detection

Singh (1989: 989) stated that: “The basic premise in using remote sensing data for change detection is that changes in land cover must result in changes in radiance values.” However, the value changes in radiance are not always produced by land cover changes. Besides different sensors and different atmospheric conditions, many seasonal changes also influence grey value

### *3 Review of change detection techniques*

changes in the image. Thus, recently DSMs have been adopted in change detection approaches, especially when only one or several land covers are of interest. Previous DSM-assisted building change detection methods can be classified into three categories: 1) DSMs are used to locate the change candidates, and these extracted change candidates are refined with spectral information; 2) DSMs are employed as change features; 3) DSMs are applied to assist building reconstruction, and are thus used as post-classification change detection.

#### **3.3.1 Change candidates**

The first approach starts by extracting ‘change candidates’ using the height information (Jung, 2004). With DSM subtraction, it is computationally easy to obtain the initial change map, which can be improved to a more precise building change map when additional information from the original image is employed. Many 3D change detection methods have started to use this approach. The changed pixels are detected through a simple subtraction of one DSM from another.

This approach is also used for DSM quality evaluation tasks (Zhang, 2005; Reinartz et al., 2006; Akca, 2007). Other studies focusing on change detection include the work of Murakami et al. (1999), who subtracted two DSMs from LiDAR in order to extract building changes, and Tong et al. (2012), who subtracted two DSM from pre- and post-earthquake IKONOS stereo pairs for building-damage detection.

However, in this category of 3D change detection, the quality of the generated DSMs determines the accuracy of the final change maps. In fact, mis-registration and significant height differences that may arise between DSMs generated from different sources often result in the detection of erroneous or irrelevant changes (false alarms). Therefore, some refinements are required based on these ‘change candidates’.

Fan et al. (1999) and Liu et al. (2003) detected house changes based on DSMs from airborne images. DSMs were computed on the digital photogrammetry workstation VirtuoZo. These “potential changed areas” were generated based on DSM subtraction. The candidate regions obtained were classified into new built, destroyed and rebuilt buildings using gradient direction histograms. These histograms were extracted from the original airborne images.

In Jung (2004), DSMs generated from airborne stereo imagery were employed to obtain an initial change map. The change regions from subtracted DSMs are used as ‘candidates’. Subsequently these candidate regions from four images (two stereo pairs) are classified into building and no-building areas based on graph features. Regions which have “building” within one dataset and “no-

building” in the other dataset are considered changed buildings. This method will fail if one building is only partly changed, or if buildings are rebuilt. Also in high density building areas, several buildings might exist in one candidate region. A false alarm will be produced if only some of them have changed.

Turker and Cetinkaya (2005) used the DSMs generated from pre- and post-earthquake stereo aerial images to detect collapsed buildings. They generated 5-meter resolution DSMs using the DSM tool SoftPlotter. The resolutions of the original pre- and post- earthquake stereo images were 0.5 meter and 0.25 meter respectively. They compared the resulting building change mask by using a threshold ( $T$ ) from 1 m to 10 m. Among them,  $T = 7$  achieved the highest KIA, while  $T = 4$  provides the highest average producer’s accuracy. However, the threshold value was shifted when other test areas were selected.

Zhu et al. (2008) detected building changes based on ADS 40 airborne stereo imagery. In their change detection procedure, initial ‘change candidates’ were derived from the height difference of the two DSMs. Height thresholding followed by low pass morphological filtering was used to eliminate the noise level. These change candidates were combined with the extracted building masks from two dates. The building masks were extracted separately for two DSMs by combining the building extraction result of two methods: Local surface normal angle transform and marker controlled watershed segmentation. Height change was analyzed only for the regions detected as buildings.

In our earlier work (Tian et al., 2010) we tried to refine the building change map using a box-fitting algorithm. To remove the noise within the initial change map, in the work of Chaabouni-Chouayakh et al. (2010) for example, post-processing steps such as morphological operations and contextual knowledge introduction have been proposed. These methods help to remove virtual changes and to preserve the real ones. When only building changes are of interest, shape features can also be used to refine the change map (Tian et al., 2011). Instead of a feature test, support vector machine was used in Chaabouni-Chouayakh and Reinartz (2011) to classify the real building changes and no-building changes based on these shape features.

#### 3.3.2 Fusion-based change detection

The second change detection approach employs fusion-based methods. Height information from DSMs is normally used as change or no-change features to highlight changes. In case the provided height information is inaccurate, it can be fused with other change features based on different data

### *3 Review of change detection techniques*

sources in order to further improve the change detection accuracy. They can be fused either at feature level or at decision level.

Champion (2007) fused an existing database, aerial images and DSMs to extract changes. Firstly, building boundaries were extracted from the DSMs; boundaries covered by vegetation were not counted. The similarity of the extracted boundaries and 2D contours from the existing database were measured to verify the database. In the second step, a new building mask was obtained by comparing the above-ground masks of two dates. The final building masks were obtained by thresholding the normalized DSMs (nDSMs).

Sasagawa et al. (2008) combined the pixel-changes generated from least square fitting of two images, and height changes detected by DSM subtraction, to generate the final change map containing three channels (Green: pixel change; Red: positive DSM change; Blue: negative DSM change). This change map was provided together with spectral imagery for manual interpretation. In their work, the generation method and the accuracy of these DSMs were not mentioned.

James et al. (2012) detected changes by using a LiDAR-DSM and historic maps. These maps provide both the historic height information and land cover boundaries, and were fused with newly obtained DSMs. However, only the height difference was considered to extract erosion, deposition and volumetric changes.

#### **3.3.3 Post-classification**

Building extraction based on DSM and satellite / airborne images is a task of great importance. Many excellent image-based building extraction methods are available, where an existing DSM can be used to improve the extraction efficiency and accuracy (Mayer, 1999; Sohn and Dowman, 2007). Based on building extraction results, building change detection can be performed by comparing the extracted building mask / boundaries to the existing building footprints, which are available in most industrial countries (Champion et al., 2008). Alternatively this method can be implemented by comparing building extraction results of two dates. We denominate this approach as ‘post-classification’. This approach has been preferred for building change detection, especially if outdated building footprints are provided.

Besides acting as reference building location status, the outdated building footprints can sometimes be used as training data (Champion et al., 2009; Matikainen et al., 2010) to assist in building extraction. Walter (2004) derived the training area from a GIS database, and used it in a maximum



likelihood classification procedure. That paper focused on extracting residential settlement objects, which are later compared with the GIS database to highlight the land use changes. Both airborne multispectral images and LiDAR data were used to classify residential objects based on the maximum likelihood method. First, pixels were classified into houses, streets, grassland and trees by using the existing GIS database, vegetation index and nDSM. The subsequent object-based classification focused on separating residential settlement objects from industrial areas. The classification results were compared with existing databases to find changes.

Different from Walter (2004), Olsen (2004) extracted building masks using a pixel-based classification. Changed buildings were extracted by performing a pixel-wise comparison with the existing map database. To refine the extracted building mask, regions with small size were removed from it. Since this method uses an existing database as training data, high classification accuracy can only be achieved if new built buildings have the same spectral information as the existing building (i.e. are built with the same or similar construction material).

In order to also extract the new object boundaries, instead of using objects provided by a GIS database, Matikainen et al. (2010) extracted initial segments from laser scanner derived DSMs based on a height thresholding. Training data were selected from these segments by combining information from existing building maps. A classification decision tree was automatically generated based on these training data. Morphological slope from DSM, NDVI and texture features from airborne images were included as features in that classification tree.

If the new buildings were built with different materials than the existing buildings, or if many buildings were removed from the outdated building maps, many false alarms might be introduced when using these maps as training data. Therefore, other methods focus on unsupervised classification methods. Heller et al. (2001) extracted changes by comparing the 3D geometry derived from images of various sensors. After matching the images, they found multiple common XY-coordinate matches. They extracted the significance level curves for the values of confidence with which the matched pixels could be compared. However, this algorithm was only focused on detecting the location of the changes: more detailed change information could not be detected using this method.

Olsen and Knudsen (2005) replaced the supervised classification method from Olsen (2004) with a rule-based method to achieve much clearer building boundaries. The absolute building height was generated from a DSM based on morphological methods. The objects in the above terrain mask

### *3 Review of change detection techniques*

were classified into building objects and no-building objects based on defined rules. The final building mask could then be compared with a prior database to obtain building changes.

Rottensteiner et al. (2007a) proposed a decision fusion-based classification method using vegetation cover maps, height values and the surface roughness from DSMs. The fusion model has been designed to separate these four classes: buildings, trees, grassland and bare soil. The detected buildings were compared with an existing map to obtain a post-classification change detection result. This building extraction method has been proven to be highly accurate and efficient, but the fusion model has to be changed when different change objects are of interest.

Waser et al. (2008) used DSMs generated from CIR aerial images for forest change detection. These DSMs were co-registered and normalized based on LiDAR data. Fractional tree / shrub covers from two dates were extracted separately based on height and a fuzzy classification. The decrease and increase of forest and other wooded areas were analyzed.

A special post-classification method was suggested by Choi et al. (2009). Instead of directly classifying the whole images, only the obtained change candidate regions were classified. These original change candidates were generated through DSM subtraction results. After refining these change candidates using morphological filters, the roughness, size and height of these remaining segments were calculated separately for each dataset. These regions were classified into ground, vegetation and building classes. Finally, a post-classification was analyzed for each candidate region. It is assumed that building reconstruction (increase or decrease in the height of buildings) can also be detected with this method. However, this method requires DSMs with very high accuracy.

### **3.4 Summary**

This chapter has presented an exhaustive review of the available literature in the field of change detection, especially 3D change detection. Many practical 2D change detection approaches have been presented and several applications using these techniques have also been reported. These 2D change detection techniques have recently been adopted for 3D change detection purposes. Previous studies have demonstrated that 3D change detection that works by fusing DSM and spectral imagery is possible and relatively good results can be achieved.

However, until now, most of these DSM-assisted change detection methods have been carried out with DSMs from LiDAR or airborne stereo imagery, which feature higher quality than DSMs

generated from satellite stereo imagery. The potential errors present in these DSMs are rarely considered. Moreover, test datasets, especially in the building change studies, have been rather limited in size, featuring simple buildings with lower density rather than complicated shaped buildings with high building density. In addition, the information that can be extracted from images has not been fully investigated.

Being limited to the quality of the Stereo-DSMs as described in Chapter 2, it is hard to achieve precise change detection results using only DSMs. Therefore, DSMs should be used in combination with the spectral information from the original stereo images. As the DSMs are generated using stereo images, spectral information from the same time and area is always available. After orthorectification of these images using the generated DSM, the ortho-image and DSM are well co-registered and can be used together. Thus, the main challenge here is how to resolve the quality restriction of DSMs by using the additional information available in the corresponding spectral images.

Although some change detection studies based on the Stereo-DSM are available, the DSMs mentioned in these papers are generated using now outdated matching algorithms. Therefore, the DSM quality is relatively low. With the recently proposed SGM matching based method, the accuracy of the generated DSM is improved considerably. Corresponding 3D change detection approaches should be designed according to the quality and character of these DSMs.

Furthermore, in the post-classification based change detection approach, the accuracy and efficiency of building extraction methods will directly influence the change detection result. A robust building extraction method is required for this purpose.

Overall, it can be stated that more work is needed to develop a useful, automatic and robust DSM-assisted change detection approach, and in particular, to be able to use DSMs from satellite stereo imagery. Satellite imagery is often easily available and cheaper than laser scanning or airborne stereo data for change monitoring in large regions, making it more suitable in applications such as emergency situations, e.g. after disasters.

In this dissertation, three approaches are proposed, which mainly differ in the way the DSMs are introduced into the process. These three approaches all have their advantages and disadvantages. All three rely on a combination of statistical evaluations, fusion processes and filtering methods. They have been developed according to the objects of interest and focus on DSMs with different

### *3 Review of change detection techniques*

quality. Hereafter these three methods are described in more detail, together with the procedural framework in which they will be used in the experimental section.

## 4. DSM-assisted change localization

### 4.1 Introduction

Correct height information can be very important when changes to specific objects are of interest, and the change to these objects can directly influence their height, as in the case of buildings and trees. Thus, using height to separate the change of interest from other changes (e.g. seasonal changes), and using height change to extract initial change candidate areas, are very efficient approaches. The continuous improvement to stereo-matching technology has greatly increased the quality of the Stereo-DSMs. However, as analyzed in Chapter 2, distance still exists when compared with the LiDAR-DSM. One of the main challenges for a change detection algorithm is the handling of potential incorrect height information in these Stereo-DSMs. Therefore, the change mask generated from DSM subtraction can only provide candidate change regions, which can be refined with further process. In this chapter, refinement methods are proposed in constructing a relationship between DSMs and satellite images. This approach is known as DSM-assisted change localization.

If two high quality DSMs from different dates are available, for instance two LiDAR-DSMs, directly subtracting the two DSMs can already produce results with high accuracy (Murakami et al., 1999). A change mask (*Mask*) can be generated after a threshold value (*T*) is applied to the height difference map (*D*).

$$Mask = D > T \quad (4-1)$$

The threshold value can be adjusted according to the requirements of the respective application. However, for DSMs generated from spaceborne stereoscopic images, such a subtraction is generally not applicable, especially in the case of different sensors, illumination and convergence angle properties. Fig. 4-1 shows the profile comparison result of a Stereo-DSM and a LiDAR-DSM. As can be seen, the Stereo-DSM (with solid line) has a much rougher surface and more blurred features. Assuming that the real 3D changes between the LiDAR-DSM and Stereo-DSM are of interest, with a simple subtraction, although there are no real changes between them, the rough surface will introduce many false alarms. This problem cannot be solved by applying a simple threshold on the difference map. These false alarms caused by blurred or even falsely located building edges cannot easily be eliminated. As illustrated on the rightmost building in Fig.

#### 4 DSM-assisted change localization

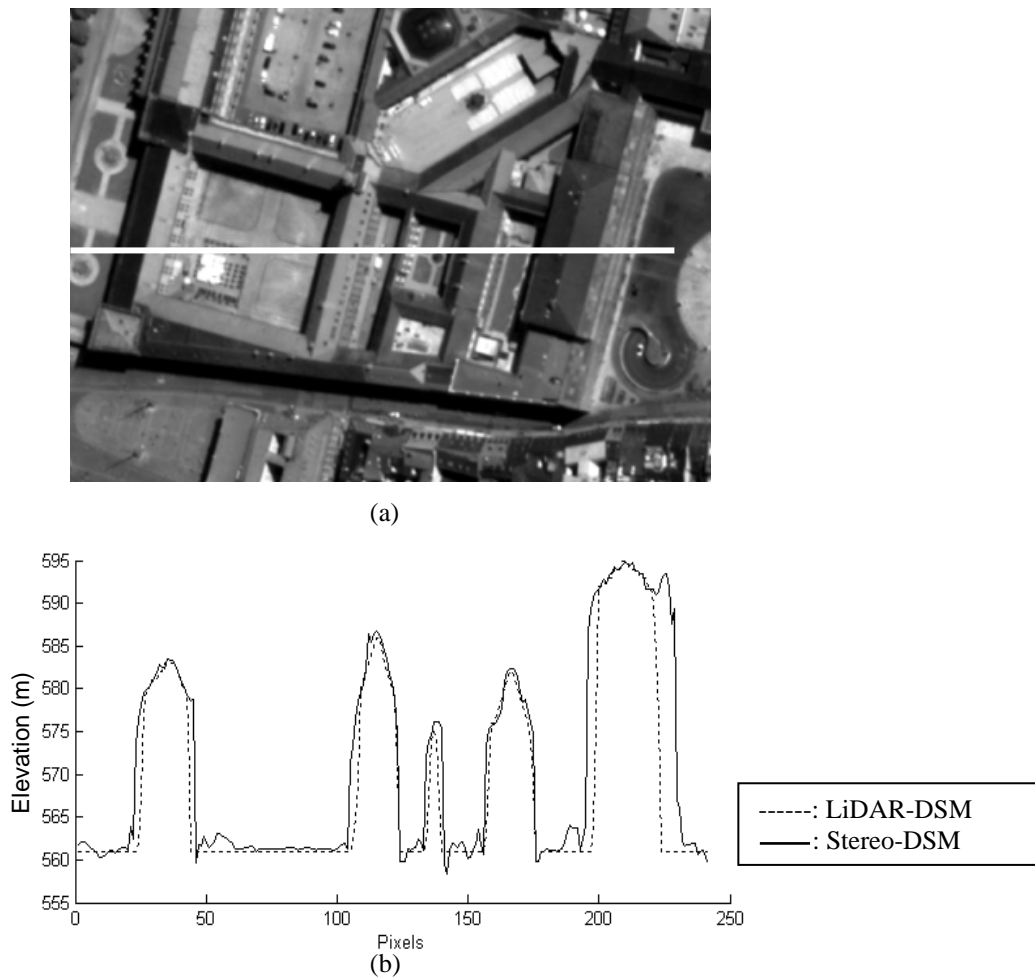


Fig. 4-1. Comparison of DSMs generated from LiDAR and stereo matching. (a) Location of the test area; (b) Profiles along the white line in (a).

4-1, this building is about 10 m wider in the Stereo-DSM than in the LiDAR-DSM, and the exceeding part shares the same height with that building (about 30 meters). Removing these false alarms and thus highlighting the real building changes is the main purpose of the presented approach.

The developed DSM-assisted change localization is a coarse-to-fine change detection method. If the original stereoscopic images also include multispectral channels, the extracted vegetation and shadow classes can be used to refine the initial change map; if not, denoising or filtering of the difference map is necessary to obtain a cleaner result. The characteristic quality of the DSM should be considered when refining the results. As illustrated in Chapter 2, in urban areas, the building boundary regions are of relatively poor quality in comparison to the building roof areas, and some noise and false height values are mixed into the correct height information. These problems can be solved via the following two processing steps: 1) improved difference image generation procedure;

2) through a refinement of the mask. If the original multi-spectral data are also provided, the land cover maps extracted from multispectral channels (e.g., shadow, vegetation) can be directly used to refine the initial height change map.

## 4.2 Robust DSM differencing

As mentioned in Chapter 3, image differencing is a traditional change detection method, because different land covers usually exhibit different radiometric properties. The change of land cover directly influences the intensity values in the images: thus, changes can be highlighted in the difference image. In theory, this approach should also be applicable for DSM-based change detection, especially for the change detection of specific objects of interest, such as buildings and forest. However, it is difficult to obtain a precise image difference map using direct DSM subtraction, as discussed above (see Fig. 4-1). Another difficulty is the co-registration of the DSMs. Even if two images or DSMs are co-registered, the true location of the corresponding pixel's center point may be anywhere within a  $3 \times 3$  pixel window surrounding that pixel (Goodchild, 1994). In addition to that, when DSMs with different resolutions are used, image resampling is necessary to make sure the pixels have the same geo-size.

Given these limitations, it is clear that directly subtracting two DSMs cannot result in the ideal change map. Considering the neighborhood pixels should improve the accuracy. Shown in Fig. 4-2 are several newly constructed buildings located in the middle of Fig. 4-2b. Fig. 4-2c features the direct subtraction results of the corresponding DSMs, showing only the positive values: all the negative values have been set to '0' for better display. As can be seen from this difference image, some unchanged buildings around them also show quite high positive change values. Moreover, some false height values exist in the middle of the rightmost building. All of these false alarms impair further change detection refinement. Therefore, this height difference map should be improved in the first step.

One possible way to refine the height change map is to first subtract the two DSMs to generate an initial change map. After that, a filter kernel can be used to filter the data using a sliding window; the value of the pixel in the center of the window should be adjusted considering all of the pixel values inside that window (Schindler, 2012). This image filtering works well for images; however, DSMs are different from images. Although the false alarms are also called 'noise' here, they are not comparable to the noise in images. The false alarms from DSM are distributed irregularly, the values cannot always be explained easily and they can be small or very large. In some cases, normal image filtering cannot remove these false alarms, and at the same time they blur the building boundaries (as shown in Fig. 4-3a). Gaussian filtering also does not improve the results (Fig. 4-3b).

#### 4 DSM-assisted change localization

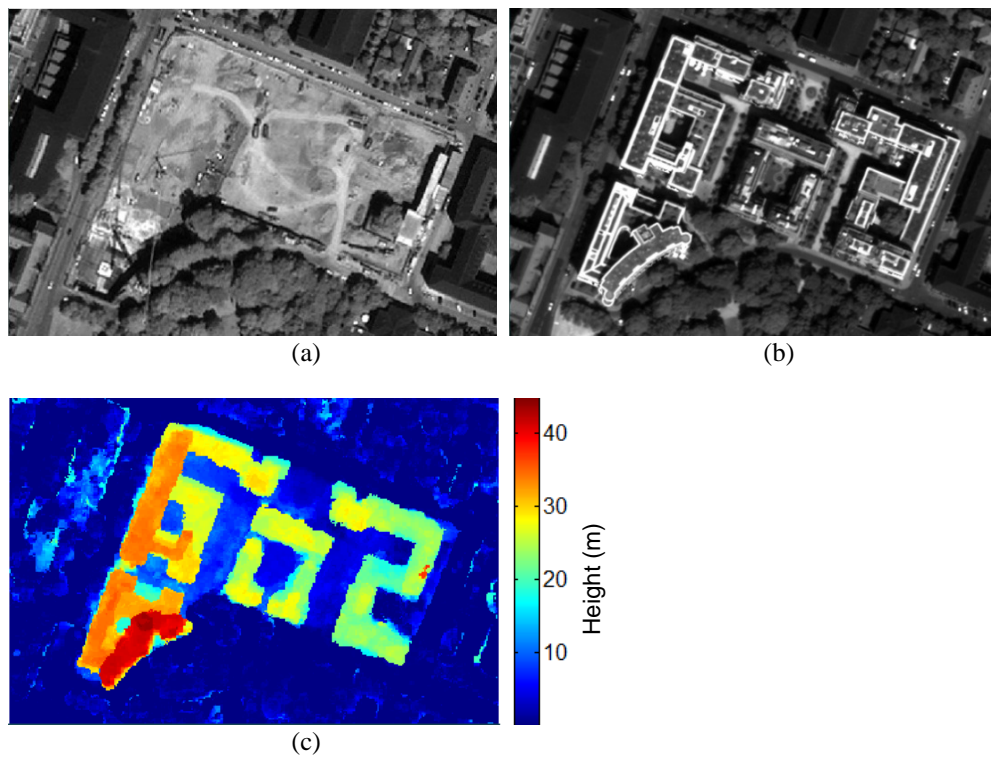


Fig. 4-2. DSM subtraction results of a building site: panchromatic images from Date1 (a) and from Date2 (b), DSMs subtraction result (c).

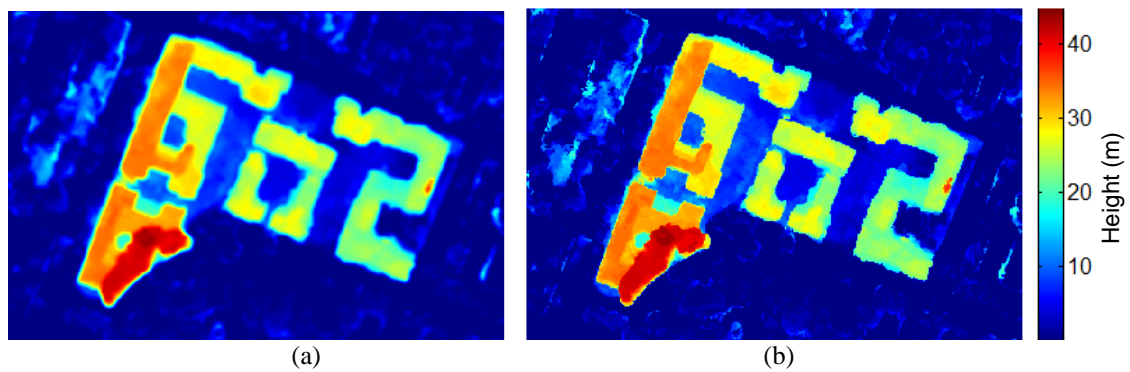


Fig. 4-3. Different image filtering results: Average filter with window size  $7 \times 7$  (a) and Gaussian filter with window size  $7 \times 7$  and  $\sigma=0.5$  (b).

Therefore, the *robust difference* measure, which was initially developed for 2D change detection, is used instead of filtering the height difference map (Castilla et al. 2009). It is designed based on the assumption that the corresponding pixels from two images should have a minimum difference in the gray value. By applying subtraction to 2D images, in addition to accurate co-registration between them, the radiometric calibration of both images is also very important. But radiometric calibration of two DSM is unnecessary, because the gray values in the DSMs directly represent elevation height. If two DSMs are generated correctly, a height difference between them can directly indicate



change. Therefore, the robust difference proposed by Castilla et al. (2009) works better in DSM-based difference map generation and has been adopted accordingly.

The robust difference between the initial DSM  $x_1$  and the second DSM  $x_2$  for the pixel  $(i, j)$  is defined as the minimum of differences computed between the pixel  $x_2(i, j)$  in the second DSM and a certain neighborhood (with window size  $(2 \times w + 1)$ ) of the pixel  $x_1(i, j)$  in the first DSM  $x_1$ . The robust positive and negative differences  $X_{Pdif}(i, j)$  and  $X_{Ndif}(i, j)$  relative to the pixel  $(i, j)$  are defined as written in Eq. (4-2) and Eq. (4-3), respectively:

$$X_{Pdif}(i, j) = \min_{\substack{p \in [i-w, i+w], \\ q \in [j-w, j+w]}} \{(x_2(i, j) - x_1(p, q)), (x_2(i, j) - x_1(p, q)) > 0\} \quad (4-2)$$

$$X_{Ndif}(i, j) = \max_{\substack{p \in [i-w, i+w], \\ q \in [j-w, j+w]}} \{(x_2(i, j) - x_1(p, q)), (x_2(i, j) - x_1(p, q)) < 0\} \quad (4-3)$$

This means that only the minimum value (greater than zero) within the defined window size in case of positive change or the maximum value in case of negative change is taken. Typically used window sizes are  $5 \times 5$  pixels up to  $15 \times 15$  pixels depending on the DSM quality and the difference in resolution between the two available DSMs. In Fig. 4-4, the image difference map generated using robust DSM differencing is shown. In Fig. 4-4a, a window size of  $7 \times 7$  is used, while in Fig. 4-4b a window size of  $15 \times 15$  is used. In contrast to Fig. 4-3, after executing the robust image difference, the noise in the background is successfully reduced, while the yellow and green areas, which are more likely to be real changed areas, are not influenced significantly, especially in the building boundary area. The window size can be adjusted according to various requirements: for example, if the two DSMs have large resolution distance, a higher window size should be used. Overall, it is a robust method, as shown in Fig. 4-4b. Here, a larger window size is adopted, which results in a cleaner background. At the same time, the building heights near the boundary area are influenced only slightly, the boundaries of the building are preserved and even the roof-error in the rightmost building is removed.

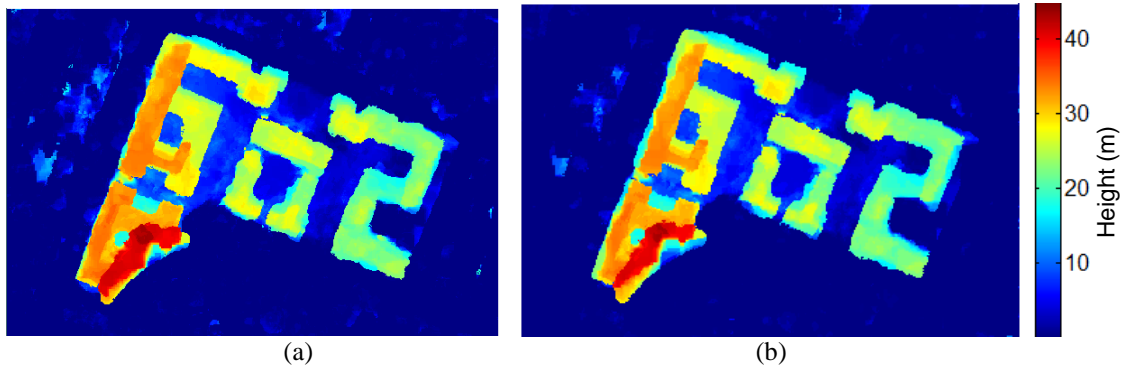


Fig. 4-4. Robust difference results with positive changes. Generated based on window size  $7 \times 7$  (a) and  $15 \times 15$  (b).

### 4.3 Change map denoising

A proper thresholding of the difference image is required to obtain the initial change mask. Automatic threshold methods can be used. But in many cases, unlike image gray value changes, the height changes have a specific physical meaning for classes like buildings or trees. In such cases, we can easily choose a fixed threshold value (e.g.,  $T = 5\text{m}$  or  $10\text{m}$ ) according to the corresponding application and the lateral resolution of the DSMs. This is more direct and robust than any other thresholding methods. However, as illustrated in section 4.1, especially adjacent to buildings, some false alarms are still obtained. These false alarms will remain in the change map after thresholding. And in some cases, these false alarms might also fuse adjacent changed buildings together. To solve this problem, morphological filters can be used to refine the mask.

#### 4.3.1 Morphological filtering of the change map

Morphological filters, which are mathematically-based image processing methods, are adopted to amend this problem. Pesaresi and Benediktsson (2001) have already used morphological filters to process remote sensing images. Morphological filtering is mainly based on two operations: erosion Eq. (4-4) and dilation Eq. (4-5).

$$\text{erode}(x, y) = \min_{(x', y') \in \text{kernel}} \text{src}(x + x', y + y') \quad (4-4)$$

$$\text{dilate}(x, y) = \max_{(x', y') \in \text{kernel}} \text{src}(x + x', y + y') \quad (4-5)$$

A morphological filter should be designed according to the object of interest. Here, buildings are taken as one example. Most of the buildings in urban areas have rectangular shapes or at least are composed of rectangular shapes. Therefore, we define a rectangular shape as the kernel.  $\text{src}$  in Eqs (4-4) and (4-5) is the cover range of all the pixels in this kernel. An example of the morphological

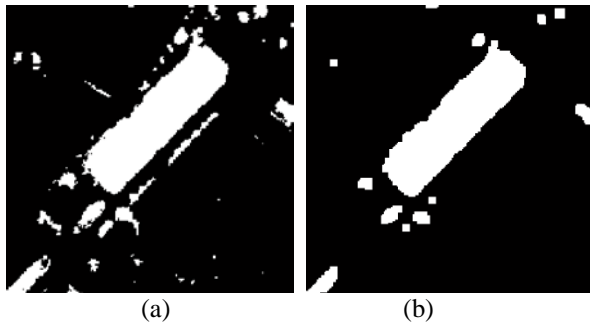


Fig. 4-5. Morphological based filtering: (a) Original mask; (b) Mask after filtering.

based method is shown in Fig. 4-5. Fig. 4-5a features the original height-based thresholding result, where noise caused by DSM failures and some land cover changes (like trees) can be seen. Fig. 4-5b shows the morphological filter result, in which the false alarms caused by this noise are considerably reduced.

### 4.3.2 Shape feature based filtering








In many cases, urban objects, like buildings, can be extracted using shape features. Although the false alarms in the DSM can have height values comparable to buildings, it is unlikely that they will have a shape and size similar to buildings. For example, the rectangular shape shown in Fig. 4-5 can easily be identified as a large building, while other smaller shapes are a result of noise. After refining the change detection map by using the shadow and vegetation land covers, it is still necessary to separate “changed buildings” from the “unchanged background”, which can be partly realized through shape analysis (Tian et al., 2011; Chaabouni-Chouayakh and Reinartz, 2011). Since buildings have more regular shapes compared to noise, we propose applying edge-based building extraction, improving the output by filtering the undesired objects based on their properties. Here, the object properties that we consider include (Tian et al., 2013a):

#### *Height:*

For our purpose, height means the average height for each object, in order to obtain only one vertical change value for each constructed/destroyed building defined by a single mask. To compute this value, we average the pixel values in the fusion result of the “difference image” belonging to the same changed object, and define this value as the vertical change of each building. Next, we exclude all pixels that have a value of 0 (no height in the changed area), as well as very low values or very high values which can be attributed to potential errors in one or both of the DSMs, so that these pixels will not be involved in the mean value calculation procedure. As displayed in Fig. 4-6, only the pixels of the middle part (gray section) of the height difference values are analyzed.

#### 4 DSM-assisted change localization

Table 4-1. Shape difference of possible buildings and artifact objects.

	Shapes			
Possible buildings				
Artifact objects				

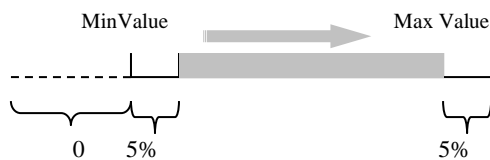


Fig. 4-6. Vertical change value evaluation.

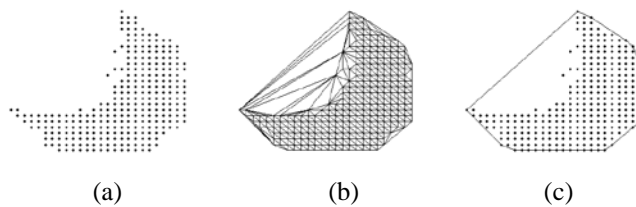


Fig. 4-7. Convex Hull calculation: (a) Pixels in original mask; (b) Generated Delaunay triangulation; and (c) Convex hull.

#### Area:

The area of the generated region is calculated by counting the number of pixels inside a single object included in the generated mask.

$$\text{Area} = \text{Number of pixels} \times (\text{Resolution})^2$$

#### Convexity:

The Convexity of a region is defined in this thesis as the area ratio of this region and the smallest convex polygon that can enclose the region. The computation of the smallest convex polygon (de Berg et al., 2008) is shown in Fig. 4-7. Fig. 4-7a shows the original mask from the difference image. After connecting all pixels in the original mask using Delaunay Triangulation (Fig. 4-7b),

the object is represented by the generated edges, which constitute the smallest convex hull of the original region (shown as the black line in Fig. 4-7c).

## 4.4 Change map refinement using land cover classification

### 4.4.1 Fusion with vegetation covers

Vegetation, including trees and grassland, forms an important component of urban areas. In urban areas with comparatively simple structures, the four classes building, trees, ground and grassland, can be obtained (Lu et al., 2006; Rottensteiner et al., 2007b) if a vegetation mask is combined with the height information from DSM. Based on this classification, building changes can be separated from tree changes by using the vegetation cover information. Alternatively, the vegetation land cover can be used to remove vegetation areas from the original DSMs, thus eliminating the influence of the vegetation on building extraction.

Vegetation has high absorption in the red band ( $\rho_{red}$ ) and low absorption in the near-infrared (NIR) band ( $\rho_{nir}$ ). This makes it easily distinguishable from other land cover classes in the images. In 1969, Jordan introduced the Ratio Vegetation Index (RVI) to extract vegetation (Jordan, 1969).

$$RVI = \frac{\rho_{nir}}{\rho_{red}} \quad (4-6)$$

RVI was improved to NDVI by Rouse et al. (1973) and Krieglner et al. (1969). NDVI is the ratio between the difference and the sum of the NIR and Red channels. The advantage of NDVI lies in its normalization of the value range. NDVI ranges from -1 to 1, while RVI ranges from 0 to infinity ( $\rho_{red} = 0$ ).

$$NDVI = \frac{\rho_{nir} - \rho_{red}}{\rho_{nir} + \rho_{red}} \quad (4-7)$$

Although many other methods have subsequently been developed for vegetation detection (Huete, 1988; Kaufman and Tanre, 1992; Qi et al., 1994; Ray, 1994), NDVI remains the most popular method due to its reliability and ease of calculation (Lu et al., 2004; Rottensteiner et al., 2007b; Awrangjeb et al., 2010). Moreover, if the vegetation cover rate is not too low, NDVI is less sensitive to the soil background and atmospheric influences compared to other methods (Ray, 1994).

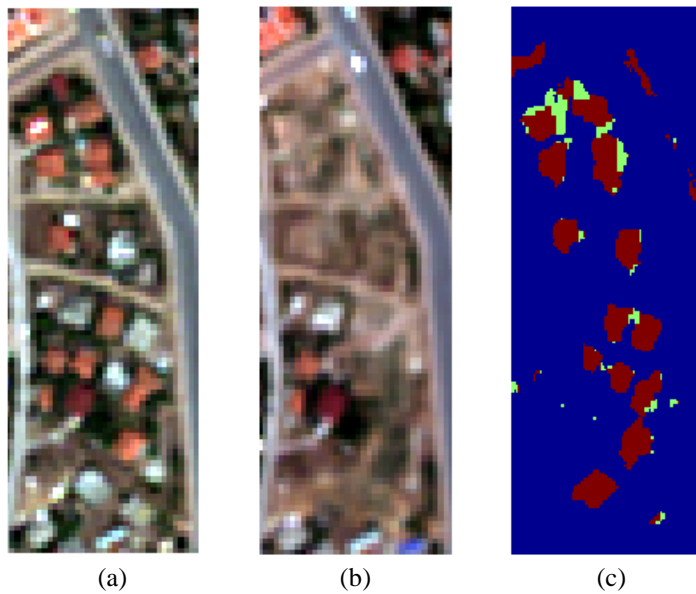


Fig. 4-8. WorldView-2 images and classified changes: (a) Before change; (b) After change; (c) Change classification results with building change (red) and vegetation change (green).

Fig. 4-8 displays an example of the classified building and vegetation changes detected using only height information and NDVI. Fig. 4-8a and 8b show WorldView-2 images before and after a change in building development. Several houses have been removed in this area (for unknown reasons). Using NDVI and height information, changes here can be classified as removed buildings (marked in red in Fig. 4-8c) and removed trees (marked in green in Fig. 4-8c).

#### 4.4.2 Fusion with shadow covers

- **DSM quality in shadow area**

As well as handling the vegetation, detecting and removing shadow influence from satellite images or DSMs is very important for urban area monitoring, especially for building change detection. In binary change detection, the very important final step consists in highlighting real positive and negative changes by getting rid of artificial change (noise and other areas which might contain irrelevant change: here, for example, not buildings). The existence of shadows causes false colour features, decreases texture information and directly influences the image matching procedure, leading to more imprecise height values for shadowed areas in the generated DSM. Moreover, the improvement in spatial resolution also leads to an increasing influence of shadows and these shadowed areas are often different for the stereo images from the two dates.

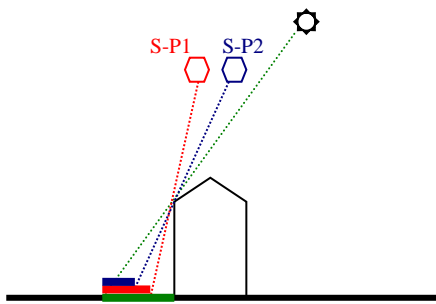
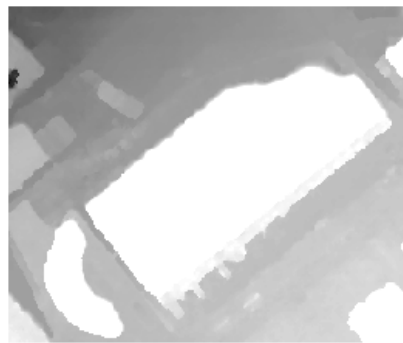


Fig. 4-9. Shadow observed in stereo data.



(a)



(b)

Fig. 4-10. DSM quality in shadow area: (a) Panchromatic image from WorldView-2; (b) DSM generated from (a) and second stereo image.

The influence of shadow on the stereoscopic images is illustrated in Fig. 4-9. The green area shown here represents the real shadow of the building on the ground surface, while the red area is the shadow that is detected from satellite position 1 (S-P1) and the blue area is the shadow that is detected from satellite position 2 (S-P2). This constellation can result in the following two problems: i) As some parts of the red area seen from S-P1 are not visible from S-P2 or even from both positions, information required for the DSM generation is missing; ii) The area observed from both positions (blue area) has relatively low radiometric values, which may cause matching failures in the DSM generation procedure.

This problem is further illustrated in Fig. 4-10: As can be seen, the shadow area, especially in the top right corner, leads to blurred boundaries in the resulting DSM (Fig. 4-10b). When analyzing the accuracy of DSMs, it can be observed that the side of the building borders with lower quality is always located at the shady side of the building (Fig. 4-10a). Even though the existence of shadow affects the quality of the data, it can also be used to acquire valuable information for building extraction (Sirmacek and Unsalan, 2011). If no other high objects are within the shadow cast, the

shadow boundaries which run parallel to the building edges can be used to refine the boundaries of the extracted buildings.

- **Shadow extraction**

Extensive reviews and evaluation of literature and methods related to shadow extraction have been conducted by Prati et al. (2003) and Sanin et al. (2012). According to their research, the selection of features is more important than the classification method in shadow detection. Shadows can be divided into static and dynamic shadows. Static shadows result from static objects, like buildings and trees, while dynamic shadows result from moving objects, like pedestrians or cars (Nadimi and Bhanu, 2004). According to the different features used, recent shadow detection methods can be divided into intensity, chromaticity, physical properties, geometry and texture-based approaches. They all have advantages depending on their application. For this work, the physical model based method developed by Makarau et al. (2011) is adopted, due to its robustness and high accuracy for remote sensing images. Another advantage of this method is that it is not restricted to the radiometric properties of the image or the illumination condition. In Fig. 4-11, the detected shadow mask based on the scene shown in Fig. 4-10 is displayed.

- **Shadow based DSM refinement**

As mentioned above, morphological filtering is a method that can be used to refine the object shape. In this work, morphological erosion and dilation have been combined with the shadow mask to refine the building mask. As has been described in Fig. 4-10, the DSMs' building boundaries on the shady side are relatively worse than on the illuminated side. Therefore, the sharp edges from the building shadows are used to refine the building boundaries.

In Fig. 4-12, the whole procedure of the shadow mask based building shape enhancement is displayed. The black line corresponds to the originally detected edges, and the blue line represents the dilated boundary, which is covered by the shadow mask. The green rectangle in Fig. 4-12 represents the refined building shape. Further erosion is performed until the original building dimensions are reached. The red polygon represents the refined building shape.





Fig. 4-11. Detected shadows based on the scene shown in Fig. 4-10.

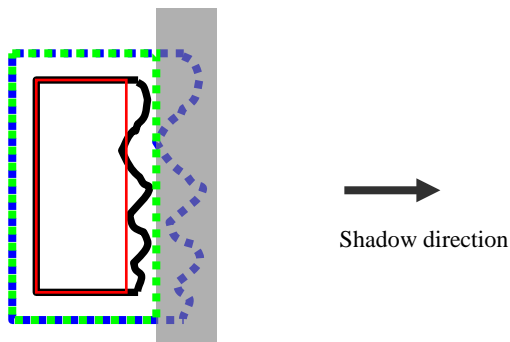


Fig. 4-12. Changed building mask and shadow mask.

## 4.5 Summary

In this chapter, DSM-assisted change localization was introduced. Both the robust image differencing and the land cover-based change map refinement are suitable for robust 3D change detection. They are not restricted by the sensor type or DSM quality and both help to improve the height change map.

Fig. 4-13 depicts the flowchart of the approach proposed in this chapter. Starting from two co-registered datasets, the robust difference, denoising and landcover-based refinement are adopted in generating the final change mask. Based on the obtained results, the method presented can alleviate some of the drawbacks of DSMs, such as unsmooth boundaries and some small potential errors.

#### 4 DSM-assisted change localization

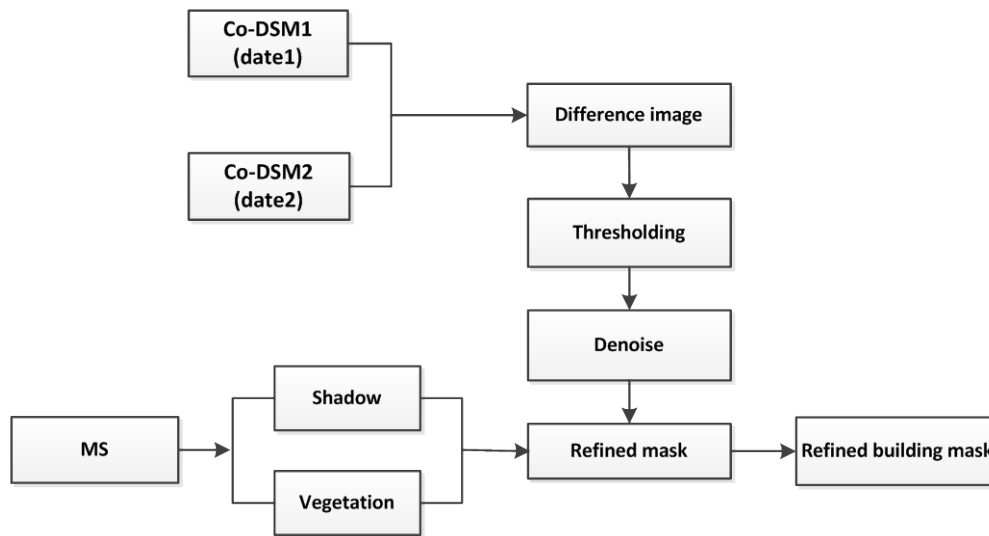


Fig. 4-13. Flow chart of the DSM-assisted change location method proposed.

However, if there are some changes that are not included in the change map, such as rebuilt buildings, which can happen for long time range change detection, these new buildings at similar locations to old buildings with similar heights might not be detected. And if the regions at the newly built building locations were originally covered with trees, the changed building mask may not be represented correctly in the initial change mask. Moreover, if any of the DSMs contain large areas of incorrect values, for example, the stereo imagery can fail to extract a whole building or one piece of forest. In that case, that building or forest will not appear correctly as change/no change in the initial change mask, and neither the robust difference nor the following refinement can cover it.

## 5. Fusion-based change detection

### 5.1 Introduction

DSM-assisted change localization approaches, as described in Chapter 4, are mainly based on refining the height difference map. If the original DEMs contain large areas of incorrect information, refinements can improve the accuracy to a certain extent, but usually cannot correct all of these errors. These incorrect height values are produced due to the influence of shadows, and some mismatched information due to the methodology used for DSM generation and eventually large different view angles of the stereo pair. The existence of these errors makes it difficult to extract initial changes precisely using only DSM subtraction. Thus, an efficient method of fusing the DSMs and ortho-rectified satellite images is proposed. This method can compensate for the disadvantages of using DSMs or spectral images alone for change detection.

The main objective of data fusion in remote sensing is to combine data from different sources, producing an improved result comprising information from all the sources. The three different types of data fusion techniques identified by identified by Hall and Llinas (1997) and Pohl and van Genderen (1998) are shown in Fig. 5-1:

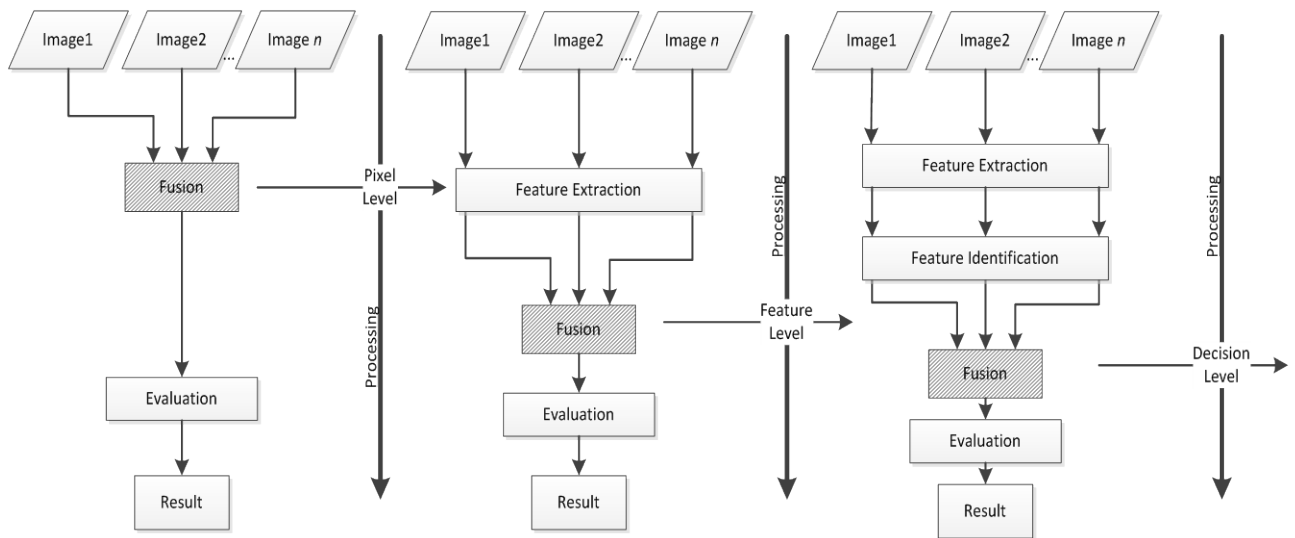


Fig. 5-1. Processing levels of image fusion (Pohl and van Genderen, 1998).

## 5 Fusion-based change detection

- Pixel / Data-level fusion
- Feature-level fusion
- Decision-level fusion

Pixel-level fusion is the method at the lowest processing level. It combines different sources of raster images into a single image, e.g. for pan-sharpening (Ehlers, 2004). The co-registration between these images plays an essential role, as inaccurate co-registration can produce artificial colors or features in the final fused result, thus influencing the later interpretation. Different pixel-based fusion methods have been reviewed in Pohl and van Genderen (1998) and Ehlers et al. (2010).

Feature-level fusion combines features extracted from two or more sources of data. Many 3D change detection methods belong to this category, since they fuse spectral features from multispectral images and height features from DSMs (Sasagawa et al., 2008; Tian et al., 2011). Until now, most of the feature fusion methods have been performed based on LiDAR data, as the LiDAR data contain highly accurate height information: thus, each value can be considered a trustable feature.

Decision fusion treats each feature or initial change detection result as one decision, so a fusion rule is needed to combine them. In change detection, these change features are usually extracted separately from various data sources, and can indicate the amount of change or change classes. The methods of decision fusion include weighted decision methods (voting mechanisms), classical inference, Bayesian inference, Fuzzy logic and Dempster-Shafer fusion (Hall and Llinas, 1997).

Since this research is focused on producing change detection maps instead of pixel-fused maps, feature fusion and decision fusion methods are developed and tested in this chapter. In both fusion-based change detection approaches, the initial change maps are generated by using DSMs and images, which is also the main difference between the fusion-based methods and the DSM-assisted change localization (Chapter 4).

As a feature fusion approach, we employ a region-based change detection method in this thesis. The line features from orthorectified panchromatic images are used to produce homogeneous regions, and then other change features extracted from images and DSMs are analyzed for these regions. For the decision fusion approach, DSM and spectral images can be treated as two parallel data sources. Both of them can give an indication of change or no-change for each pixel. Fusion rules should be built for each class. Since this decision fusion focuses on evaluating the efficiency

of the proposed indicators and the decision fusion method: thus, it is built on the pixel level. Moreover, the buildings are taken as objects of interest for the decision fusion model generation.

## 5.2 Feature fusion-based change detection

### 5.2.1 Feature fusion

DSMs generated from stereo images as described in Chapter 2 exhibit relatively low quality in boundary areas. Fig. 5-2 displays a comparison of the boundary sharpness between a panchromatic image and the corresponding DSM. Fig. 5-2a shows the orthorectified panchromatic image; Fig. 5-2b shows the corresponding DSM. As can be seen here, buildings in panchromatic image have much clearer and straighter boundaries than in the DSM. Therefore, using boundary features from panchromatic images is more accurate than from DSMs. Poli and Soille (2012) suggest using segmentation results from VHR images (e.g. WorldView-2) to refine a Cartosat-1 DSM. In many cases, the original Cartosat-1 image itself also provides reasonable segmentation results. Moreover, applications using region-based change detection methods can improve the change detection accuracy, especially by removing the “pepper and salt” effect in the 2D image-based change detection methods. Therefore, instead of directly fusing pixel-based features from the two image sources, we extract regions from the panchromatic images.

As shown in Fig. 5-3, a two-step region-based change detection procedure is proposed according to the character of the DSMs produced using stereo matching. In the first step, segmentation on orthorectified Cartosat-1 images is performed to obtain initial regions. Then regions from two dates (Date1 and Date2) are combined to obtain an initial segmentation map. To correct the over-segmentation resulting from the region combination, a region merging strategy is proposed to reach a reasonable segmentation level.

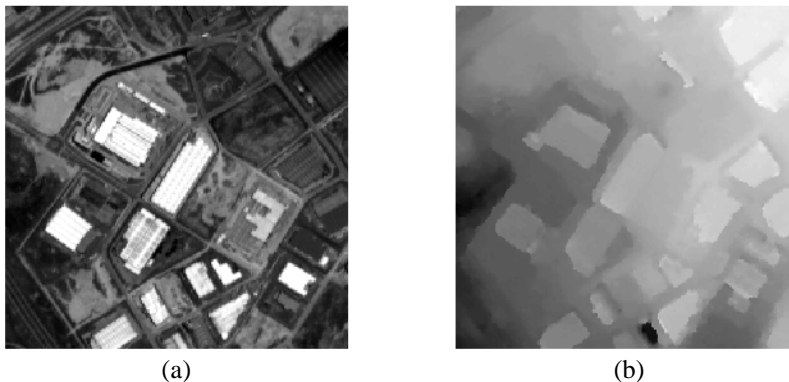


Fig. 5-2. Boundary sharpness comparison between panchromatic image (a) and generated DSM (b).

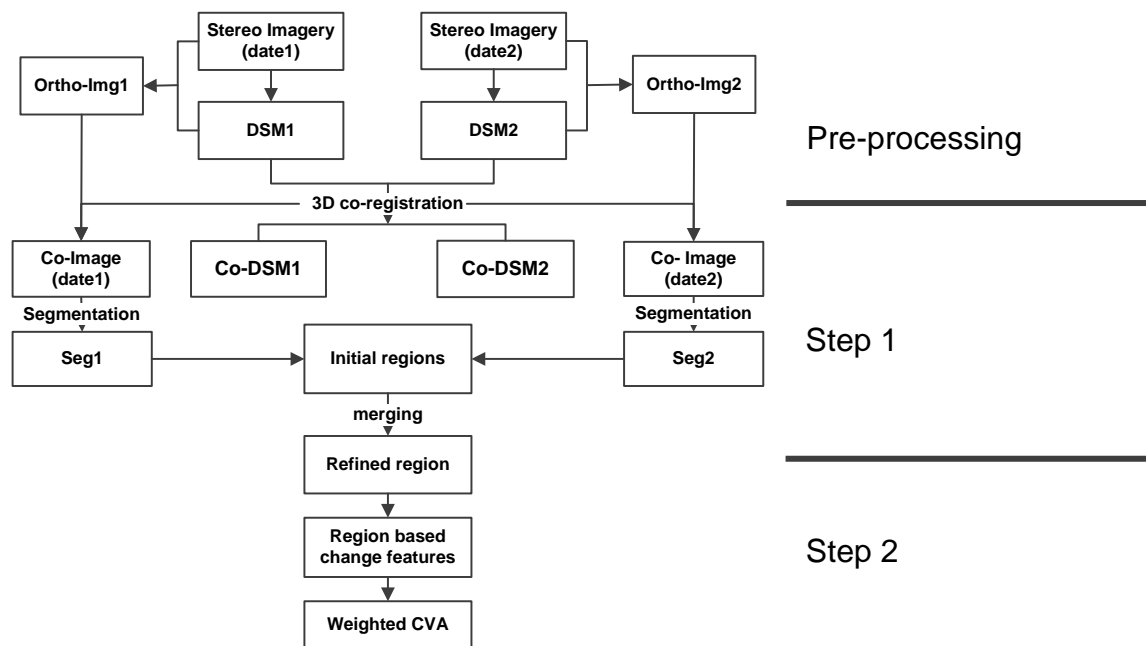


Fig. 5-3. Flow chart of the feature fusion approach (Tian et al., 2013b).

### 5.2.2 Mean-shift segmentation

Image segmentation is the process of partitioning an image into groups of pixels that are spectrally similar. Here, the objective of segmentation is to produce small units that have different spectral characteristics in comparison to the areas nearby. A proper level of segmentation in high resolution satellite images is difficult to reach due to mixed pixels, spectral similarity between different land covers and the textured appearance of specific land covers. Many segmentation methods have been introduced in computer vision, like watershed (Vincent and Soille, 1991), level-set, mean-shift (Comaniciu and Meer, 2002) and several more. All approaches attempt to reach an appropriate segmentation level by adjusting one or more parameters based on one segmentation algorithm (Vincent and Soille, 1991; Meyer and Beucher, 1990; Comaniciu and Meer, 2002; Melendez et al., 2011). Since urban areas typically consist of different land covers (e.g. buildings, roads, shadows, trees), it is difficult to obtain adequate segments for all land covers using a single segment scale, since different kinds of objects require different segmentation levels. Over- and under-segmentation typically appear together. Some methods perform multi-scale segmentation. Since image splitting is more difficult to control, much work has been performed based on merging over-segmentation results (Haris et al., 1998; Ning et al., 2010; Nock and Nielsen, 2004).

Here, over-segmentation is performed using the mean shift implementation from the EDISON library (Comaniciu and Meer, 2002) to obtain small change units. Since segmentation is not the focus of this thesis, the details will not be provided here, but it can be said that for our purpose, small segments are an important pre-requisite. The segmentation could also be performed using other methods, which can lead to over-segmentation. After segmentation, two region maps are obtained from the two corresponding datasets. An intersection of the two region maps is performed to capture all possible change regions.

Fig. 5-4 presents the intersection procedure of two segments from two dates (before change and after change). The example shown in Fig. 5-4c is the ideal intersection result, on which only the changed areas produce new segments. However, in many situations, the obtained result is less ideal. As shown in Fig. 5-4d, some unexpected segments, like the segment textured with the oblique grid, are introduced because of the displaced boundaries. This problem arises frequently in change detection due to factors such as imperfect co-registration accuracy between two datasets. Many small segments composed of only one or two pixels are produced, primarily at the edges of the segmentation results of the two data (as shown in Fig. 5-5). Due to suboptimal registration or radiometric differences (e.g. different shadow cast) caused by seasonal or sun angle differences, border regions of segment maps from two dates cannot be matched perfectly. Moreover, DSMs usually feature lower quality (height values are less accurate) at segment / object edges compared to the center of segment / object. Therefore, these small regions introduce false alarm to the final results, and appropriate region merging is required before any further steps are performed. A detailed requirement of initial regions for change detection was proposed by Bruzzone and Prieto (2000). However, they do not specify how to reach this requirement. In the following section, a new region merging procedure is proposed to solve this problem.

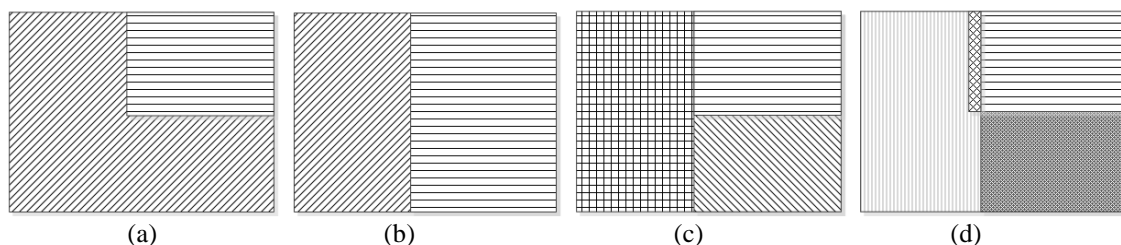


Fig. 5-4. Segmentation intersection procedure: Segments from Date1 (a) and from Date2 (b); Ideal intersection result (c) and actual intersection result with displaced edges (d).

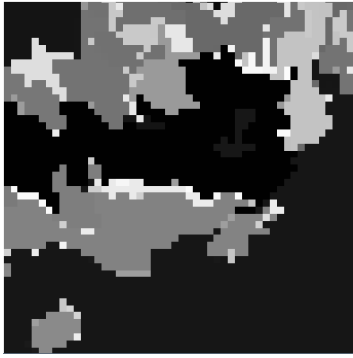


Fig. 5-5. Small regions at the borders resulting from imperfect registration.

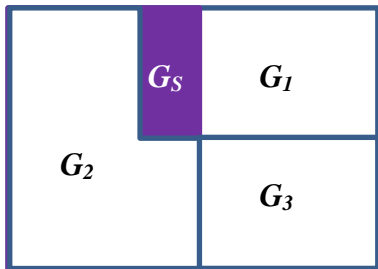


Fig. 5-6. Merging problem statement.

### 5.2.3 Greedy strategy for small region merging

In order to merge these small regions without influencing the original segmentation result, only very small regions (size less than 10 pixels, for example) are considered in the merging procedure. As shown in Fig. 5-6, the purple region  $G_s$  is the small region which has to be merged. Each  $G_s$  is surrounded by several regions, which are hereafter referred to as candidate regions  $G_i$  ( $i=1, 2, \dots, k$ ,  $k$  is the total number of the neighbors). The objective of this step is to merge each small region  $G_s$  into one of the candidate regions. To achieve this, the similarity between these small regions and the candidate regions is calculated. Following the greedy strategy (Cormen et al., 2001), the candidate with the minimum heterogeneity increase after merging it with the region  $G_s$  is selected. An energy minimization model composed of four energy functions is proposed for this step. The similarities between the regions are measured for both the image intensity difference map  $D$  (Eq. (5-1)) and the height difference map  $H$  (Eq. (5-2)).

$$D = |f_1 - f_2| \quad (5-1)$$

$$H = |h_1 - h_2| \quad (5-2)$$



Here  $f_1$  and  $f_2$  represent the pixel values of the panchromatic images from Date 1 and Date 2, and  $h_1$  and  $h_2$  are the two DSMs from Date 1 and Date 2. Merging  $G_s$  into one of the  $G_i$  regions will produce  $G_{i'}$ . We define the following energy functions to measure the distance between  $G_i$  and  $G_{i'}$ . The first function (Eq. (5-3)) is the area weighted mean value distance based on  $D$ .

$$E_1 = \omega * |\mu(G_{i'}) - \mu(G_i)| \quad (5-3)$$

where  $\mu(G_i)$  is the average pixel value of region  $i$ , the pixel values are extracted from the image intensity difference map  $D$ . The energy term is multiplied by a weight  $\omega$  because the larger  $G_i$  is, the less likely it is to change after merging with  $G_s$ . For this work, the size of  $G_i$  (total pixel number) is used as the weight  $\omega$ .

The intensity range of each region is chose as the second energy function.

$$E_2 = |Maxmin(G_{i'}) - Maxmin(G_i)| \quad (5-4)$$

Where  $Maxmin(G_i)$  is the difference between the maximum and minimum value of region  $G_i$ . Also  $E_2$  is computed based on the  $D$  image. Similar energy functions are computed based on  $H$  (shown in Eq. (5-2)), denoted as  $E_3$  and  $E_4$ .  $E_3$  is the weighted mean value using average height values and  $E_4$  is the height range of each region.

Next the candidate regions are ranked four times according to the four similarity energy functions separately: the smaller the value of the energy function, the higher the rank  $O$  it is assigned.  $O_j^i$  represent the rank order of  $G_i$  according to the energy feature  $j$ . A ranking matrix  $\Omega$  is generated recording the ranks of the candidates.

$$\Omega = \begin{Bmatrix} O_1^1 & O_2^1 & O_3^1 & O_4^1 \\ O_1^2 & O_2^2 & O_3^2 & O_4^2 \\ \dots & \dots & \dots & \dots \\ O_1^k & O_2^k & O_3^k & O_4^k \end{Bmatrix} \quad (5-5)$$

A ranking vector  $\{O_1^i; O_2^i; O_3^i; O_4^i\}$  is generated for each candidate  $i$  ( $i = 1, \dots, k$ ). The candidate with the highest sum of ranking  $Min(\sum_{j=1}^4 O_j^i)$  is chosen as the object to ‘merge to’ (since a high rank results in a low value). To speed up the merging procedure, all of the neighborhood candidates should have more pixels than the defined size of the small regions. For instance, when

a region with fewer than ten pixels has to be merged with one of its neighbours, only the neighbourhood regions larger than ten pixels size are considered as candidates.

### 5.2.4 Region-based change detection

Using the produced regions described in the previous section, it is possible to apply region-based change detection methods. The main difference between pixel-based and region-based change detection is the fact that pixel-based change detection operates on single pixels, while region / object-based change detection methods operate on a group of pixels belonging to one region / object. Change features should thus be extracted for each region. The region-based features can be divided into two groups: the average of existing pixel-based features and textures analysis using each region as the smallest unit.

- **Region-based change features based on average difference**

Radke et al. (2005) reviewed numerous change features. The image grey value difference is still an essential feature of change detection and is easy to implement. The first step of this approach is to generate an absolute-valued difference image map  $D$  (Eq. (5-1)), as we have already used in the region merging procedure.

In our published work (Tian et al., 2013b), the multi-level feature sets (Celik, 2009) are adopted. These features are obtained by considering a series of neighborhoods with various window sizes  $S$   $\{S=1, 3; 5; 7; 9\}$  surrounding the corresponding pixels specified by

$$v = \frac{1}{S \times S} \sum_{p=1}^{S \times S} D^p \quad (5-6)$$

This can also be considered as the difference map at different resolution levels. With this multi-level representation of changes, the neighborhoods of the corresponding pixels are included, highlighting the behavior of changes for various land cover classes.

- **Change detection**

In this section, the region-based change vector ( $v$ ) is described by using the mean value of the change features from DSMs and images ( $v_i^r$  is the  $i^{\text{th}}$  change vector for region  $r$ ;  $r = 1, 2, \dots, n_r$ ;  $n_r$  is the total number of regions.). Since the Euclidean distance is adopted for the CVA in this work, a standardization of the change vector is preferable when the change features within the change vectors are in the same range. All of the change vectors are normalized by dividing them by their standard deviation ( $\sigma(v_i)$ );  $v_i$  is the  $i^{\text{th}}$  change vectors of all regions;  $i = 1, 2, \dots, l$ ;  $l$  is the total

number of region-based change vectors). Moreover, weights  $\rho_i$  are used to balance the influence of the change vector from the images and from the DSMs. In this paper, we use five region-based change vectors from the image difference map (multi-level features) and only one region-based change vector from DSM (see also section 3.2.1). Weight  $\rho = 0.1$  is used for the change vectors from images, and  $\rho = 0.5$  for the change vector from DSM. This weight can also be manually adjusted if the DSMs exhibit obviously lower or higher quality. The formulated weighted multilevel CVA for each region ( $r$ ) is

$$\|\Delta V^r\| = \sqrt{\rho_1 \left( \frac{v_1^r}{\sigma(v_1)} \right)^2 + \rho_2 \left( \frac{v_2^r}{\sigma(v_2)} \right)^2 + \dots + \rho_l \left( \frac{v_l^r}{\sigma(v_l)} \right)^2} \quad (5-7)$$

Here, a larger value of  $\|\Delta V^r\|$  indicates a higher probability of change.

### 5.3 Decision fusion-based change detection

#### 5.3.1 Decision fusion

The general term ‘decision fusion’ means in our case the fusion of different change indicators. These change indicators can be extracted directly from images, and can also be a change detection result from selected independent methods. The aim of fusing decisions of separate indicators is to increase the overall performance. The advantage of decision fusion is its feasibility, low computational complexity and robustness to the removal or addition of individual indicators (AUG Signals, 2013). Benediktsson and Kanellopoulos (1999) combined three voting systems - majority voting, complete agreement and consensus-neural network - to fuse the classification results from neural network and Gaussian maximum likelihood. However, the classification results from both classification methods directly influence the final fusion results. Fuzzy logic is also a popular decision fusion method. As a form of probabilistic logic, it can combine all logic variables with values between 0 and 1 (Zadeh, 1965). For fuzzy logic-based classification, membership functions have to be defined for each class in the training procedure (Vögtle and Steinle, 2003). It provides a large number of combination operators but is said to be a weak method in modelling uncertain information (Le Hégarat-Masclé et al., 1997). Fuzzy logic assumes that the information from each indicator is 100% certain, which is not true in many cases (Rottensteiner et al., 2005). In object-focused change detection, it is usually not possible to decide with 100% certainty if an extracted feature based on only one data source has changed or not. Usually, only a probability of change can be provided for each feature. This probability can be determined according to the quality of the DSM and the quantity of change in height. Correspondingly, the changes detected in

spectral images all represent potential changes. When only the change of some objects is of interest, the radiometric changes obtained from satellite spectral images do not directly indicate changes, but change probabilities. In this work, Dempster-Shafer fusion is used for the purpose of fusing uncertain information.

### 5.3.2 Dempster-Shafer fusion

Dempster-Shafer (DS) fusion was introduced in the 1960s by Dempster (1968) and extended by Shafer (1976) for the purpose of handling imprecise and incomplete information as well as solving conflicting information among different inputs. Another advantage of DS fusion is that it can handle not only single classes but also the union of several classes (Le Hégarat-Masclé et al., 1997). This is particularly useful when change detection is performed for specific objects of interest. For example, NDVI can indicate increasing or decreasing vegetation. Here, vegetation would be a combination of trees, grass and bush classes instead of a single class. Moreover, DS can handle the “mixed” pixel problem, since it can deal with any union of classes (Le Hégarat-Masclé et al., 1997). Le Hégarat-Masclé et al. (1997) and Rottensteiner et al. (2007) have adopted Dempster-Shafer fusion for classification of multisource remote sensing images. DS fusion rules have to be designed according to the object of interest. Therefore, in this work, we take building change as one example: only the indicators related to buildings are of interest.

DS theory is a belief function-based combination method. Each indicator gives different certainty to the decision class ( $A \in 2^\Theta$ ),  $2^\Theta$  represent the object classes of interest (such as the building change class) and all possible combinations of them. The certainties of the decision class  $A$  are called probability masses ( $m(A)$ ), and have values between 0 and 1.

$$\begin{cases} m(\phi) = 0 \\ \sum_{A \in 2^\Theta} m(A) = 1 \end{cases} \quad (5-8)$$

When  $p$  indicators are considered, each indicator provides a probability for the classes in  $B$  ( $B \in 2^\Theta$ ), which have a fuzzy relationship with  $A$ . The fused certainty of  $A$  can be represented with DS fusion theory (Le Hégarat-Masclé and Seltz, 2004) based on:

$$m(A) = \frac{\sum_{B_1 \cap \dots \cap B_p = A} \prod_{i=1}^p m_i(B_i)}{1 - K} \quad (5-9)$$

$i$ : indicator, and  $1 \leq i \leq p$ ,

$m_i$ : probability mass

$p$ : amount of indicators

$$K = \sum_{B_1 \cap \dots \cap B_p = \emptyset} \prod_{i=1}^p m_i(B_i) \quad (5-10)$$

$K$  is a measure of conflict among different indicators, meaning that a decision cannot be, for example, change and no-change at the same time.  $m_i(B_i)$  represents mass functions of indicator  $i$  to class  $B$ . The probability masses shown in Eq. (5-8) are the original belief certainties. For instance, if one pixel has less than 1 m height change, then it is given a small probability, indicating low significance of real building changes, since this is in the range of DSM-noise. In contrast, if it has more than 10 m of height change, it will have a high probability of indicating actual building change. In Rottensteiner et al. (2007), a cubic parabola was used to model the mass function for the probability between these extremes, which also works well in DSM-assisted classification. In that model, several experimental based thresholds are needed for each indicator. Therefore, for this investigation, the sigmoid curve is preferred (Tian et al., 2013a) and use it to obtain a similar ‘S’ shape with fewer parameters.

$$P_i(x) = \frac{0.99}{1 + e^{-\frac{x-T}{\tau}}} \quad (5-11)$$

In Eq. (5-11),  $x$  is the original value of each indicator. For our purpose, we add two parameters  $T$  and  $\tau$  to control the symmetry point and the shape of the sigmoid function. The symmetry point (black point) in Fig. 5-7 indicates a certainty of 50%. Obtaining the parameter  $T$  for this value can be treated as a threshold selection problem. For pixels having a value ( $x$ ) near the selected threshold value ( $x \approx T$ ), the obtained probability of  $P_i$  is about 0.5. It indicates 50% of certainty to be assigned to the correct class of interest, which means that these values also have a high risk of raising a false alarm. The more a value deviates from the threshold, the higher is the probability that it is assigned to the correct class.

In image processing, much research has been performed focused on the selection of an appropriate gray level threshold to separate the object of interest from the background. The threshold value can be provided manually or calculated automatically by adjusting the relationship of the object of interest and the background. If the gray values in the images have well understood physical meanings, like the height values in DSM, it is not difficult to choose a threshold value. If not, the automatic method is more robust. This work employs the Minimum Error thresholding method

## 5 Fusion-based change detection

proposed by Kittler and Illingworth (1986) to select  $T$ , because of its efficiency and robustness. This thresholding method assumes that both objects and the background in the image follow a Gaussian distribution. The threshold  $T$  is chosen to optimally separate these two distributions, which means that the pixels at this threshold point have a probability of 50% of being classified as an object.  $\tau$  is used to control the slope of this distribution, which can be estimated using a given sample. According to the rule of the DS fusion theory, no indicator can give 100% certainty: therefore, a maximum value of 0.99 is chosen for this model.

The main advantage of sigmoid curve distribution is that it can expand the middle part of the original feature values. After transforming the obtained values, they would be less sensitive to the selected threshold value. Fig. 5-8 displays the original height change indicators and the generated mass function. The values have been stretched from [-20 m, 23.9 m] to [0, 0.99]. In this example, the sample datum ( $P(x) = 0.1 | x = 1$ ) is used to calculate  $\tau$ : therefore, all of the values below 1 will automatically obtain a change probability between 0 and 0.1. As can be seen in Fig 5-8b, the pixels with height changes more than 6~7 m obtain a higher change probability (more than 0.8). The middle values are mainly located in the building boundaries or shadow areas.

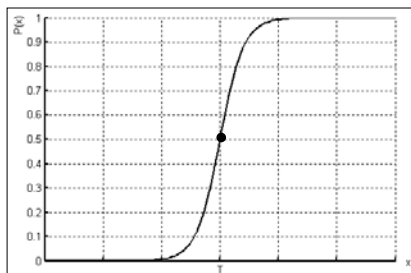


Fig. 5-7. The probability mass function.

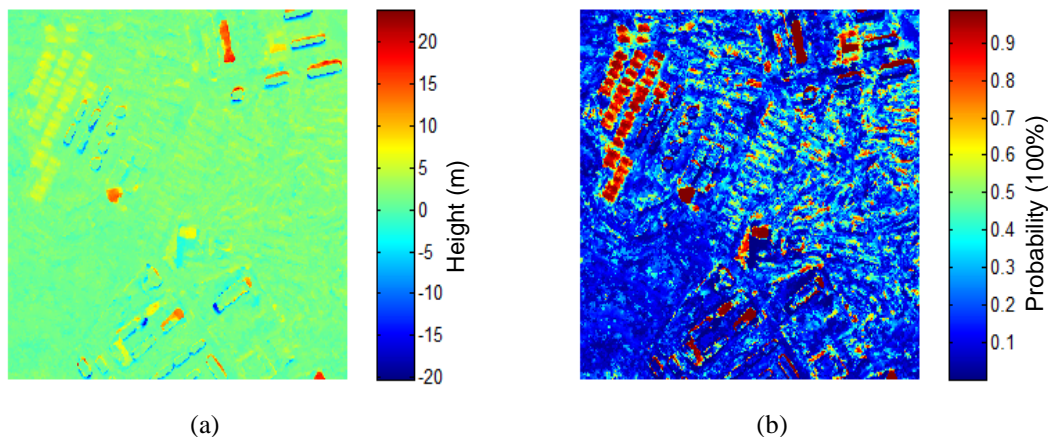


Fig. 5-8. Comparison of the change map before and after sigmoid curve stretching: (a) Original height change map; (b) Change indicator stretched with Sigmoid curve.

### 5.3.3 Change and no-change indicators

DS fusion methods from Le Hégarat-Masclé et al. (1997) and Rottensteiner et al. (2007) are focused on classification: thus, the indicators are defined with a focus on different land cover classes. In this section, only building changes are of interest. Therefore, we have classified all of the features we can get from images and DSMs to building change indicators and no-building change indicators. Change features are all of the features that can be used to indicate the changes for the specific object, like height changes derived from a DSM or dissimilarity from spectral images. In Chapter 4, features that can be extracted from DSMs and images for change map generation and refinement were illustrated. Features have been used mainly separately in improving change accuracy in that approach. If one feature exhibits large errors, false alarms will be produced in the final change result. Therefore, in this section, these features have been classified to change indicators and no-change indicators. They are fused together to obtain a change map based on DS fusion theory.

As explained in Chapter 4, radiometric information for the same land cover class sometimes differs considerably due to the different acquisition circumstances. Instead of comparing gray values directly, we use an information similarity measure to highlight building changes. Information similarity measures have been widely used in the image processing community, for example for image registration (Suri and Reinartz, 2010, Reinartz et al., 2011) and change detection (Inglada and Mercier, 2007; Gueguen et al., 2011). Inglada and Mercier (2007) extended Kullback-Leibler (KL) divergence for multi-temporal change detection based on the evolution of the local statistics of the image between two dates. The local statistics are estimated by using one-dimensional Edgeworth series expansion, which approximates probability density functions in the neighborhood of each pixel in the image. Bovolo et al. (2008) adopted this method for object-based change detection by computing the KL divergence of the two corresponding objects derived using image segmentation. In our work (Tian et al., 2012a), KL divergence was used to obtain a change indicator from images. A brief introduction to the calculation procedure is provided below.

Assuming that the local neighborhood of two corresponding pixels is considered as two random variables  $X$  and  $Y$ , with  $f_X$  and  $f_Y$  being the marginal Probability Density Function (PDF), then the KL divergence of these two random variables is defined as

$$K(Y|X) = \int \log \frac{f_X(x)}{f_Y(x)} f_X(x) dx \quad (5-12)$$

## 5 Fusion-based change detection

The remaining task is to estimate the marginal distributions, which can be achieved using several approaches. The simplest way is to make use of histograms. However, to do that, a large number of samples are required for proper estimation (Inglada and Mercier, 2007), meaning that a relatively large window size would have to be used. Here we use the Cumulant-based KL approximation for PDF estimation as proposed in Inglada and Mercier (2007). This method works under the assumption that the PDF is close to a Gaussian distribution. In this case, the marginal density function  $f_X(x)$  can be expressed by the Edgeworth series expansion, which relies only on the first four cumulants (Cui et al., 2011). The cumulants can be estimated from moment  $\mu_{X;i}$ , which is the centered moment of order  $i$  of variable. The Edgeworth series expansion can be expressed as described in Inglada and Mercier (2007) and Lin et al. (1999):

$$\begin{aligned}
 KL_{Edgeworth}(X, Y) = & \frac{1}{12} \frac{\kappa_{X;3}^2}{\kappa_{X;2}^3} + \frac{1}{2} \left( \log \frac{\kappa_{Y;2}}{\kappa_{X;2}} - 1 + \frac{1}{\kappa_{Y;2}} (\kappa_{X;1} - \kappa_{Y;1} + \kappa_{X;2}^{1/2})^2 \right) \\
 & - (\kappa_{Y;3} \frac{a_1}{6} + \kappa_{Y;4} \frac{a_2}{24} + \kappa_{Y;3}^2 \frac{a_3}{72}) - \frac{1}{2} \frac{\kappa_{Y;3}^2}{36} (c_6 - 6 \frac{c_4}{\kappa_{X;2}} + 9 \frac{c_2}{\kappa_{Y;2}^2}) \\
 & - 10 \frac{\kappa_{X;3} \kappa_{Y;3} (\kappa_{X;1} - \kappa_{Y;1}) (\kappa_{X;2} - \kappa_{Y;2})}{\kappa_{Y;2}^6}
 \end{aligned} \tag{5-13}$$

Where:

$$\begin{aligned}
 \kappa_{X;1} &= \mu_{X;1}; \kappa_{X;2} = \mu_{X;2} - \mu_{X;1}^2; \kappa_{X;3} = \mu_{X;3} - 3\mu_{X;2}\mu_{X;1} + 2\mu_{X;1}^3 \\
 \alpha_1 &= c_3 - 3 \frac{\alpha}{\kappa_{Y;2}}; \alpha_2 = c_4 - 6 \frac{c_2}{\kappa_{Y;2}} + \frac{3}{\kappa_{Y;2}^2}; \alpha_3 = c_6 - 15 \frac{c_4}{\kappa_{Y;2}} + 45 \frac{c_2}{\kappa_{Y;2}^2} - \frac{15}{\kappa_{Y;2}^3} \\
 c_2 &= \alpha^2 + \beta^2; c_3 = \alpha^3 + 3\alpha\beta^2; c_4 = \alpha^4 + 6\alpha^2\beta^2 + 3\beta^4 \\
 c_6 &= \alpha^6 + 15\alpha^4\beta^2 + 45\alpha^2\beta^4 + 15\beta^6 \\
 \text{and } \alpha &= \frac{\kappa_{X;1} - \kappa_{Y;1}}{\kappa_{Y;2}}; \beta = \frac{\kappa_{X;2}^{1/2}}{\kappa_{Y;2}}
 \end{aligned}$$

Because of the asymmetric property of KL divergence, the symmetric KL divergence

$$KLD = KL_{Edgeworth}(X, Y) + KL_{Edgeworth}(Y, X) \tag{5-14}$$

is used as one change indicator in this work.

Theoretically, the probability density function is always positive. However, negative function values can occur when it is approximated by Edgeworth series expansion because it relies on a Hermite polynomial, which has fluctuating properties. In addition, the KL divergence is



approximated using only the first four order statistics. In this case, some unusually large values or negative values can be obtained when calculating KL divergence. To avoid the influence of these values on further processing steps, in this approach, all of the negative values have been set to '0', and the 1% very large values have been also removed from the KL divergence result.

In contrast to change indicators, no-building change indicators show the possibility of low probability of being the change of interest. In this work, these no-building change indicators are introduced to our DS fusion model because of the characteristics of the data used and they are very important to achieve reliable results. This can also be seen as one major contribution in the overall workflow. As presented in Fig. 5-8b, after being stretched to a sigmoid curve distribution, some small height changes (3 ~ 5 m) have change probabilities of approximately 50%. DSM changes of this magnitude are caused mainly by errors in the DSMs or the growth of vegetation. In urban areas with simple building structures, as described in Chapter 4, vegetation and shadow areas can be successfully detected in the multispectral data in order to substantially reduce their influence on the final results. However, for urban areas with more complicated building structures, the differentiation can be more difficult, as, for example, vegetation can also be found on the roofs of buildings. Therefore, these features are only used to calculate a probability of no-building indication, while not making an absolute decision. The vegetation cover and shadow cover extraction methods, also applied here, have been introduced in section 4.4.2.

#### 5.3.4 Two-step DS fusion model

After obtaining all of these change indicators and stretching the change maps to sigmoid curve distributions, the DS fusion model has to be designed. This processing step is divided into two tasks. First, the building change indicators are fused. Second, the no-building change indicators are fused with the result of the first step.

As mentioned in section 5.3.2, DS fusion is able to handle not only single classes but also the union of several classes. Three classes are considered in urban areas, building changes ( $B$ ) land surface changes ( $S$ ) and no-change ( $N$ ). The hypotheses to be considered in DS formulation are:  $\emptyset$  (whose mass is null according to (5-8), and therefore, it does not need to be considered here), the simple hypothesis  $B$ ,  $S$  and  $N$ , and the compound hypothesis,  $B \cup S$ ,  $S \cup N$ ,  $B \cup N$ ,  $\Theta = B \cup S \cup N$ .

## 5 Fusion-based change detection

Table 5-1. Decision model for the six classes.

Type	$\Delta H$	$\Delta SM$	Combined Mass
$B$	$P_{\Delta H}$	0	$\frac{P_{\Delta H} \times P_{\Delta SM}}{1 - K}$
$S$	0	0	$\frac{(1 - P_{\Delta H}) \times P_{\Delta SM}}{1 - K}$
$N$	0	$1 - P_{\Delta SM}$	$\frac{(1 - P_{\Delta H}) \times (1 - P_{\Delta SM})}{1 - K}$
$B \cup S$	0	$P_{\Delta SM}$	
$S \cup N$	$1 - P_{\Delta H}$	0	
$B \cup N$	0	0	

Where  $B$ : Building change;  $S$ : Land surface change;  $N$ : No-change;  $\Delta H$ : Height change;  $\Delta SM$ : Similarity; and  $K = P_{\Delta H} \times (1 - P_{\Delta SM})$

Table 5-1 shows the pixel-based decision model used in this work. The purpose of this procedure is to distinguish  $B$  from  $S$  and  $N$ . The height changes ( $\Delta H$ ) indicate building change. We assign a probability mass  $P_{\Delta H}$  to the class  $B$ , and  $(1 - P_{\Delta H})$  to the combined class  $S \cup N$ .  $P_{\Delta H}$  is calculated according to the sigmoid curve as described in section 5.3.2. We suppose that for new, demolished or changed buildings, as well as height change, some spectral change is also likely. The KL divergence similarity ( $\Delta SM$ ) indicates building change and surface change, which are represented as the combined class  $B \cup S$ . We assign a probability mass  $P_{\Delta SM}$  to this combined union, and  $(1 - P_{\Delta SM})$  to  $N$ . In Table 5-1,  $K$  is the measure of conflict among different indicators, shown in Equation (5-9).

The second step is the fusion of the no-building change indicators to the building change probability map. Here we employ the result from Step One as one change indicator (i), and fuse it with the other two no-building change indicators separately based on Eq. (5-9). The no-building change indicators are considered and accepted in the fusion procedure when they have mass values  $m(A)$  greater than threshold  $T$ . According to the decision rule mentioned in Le Hégarat-Masclé et al. (1997),  $T = 50\%$  is used in this procedure. Three different cases are considered in the fusion procedure:

- 1)  $m(i) > T$ , while  $m(j) < T \Rightarrow$  Building change
- 2)  $m(i) < T$ , while  $m(j) > T \Rightarrow$  No-building change

3)  $m(i) > T$ , and  $m(j) > T \Rightarrow$  DS fusion

In case 1) of the fusion rules described above, if corresponding pixels exhibit a high certainty of building change  $m(i)$  and a low certainty of being vegetation or shadow  $m(j)$ , the decision is “building change”. If both indicators are higher than the assumed threshold ( $T$ ), as for case 3), the DS fusion is performed based on Eq. (5-9).

A building change probability map is generated after these two fusion steps. This intermediate result might already be important for change indication in a semi-automatic procedure (not considered here) because it can save time and manual work if an interpreter uses this initial information. Here, our goal is to identify the real changed buildings using a fully automated procedure. Therefore, a thresholding of the building change probability map is needed to receive an initial building change mask before the object-based building extraction. Since a further refinement is considered after we obtain the pixel-based building change map, in order to leave more candidates for the subsequent object filtering, a relatively lower threshold should be given.

After generating the building change mask, it is still necessary to separate “changed building” from false change alarms. This is the drawback of the pixel-based change detection method. Therefore, shape features or other object-based features as described in section 4.3.2 can be adopted to extract the real changed buildings and improve the change detection accuracy. Height, area (size) and convexity are three of the most important features of buildings used in our work (Tian et al., 2012a).

## 5.4 Summary

This chapter presents two fusion-based 3D change detection methods. Unlike the approach proposed in Chapter 4, both the methods presented here have used the information from satellite images to obtain the initial change maps. In section 5.2, the feature fusion method is described. This method is based on a better understanding of the quality characters of the Stereo-DSMs. Since the panchromatic images have much sharper boundaries than DSMs generated from them, it would be very helpful to use the regions provided by panchromatic images to design a region-based change detection workflow. In this procedure, a region merging strategy is proposed to handle the small regions produced from combining segmentation results of two dates. In section 5.3, a decision fusion method is described. This decision fusion model, fusing both the satellite images and Stereo-DSMs, provides a certain probability for whether or not an object of interest has changed. In this method, height changes from DSMs and KL Divergence from panchromatic

### *5 Fusion-based change detection*

images are used as change indicators, while shadow and vegetation maps from multi-spectral channels are used as no-building change indicators. These indicators are fused in two steps based on DS fusion theory.

## **6. Post-classification change detection using GIS data**

### **6.1 Introduction**

DSM-assisted change localization and fusion-based DSM-assisted change detection were presented in Chapter 4 and Chapter 5 respectively. All of these methods are based on directly comparing images to their corresponding DSM. In the past few years, 2D building footprints (mainly GIS vector data) have been completed in most industrialized countries (Champion et al., 2008). Along with the trend of rapidly growing cities, urban areas are also changing more frequently than before: it is therefore necessary to update these building footprints regularly. Although these footprints are usually updated manually, developing a post-classification building change detection procedure can save manual work and some processing time. Also, the use of pre-knowledge from GIS information makes the locating of changed buildings easier and more accurate. In order to update the existing building footprint, not only can destroyed buildings be marked and removed, but new buildings are expected to be detected and their footprint to be added. Digitizing a satellite image requires a great deal of time, especially when many changes occur, such as after disasters. Therefore, a robust building boundary extraction method is proposed for this building change detection approach. The extracted new building boundaries are finally used to update the original footprints.

### **6.2 Change classification**

#### **6.2.1 nDSM generation**

DSMs contain object elevation values. If the terrain in the test area has a large slope, it is not easy to obtain a building mask directly from a DSM. Thus it is necessary to separate the terrain height from the DSM. The normalized DSM (nDSM), which contains the absolute height of the non-ground area, can directly assist building extraction. Therefore, the first step of DSM-assisted building change detection is to remove the underlying elevation by subtracting DTM from the DSM. The DTM can be provided from other sources, or it can be calculated based on a suitable filtering methodology (Arefi et al., 2011) using the DSM. Another method uses grey-scale morphological opening, where the lowest point inside one window is defined as the ground terrain, and by this, a DTM can be generated (Weidner and Förstner, 1995) depending on the selected window size.

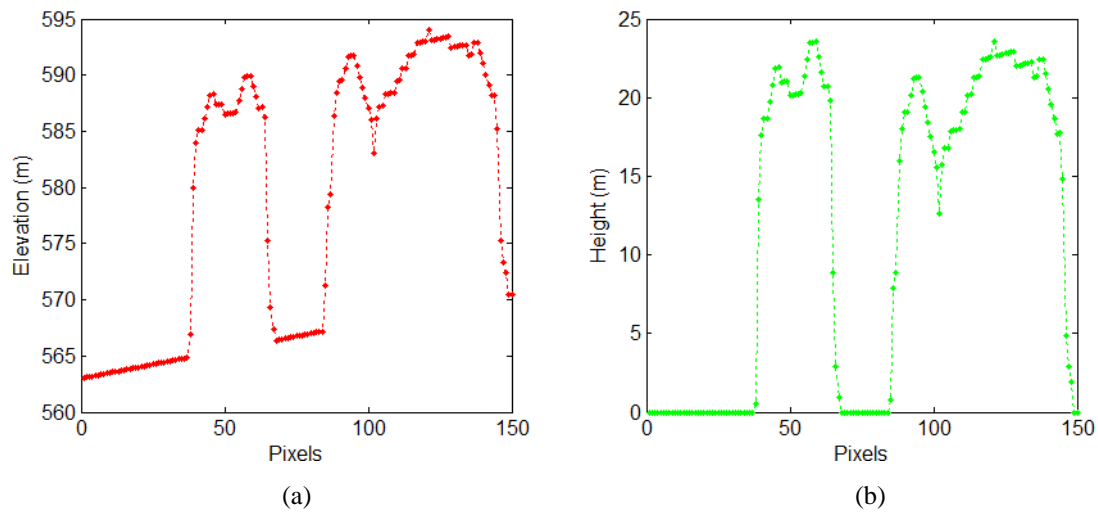


Fig. 6-1. DSM vs nDSM. nDSM (b) is generated from DSM (a).

As shown in Fig. 6-1, the nDSM (Fig. 6-1b) generated from the DSM (shown in Fig. 6-1a) contains the absolute height of each building. Even though nDSMs generated from morphological filters have the disadvantage of requiring a proper window size, they are adopted in this work due to their efficient and easy implementation.

### 6.2.2 Destroyed building detection

When a building footprint is provided together with stereo imagery, a co-registration between them is necessary. Automatic co-registration of the building footprint with satellite images is not an easy task. In this work, since we are not focusing on image registration, a manual registration is used. With the help of the existing building footprint, the destroyed buildings can be highlighted by extracting the average height within each building mask. Here, the average height of each building mask is calculated based on local values of the nDSM.

Fig. 6-2a shows the original DSM, which contains the terrain height. The extracted nDSM based on the method illustrated in section 6.2.1 is displayed in Fig. 6-2b. Fig. 6-2c features the raster image of building footprints. After giving each building mask the average height value from the nDSM, the resulting 3D building footprint is presented in Fig. 6-2d. As can be seen here, one destroyed building (marked with a white circle) is shown in light blue color, which corresponds to a small height value of about 4 m. This building can be easily identified as a destroyed building. The low height value might be introduced from vegetation cover or from reconstruction of the building.

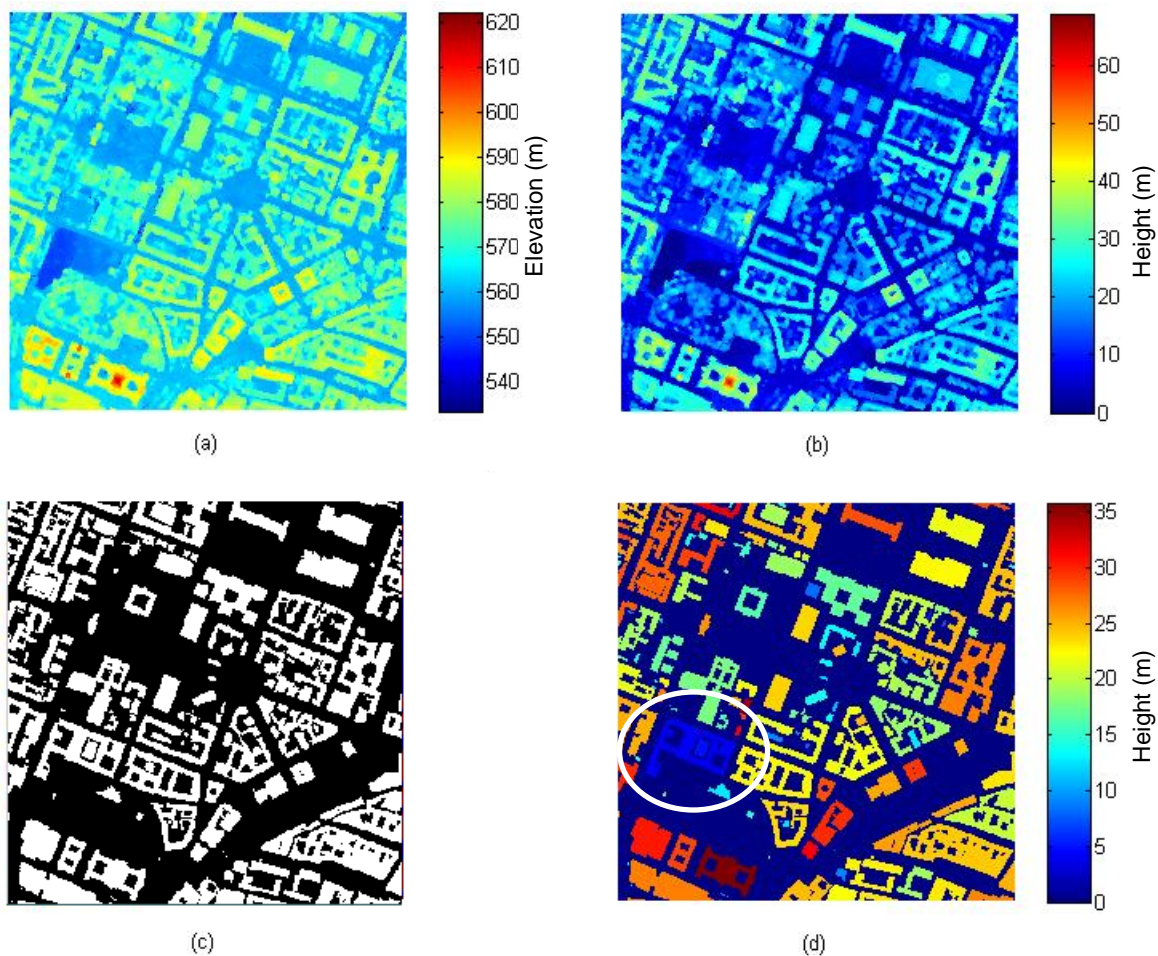


Fig. 6-2. Footprint-assisted destroyed building extraction: (a) Original DSM; (b) Calculated nDSM; (c) Outdated building footprint; (d) Resulting building footprint, the changed building is marked with white circle.

### 6.2.3 New-built building location detection

Extracting newly constructed buildings is much more challenging than detecting destroyed buildings. The accurate building boundaries of the new buildings are needed to update the existing building footprint. To detect newly constructed buildings, two processing steps are required. First, the locations of changed buildings have to be detected; next, the building boundary can be extracted.

Building location detection is not an easy task, especially in very dense urban areas. Some building change detection methods have skipped this step and instead manually specify the building location (Cui et al., 2012). This problem can be solved automatically when height information is available. Based on nDSM, the high-rise features can be extracted after setting a pre-defined threshold (Sohn

and Dowman, 2007). In urban areas, high-rise features are mainly buildings and trees. Again, among them, trees can be removed using NDVI. Fig. 6-3b shows the original DSM generated from the stereo image; one scene of the stereo pair is shown in Fig 6-3a. After the trees are removed, the nDSM containing mainly buildings is shown in Fig. 6-3c. A threshold of  $T = 5\text{ m}$  was used in this example to obtain the initial building mask. It has been refined with morphological erosion and dilation. After removing small masks composed of only a few pixels, the resulting automatically generated building mask, which can be used to compare with the building footprints from GIS data, is displayed in Fig. 6-3d. The initial building mask can only show the possible locations of each building. If vegetation areas are located on the roofs of some buildings, only the buildings without vegetation cover are shown in the initial building mask.

To compare the mask with the outdated building footprint, a comparison rule is required. Matikainen et al. (2003) use the cover rate of each corresponding mask to extract changed buildings. However, when comparing the Stereo-DSM mask and the building footprint, the mis-registration directly influences the cover-rate. Therefore, the robust-differencing described in section 4.2 is used here for mask comparison. The changed parts are projected to the mask from the Stereo-DSM to calculate the cover rate.

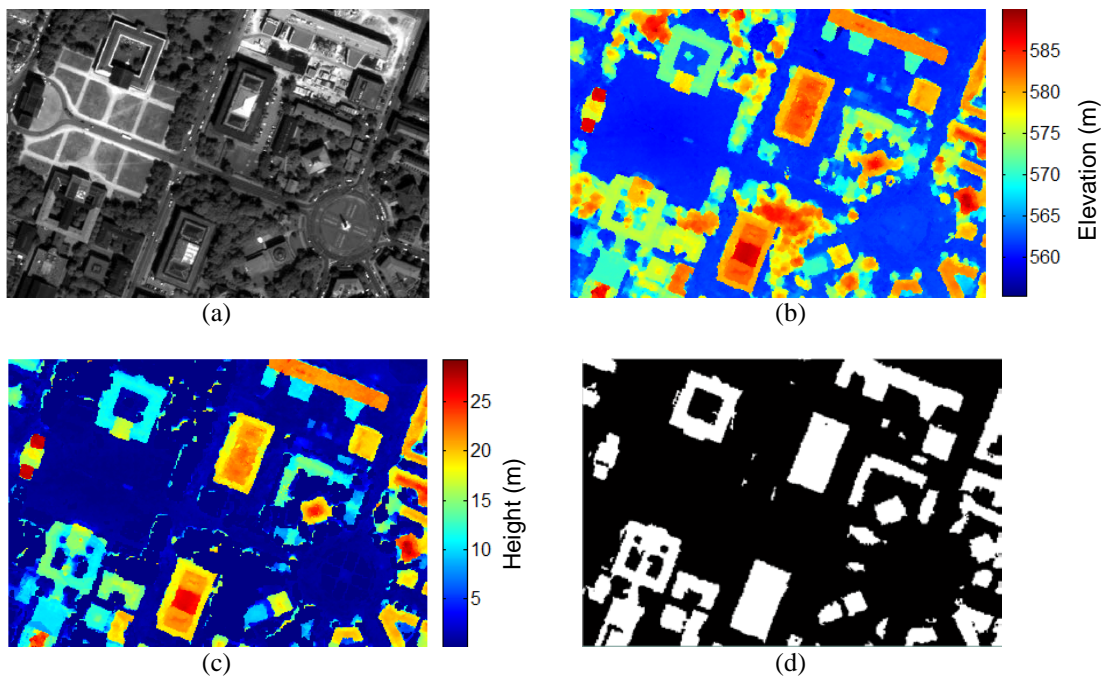


Fig. 6-3. Building location detection procedure: (a) Panchromatic image; (b) Original DSM; (c) nDSM after removing vegetation; (d) Cleaned initial building mask.



Comparison rules are created by referencing the rules from Matikainen et al. (2003).

- Over 80% of the building mask in the Stereo-DSM mask is covered by the changed part=> new building.
- 5-80% of the building mask is covered by the changed part=> rebuilt building.
- Less than 5% of the building mask is covered by the changed part=>non-changed building.

As buildings in urban areas are often rebuilt to a similar shape, or only enlarged, a smaller cover rate value is used as the threshold to separate the non-changed buildings and rebuilt buildings. After classifying these buildings, the corresponding rebuilt and newly built buildings are recorded in the building mask: these buildings will be further processed. The accurate boundaries of these buildings are required in order to update the outdated building footprint.

### 6.3 Building reconstruction

In order to automatically update the building footprint, a high quality robust building boundary extraction method is required. Automatic building boundary extraction has been one of the most challenging problems for VHR remote sensing data in recent years. In addition to 2-D information from multi-spectral satellite images, height information from DSMs has received increasing attention for automatic building extraction. A detailed workflow for DSM-assisted building extraction has been proposed in Baltasvias et al. (1995). Restricted to the DSM quality at that time, only a coarse model of buildings could be generated. Weidner and Förstner (1995) worked out a building extraction method based on only the DSM. Ridley et al. (1997) used airborne data to simulate the IKONOS stereo data for land cover extraction, from which only 73% and 86% of buildings could be interpreted correctly, while no automatic method was mentioned in that paper. After IKONOS imagery became commercially available, more detailed 3D building models were extracted, but this work was still based on manual interpretation (Fraser et al., 2002). Noronha and Nevatia (2001) proposed a building detection method based on multiple airborne data; however, only simple building shapes were considered.

Although many methods have been proposed and tested on stereo imagery, they either require a lot of manual work or they are only suitable for simple building shapes. A typical European city like Munich, Germany features more complex shaped buildings than rural, suburban or industrial regions. Therefore, more work is required for automatic building extraction from spaceborne stereo imagery, especially in high density building areas with complex building shapes. In the

LiDAR-DSM-based building extraction method, the quality of the extracted buildings' boundaries depends on the density of original LiDAR point clouds. Methods developed for airborne multi-spectral imagery and LiDAR-DSM can hardly be directly adopted for, for example, WorldView-2 stereo imagery, because of the potential incorrect height information and blurry boundaries of the generated DSM.

Here a novel efficient building extraction method is proposed. It combines height from DSMs, line segments from the panchromatic image together with building class probability from random forest classification based on the multispectral channels. A knowledge-based training data extraction method is introduced here for supervised classification. The flow chart of the building extraction procedure is shown in Fig. 6-4: the blue frame is the building location detection procedure described in section 6.2.3. Using the obtained changed building location, we will focus on building boundary extraction in the following. Firstly, building line segments are obtained from panchromatic images. As well as the direction of the building, rectangular segments can also be extracted based on these lines. These segments are classified from the automatic classification results, while the training data used in the classification procedure are automatically generated using a knowledge-based method.

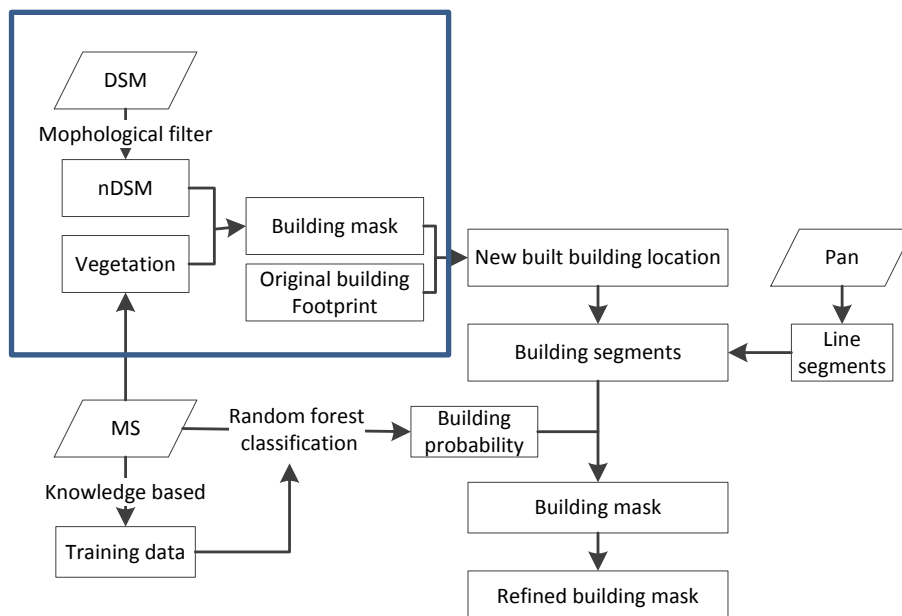


Fig. 6-4. Flowchart of the proposed method (Pan: Panchromatic image; MS: Multispectral image).

### 6.3.1 Hough Line-based segmentation

The Hough transform, which was first proposed by Paul Hough (Duda and Hart, 1972), is a common feature extraction technique used for computer vision and digital image processing. It was initially designed for line extraction and later extended to the detection of other shapes, like circles and ellipses. In this study, we aim to extract lines and rectangular segments in and around each building.

The Hough line extraction procedure is based on transforming the plane coordinates to polar coordinates. As shown in Fig. 6-5,  $x$  and  $y$  are the pixel coordinates in the image space. When representing them in polar coordinates, the parameter  $r$  is the distance between the line and the origin.  $\theta$  ( $\theta \in [0, 2\pi)$ ) is the angle of the vector from the origin to the nearest points in the line. The equation of the line can be written as:

$$y = \left(-\frac{\cos \theta}{\sin \theta}\right)x + \left(\frac{r}{\sin \theta}\right) \quad (6-1)$$

which can be rearranged to  $r = x \cos \theta + y \sin \theta$

The whole procedure of Hough transform-based line detection is shown in Fig. 6-6. The Canny lines (shown in Fig. 6-6b) extracted from the panchromatic image are transformed using a Hough transform. Fig. 4c displays all of the detected Hough lines. Using the Hough transform, the lines with the same  $\theta$  are intersected as explained in Cui et al. (2012). The pixels with the highest values in the transformed image represent the main direction ( $\theta_h$ ) of the building. Here it is assumed that each building has only two main directions, and that these two directions are perpendicular to each other. Then the other direction of the building is:

$$\begin{cases} \theta_h + 90^\circ & (\text{if } \theta_h < 90^\circ) \\ \theta_h - 90^\circ & (\text{if } \theta_h \geq 90^\circ) \end{cases} \quad (6-2)$$

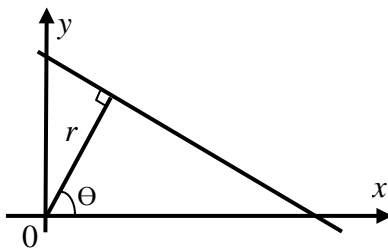


Fig. 6-5. Line representation.

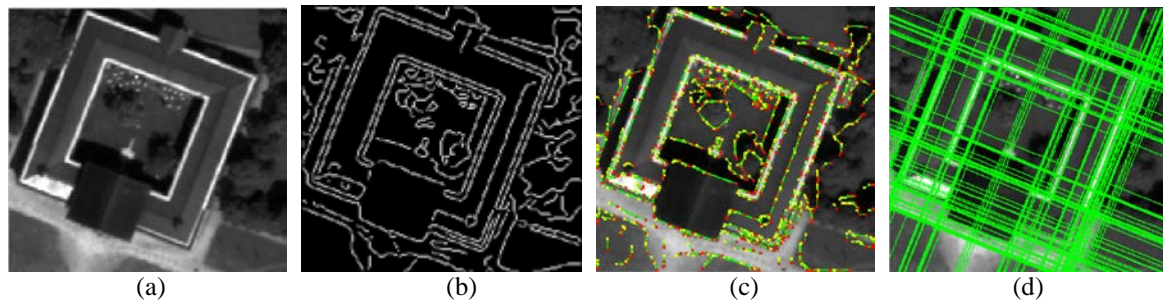


Fig. 6-6. Building boundary detection procedure: (a) Panchromatic image; (b) Canny edges; (c) Hough lines; (d) Extended Hough lines in main directions.

With the two main directions, the extracted Hough-lines can be refined. The lines parallel or near parallel (within  $\pm 5^\circ$ ) to these two directions are kept. The line segments extended to lines are shown in Fig. 6-6d. After obtaining the Hough lines, Cui et al. (2012) and Grigillo et al. (2012) have tried to use the intersection points to extract building boundaries. However, these methods would be very difficult to apply in very complex buildings or those without rectangular corners, like buildings with holes in the middle or with curved boundaries.

Separating buildings into 2D polygons has been shown to be an efficient method in building extraction (Schimid and Zisserman, 1997; Zebedin et al., 2008). After obtaining the rectangular segments, Zebedin et al. (2008) used only the height information based on a Graph-cut to merge those segments belonging to buildings. The DSM used in this work is not as accurate as the one used by Zebedin et al. (2008), but corresponding spectral information is available: thus, in this work, instead of a Graph-cut, we applied a classification method instead to classify the obtained rectangular segments into buildings and no-buildings. Besides height information from DSM, the spectral channels and texture information can be used as features in the classification procedure.

### 6.3.2 Automatic training data selection

Many supervised classification techniques are used to classify remote sensing images, from maximum likelihood to support vector machine or random forest. Their efficiency and accuracy have been demonstrated in several publications (Dempster et al., 1977; Bruzzone et al., 2006). Moreover, when specific objects are of interest (in this case, buildings), it is more logical to use supervised classification. However, these methods cannot be directly used for automatic classification, as they require training data. Training data selection is time consuming, and incorrect or insufficient training data (e.g. number of classes) may weaken the finally accuracy. Using IKONOS stereo images, Shaker et al. (2011) employed manually selected training data for building classification. Manual training data selection requires quite a lot of work, especially when

buildings' roofs are in various colours and are built with various materials. Moreover, a fully automatic building extraction is more desirable, especially if VHR satellite images, like WorldView-2 data, are available. Therefore, we adopted the knowledge-based classification method to automatically select training samples.

Knowledge-based decision rules can make the classification easy and fast, if the knowledge is accurate and detailed enough (Baltsavias, 2004). However, normally only part of the knowledge can be obtained in advance, which complicates the class definition. Rule-based classification is one of the typical knowledge-based methods where rules are presented with one or more decision trees (Huang and Jensen, 1997; Lawrence and Wright, 2001). However, each root of the decision tree needs at least one threshold and the accuracy of these thresholds directly influences the classification result. The pixels with much higher or lower values compared to the threshold usually produce more correct classification results; values close to the threshold may lead to false alarms.

Therefore, in this work, existing knowledge is used to build an incomplete decision tree. This incomplete decision tree means it cannot classify all of the pixels in an image to defined classes. Only the pixels with high probability of belonging to a defined class are labelled. For example, only the pixels with a height more than 10m and a high NDVI value are classified as 'tree'. It is difficult to define a 'high' value for the NDVI. Thus, the NDVI and shadow values are projected to a sigmoid curve to obtain the Belonging-Certainly Value (BCV) from 0% to 100% using the same method described in section 5.3.2. The training data selection rules are listed in Table 6-1. Thresholds ( $T = 80\%$ ) are selected by long-term experience; since the vegetation and shadows are based on BCV, the thresholds can be easily used for all data.

With the proposed uncompleted decision tree, most of the pixels can be classified into an appropriate class. The pixels that cannot be classified are mainly located in the boundary regions: for example, the pixels between buildings and shadows, or grass and roads. Fig. 6-7 shows an example of the training data generated. After the rule-based labelling, segments with size smaller than 100 pixels have been removed from each class layer. As can be seen here, the main part of buildings, the centre of the roads and trees are extracted successfully. The unlabeled pixels, named as background pixels, marked in blue, are remains on the building and tree edges. In these regions, even manual interoperation cannot properly label these pixels. A more intelligent supervised classification is required to finally classify these pixels.

Table 6-1. Knowledge-based training data selection rules.

Classes	Rules		
	Height (m)	BCV_veg [%]	BCV_shadow [%]
Trees	>10	>80	<80
Grass	<2	>80	<80
Building	>10	<50	-----
Shadow	<10	-----	>80
Ground	<2	<50	<50

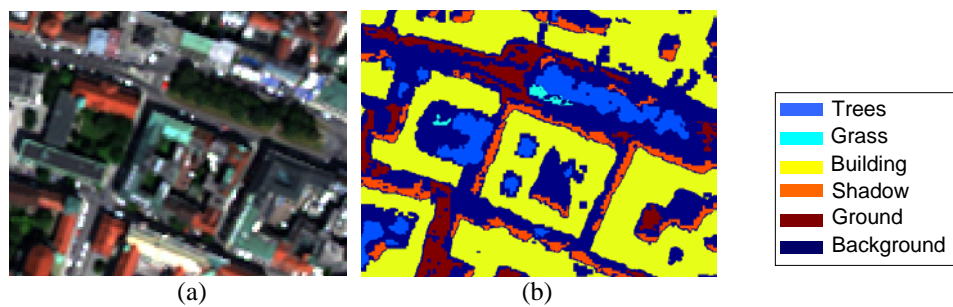


Fig. 6-7. Generated training data (with five classes) (b) based on WorldView-2 (a).

### 6.3.3 Random forest-based segments classification

Random forest is a robust and powerful machine learning classifier, and is able to process large datasets (Breiman, 2001). The random forest consists of several decision trees, also called classification trees. These classification trees are “grown” based on training sets. Random forest builds the tree nodes randomly with random features. Each decision tree is used separately to classify the satellite image of interest. By summarizing these classification results, one classification map and several membership maps for each class can be obtained.

In this work, random forest classification is adopted to extract the building class based on the generated training data. All of the channels from multispectral images, together with panchromatic image and DSM, are used as input features. Since only the building class is of interest, instead of directly using the classification result, the probability map of building classes is used in the later process. In the probability map, the pixel value means BCV to the building class. To combine the classification result with the 2D segment obtained from the method described in section 6.3.1, these probability values from each pixel are averaged to each 2D segment. In the end, the final building mask is generated after thresholding.

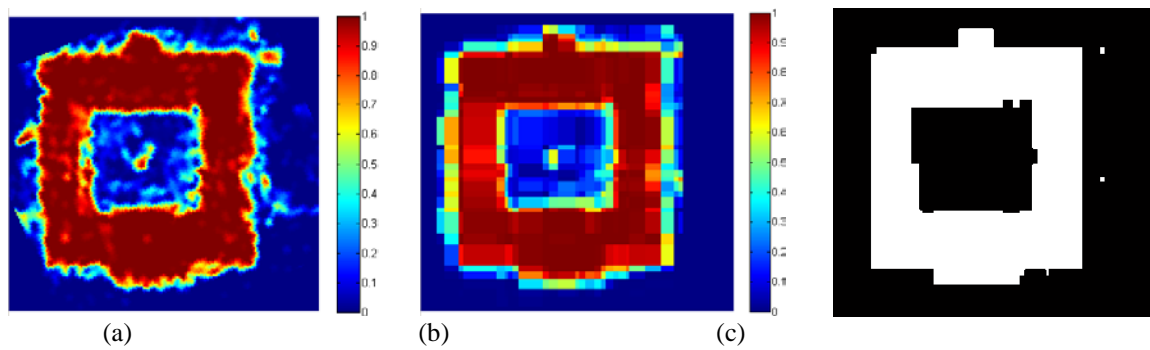


Fig. 6-8. Building footprint extraction procedure: (a) Original building class probability map; (b) result after fitting values in (a) to the obtained rectangular segments; (c) Generated building mask.

Fig. 6-8 shows an example of the steps in the building outline extraction procedure. Fig. 6-8a is the obtained building class probability map; pixels with red color have high probability of belonging to a building. The pattern shown closely resembles a building shape with blurred edges. After fitting these values to the extracted rectangular segments (Fig. 6-8b), the building shapes are more regular than before. As displayed in Fig. 6-8c, the final building mask can be obtained after thresholding. Finally, small mistakes remaining in the refined building mask are eliminated by using morphological filters. With the obtained building mask, building boundaries can be easily extracted. The extracted building boundaries can be added to the outdated building footprint map, thus updating it.

## 6.4 Summary

In this chapter, a post-classification approach is described. When comparing the satellite images with the outdated building footprint, it is not necessary to check all buildings one by one. The changed building locations are detected using mask comparison. In section 6.2, a fast change classification method is proposed by referring to existing studies. After locating the changed buildings, section 6.3 focuses on reconstructing these building locations. For this purpose, a novel automatic building extraction method is proposed. 2D segments are extracted using Hough lines and these segments are further classified using the supervised classification method “random forest”. To avoid manually selecting necessary training samples, a knowledge-based training data selection approach is proposed for the classification. Finally, the changed buildings are extracted by combining the 2D segments and the classification results.





## 7. Experiments

In this chapter, the experiments performed with the three proposed DSM-assisted change detection methods described in Chapters 4 - 6 are presented. The goal of these experiments is to examine the performance of the proposed approaches for different stereo images. In this thesis, four test areas with different kinds of sensor combinations and change objects are investigated. The required DSMs are generated using stereo imagery with the method described in Chapter 2. The four selected test areas are located in Dong-an, Munich, Istanbul and Oberammergau respectively, shown in Fig. 7-1. Each dataset consists of two pairs of stereo images captured at different dates, generated DSMs and manually extracted change reference data.



Fig. 7-1. Locations of the four test areas.

## 7 Experiments

The four test areas represent four different scenarios. Among them, the Dong-an test set features a typical industrial area, acquired by the IKONOS satellite with 1 m Ground Sample Distance (GSD). The Munich test set was captured using two different sensors: IKONOS and WorldView-2. This test area is located in the centre of Munich, showing a dense urban area. The other two datasets were both captured by the Indian satellite Cartosat-1, with 2.5 m GSD. One test area is located in an industrial area (Istanbul), and the other in a forest area (Oberammergau, Bavaria). The evaluation procedure includes both the visual inspection of the change maps and the quantitative evaluation of the change maps and masks. In the following, the test areas and experimental results will be described in more detail.

### 7.1 Experimental set-up and quality assessment methods

#### 7.1.1 Experimental set-up

The experiments are conducted by comparing the performance of the change detection methods described in Chapters 4, 5 and 6 for each test site. Based on these three approaches, four methods are adopted here:

- Method I: DSM-assisted change localization. In this method, change maps are generated using the robust change difference. Change masks are refined using morphological filtering, knowledge about shadow location and the vegetation mask. Following this, shape-based filtering is employed, as illustrated in Chapter 4.
- Method II: Feature fusion method as described in section 5.2. After the original regions are extracted from orthorectified panchromatic images, the proposed region merging approach is applied to obtain the initial homogeneous regions. Change features from DSMs and images are combined to generate the initial change map by thresholding and a change mask is generated. After removing the small regions (noise) in that mask, the final change mask is generated.
- Method III: Decision fusion method as described in section 5.3. Even though this method is designed for building change detection, the change indicator fusion step can also be used for forest change detection, but not the no-change indicators. Since Cartosat-1 data contain only panchromatic channels, the desired no-change indicators cannot be extracted. Thus, in the first two test areas, the change map is generated using two fusion procedures, while for the other two test areas only the first fusion is applied. In this method, the initial change

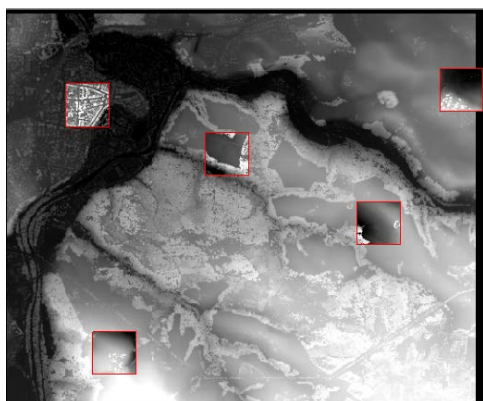


Fig. 7-2. Example of selected regions used for 3D-shift determination.

mask is generated by thresholding. The refined change mask is generated by using the shape features.

- Method IV: Post classification as described in Chapter 6. This method is only suitable for building change detection and needs a building footprint map as *a priori* knowledge. Thus this method could only be tested in the Munich test area.

DSMs for all four test areas are generated using SGM, followed by a 3D co-registration as described in Chapter 2. In cases where a large study site is used, instead of using the whole images, some regions, distributed evenly over the entire scene (shown in Fig. 7-2), which exhibit relatively flat terrain and low variation, are chosen for the 3D co-registration. Then, the values from these regions are averaged to compute a final shift value. The obtained shift values are applied to both DSMs and orthorectified satellite images.

### 7.1.2 Evaluation method

In order to evaluate the effectiveness of the proposed methods for the detection of the position and size of the changed objects and the overall change situation, the results obtained are compared with reference data. The evaluation is performed for each test area with the four described methods. The assessment of the change detection results is carried out at both pixel and object level. At the pixel level, the change map and change masks are evaluated separately.

- Pixel-based evaluation

For the pixel-level evaluation, the results are displayed in terms of Receiver Operating Characteristics (ROC) curve analysis (Hand and Till, 2001). We use the area under the ROC curve to evaluate the quality of each change index and the generated result. For each threshold between

## 7 Experiments

the maximum and minimum value in the result, we calculate the percentage of true positives (the building change pixels that are correctly detected as building change) and the percentage of false positives (no-building change pixels that are falsely detected as building change pixels). The ROC curve shows the relationship of true positive against false positive. The area under the ROC curve (AUC) is used to measure the ability of single or combined features and the difference map to detect the real building changes (Hand and Till, 2001).

In the pixel-based change mask evaluation, in order to compute the agreement of the change detection mask with the reference change mask, the following measures are involved.

- a) True Positive ( $TP$ ): the number of changed pixels correctly detected as changed.
- b) True Negative ( $TN$ ): the number of unchanged pixels correctly detected as unchanged.
- c) False Positive ( $FP$ ): the number of unchanged pixels incorrectly detected as changed.
- d) False Negative ( $FN$ ): the number of changed pixels incorrectly detected as unchanged.
- e) Overall Accuracy:  $OA = \frac{TP+TN}{N} \times 100\%$
- f) Kappa Index of Agreement (Congalton, 1991):  $KIA = \frac{Pr(a)-Pr(e)}{1-Pr(e)}$

Where  $Pr(a)$  is the relative observed agreement between the extracted results and reference data, it is calculated in the same way as  $OA$ , while

$Pr(e) = \frac{(TP+FP)*(TP+FN)+(FN+TN)*(FP+TN)}{N*N}$  represents the hypothetical probability of the agreement between the extracted result and reference data;

$N$  is the total number of pixels.

- Object-based evaluation

Concerning building change detection accuracy, correctly detected building numbers are in some cases more important. As a higher level of the analysis, the changed buildings are treated as single objects without consideration of their size. Only the effectiveness of the detection of distinct changed areas (buildings) in the change map is considered in the assessment. Therefore, four parameters are measured to evaluate the object-based change detection result:

- a) True detected number ( $TDN$ ): The number of changed objects that are correctly detected as changed.
- b) True detected rate ( $TD$ ): The number of true detected objects as a percentage  $TD = TDN / N_R \times 100$ .
- c) False detected number ( $FDN$ ): The number of unchanged objects that are incorrectly detected as changed.

- d) False detected rate (*FD*): The number of falsely detected objects as a percentage  $FD = FDN / N_D \times 100$ ,

where  $N_R$  and  $N_D$  are the total number of changed objects of the reference data and within the change map, respectively.

## 7.2 Dong-an test area, North Korea

### 7.2.1 Data introduction

The first dataset exhibits very high resolution, same sensor image data acquired over an industrial area. It consists of two 1 m resolution IKONOS stereo images, which were acquired on February 12, 2006 and May 2, 2011 respectively. The dataset represents an industrial region in Dong-an, North Korea with an area of  $0.6 \times 0.6 \text{ km}^2$  (equivalent to  $600 \times 600$  pixels). It contains mainly well separated, regularly shaped smaller size buildings. The panchromatic images and DSMs from both dates can be seen in Fig. 7-3 (a-d). Fig 7-3e shows the digitized reference change map: green objects represent newly built buildings, red objects represent destroyed buildings and blue objects represent rebuilt buildings. The reference data include high buildings, normal houses and tents. A tent with an area size less than  $100 \text{ m}^2$  and low height cannot be detected accurately enough with 1 meter resolution satellite images and DSMs from stereo matching. Thus in this research, only buildings larger than  $100 \text{ m}^2$  are considered in the reference data, procedures and experiments.

## 7 Experiments

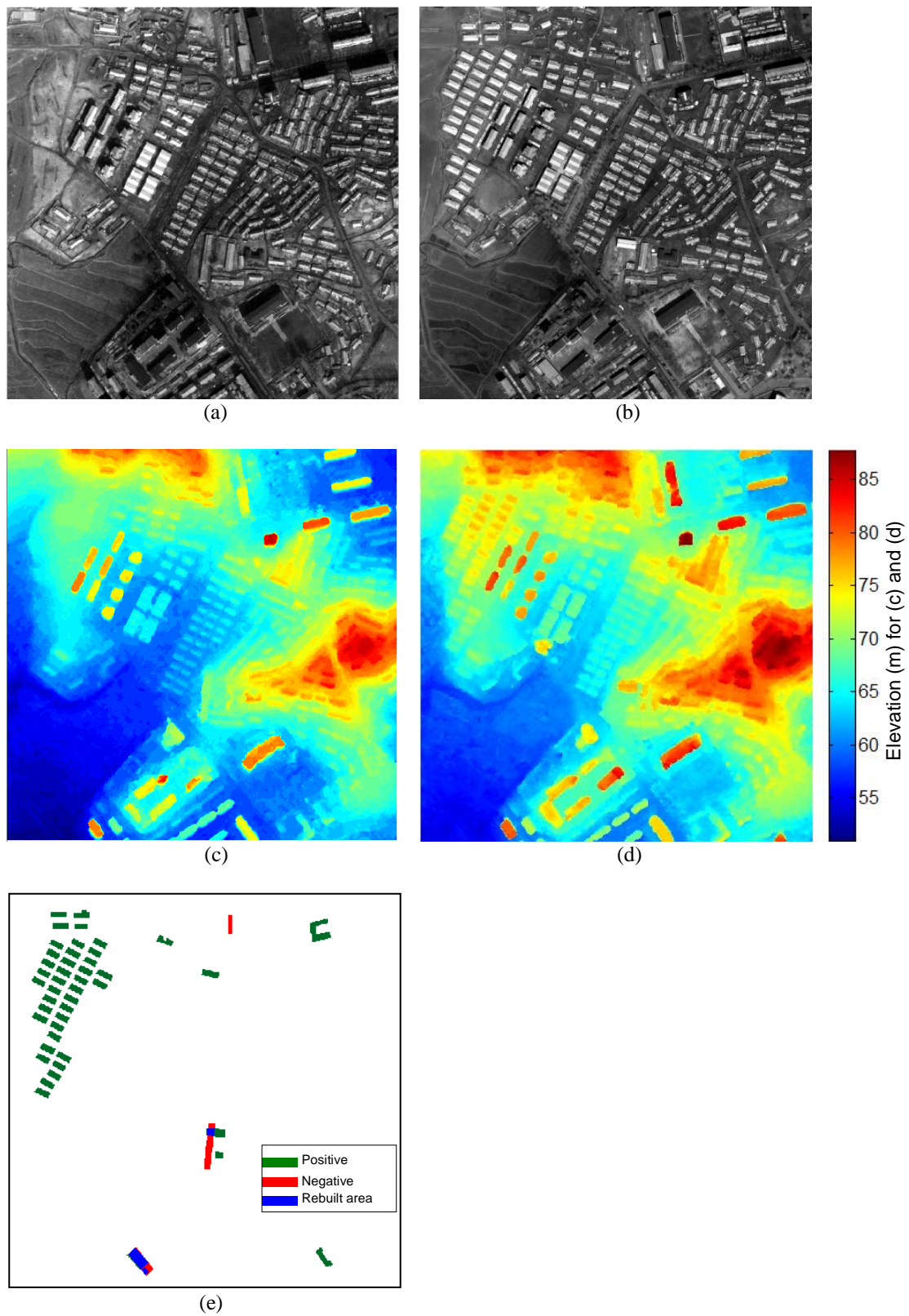
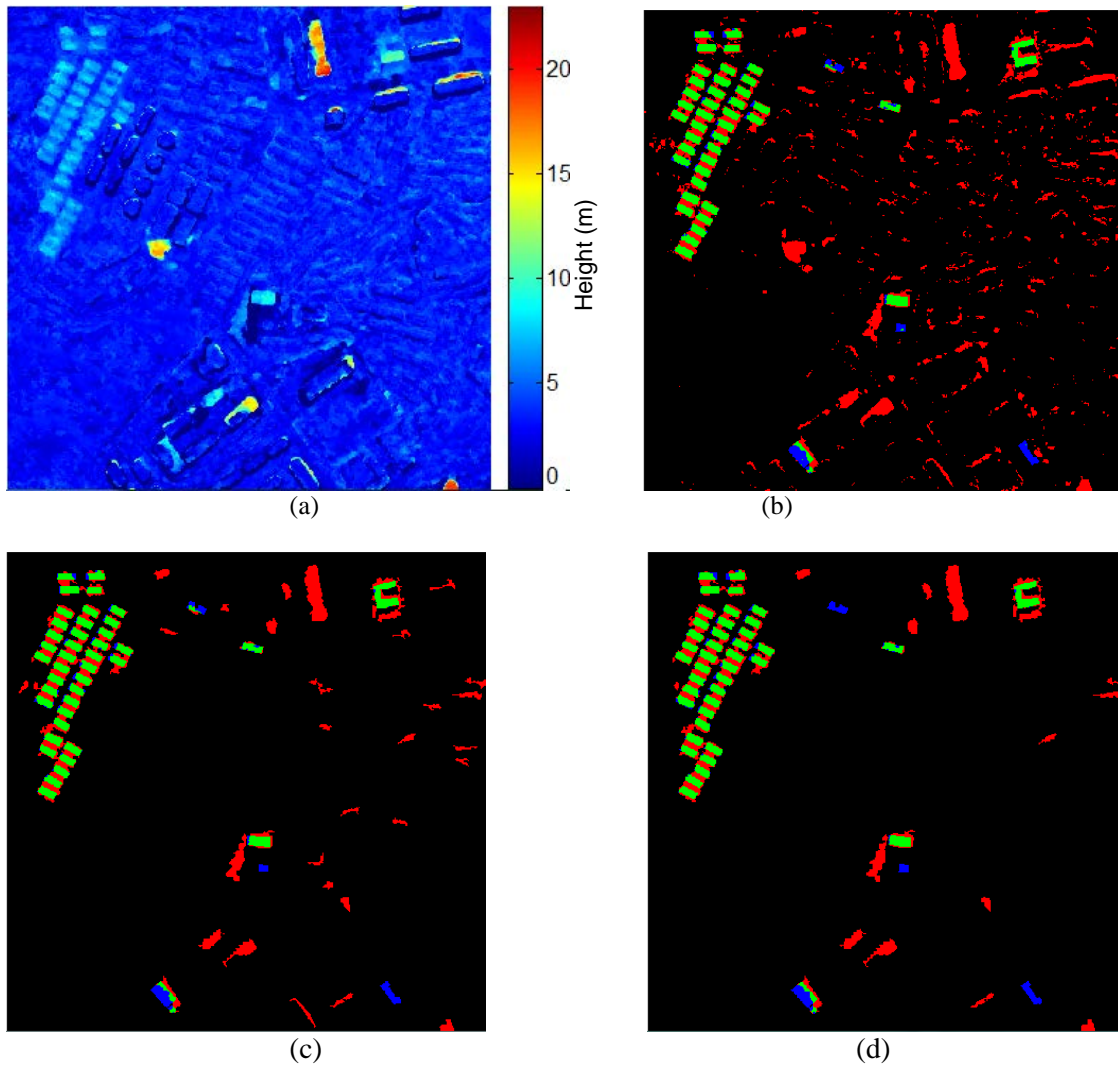


Fig. 7-3. Dataset in the Dong-an test area. Orthorectified panchromatic image from date 1 (a) and date 2 (b), DSM from date 1 (c) and date 2 (d), and reference change map (e).

### 7.2.2 Results

Due to the lack of available building footprints, only the first three methods are tested. The results are shown in Fig. 7-4. In order to present the accuracy of the mask, the detected change mask is overlaid with the change reference map. Although both positive changes (newly built buildings) and negative changes (destroyed buildings) are detected in this test area, only the positive change detection results are evaluated in this thesis, since only a very small number of negative changes are included in this dataset. As the two DSMs used in the test area are from the same sensor, a small window size is used for the robust DSM difference. Fig. 7-4a shows the robust height difference map generated with a  $3 \times 3$  pixel window size. Fig. 7-4b features the height-based thresholding result. A pre-designed threshold ( $T = 5$  m), with general validity, is used in this procedure. After refining the change mask by using shadow and vegetation masks, as well as a series of morphological refinements, the result is shown in Fig. 7-4c. The true detected pixels are shown in green; false alarms (no-change pixels falsely detected as change) are displayed in red; the blue pixels represent the missed alarms (changed pixels falsely detected as no-change).

The change map and change masks resulting from Method II are presented in Fig. 7-5. Fig. 7-5a is the generated change map, which has been transformed to sigmoid distribution for better understanding. After the transformation, all of the pixels in the change map are in the value range 0-1. Therefore, it can be used as a change probability map. In this experiment, we used ( $T = 40\%$ ) to generate the change mask displayed in Fig. 7-5b. To remove the noise present in this scene, only the change masks with a size larger than 100 pixels ( $100 \text{ m}^2$ ) are preserved in the final change mask (shown in Fig. 7-5c). Both change masks are overlaid with reference data. Compared to the change masks displayed in Fig. 7-4 from Method I, the changed buildings are better separated from each other.



	Reference data		
		Change	No-Change
Change Mask			
	Change	<b>True detected</b>	<b>False alarm</b>
	No-Change	<b>Missed alarm</b>	

Fig. 7-4. Change map and masks generated in the Dong-an Test area with Method I:

- (a) Robust difference map with positive changes;
  - (b) Initial change mask overlaid with reference data ( $T = 5$  m);
  - (c) Mask refined using morphological filtering and the land cover mask;
  - (d) Refined mask overlaid with reference data ( $T_{H_{region}} = 5m$ ;  $T_{size_{region}} = 100$  pixels;  $T_{convex_{region}} = 0.5$ );
- (b) - (d) Green: true detected; Red: false alarm; Blue: missed alarm.



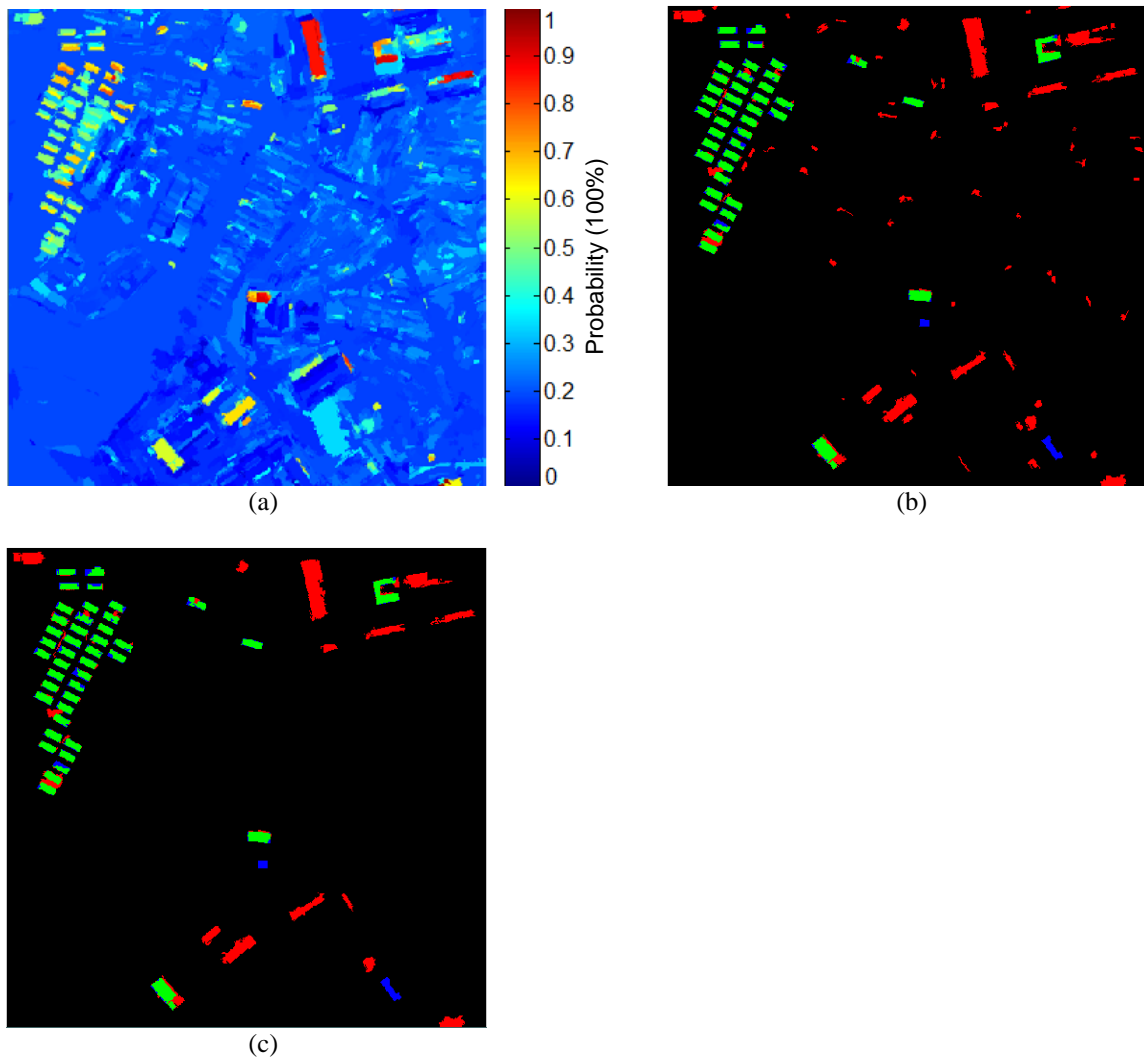


Fig. 7-5. Change map and masks generated in the Dong-an test area with Method II:

- (a) Region-based change probability map;
  - (b) Initial change mask overlaid with reference data ( $T = 0.4$ );
  - (c) Refined mask overlaid with reference data ( $T_{size_{region}} = 100 \text{ pixels}$ );
- (b) - (c) Green: true detected; Red: false alarm; Blue: missed alarm.

The decision fusion based Method III has also been applied on the same dataset. After two steps of DS fusion, Fig. 7-6a shows the detected change map. The change masks after thresholding and after shape feature based refinement are shown in Fig. 7-6b and Fig. 7-6c.

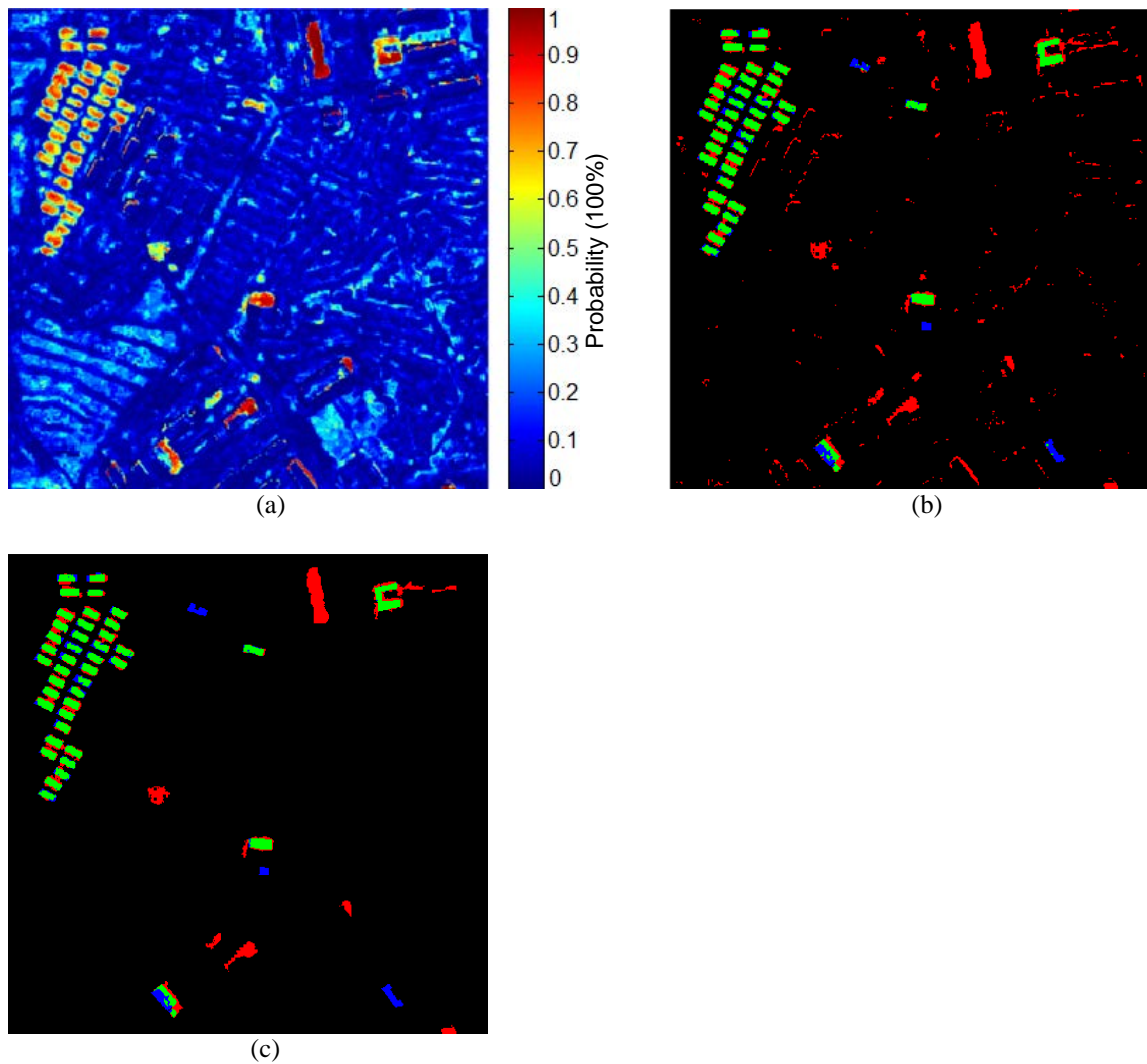


Fig. 7-6. Change map and masks generated in the Dong-an test area with Method III:

- (a) Change map after two steps of decision fusion;
- (b) Initial change mask overlaid with reference data ( $T = 0.5$ );
- (c) Refined mask overlaid with reference data ( $T_{H_{region}} = 5m$ ;  $T_{size_{region}} = 100 \text{ pixels}$ ;  $T_{convex_{region}} = 0.5$ );
- (b) - (c) Green: true detected; Red: false alarm; Blue: missed alarm).

### 7.2.3 Evaluation of results

The goal of this experiment was to evaluate and compare the effectiveness of the proposed frameworks for the VHR stereo imagery from the same sensor. The change detection results from all three methods are compared to the reference data in the evaluation procedure. Table 7-1 comprises the quality comparison of the change maps generated from 2D features and the three proposed methods. Three 2D features: Absolute image difference, KL divergence and multi-level features and height difference as 3D feature are considered. They have been used as intermediate results for change features in the proposed frameworks. AUC has been adopted to compare them

Table 7-1. AUC comparison in the Dong-an test area.

Method	AUC
Absolute image difference	0.8132
KL divergence	0.9122
Multi-level	0.8354
Height Difference	0.8993
I	0.9299
II	0.9299
III	0.9621

with the change maps displayed in Fig. 7-4a, 7-5a and 7-6a. A higher AUC value indicates a better quality of change map, meaning that it is more capable of providing a high accuracy change mask. As can be seen in Table 7-1, KL divergence achieves the best 2D change result in this test area. But Method I, II, III show higher values than the four intermediate results. Although the result from Method II seems visually much better than Method I, which only uses the DSMs, the AUC has not improved (and is coincidentally the same as for Method I). The decision fusion-based Method III, on the other hand, has substantially improved the AUC value.

The improvement in change masks in each step has been evaluated for both the pixel-based and the object-based method and the results are recorded in Table 7-2 and Table 7-3 respectively. The meaning of the evaluation parameters is explained in section 7.1.2. For Method I, according to the results listed in Table 7-2, the resulting accuracy in both OA and KIA are improved when more information is used. Among them, the decision fusion method achieves the best accuracy.

It is noteworthy that in Table 7-3, although Method I and Method II share the same false detected building numbers, the false detection rates differ considerably. That can be explained by the total number of detected buildings. In the change mask from Method II, all of the small buildings in the upper left corner of the images are well separated. For Method I, the initial mask is generated by thresholding the height difference map. As this is restricted by the DSM quality, these buildings cannot be separated correctly. Due to the lower building heights and high building density, the shadow mask in this scene cannot be generated optimally: therefore, after refinement, many buildings are still fused together in the final change mask; they are counted as one building in the evaluation procedure. That is why the change mask extracted from Method I has fewer buildings than from Method II, which influences the false detected rate.

## 7 Experiments

Table 7-2. Pixel-based change masks evaluation in the Dong-an test area.

Method	Description	Overall Accuracy (OA)	Kappa Index of Agreement (KIA)
I	Initial mask	0.9424	0.4227
	Land cover refinement	0.9650	0.5540
	Shape feature based refinement	0.9688	0.5817
II	Initial mask	0.9711	0.5951
	Remove small size regions	0.9765	0.6452
III	Initial mask	0.9705	0.5866
	Shape feature based refinement	0.9791	0.6686

Table 7-3. Region-based change masks evaluation in the Dong-an test area.

Method	True Detected (Object)		False Detected (Object)	
	Number	Rate [%]	Number	Rate [%]
I	42	93.33	13	56.52
II	43	95.56	13	26.53
III	42	93.33	7	17.95

Consequently, based on the evaluation result of the Dong-an test region we can conclude, that in general Method III works better for VHR resolution imagery. However, Method II produces sharper boundaries than Method I and Method III. If one of the input DSMs exhibits large regions with height errors, like the false-alarm-building in the upper-right corner in Fig. 7-4, all of these three methods will produce false alarm.

### 7.3 Munich test area, Germany

#### 7.3.1 Data introduction

The second dataset consists of two VHR images from different sensors acquired in the centre of Munich, Germany ( $1.3 \times 1.3 \text{ km}^2$ ). As can be seen in Fig. 7-7, this dataset features a typical urban environment, here the historic city of Munich. The building heights in this scene range from 5 up to 100 meters.

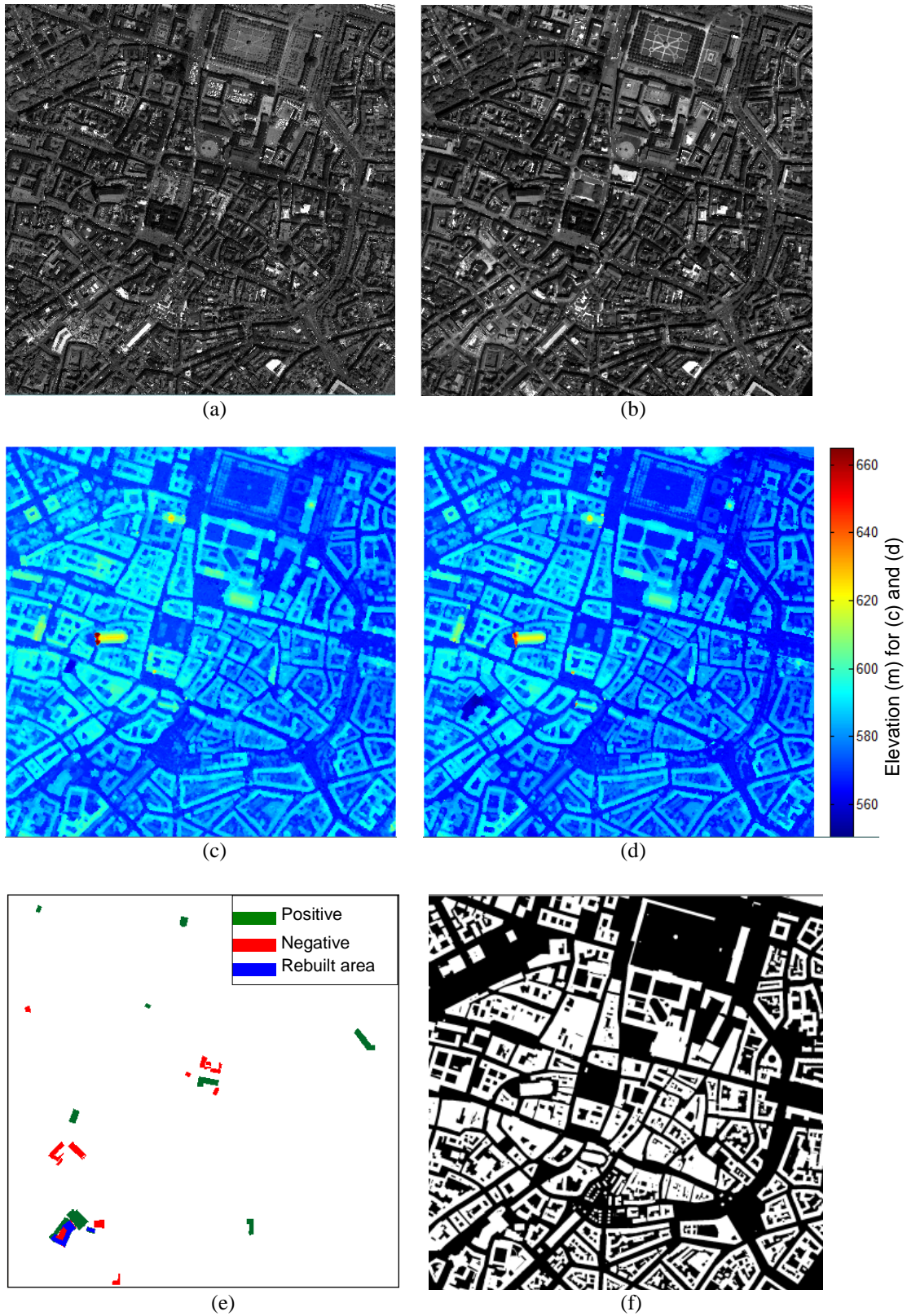


Fig. 7-7. Dataset in the Munich test area. Orthorectified panchromatic image from date 1 (a) and date 2 (b), DSM from date 1 (c) and date 2 (d), reference change map (e) and building footprint from 2003 (f).

## 7 Experiments

The first pair of stereo images was captured by IKONOS on July 15, 2005, while the other pair features WorldView-2 data acquired on July 12, 2010. Since the datasets have different resolutions, the WorldView-2 images are down sampled to 1 m resolution, to be equal to the IKONOS data. The test area contains more complicated building shapes as well as buildings that cannot easily be separated from each other. As well as the satellite image data, the building footprint from 2003 is also provided (shown in Fig. 7-7f): thus, Method IV, based on post-classification, can be tested here.

### 7.3.2 Results

All of the four methods proposed in this thesis are tested and compared in this test area. The results obtained are displayed in a similar way to the results from the Dong-an test site. All positive changes are chosen to display and evaluate the detected results. To better visualize the quality of the extracted changed masks, all of them have been overlaid with the reference map. Three colours are shown.

In Fig. 7-8, the generated change map and final change detection masks from Method I are shown. Due to the multi-sensor differences, a large window size of  $5 \times 5$  pixels is used for the computation of the robust difference. The resulting change map is displayed in Fig. 7-8a, where green pixels highlight large height change values (change more than 25 m) and blue pixels indicate no change or small height change values.

Fig. 7-9 presents the change detection result based on the region-based method. Fig. 7-9a shows the change map. The advantage of the region-based method is not as obvious as in the industrial area. Although most of the important building changes are highlighted correctly, some no-building change areas are also marked with a high change probability. After thresholding ( $T = 0.8$ ), one alarm, marked with a white circle, is still present in the change mask (Fig. 7-9b). Since it exhibits a relatively large size and regular shape, the false alarm is also shown in the final change mask (Fig. 7-9c). But the other five large changed buildings are correctly detected, as displayed in Fig. 7-9c in green.

The building change detection results based on Method III are depicted in Fig. 7-10. Fig. 7-10a shows the DS fusion result, including all the change indices and no-change indices. The values in the image represent the probability of each pixel containing building change. The changed buildings are mostly highlighted in red. Fig. 7-10b shows the thresholding result. The green masks represent possibly changed buildings. A combined measure of area, height and convexity is applied to extract the really changed buildings, as shown in Fig. 7-10c. The final change mask

obtained shows better quality than the results from the other two methods. Only three false alarm small building changes remain in the final change mask.

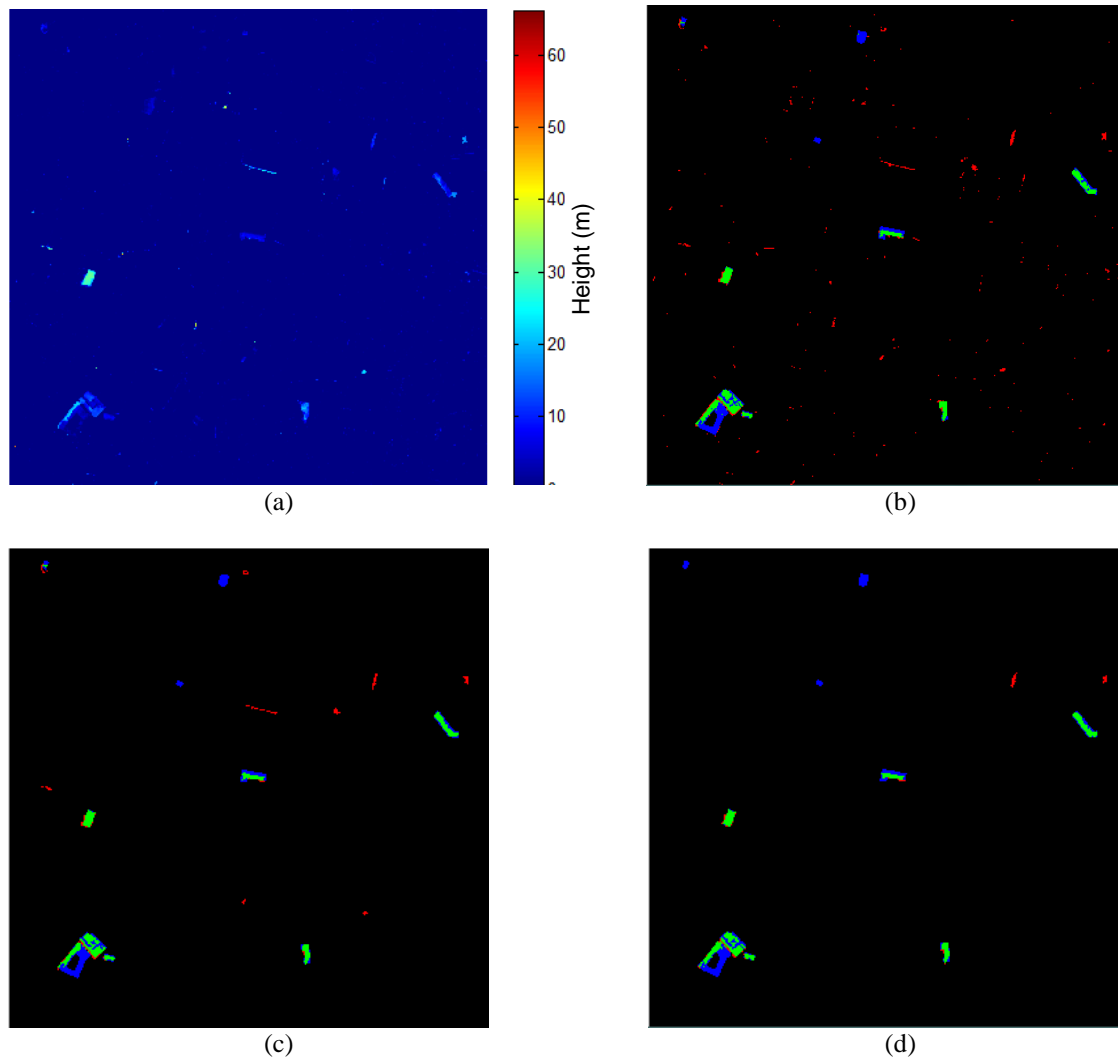


Fig. 7-8. Change map and masks generated in the Munich test area with Method I:

- (a) Robust difference map with positive change;
  - (b) Generated change candidates after thresholding overlaid with reference data ( $T = 5$  m);
  - (c) Refined mask based on morphological filtering and the land cover mask;
  - (d) Refined mask overlaid with reference data ( $T_{H_{region}} = 5m$ ;  $T_{size_{region}} = 200$  pixels;  $T_{convex_{region}} = 0.4$ );
- (b) - (d) Green: true detected; Red: false alarm; Blue: missed alarm.

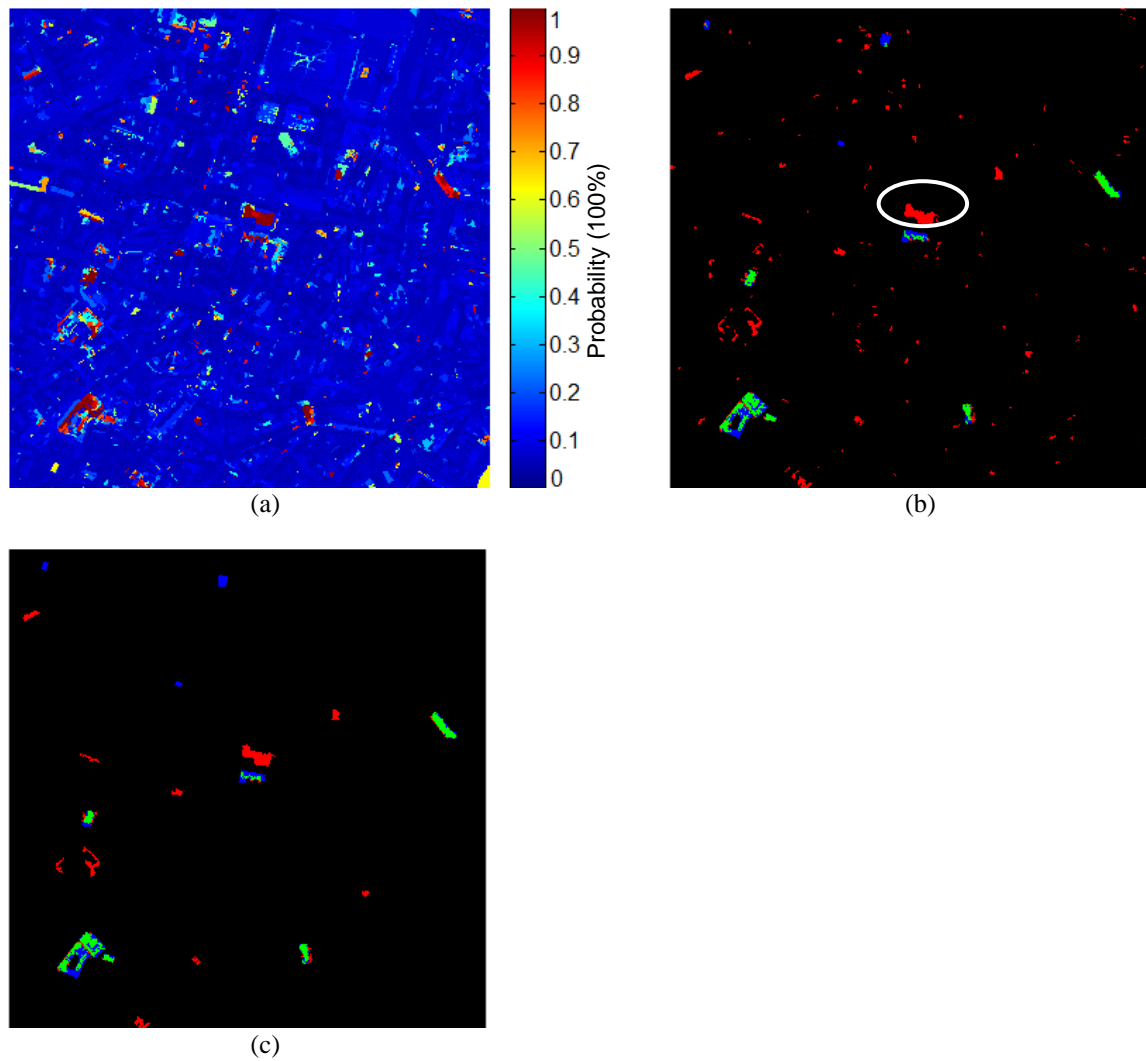


Fig. 7-9. Change map and masks generated in the Munich test area with Method II:

- (a) Region based change probability map;
  - (b) Initial change mask overlaid with reference data ( $T = 0.8$ );
  - (c) Refined mask overlaid with reference data ( $T_{size_{region}} = 200 \text{ pixels}$ );
- (b) - (c) Green: true detected; Red: false alarm; Blue: missed alarm.



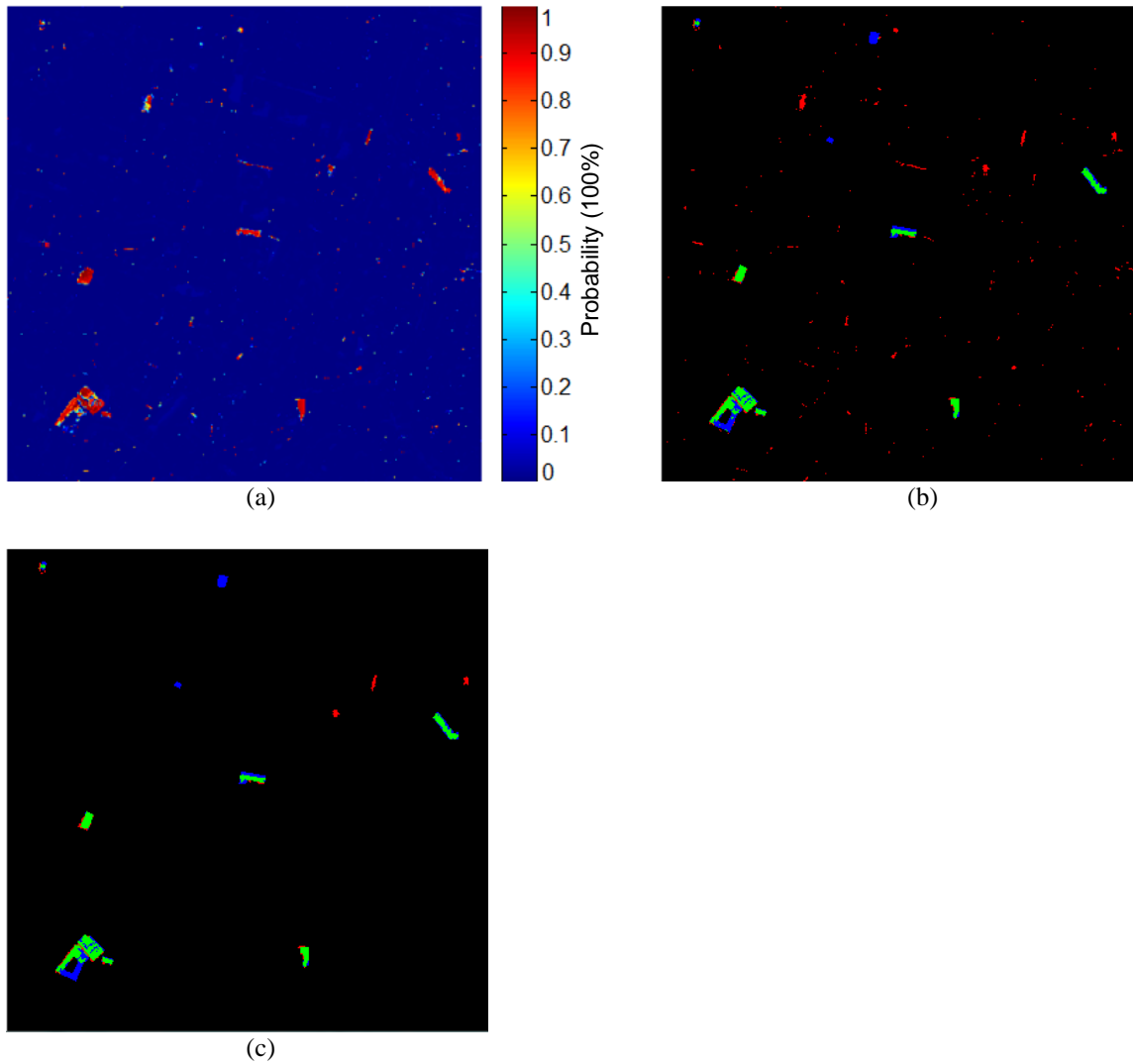


Fig. 7-10. Change map and masks generated in the Munich test area with Method III:

(a) Change map after two steps of decision fusion;

(b) Initial change mask overlaid with reference data ( $T = 0.4$ );

(c) Refined mask overlaid with reference data ( $T_{H_{region}} = 5m$ ;  $T_{size_{region}} = 200 \text{ pixels}$ ;  $T_{convex_{region}} = 0.45$ );

(b) - (c) Green: true detected; Red: false alarm; Blue: missed alarm).

The experiment design for Method IV differs from the other experiments. The dataset from WorldView-2 has been directly compared with the building footprint presented in Fig. 7-7f. First, the initial building mask is generated, after the lower objects (below 15 meters) and trees have been removed from the nDSM. By comparing this building mask with the outdated building footprint, the possible changed buildings obtained are marked in Fig. 7-11a. The green masks represent new buildings, while the blue masks show the reconstructed buildings (partly new built). The change reference data are overlaid with that extracted change mask, and the results are shown in Fig. 7-11b. Since the outdated building footprint was extracted years before the IKONOS scene was acquired, several destroyed buildings in IKONOS images are still present in the building

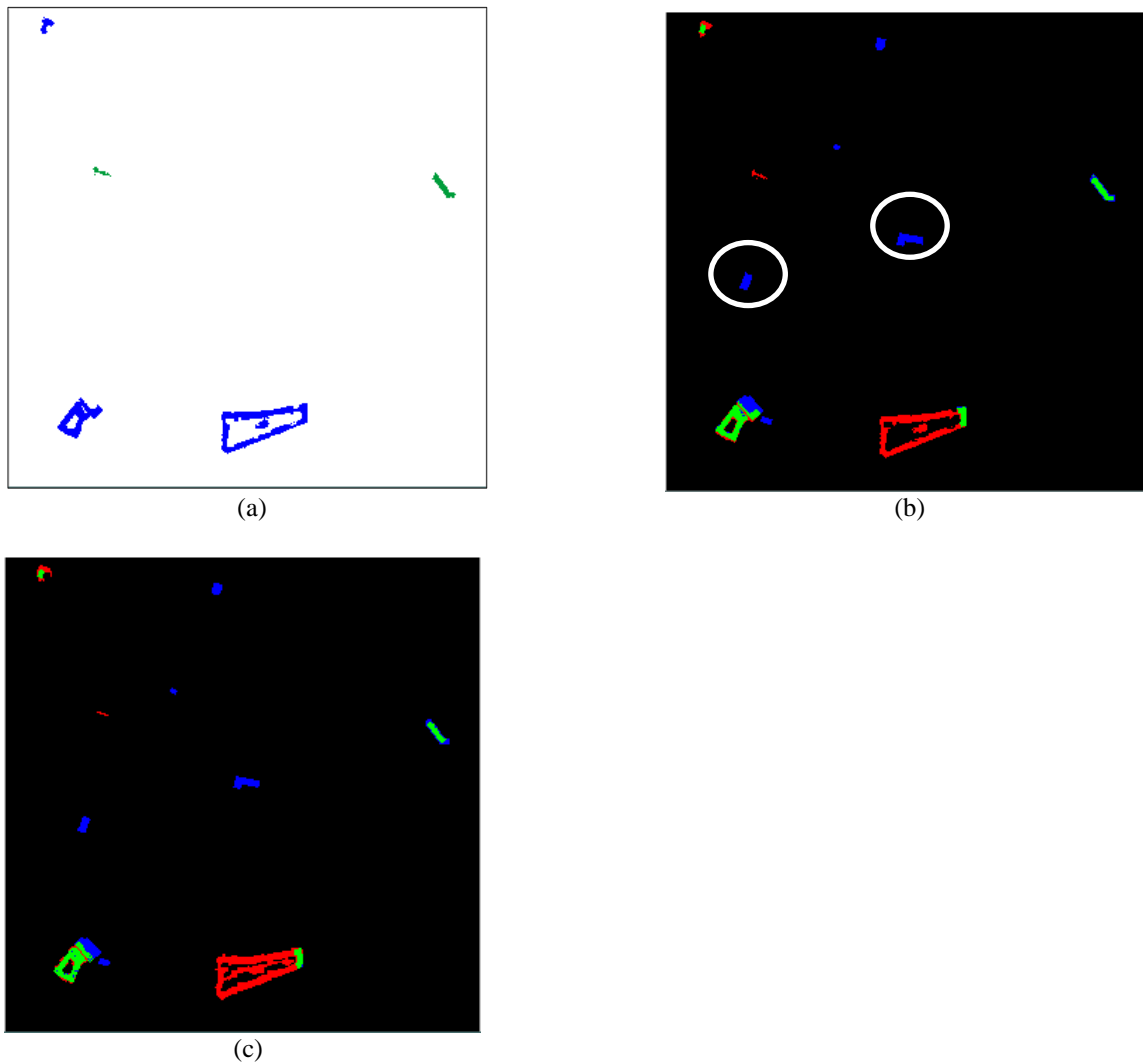


Fig. 7-11. Change masks generated in the Munich test area with Method IV:

- (a) Extracted initial change mask, new built buildings (green), partly rebuilt buildings (blue);
- (b) Change mask overlaid with reference data;
- (c) Extracted building masks overlaid with reference data;
- (b) - (c) Green: true detected; Red: false alarm; Blue: missed alarm.

footprints (marked with white circles in Fig. 7-11b). As it is the centre of Munich, the newly built buildings often share the same shape as the destroyed ones. These areas cannot be counted as building changes in the sense of the presented investigations. Fig. 7-11c shows the final refined building masks.

### 7.3.3 Evaluation of results

The AUC evaluation results shown in Table 7-4 highlight the advantage of the fusion-based method. Since images of two dates from multi-sensor imagery are adopted in this test area, none

of the 2D features leads to good results. However, fusing these features with height changes largely improve the change detection accuracy. In this test area, the region-based method produces a change map with larger AUC than the decision fusion method. However, the change mask of the region-based method has relatively low OA and KIA (as shown in Table 7-5). That can be explained by several negative changes or normal land cover changes in this test area. Fig. 7-12 shows more detailed information about a building in the centre of the test area in Munich. Fig. 7-12a is the panchromatic image from date 1 (IKONOS). Fig. 7-12b is the panchromatic image from date 2 (WorldView-2). In the WorldView-2 data, some parts of the building were removed and the center of that building has become much brighter. Thus, directly using the intensity value from the panchromatic image, as Method II does, will introduce false alarms.

Table 7-4. AUC comparison in the Munich test area.

Method	AUC
Absolute image difference	0.7952
KL divergence	0.5025
Multi-level	0.8571
Height difference	0.8695
I	0.8745
II	0.9270
III	0.9025

Table 7-5. Pixel-based change masks evaluation in the Munich test area.

Method	Description	Overall Accuracy (OA)	Kappa Index of Agreement (KIA)
I	Initial mask	0.9931	0.5396
	Land cover refinement	0.9949	0.6126
	Shape feature based refinement	0.9954	0.6348
II	Initial mask	0.9880	0.4228
	Remove small size regions	0.9917	0.5044
III	Initial mask	0.9927	0.5700
	Shape feature based refinement	0.9957	0.6899

## 7 Experiments

Table 7-6. Object-based change masks evaluation in the Munich test area.

Method	True Detected (Object)		False Detected (Object)	
	Number	Rate [%]	Number	Rate [%]
I	7	70	2	22.22
II	7	70	10	62.50
III	8	80	3	30

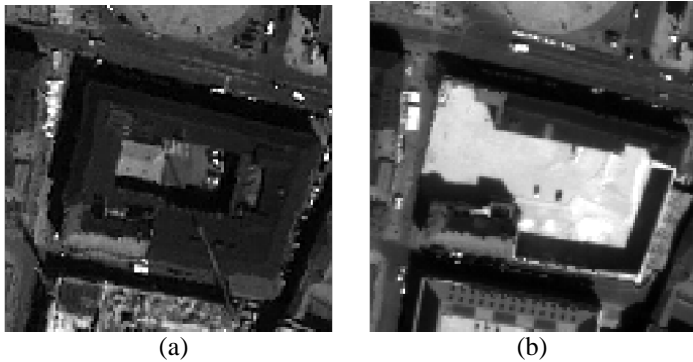


Fig. 7-12. An example of non-positive building change, consisting of the panchromatic image from IKONOS date 1 (a) and from WorldView-2 date 2 (b).

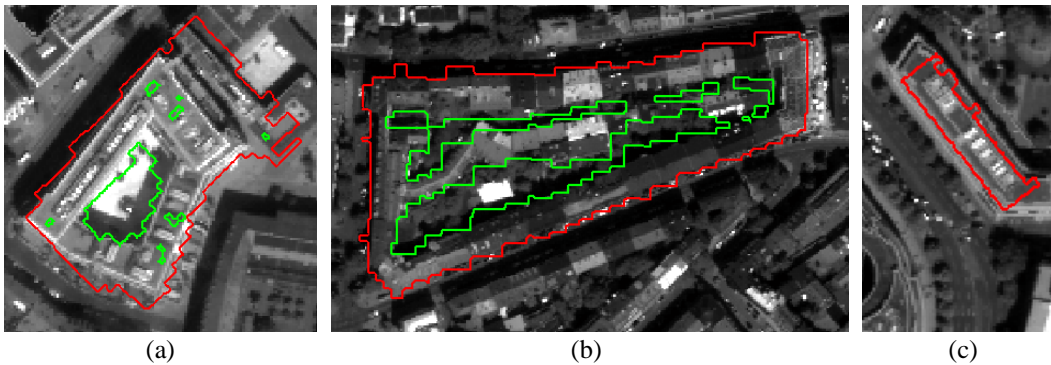


Fig. 7-13. Three extracted changed buildings. Red polygons = outside boundary of each building; Green polygons = inside boundary of each building.

Table 7-4 and Table 7-5 illustrate the change mask evaluation results. According to the pixel-based change mask evaluation result, Method III obtains the best change mask. The KIA obtained from Method III is little higher than Method I, but much higher than Method II. It is worth noting that the land cover and shape feature-based refinement has improved the KIA from 0.5396 to 0.6348 in Method I. The object-based evaluation results listed in Table 7-6 again show the weakness of Method II in the Munich test area.

Unlike the first three change detection methods, Method IV aims at automatically obtaining the correct precise building boundaries. The extracted building mask for this scene is shown in Fig. 7-13. To evaluate its accuracy, we have overlaid the boundaries of these buildings to the WorldView-2 panchromatic images. Three of the changed buildings have been selected and are displayed in Fig. 7-13. These are the original building boundaries without any refinement. As can be seen here, the main building shapes have been extracted clearly. However, some small errors remain, as shown in Fig. 7-13a where small holes remain in the middle of the buildings. Also, the boundaries for all three changed buildings are not so straight. These problems will probably be refined in our future work.

The evaluation result of this test set proves again that Method III is able to produce a better change detection result in VHR imagery than Method I and Method II. This can be explained by the fact that it is difficult to achieve a proper segmentation in VHR imagery. High resolution data feature sharp boundaries, but also obstructive information for building change detection, like windows on building roofs, seen from different perspectives or cars on the road. This information influences the segmentation procedure, thus affecting the application of Method II.

## 7.4 Istanbul test area, Turkey

### 7.4.1 Data introduction

To study the behaviour of the proposed method for HR stereo imagery, a third test area acquired by Cartosat-1 is selected. As illustrated in Chapter 2, the DSM generated from Cartosat-1 exhibits limited accuracy in urban areas, especially for building boundaries. Without the availability of multi-spectral channels (Cartosat-1 acquires only panchromatic data), it is quite difficult to use these data for urban building change detection. Therefore, an industrial area with large buildings in the eastern part of Istanbul was selected for this experiment. Cartosat-1 images were acquired with a time difference of three years: the first was acquired on August 14, 2008 and the second was acquired on May 08, 2011. In this industrial area, the focus of change detection lies on newly constructed buildings. Fig. 7-14 shows the original panchromatic images (a) and (b), the generated DSMs (c) and (d), and the change reference data (d). For this test area, only the changed regions with a size greater than 40 pixels ( $1000 \text{ m}^2$ ) are considered. It covers a total area of  $2.5 \text{ km} \times 2.5 \text{ km}$  and is characterized as an industrial area with low-rise buildings in high density.

## 7 Experiments

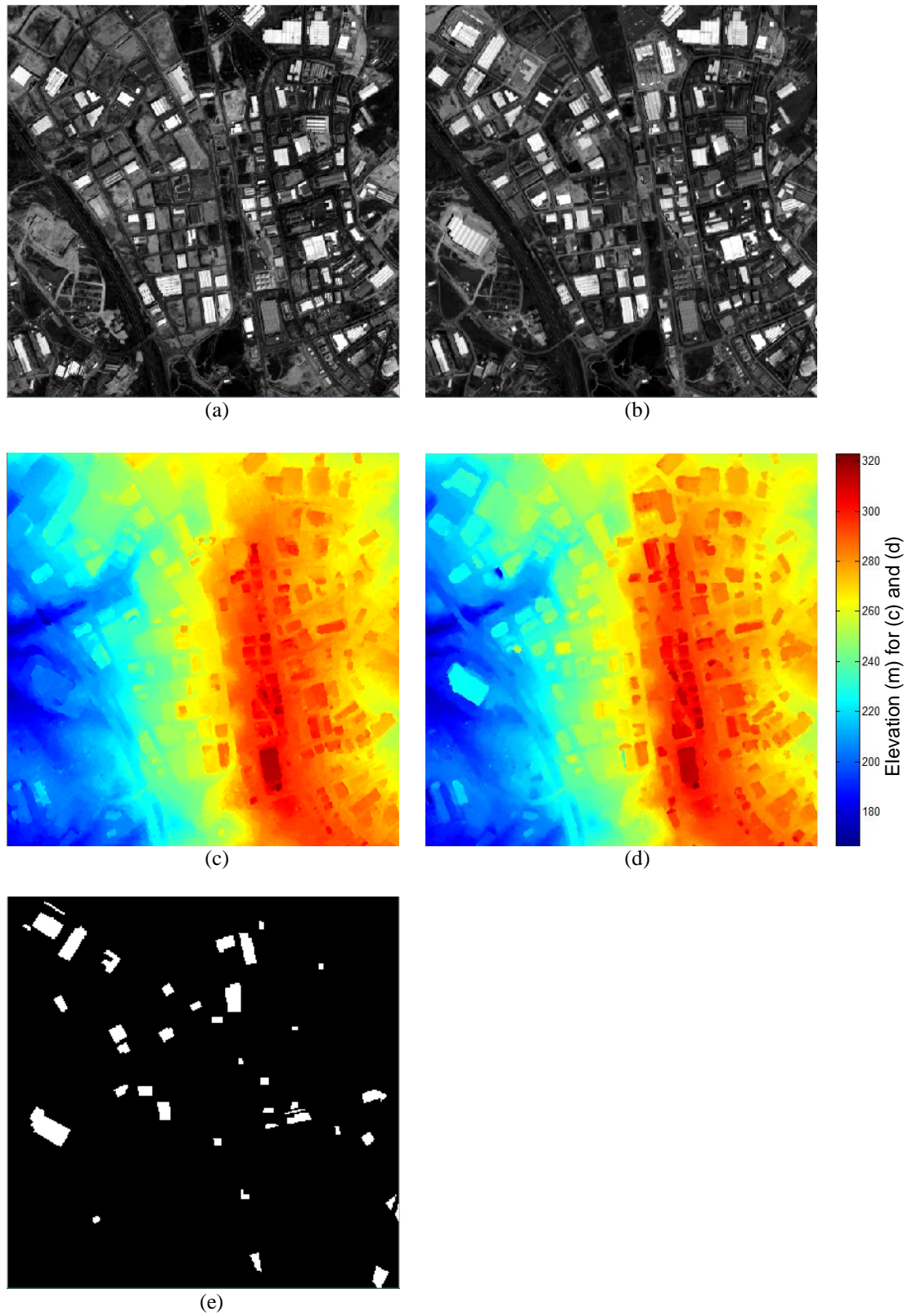


Fig. 7-14. Dataset in the Istanbul test area. Orthorectified panchromatic image from date 1 (a) and date 2 (b), DSM from date 1 (c) and date 2 (d), and reference change map showing newly built buildings (e).

Since no multi-spectral channels are available, it is not possible to apply the full workflow of Method I and Method II in this test region. Therefore, in Method I, the land cover-based refinement procedure is skipped. The initial change mask is generated only through refinement using shape features. For Method III, since the no-building change indicators are not available, only the first fusion step (change feature fusion) is processed.

### 7.4.2 Results

The change detection results from Method I are presented in Fig. 7-15. Among them, Fig. 7-15a shows the robust change difference map generated with a window size of  $3 \times 3$  pixels. After thresholding ( $T = 5 m$ ), the resulting mask overlaid with the reference map is presented in Fig. 7-15b. Fig. 7-15c is the shape-based refinement result. In this test, we have used the threshold values  $T_{H_{region}} = 5m$ ;  $T_{size_{region}} = 40 \text{ pixels}$ ;  $T_{convex_{region}} = 0.55$ .

Method II is a region-based change detection method. Compared to Fig. 7-15, the results shown in Fig. 7-16 are less noisy and the detected buildings exhibit much sharper boundaries. Fig. 7-16a depicts the change map, which is automatically stretched to a change probability map. Fig. 7-16b is the initial change mask overlaid with the reference change map. After removing the regions with a size of less than 40 pixels ( $1000 \text{ m}^2$ ), the refined mask is displayed in Fig. 7-16c. The obtained result notably improves compared to the results from Method I. A problematic area appears in the top middle of the test area and some false alarms remain around the buildings. These false alarms, represented in red, can be seen in Fig. 7-16c.

Without a multi-spectral channel, only the first step of Method III can be processed. Fig. 7-17a shows the obtained fusion result based on the two building-change indicators. The initial change mask produced after thresholding is shown in Fig. 7-17b. A lower threshold value is used here, because in this industrial area, the building height is relatively low: thus, the construction of new buildings is unlikely to result in large height differences. To refine the initial change mask, the same shape features used for Method I ( $T_{H_{region}} = 5m$ ;  $T_{size_{region}} = 40 \text{ pixels}$ ;  $T_{convex_{region}} = 0.55$ ) are also used here. The refined result is shown in Fig. 7-17c.

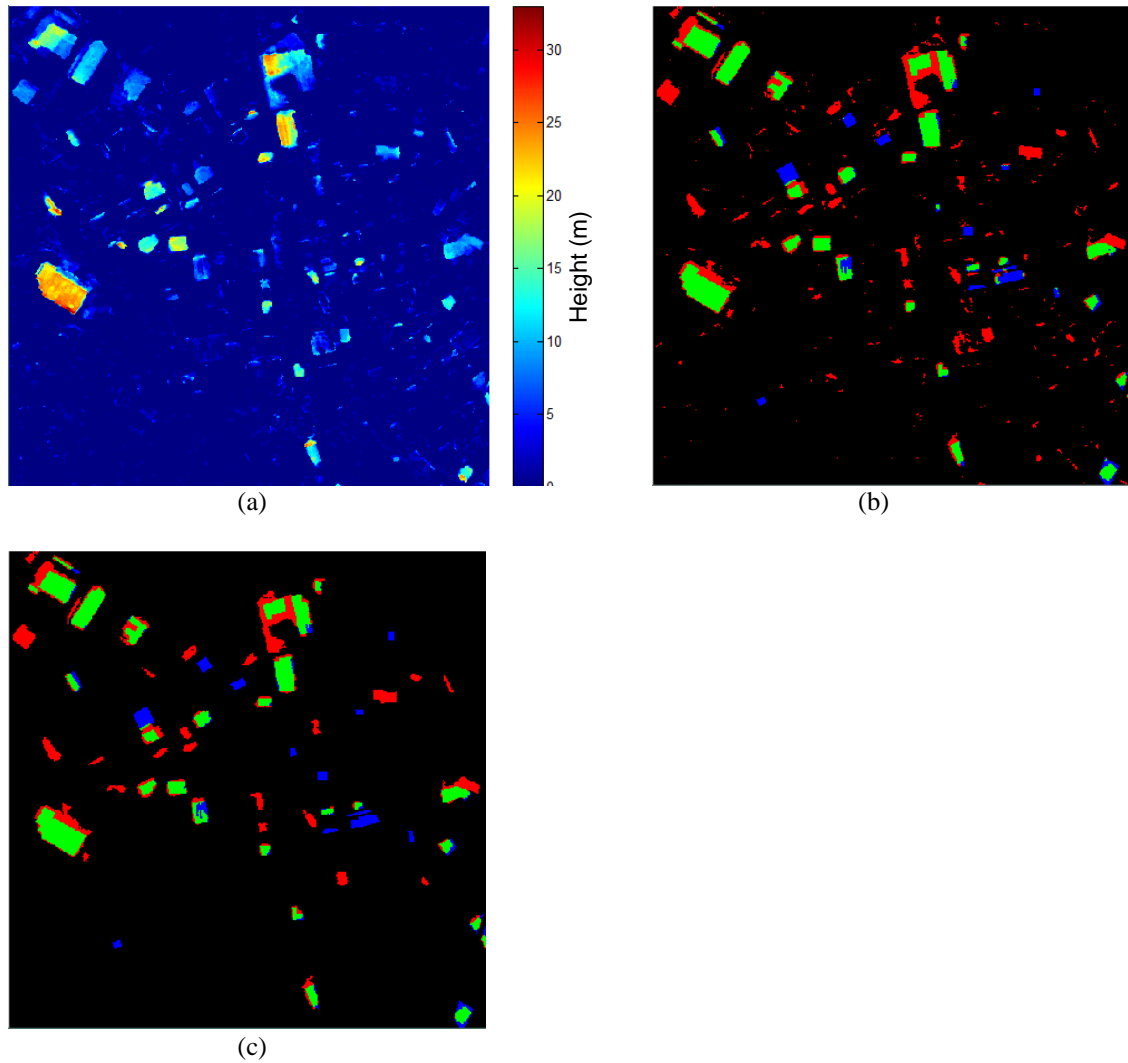


Fig. 7-15. Change map and masks generated in the Istanbul test area with Method I:

(a) Robust difference map with positive change;

(b) Initial change mask overlaid with reference data ( $T = 5\text{m}$ );

(c) Refined mask overlaid with reference data ( $T_{H_{region}} = 5\text{m}$ ;  $T_{size_{region}} = 40\text{ pixels}$ ;  $T_{convex_{region}} = 0.55$ );

(b) - (c) Green: true detected; Red: false alarm; Blue: missed alarm.



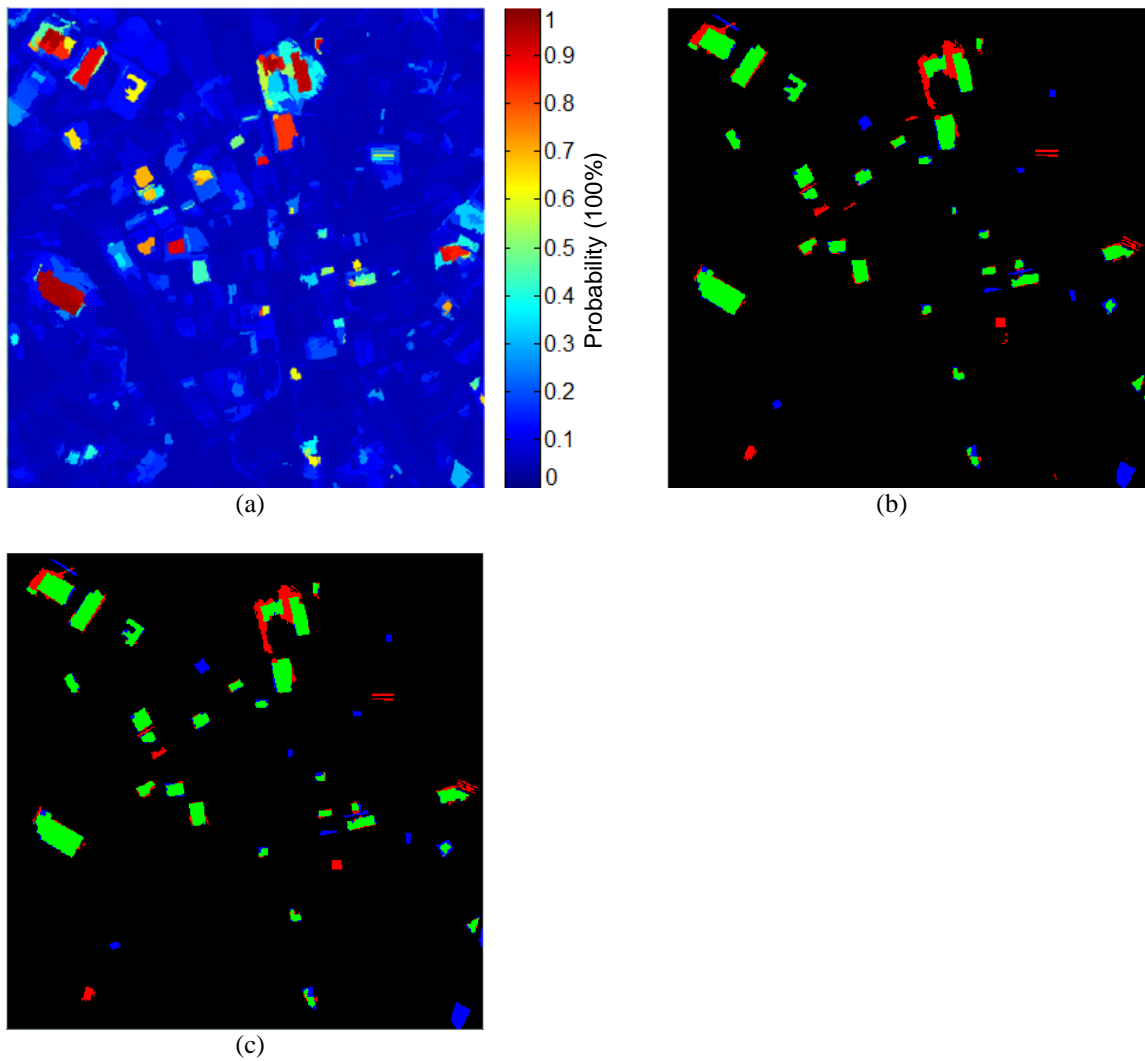


Fig. 7-16. Change map and masks generated in the Istanbul test area with Method II:

- (a) Region-based change probability map;
  - (b) Initial change mask overlaid with reference data ( $T = 0.4$ );
  - (c) Refined mask overlaid with reference data ( $T_{size_{region}} = 40 \text{ pixels}$ );
- (b) - (c) Green: true detected; Red: false alarm; Blue: missed alarm.

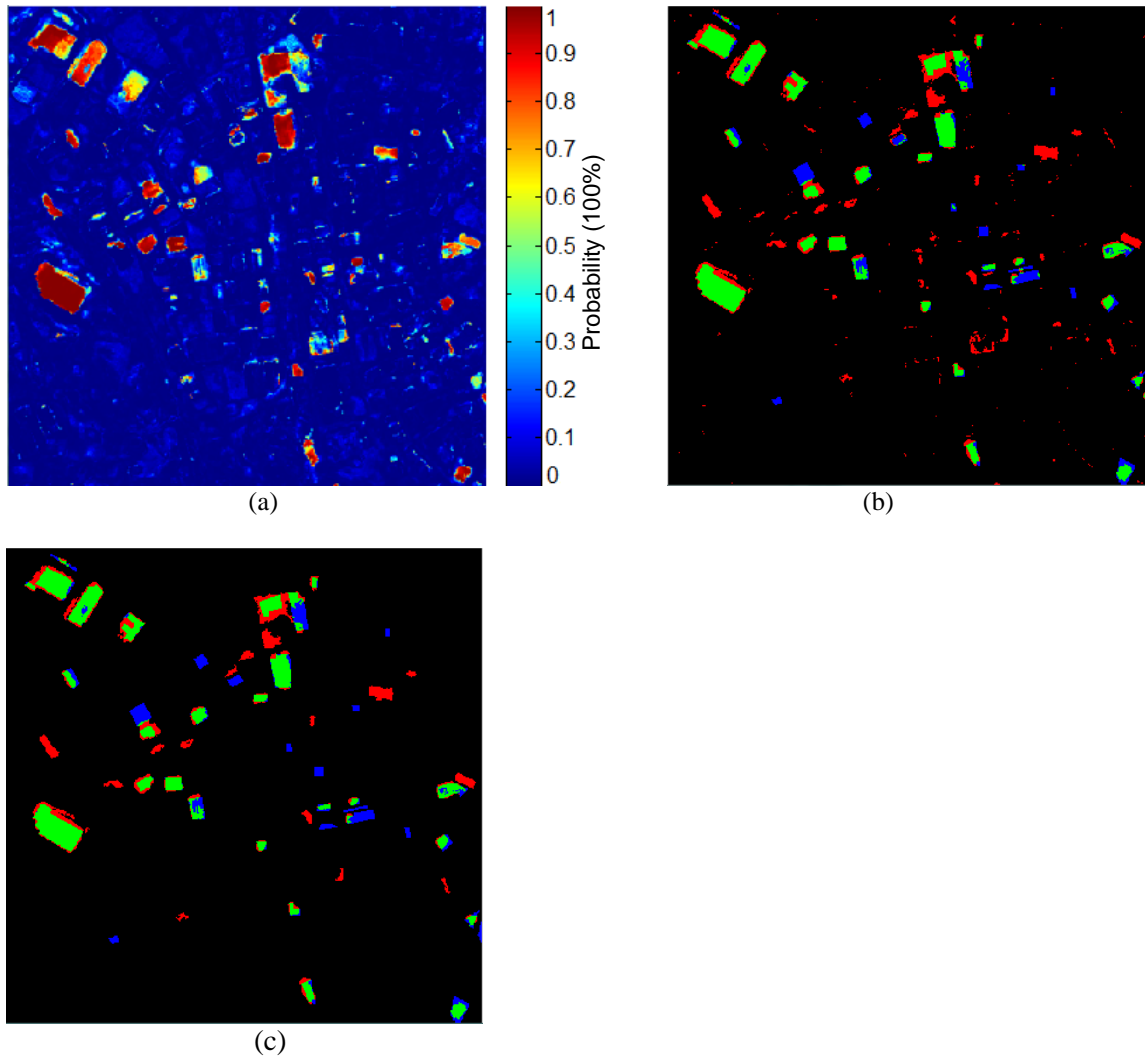


Fig. 7-17. Change map and masks generated in the Istanbul test area with Method III:

- (a) Change map after one step of decision fusion;
  - (b) Initial change mask overlaid with reference data ( $T = 0.4$ );
  - (c) Refined mask overlaid with reference data ( $T_{H_{region}} = 5m$ ;  $T_{size_{region}} = 40 \text{ pixels}$ ;  $T_{convex_{region}} = 0.55$ );
- (b) - (c) Green: true detected; Red: false alarm; Blue: missed alarm.

### 7.4.3 Evaluation of results

The AUC obtained from all three methods are listed in Table 7-7. The quality of the DSMs from HR stereo imagery (Cartosat-1) is not as high as from VHR stereo images for buildings. Even though an industrial area with large size buildings is selected, the AUC obtained from Method I (robust height difference) is lower than from absolute images difference and multi-level features. In this case after fusing the 2D change features and height changes, the AUCs improve significantly for Method II and Method III. Method II produces the most accurate change map for this test area.

Table 7-7. AUC comparison in the Istanbul test area.

Method	AUC
Absolute image difference	0.9280
KL Divergence	0.8508
Multi-level	0.9511
Height Difference	0.8993
I	0.9065
II	0.9667
III	0.9432

Table 7-8. Pixel-based change masks evaluation in the Istanbul test area.

Method	Description	Overall Accuracy (OA)	Kappa Index of Agreement (KIA)
I	Initial mask	0.9586	0.5582
	Shape feature based refinement	0.9655	0.6360
II	Initial mask	0.9821	0.7577
	Remove small size regions	0.9823	0.7578
III	Initial mask	0.9684	0.6106
	Shape feature based refinement	0.9720	0.6365

Table 7-9. Object-based change masks evaluation in the Istanbul test area.

Method	True Detected (Object)		False Detected (Object)	
	Number	Rate [%]	Number	Rate [%]
I	28	71.79	21	45.65
II	29	74.36	5	16.67
III	28	71.79	14	35.90

## 7 Experiments

Tables 7-8 and 7-9 show the change mask evaluation results. Since the change map from Method II exhibits a higher AUC value than the other two methods, it leads to the most accurate change mask. As shown in Table 7-8, the initial change mask already shows better OA and KIA values than the final masks from the other methods. Table 7-9 has again proven the advantages of Method II. For the Cartosat-1 images, it is much easier to achieve a proper segmentation level than using the much higher resolution images from IKONOS and WorldView-2. Thus, few false alarm objects remain in the obtained change mask.

According to the object evaluation result (Table 7-9), Method II and III (fusion-based methods) perform better in this test area. This proves that the original images can be very helpful in the change detection procedure, when the DSM quality is not good enough. Besides extracting changed buildings with sharp boundaries, Method II has also correctly detected more changed buildings and the false alarms drop significantly in comparison to the other two methods. This can be explained by the appropriate segmentation result. This test set is located in an industrial area with simple building shapes. It is much easier to get a proper segmentation result in this test area compared to more complex structures as for example in the Munich test area. Furthermore, due to the lack of multi-spectral channels, only the first fusion step of Method III can be performed.

## 7.5 Oberammergau test area, Germany

### 7.5.1 Data introduction

To evaluate the performance of our proposed approaches in a forest area, the fourth dataset used was acquired over a forest region near Oberammergau in Bavaria, Germany. With an extension of 2.25 km<sup>2</sup>, as shown in Fig. 7-18, this region is characterized by a combination of forest and crop change. The first scene is from May 12, 2008 and the second is from May 25, 2009. Both have a GSD of 2.5 m and the test images have a size of 600 × 600 pixels. Both scenes were captured by Cartosat-1, which, as illustrated before, has only a panchromatic band within each of the two stereo cameras. The two generated DSMs are displayed in Fig. 7-18c and 7-18d. In this test area, with one year's time difference, it can be assumed that there are not many positive changes (e.g., new trees, tree growth). Most of the changes are a result of deforestation: therefore, in this test area, only the negative changes are of interest. The manually extracted reference data are shown in Fig. 7-18e.

7.5 Oberammergau test area, Germany

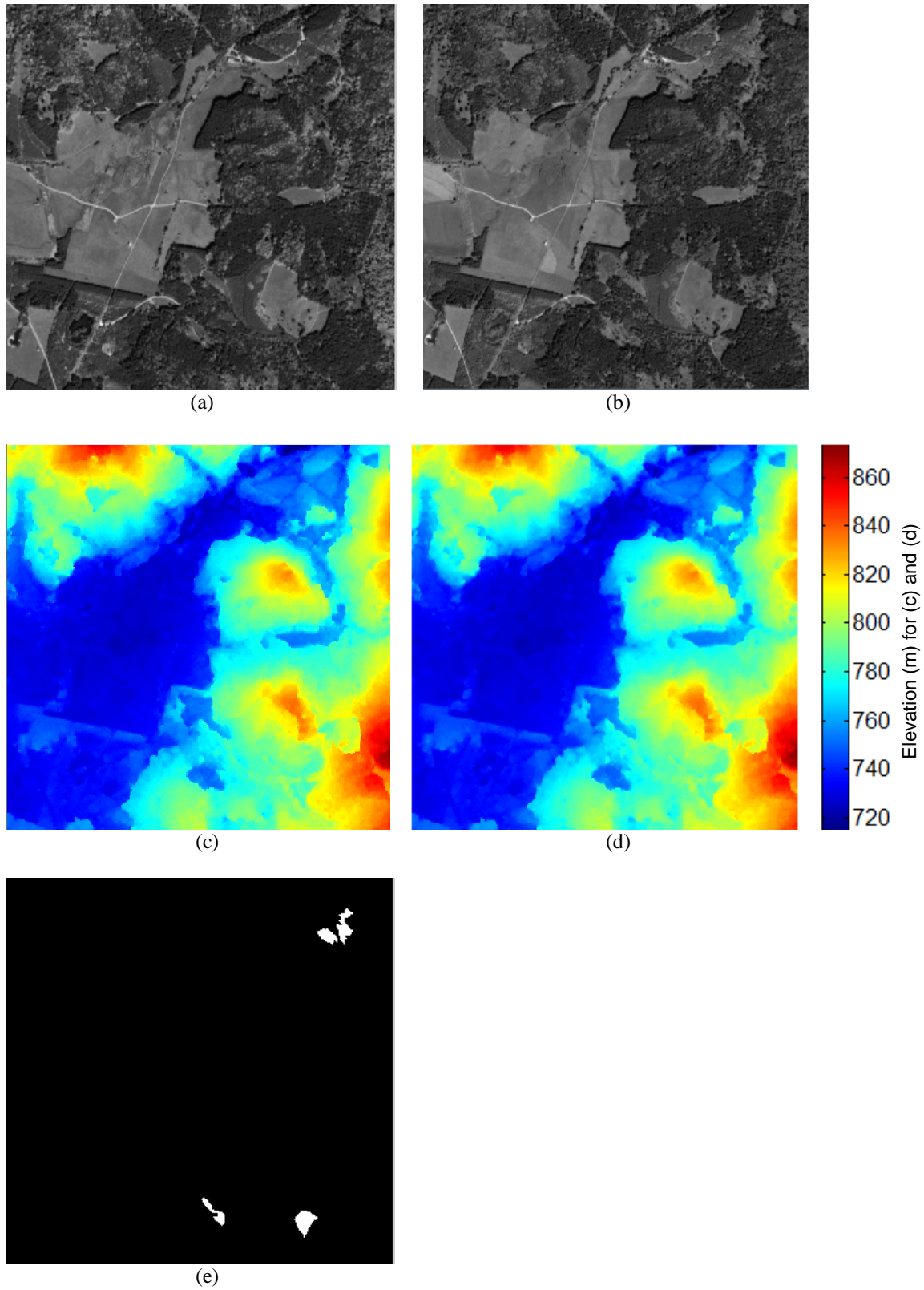


Fig. 7-18. Dataset in the Oberammergau forest test area. Orthorectified panchromatic image from date 1 (a) and date 2 (b), DSM from date 1 (c) and date 2 (d), and reference change map (e).

### 7.5.2 Results

To evaluate the performance of our proposed methods for forest areas, the methods are again applied separately. Due to the lack of spectral methods and less meaningful shape features, only part of Methods I and III can be applied. In addition, as the experiments for this test area are focused on forest change detection and the fusion model of Method III is designed for building changes, only the first fusion step can be used to detect forest change. In this test area, the imagery of two dates are from same HR sensor, in the experiments direct height difference is adopted instead of robust difference for Method II and III due to the better performance.

Fig. 7-19 describes the change detection result from Method I. The change map shown in Fig. 7-19a is generated from robust height difference. Since the negative changes are of interest, smaller values, shown in blue, indicate larger change. After selected the pixels with height decreases of more than 10 meters, the result is shown in Fig. 7-19b. The small size regions remaining have been removed to generate the final change mask (shown in Fig. 7-19c).

Fig. 7-20 shows the change detection results obtained using Method II. As can be seen, all of the real forest changes have been detected successfully. After thresholding, very few false alarms remain in the change mask (shown in Fig. 7-20b), and these can be easily removed by filtering through size constraints. The final change detection result is displayed in Fig. 7-20c. As can be seen, all of the four changed forest areas have been correctly detected, and three of them are very close to their reference boundaries.

Since in Method III, the direct height difference is used instead of the robust difference, the change map obtained contains comparatively more noise. This noise is mainly located near the forest boundary regions. However, this noise does not greatly influence the final change detection result. As presented in Fig. 7-21b, all the real changed regions are successfully detected, and only some noise-based false alarms remain. But these false alarm regions are mostly small in size, and after the removal of regions less than 100 pixels (2500 m<sup>2</sup>), the results are much better (as shown in Fig. 7-21c), but still two areas with wrongly detected changes are presented.

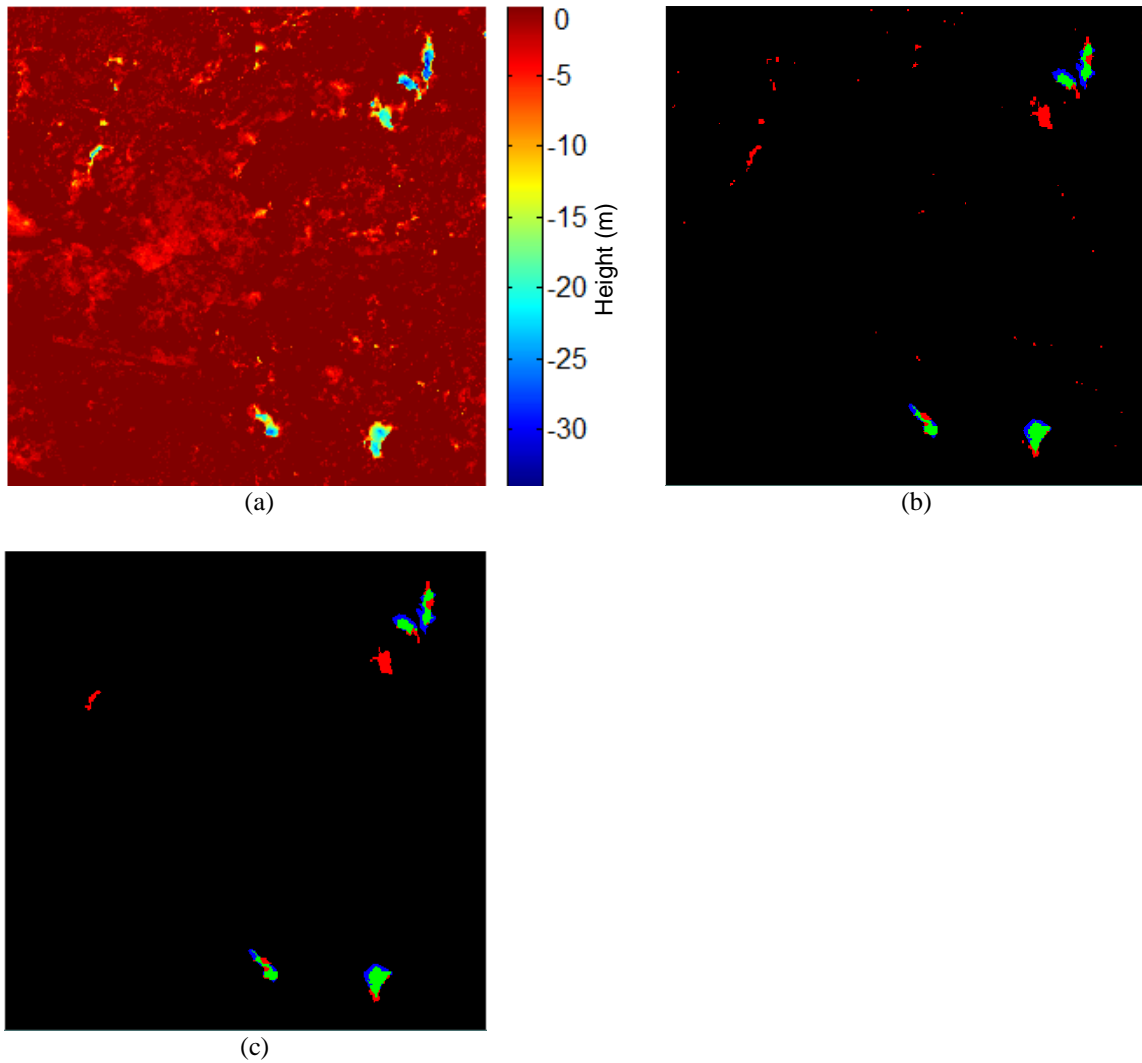


Fig. 7-19. Change map and masks generated in the Oberammergau test area with Method I:

(a) Robust difference map with negative changes;

(b) Initial change mask overlaid with reference data ( $T = 10$  m);

(c) Refined mask overlaid with reference data ( $T_{size\_region} = 100$  pixels);

(b) - (c) Green: true detected; Red: false alarm; Blue: missed alarm.

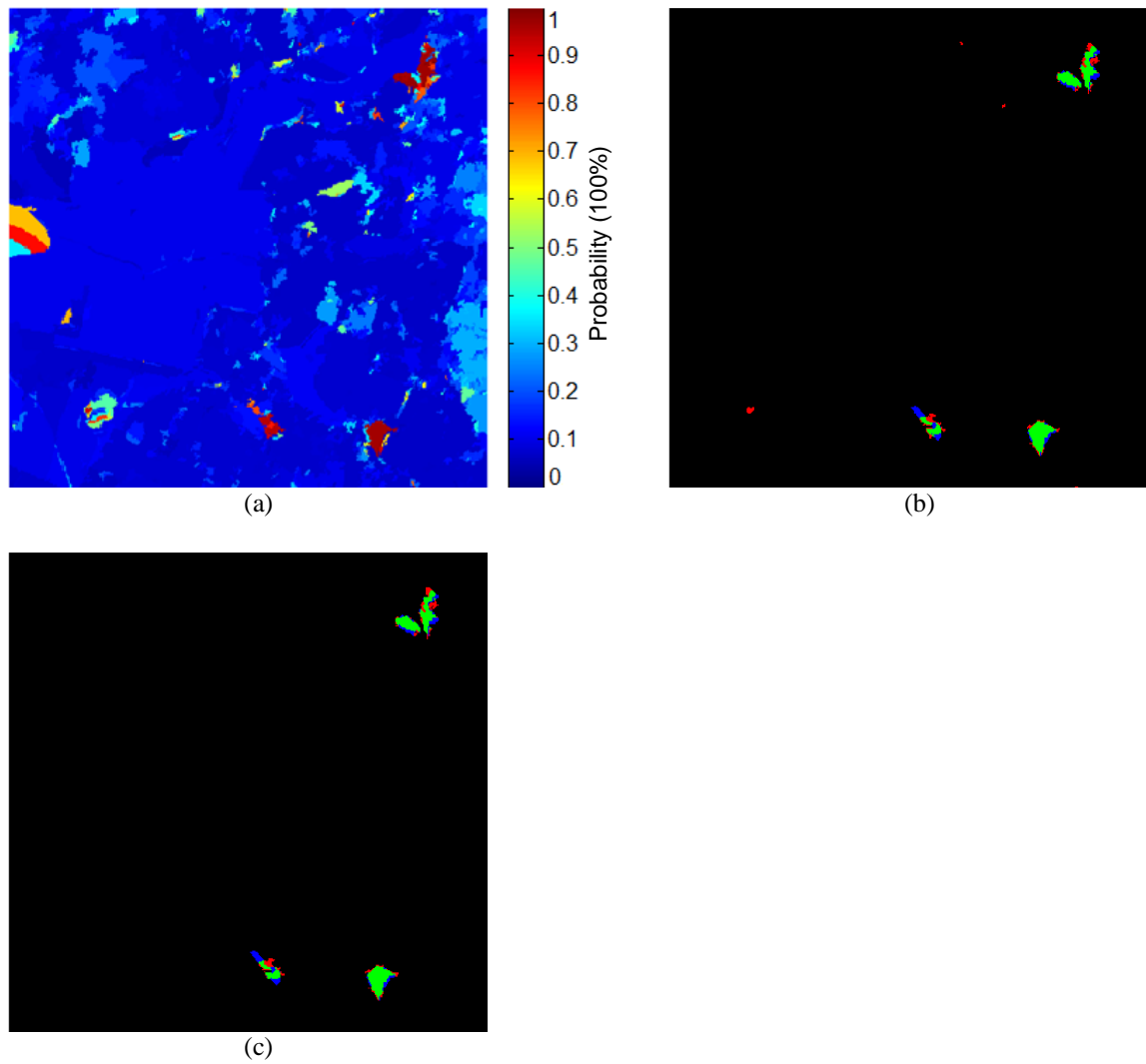


Fig. 7-20. Change map and masks generated in the Oberammergau test area with Method II:

- (a) Region-based change probability map;
  - (b) Initial change mask overlaid with reference data ( $T = 0.9$ );
  - (c) Refined mask overlaid with reference data ( $T_{size_{region}} = 100 \text{ pixels}$ );
- (b) - (c) Green: true detected; Red: false alarm; Blue: missed alarm.



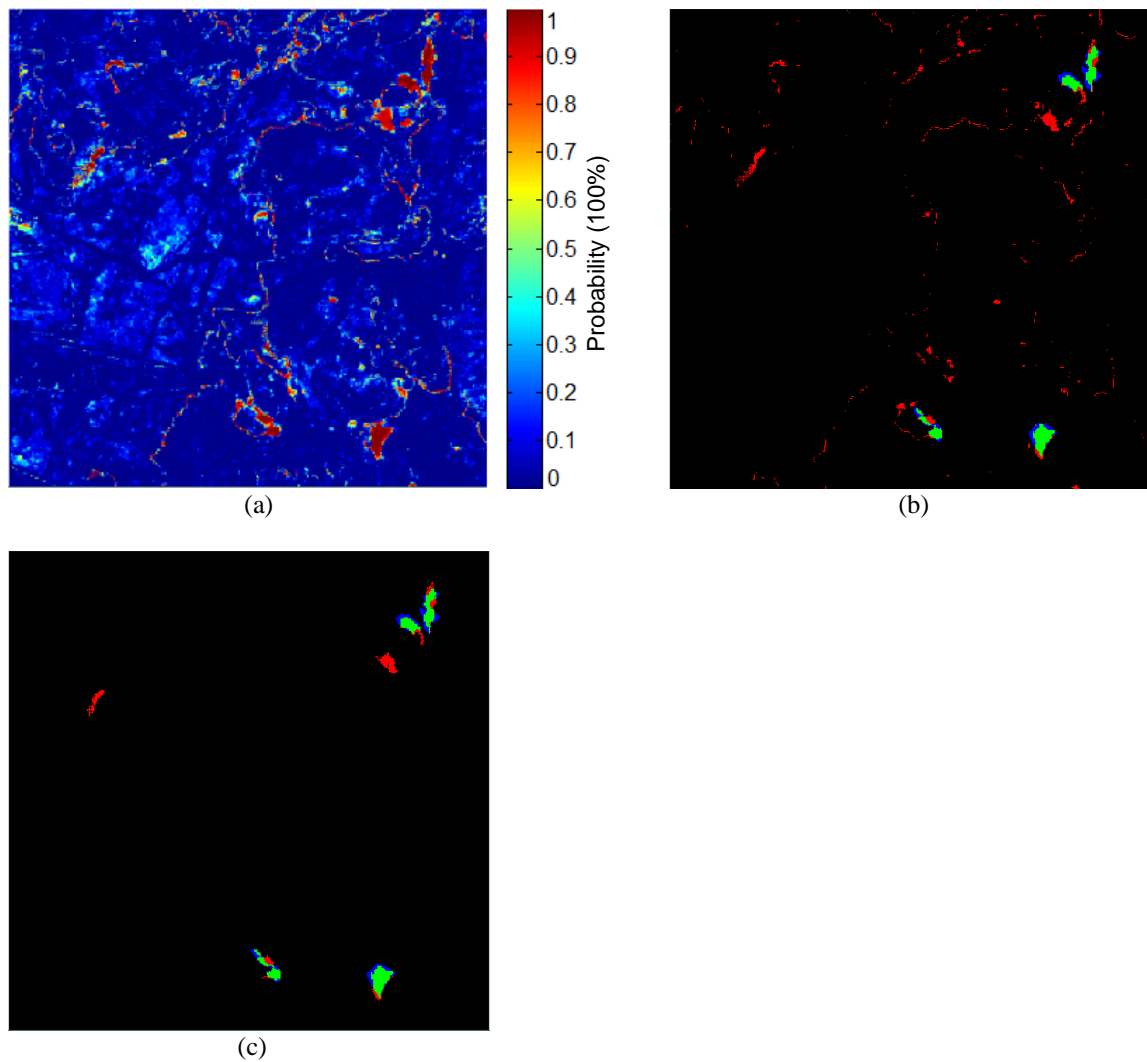


Fig. 7-21. Change map and masks generated in the Oberammergau test area with Method III:

- (a) Change map after one step of decision fusion;
  - (b) Initial change mask overlaid with reference data ( $T = 0.9$ );
  - (c) Refined mask overlaid with reference data ( $T_{size\_region} = 100 \text{ pixels}$ );
- (b) - (c) Green: true detected; Red: false alarm; Blue: missed alarm.

### 7.5.3 Evaluation of results

As the AUC comparison result in Table 7-10 shows, DSM from HR stereo imagery is advantageous in forest areas in contrast to urban areas. AUC from direct height differencing is higher than from absolute image difference and KL divergence. It has to be noted that Method 2 (region-based) does not result in a higher AUC, the pixel based method with same features reaches higher AUC values, these results are shown in Tian et al. (2013b). Even if the change map from Method III, which looks more noisy (in Fig. 7-21a) than the change map from Method II, the AUC

## 7 Experiments

Table 7-10. AUC comparison in the Oberammergau test area.

Method	AUC
Absolute image difference	0.9062
KL divergence	0.8377
Multi-level	0.9357
Height difference	0.9231
I	0.8961
II	0.9316
III	0.9505

Table 7-11. Pixel-based change masks evaluation in the Oberammergau test area.

Method	Description	Overall Accuracy (OA)	Kappa Index of Agreement (KIA)
I	Initial mask	0.9931	0.5347
	Remove small size regions	0.9945	0.5879
II	Initial mask	0.9967	0.7472
	Remove small size regions	0.9970	0.7633
III	Initial mask	0.9899	0.4603
	Remove small size regions	0.9952	0.6450

in this case is still higher than for Method II. However, after further processing (thresholding and refinement), Method II obtains the most accurate change mask. This improvement is further confirmed by the increase of OA and KIA, presented in Table 7-11.

Among the tested approaches, Method II (the region-based approach) achieves the most precise final change masks. In fact, the initial change mask of Method II already exhibits higher OA and KIA values than the final masks of the other two methods. This demonstrates that the region-based method is more appropriate for change detection in forest areas. These results are similar to those achieved in the Istanbul test set, although for these two test sets different change objects are of interest. This proves again that Method II is more appropriate for HR imagery, especially when only panchromatic imagery is available.

## 7.6 Discussion and summary

In this chapter, it has been shown that the proposed approaches for DSM-assisted change detection can be applied for building as well as for forest change detection. In the quality inspection procedure, the manually extracted reference data have been used as ground truth. The evaluation includes both visual inspection and a qualitative comparison to the provided reference data. In the experimental procedure, four types of data groups are selected for different change objects of interest. To prove the advantages of DSM-assisted change detection, the change maps obtained from the proposed methods have been compared with 2D change features. As can be seen from the evaluation results, the proposed methods perform in nearly all cases better than the 2D features alone. It has to be noted, that although the accuracies of these 2D change features are various and often low, they are important, since fusing them with height changes leads to improved change detection accuracy.

In the following, some problems which have arisen during our investigations are discussed.

As evident from the visual inspection of the obtained results, most of the changes can be correctly extracted. The reasons for high false alarm rates in some cases can be divided into three categories:

- One or both of the DSMs exhibit large area matching errors;
- A reconstructed building is not detected properly;
- Changed objects which are not of interest lead to large spectral changes in satellite images from the two dates.

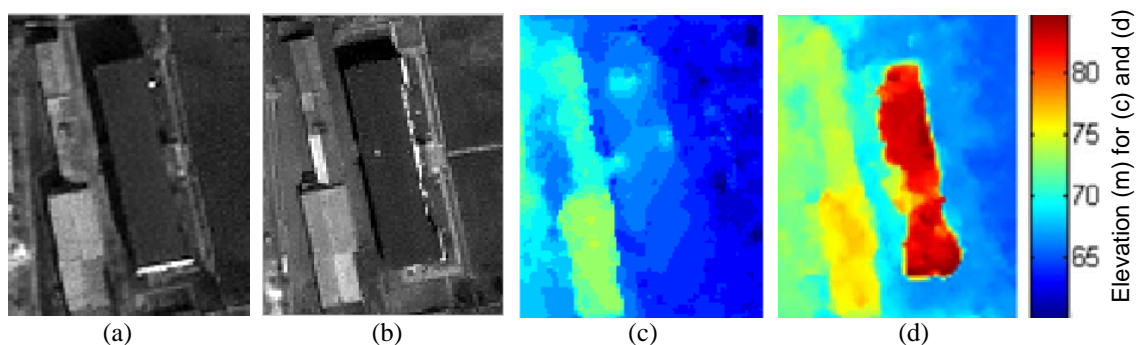


Fig. 7-22. Example of a mis-detected building change: Panchromatic image from date 1 (a) and data 2 (b); DSM of date 1 (c) and date 2 (d).

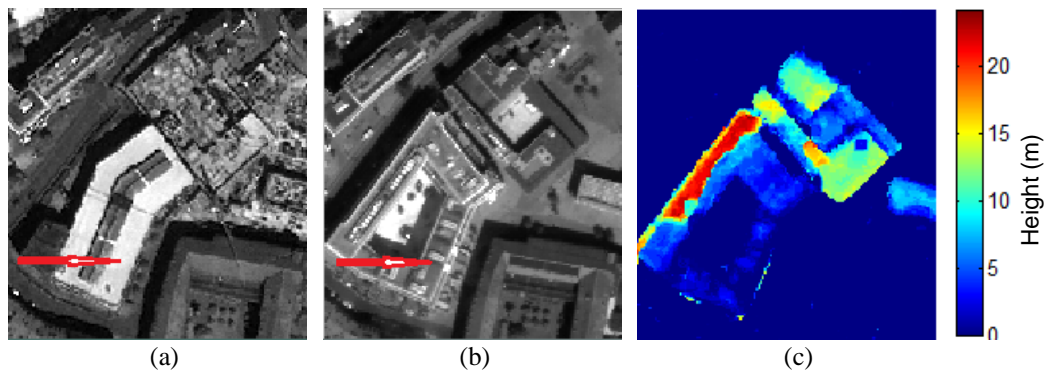


Fig. 7-23. Example of a reconstructed building: Panchromatic image from date 1 (a) and data 2 (b); and height change map (c). The red arrow points to the reconstructed part.

An example for the first category of false alarms is illustrated in Fig. 7-22. In the panchromatic images shown in Fig. 7-22a and 7-22b, it can be seen clearly that the buildings have not changed. Due to matching errors of the stereo images from date 1, the large building in the centre of the images is missing in the corresponding DSM (Fig. 7-22c). However, in the DSM from the later date (Fig. 7-22d) the building is represented correctly. These kinds of mistakes cannot be handled by any of the three proposed methods. Also, for the Istanbul test areas, incorrect heights have been computed for several industrial buildings. These DSM errors lead to several mis-detections in the final change mask.

The second category of false alarms arises when buildings are reconstructed, since the height can also be partly influenced in this situation. Within a large time period, several locations might show this kind of change. One example can be seen in the centre of Fig. 7-23a. The displayed building was reconstructed with a new building shape (see Fig. 7-23b). Using only the height difference map, only half of the changes can be detected (as shown in Fig. 7-23c). By combining the change information from the DSM and the multispectral images, some small changed areas are extracted, as shown in Fig 7-9a and 7-10a. But since they are relatively small and not connected with the other parts of the buildings, they are not kept in the change mask after the object-based filtering.

The last category of false alarms occurring in the proposed approaches appears if objects that are not of interest show high probability of change. This problem exists only for Methods II and III, since in these two methods, the orthorectified satellite images are also adopted in the initial change map generation procedure. One example is shown in Fig. 7-12. The spectral information in the ground area which was surrounded by the building of interest has changed considerably. This kind of change in image can increase the change probability of our final change map. Thus false alarms are introduced.

According to the above discussion and the evaluation results provided, Method I is relatively simple and requires a limited amount of computation, the results are quite robust, and the method can achieve useful change detection results. As a prerequisite of successful change detection, the DSMs should be of relatively good quality, e.g. DSMs from VHR multiple stereo imagery, and the scenes should not feature too many changes through reconstruction of buildings in the same area. Taking the Munich test area as an example, Method I can achieve better results than Method II. However, when DSMs are not very accurate, as in the Istanbul test area, Method II achieves much higher accuracy than Method I.

On the other hand, the region-based method (Method II) does not work so well in dense urban areas. This can be explained by the fact that in dense urban areas, it is difficult to achieve a proper segmentation level. Several false alarms might also occur in small regions, and these small regions can be falsely detected as changes. However, this method works very well for high resolution satellite images such as those from Cartosat-1, as could be shown in the Istanbul and Bavaria test sites. This is probably due to the fact that it is much easier to obtain a good segmentation result in these two test areas.

Method III, decision fusion, is a relatively robust change extraction method. The adopted DS fusion is especially powerful in dealing with uncertain information, which fits well with our problem. The DSM and the image can together provide a decision to determine whether an object of interest has changed. The fusion of multiple possible change and no-change indicators eventually highlights the real changes. Moreover, for this approach, a statistically based similarity measurement is adopted, which is more robust than using the intensity values directly.

The post-classification approach (Method IV) has only been tested in the Munich area, as it is the only area where an outdated building footprint is available. The real positive changes in this test area are quite complex: most of the changed buildings are still undergoing reconstruction, while some new buildings are only partly finished. All of these problems have noticeably increased the complexity of the detection procedure. The proposed building extraction method has also been tested in our work (Tian and Reinartz, 2013), and it performs well even with more complicated buildings.

A major problem that remains is the selection of appropriate threshold values, which still influences the results in several procedures. A manually selected threshold value is used to obtain the change masks. A bigger or smaller value might slightly influence the number of missed or detected false alarms. In the shape feature-based refinements, we have also used several threshold

## *7 Experiments*

values. The selection of a proper threshold value is not trivial and we cannot guarantee that the values used here are the best choice. Higher accuracy might be achieved when using other threshold values, but this improvement is limited. An analysis of how the shape feature threshold values can influence the final masks is presented in Tian et al. (2013a). More automatic threshold selection methods are a topic of interest for potential future research in this field.

## **8. Conclusions and outlook**

Blair and Bar-Shalom (1996: 450) pose the question “Does more data always mean better estimates?” The answer should be ‘yes’ if the data contain mainly correct information. Moreover, correct information should be distinguished from incorrect information using the proper methodology. With respect to the research topic of this thesis, the DSMs from spaceborne stereo matching contain height information, which is essential for change detection. However, the height information provided is not always correct. The goal of this thesis was to find out if fusing imperfect DSMs with other data or information can improve the final change detection accuracy and how these data can be used properly.

### **8.1 Summary of achievements**

In this thesis, three 3D change detection approaches have been addressed by considering combined usage of DSMs generated from stereo imagery and the corresponding ortho-rectified satellite images for change detection. Starting with the DSM generation methodology and DSM quality evaluation, these approaches are proposed, taking into account the advantages and drawbacks of these DSMs. The approaches differ in the way the DSMs are introduced within the overall procedure.

- **DSM quality evaluation**

Before addressing the problem of DSM-assisted change detection, a quality evaluation of DSMs based on satellite stereo images from different sensors with varying resolutions is presented. The quality has been evaluated based on several land covers in urban and forest regions. As illustrated in Chapter 2, in urban regions, the quality of DSMs is more sensitive to the image resolution. The main drawbacks of the Stereo-DSM are the blurred boundaries of objects and potential errors produced in the stereo matching procedure.

- **DSM-assisted change localization**

The first change localization approach uses a DSM subtraction result to obtain initial change candidates. Additional land cover information, such as vegetation, shadow etc., can be used to refine the change map according to the characteristics of the objects of interest. The advantage of

this method lies in its efficiency, robust results and ease of use. Especially after disasters, this approach can provide a fast building damage estimate.

- Fusion-based change detection

For full use of all of the change information contained in the original panchromatic or multi-spectral images and the height information, both the feature fusion and the decision fusion method are proposed. Considering the blurred boundaries of some objects in the DSMs, a region-based change detection method is tested. It fuses the line features from panchromatic images and change features from DSMs. Moreover, a segment merging strategy is proposed to generate the initial region unit. Weighted CVA is adopted for the change map generation.

For the decision fusion procedure, we use the DS fusion theory to extract probabilities for real building changes. The sigmoid function is employed in the automatic initial mass function generation procedure. Furthermore, the whole procedure works unsupervised and the probability calculation is much faster than model-based clustering methods. The generated probability maps can also be used effectively in a semi-automatic procedure.

- Post-classification change detection

Post-classification is a highly efficient method for building change detection when outdated building footprints are available. This approach consists of two steps: Firstly, the changed building locations are detected through initial building mask comparison. Secondly, the destroyed buildings and newly built buildings are processed separately. The destroyed buildings can be identified easily and removed from the outdated building footprint. The new building boundaries have to be detected. For this purpose, a robust building reconstruction method is proposed and evaluated in Chapter 6. This approach automatically splits the images around the changed buildings into 2D rectangular segments by using the Hough transform. The probability of each segment belonging to a building is calculated from random forest estimations. An automatic training data selection strategy is designed for this supervised classification. Finally, the 2D segments and classification results are combined to obtain the final building boundaries.

## 8.2 Conclusion

To conclude, according to the objectives stated in Chapter 1, this thesis has contributed to DSM-assisted change detection for spaceborne optical data in the following fields:



- Using the three proposed approaches, it has been shown that DSMs generated from satellite stereo imagery are suitable for automatic and semi-automatic 3D change detection for objects like buildings and forest. It has been proven that the accuracy of change detection is improved in all approaches after introducing DSMs to the process chain.
- As DSMs generated from spaceborne stereo data contain potential errors, using only these DSMs to extract the change mask would result in numerous false alarms. The original images used to create the DSMs can and should be involved in the change detection procedure to improve the result. It has been shown that full use should be made of features from original satellite images in order to cope with the potential incorrect information from DSMs. Shortcomings may appear if one of the DSMs does not meet the minimum required quality (missing objects, large errors) and does not represent the real situation, as a result of which the result is negatively influenced.
- It has been proven that in urban areas, VHR spaceborne imagery can provide high accuracy DSMs. In this case, directly refining the change mask using DSM subtraction can already extract the important changes. Further methods can only improve the accuracy slightly. However, the fusion-based methods improve the results considerably when DSMs of low quality are used.
- It has been shown that the decision fusion method is advanced in VHR imagery, but for high resolution stereo imagery, the region-based method can normally obtain a clear change mask with high accuracy. When an outdated building footprint is available, the post-classification method is more efficient for building change detection and makes it possible to detect rebuilt areas.

### 8.3 Outlook

Three main approaches have been demonstrated for DSM-assisted change detection and all of them have the ability to highlight changes and provide useful change masks. The most outstanding feature of these approaches is the robust fusion of the change and no-change features from DSMs and original satellite images. Future investigations for each approach should focus on adopting more change features for the fusion models or refining the change detection algorithms. The following possible future research topics could be of interest:

- Investigating the possibilities of using more terrain features from DSMs and more texture features from images. These features can be adopted in many stages of change detection: for instance, change feature calculation, change indicator modelling, segmentation, segmentation merging, etc.

## *8 Conclusions and outlook*

- The quality of the generated DSM is more sensitive in the DSM-assisted localization approach. Therefore, appropriate DSM refinement processing might help to increase the accuracy of change detection.
- A basic DS fusion model was designed for the decision fusion-based change detection. If more indicators can be extracted, like land cover classes, roughness and texture features, the DS fusion model can also be refined to improve the accuracy of change detection.
- Improving the performance of each approach by employing more 2D change detection methods, e.g. kernel change detection, IR-MAD etc. Due to the limited time available for this research, limited 2D change detection algorithms are tested. More linear or kernel change detection algorithms could be introduced to 3D processing.
- A more detailed introduction to 3D shape features could be of interest. During the last step of building change detection, a number of shape features are adopted. More 2D as well as 3D shape features could be used to improve the overall change detection result.
- A robust automatic thresholding method could be helpful. In the existing framework, several threshold values are still required. It is difficult to establish one automatic threshold method that can be used for all purposes.

## **Bibliography**

Aguirre-Gutiérrez, J., Seijmonsbergen, A.C., and Duivenvoorden, J.F., 2012. Optimizing land cover classification accuracy for change detection: a combined pixel-based and object-based approach in a mountainous area in Mexico. *Applied Geography*, vol. 34, pp. 29-37.

Akca, D., 2007. Matching of 3D surfaces and their intensities. *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 62, no. 2, pp. 112-121.

Alberga, V., 2009. Similarity measures of remotely sensed multi-sensor images for change detection applications. *Remote Sensing*, vol. 1, no. 3, pp. 122-143.

Arefi, H., d'Angelo, P., Mayer, H., and Reinartz, P., 2011. Iterative approach for efficient digital terrain model production from CARTOSAT-1 stereo images. *Journal of Applied Remote Sensing*, vol. 5, no.1, pp. 053527.

Awrangjeb, M., Ravanbakhsh, M., Fraser, C.S., 2010. Automatic detection of residential buildings using LIDAR data and multispectral imagery. *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 65, no. 5, pp. 457-467.

AUG signals, 2013. Data Fusion techniques: Image Fusion and algorithm Fusion. <http://www.augsignals.com/userfiles/file/AUG%20Signals%20-%20Fusion%20Techniques%20-%20Technical%20Brief.pdf>, (accessed 21th January, 2013).

Baltsavias, E.P., Mason, S., and Stallmann, D., 1995. Use of DTMs/DSMs and orthoimages to support building extraction. In: A. Grün, O. Kuebler, P. Agouris (Eds.), *Automatic Extraction of Man-Made Objects from Aerial and Space Images*, Birkhauser Verlag, Basel, pp. 189–198.

Baltsavias, E.P., 2004. Object extraction and revision by image analysis using existing geodata and knowledge: current status and steps towards operational systems. *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 58, no. 3, pp. 129-151.

## *Bibliography*

Benediktsson, J. and Kanellopoulos, I., 1999. Classification of multisource and hyperspectral data based on decision fusion. *IEEE Transactions on Geoscience and Remote Sensing*, vol. 37, no. 3, pp. 1367-1377.

Bentabet, L., Jodouin, S., Ziou, D., and Vaillancourt, J., 2003. Road vectors update using SAR imagery: A snake based method. *IEEE Transactions on Geosciences and Remote Sensing*, vol. 41 no. 8, pp. 1785-1803.

Blair, W.D. and Bar-Shalom, Y., 1996. Tracking maneuvering targets with multiple sensors\_does more data always mean better estimates? *IEEE Transactions on Aerospace and Electronic Systems*, vol. 32, no. 1, pp. 450-456.

Blaschke, T., 2005. Towards a framework for change detection based on image objects. *Göttinger Geographische Abhandlungen*, vo. 113, pp. 1-9.

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

Bovolo, F., 2006. Advanced methods for automatic change detection in multitemporal remote sensing images acquired by SAR and multispectral sensors. PhD Thesis, University of Trento, Italy.

Bovolo, F., and Bruzzone, L., 2007. A theoretical framework for unsupervised change detection based on change vector analysis in the polar domain. *IEEE Transactions on Geoscience and Remote Sensing*, vol. 45, no. 1, pp. 218-236.

Bovolo, F., Bruzzone, L., and Marconcini, M., 2008. A novel approach to unsupervised change detection based on a semisupervised SVM and a similarity measure. *IEEE Transactions on Geoscience and Remote Sensing*, vol. 46, no. 7, pp. 2070-2082.

Bovolo, F., 2009. A multilevel parcel-based approach to change detection in very high resolution multitemporal images. *IEEE Geoscience and Remote Sensing Letters*, vol. 6, no. 1, pp. 33-37.

Bowen, H.S., 2002. Absolute radiometric calibration of the IKONOS sensor using radiometrically characterized stellar sources. *International archives of photogrammetry remote sensing and spatial information sciences*, vol. 34, pp. 7-13.

Bouziani, M., Goita, K., and He, D.-C., 2010. Automatic change detection of buildings in urban environment from very high spatial resolution images using existing geodatabase and prior knowledge. *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 65, no. 1, pp. 143-153.

Breiman, L., 2001. Random Forest. *Machine Learning*, vol. 45, no. 1, pp. 5-32.

Brown, L.G., 1992. A survey of image registration techniques. *ACM computing surveys (CSUR)*, vol. 24, no. 4, pp. 325-376.

Bruzzone, L., and Serpico, S.B., 1997. An iterative technique for the detection of land-cover transitions in multitemporal remote-sensing images. *IEEE Transactions on Geoscience and Remote Sensing*, vol. 35, no. 4, pp. 858-867.

Bruzzone, L., and Prieto, D.F., 2000. An adaptive parcel-based technique for unsupervised change detection. *International Journal of Remote Sensing*, vol. 21, no. 4, pp. 817-822.

Bruzzone, L., Chi, M., and Marconcini, M., 2006. A novel transductive SVM for the semisupervised classification of remote-sensing images. *IEEE Transactions on Geoscience and Remote Sensing*, vol. 44, no. 11, pp. 3363-3373.

Camps-Valls, G., and Bruzzone, L., 2009. *Kernel methods for remote sensing data analysis*. NJ, USA: J. Wiley & Sons.

Canty, M.J., and Nielsen, A.A., 2008. Automatic radiometric normalization of multitemporal satellite imagery with the iteratively re-weighted MAD transformation. *Remote Sensing of Environment*, vol. 112, no. 3, pp. 1025-1036.

Castilla, G., Guthrie, R.H., and Hay, G.J., 2009. The Land-cover Change Mapper (LCM) and its application to timber harvest monitoring in western Canada. *Photogrammetric Engineering and Remote Sensing*, vol. 75, no. 8, pp. 941-950.

Celik, T., 2009. Unsupervised change detection in satellite images using principal component analysis and k-means clustering. *IEEE Geoscience and Remote Sensing Letters*, vol. 6, no. 4, pp. 772-776.

## *Bibliography*

Chaabouni-Chouayakh, H., Krauss, T., d'Angelo, P., and Reinartz, P., 2010. 3D change detection inside urban areas using different digital surface models. PCV 2010 - ISPRS Technical Commission III Symposium on Photogrammetry Computer Vision and Image Analysis, 1-3 September 2010, Paris, France, 6 pages (on CD-ROM).

Chaabouni-Chouayakh, H., and Reinartz, P., 2011. Towards automatic 3D change detection inside urban areas by combining height and shape information. *Photogrammetrie Fernerkundung Geoinformation*, no.4, pp. 205-217.

Champion, N., 2007. 2D building change detection from high resolution aerial images and correlation digital surface models. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 36, part 3, pp. 197-202.

Champion, N., Matikainen, L., Rottensteiner, F., Liang, X., and Hyypä, J., 2008. A test of 2D building change detection methods: Comparison, evaluation and perspectives. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 37, pp. 297-305.

Champion, N. and Rottensteiner, F. and Matikainen, L. and Liang, X. and Hyypä, J., and Olsen, B.P., 2009. A test of automatic building change detection approaches. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 38, Part 3/W4, pp.145-150.

Chavez Jr, P.S., 1988. An improved dark-object subtraction technique for atmospheric scattering correction of multispectral data. *Remote Sensing of Environment*, vol. 24, no. 3, pp. 459-479.

Chen, J., Gong, P., He, C., Pu, R., and Shi, P., 2003. Land-use / land-cover change detection using improved change-vector analysis. *Photogrammetric Engineering and Remote Sensing*, vol. 69, no. 4, pp. 369-379.

Choi, K., Lee, I., and Kim, S., 2009. A feature based approach to automatic change detection from Lidar data in urban areas. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 38, Part 3/W8, pp. 259-264.

Comaniciu, D., and Meer, P., 2002. Mean shift: a robust approach toward feature space analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 5, pp. 603-619.

Congalton, R.G., 1991. A review of assessing the accuracy of classifications of remotely sensed data. *Remote sensing of Environment*, vol. 37, no.1, pp. 35-46.

Coppin, P.R., and Bauer, M.E., 1996. Digital change detection in forest ecosystems with remotely sensed imagery. *Remote Sensing Reviews*, vol. 13, no. 3-4, pp. 207-234.

Cormen, T.H., Leiserson, C.E., Rivest, R.L., and Stein, C., 2001. *Introduction to Algorithms* (2<sup>nd</sup> ed.). Chapter 16 "Greedy Algorithms", MIT Press and McGraw-Hill.

Cui, S., Datcu, M., and Gueguen, L., 2011. Information theoretical similarity measure for change detection. In: Stilla U, Gamba P, Juergen C, Maktav D (Eds.) *Proceedings JURSE 2011*, Munich.

Cui, S., Yan, Q., and Reinartz, P., 2012. Complex building description and extraction based on Hough transformation and cycle detection. *Remote Sensing Letters*, vol. 3, no. 2, pp. 151-159.

Dai, X. and Khorram, S., 1998. The effects of image misregistration on the accuracy of remotely sensed change detection. *IEEE Transactions on Geoscience and Remote Sensing*, vol. 36, no. 5, pp. 1566-1577.

d'Angelo, P., Lehner, M., Krauss, T., Hoja, D., and Reinartz, P., 2008. Towards automated DEM generation from high resolution stereo satellite images. *International Society for Photogrammetry and Remote Sensing*, vol. 37, Part B4, pp. 1137-1342.

d'Angelo, P., 2010. Image matching and outlier removal for large scale DSM generation. *Convergence in Geomatics. CGC & ISPRS . ISPRS Symposium Commission I*, 15.-18. June 2010, Calgary, Canada.

de Berg, M., Cheong, O., van Kreveld, M., and Overmars M., 2008. *Computational Geometry: Algorithm and application*, Third Edition, Springer-Verlag.

Dempster, A.P., 1968. A generalization of Bayesian Inference. *Journal of the Royal Statistical Society, Series B*, vol. 30, no. 2, pp. 205-247.

Dempster, N.M., Laird, A.P., and Rubin, D.B., 1977. Maximum likelihood from incomplete data via the EM algorithm, *Journal of the Royal Statistical Society, Series B*, vol. 39, no. 1, pp. 1-38.

## *Bibliography*

Desclée, B., Bogaert, P., and Defourny, P., 2006. Forest change detection by statistical object-based method. *Remote Sensing of Environment*, vol. 102, no. 1-2, pp. 1-11.

Dronova, I., Gong, P., and Wang, L., 2011. Object-based analysis and change detection of major wetland cover types and their classification uncertainty during the low water period at Poyang Lake, China. *Remote Sensing of Environment*, vol. 115, no. 12, pp. 3220-3236.

Du, Y., Teillet, P.M., and Cihlar, J., 2002. Radiometric normalization of multitemporal high-resolution satellite images with quality control for land cover change detection. *Remote Sensing of Environment*, vol. 82, no. 1, pp. 123-134.

Duda, R.O., and Hart, P.E., 1972. Use of the Hough transformation to detect lines and curves in pictures. *Communications of the ACM*, vol. 15, no. 1, pp. 11-15.

Durieux, L., Lagabrielle, E., and Nelson, A., 2008. A method for monitoring building construction in urban sprawl areas using object-based analysis of Spot 5 images and existing GIS data. *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 63, no. 4, pp. 399-408.

Duveiller, G., Defourny, P., Desclée, B., and Mayaux, P., 2008. Deforestation in Central Africa: Estimates at regional, national and landscape levels by advanced processing of systematically-distributed Landsat extracts. *Remote Sensing of Environment*, vol. 112, no. 5, pp. 1969-1981.

Ehlers, M., 2004. Spectral characteristics preserving image fusion based on Fourier domain filtering. *Proceedings of SPIE, Maspalomas, Spain*, vol. 5574, pp. 1 - 13.

Ehlers, M., Klonus, S., Åstrand, P.J., and Rosso, P., 2010. Multi-sensor image fusion for pansharpening in remote sensing. *International Journal of Image and Data Fusion*, vol. 1, no. 1, pp. 25-45.

Fan, H., Zhang, J., Zhang, Z., and Liu, Z., 1999. House change detection based on DSM of aerial image in urban area. *Geo-spatial Information Science*, vol. 2, no. 1, pp. 68-72.

Fraser, C.S., Baltsavias, E.P, and Gruen, A., 2002. Processing of IKONOS imagery for submetre 3D positioning and building extraction. *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 56, no. 3, pp. 177-194.



Fung, T., 1990. An assessment of TM imagery for land-cover change detection. *IEEE Transactions on Geoscience and Remote Sensing*, vol. 28, no. 4, pp. 681–684.

Goodchild, M.F., 1994. Integration GIS and remote sensing for vegetation analysis and modeling: methodological issues. *Journal of Vegetation Science*, vol. 5, no.5, pp. 615-626.

Green, K., Lempka, D., and Lackey, L., 1994. Using remote sensing to detect and monitor land-cover and land-use change. *Photogrammetric Engineering and Remote Sensing*, vol. 60, no. 3, pp. 331-337.

Grigillo, D., Fras, M.K., and Petrovič, D., 2012. Automated building extraction from IKONOS images in suburban areas. *International Journal of Remote Sensing*, vol. 33, no. 16, pp. 5149-5170.

Grodecki, J., and Dial, G., 2003. Block adjustment of high-resolution satellite images described by rational polynomials. *Photogrammetric Engineering and Remote Sensing*, vol. 69, no. 1, pp. 59-68.

Grodecki, J., Dial, G., and Lutes, J., 2004. Mathematical model for 3D feature extraction from multiple satellite images described by RPCs. *Proceedings of ASPRS 2004 Conference*, Denver, Colorado, USA, May 23-28, 13 pages (on CD-ROM).

Grohman, G., Kroenung, G., and Strebeck, J., 2006. Filling SRTM voids: The delta surface fill method. *Photogrammetric Engineering and Remote Sensing*, vol. 72, no. 3, pp. 213-216.

Gruen, A., and Akca, D., 2005. Least squares 3D surface and curve matching. *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 59, no. 3, pp. 151-174.

Gueguen, L., Soille, P., and Pesaresi, M., 2011. Change detection based on information measure. *IEEE Transactions on Geoscience and Remote Sensing*, vol. 49, no. 11, pp. 4503-4515.

Hall, O., and Hay, G.J., 2003. A multiscale object-specific approach to digital change detection. *International Journal of Applied Earth Observation and Geoinformation*, vol. 4, no. 4, pp. 311-327.

Hall, D., and Llinas, J., 1997. An introduction to multisensor data fusion. *Proceedings of the IEEE*, vol. 85, no. 1, pp. 6-23.

## *Bibliography*

Hand, D.J., and Till, R.J. 2001. A simple generalization of the area under the ROC curve to multiple class classification problems. *Machine Learning*, vol. 45, no.2, pp. 171–186.

Haris, K., Efstratiadis, S., Maglaveras, N., and Katsaggelos, A., 1998. Hybrid image segmentation using watersheds and fast region merging. *IEEE Transactions on Image Processing*, vol. 7, no. 12, pp. 1684-1699.

Hayes, D.J., and Sader, S.A., 2001. Comparison of change-detection techniques for monitoring tropical forest clearing and vegetation regrowth in a time series. *Photogrammetric Engineering and Remote Sensing*, vol. 67, no. 9, pp. 1067-1075.

Heller, A.J., Leclerc, Y.G., and Luong, Q.-T., 2001. Framework for robust 3D change detection. *International Symposium on Remote Sensing, SPIE, Toulouse, France*, pp. 639-649.

Huete, A. R., 1988. A Soil-Adjusted Vegetation Index (SAVI). *Remote Sensing of Environment*, vol. 25, no. 3, pp. 295-309.

Hirschmüller, H., 2008. Stereo processing by semiglobal matching and mutual information. *IEEE Transactions on Pattern Analyses and Machine Intelligence*, vol. 30, no. 2, pp. 1-14.

Höhle, J., and Höhle, M., 2009. Accuracy assessment of digital elevation models by means of robust statistical methods. *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 64, no. 4, pp. 398-406.

Höhle, J., and Potuckova, M., 2011. Assessment of the quality of Digital Terrain Models, European Spatial Data Research, EuroSDR. <http://www.eurosd.net/publications/60.pdf>.

Huang, X., and Jensen, J.R., 1997. A machine-learning approach to automated knowledge-base building for remote sensing image analysis with GIS data. *Photogrammetric Engineering and Remote Sensing*, vol. 63, no. 10, pp. 1185-1193.

Im, J., and Jensen, J.R., 2005. A change detection model based on neighborhood correlation image analysis and decision tree classification. *Remote Sensing of Environment*, vol. 99, no. 3, pp. 326-340.

Im, J., Jensen, J.R., and Tullis, J.A., 2008. Object based change detection using correlation image analysis and image segmentation. *International Journal of Remote Sensing*, vol. 29, no. 2, pp. 399-423.

Inglada, J., and Mercier, G., 2007. A new statistical similarity measure for change detection in multitemporal SAR images and its extension to multiscale change analysis. *IEEE Transactions on Geoscience and Remote Sensing*, vol. 45, no. 5, pp. 1432-1445.

James, L.A., Hodgson, M.E., Ghoshal, S., and Latiolais, M.M., 2012. Geomorphic change detection using historic maps and DEM differencing: The temporal dimension of geospatial analysis. *Geomorphology*, vol. 137, no.1, pp. 181-198.

Jensen, J.R., 1996. *Introductory Digital Image Processing: A Remote Sensing Perspective*, Second Edition, Prentice Hall, Upper Saddle River, New Jersey.

Jordan, C.F., 1969. Derivation of leaf-area index from quality of light on the forest floor. *Ecology*, vol. 50, no. 4, pp. 663-666.

Jung, F., 2004. Detecting building changes from multitemporal aerial stereopairs. *ISPRS Journal of Photogrammetry and Remote sensing*, vol. 58, no. 3, pp. 187-201.

Kaufman, Y.J., and Tanre, D., 1992. Atmospherically resistant vegetation index (ARVI) for EOS-MODIS. *IEEE Transactions on Geoscience and Remote Sensing*, vol. 30, no. 2, pp. 261-270.

Kittler, J., and Illingworth, J., 1986. Minimum error thresholding. *Pattern Recognition*, vol. 19, no. 1, pp. 41-47.

Kriegler, F.J., Malila, W.A., Nalepka, R.F., and Richardson, W., 1969. Preprocessing transformations and their effects on multispectral recognition. In: *Proceedings of the Sixth International Symposium on Remote Sensing of Environment*, University of Michigan, Ann Arbor, MI, pp. 97-131.

Lawrence, R.L., and Wright, A., 2001. Rule-based classification systems using classification and regression tree (CART) analysis, *Photogrammetric Engineering and Remote Sensing*, vol. 67, no. 10, pp.1137-1142.

## *Bibliography*

Le Hégarat-Masclé, S., Bloch, I., and Vidal-Madjar, D., 1997. Application of Dempster-Shafer evidence theory to unsupervised classification in multisource remote sensing. *IEEE Transactions on Geoscience and Remote Sensing*, vol. 35, no. 4, pp. 1018–1031.

Le Hégarat-Masclé, S., and Seltz, R., 2004. Automatic change detection by evidential fusion of change indices. *Remote Sensing of Environment*, vol. 91, no. 3, pp. 390-404.

Lillesand, T.M., and Kiefer, R.W., 1987. *Remote Sensing and Image Interpretation*, Second Edition, John Wiley & Sons, Inc. Toronto.

Lin, J.-J., Saito, N., and Levine, R.A., 1999. Edgeworth expansions of the Kullback-Leibler information. Technical Report, Division of Statistics, University of California, Davis, submitted to *J. Amer. Statist. Assoc.*

Liu, Z., Zhang, J., Zhang, Z., and Fan, H., 2003. Change detection based on DSM and image features in urban areas. *Geo-Spatial Information Science*, vol. 6, no. 2, pp. 35-41.

Liu, H., and Zhou, Q., 2004. Accuracy analysis of remote sensing change detection by rule-based rationality evaluation with post-classification comparison. *International Journal of Remote Sensing*, vol. 25, no. 5, pp. 1037-1050.

Lu, D., Mausel, P., Brondizio, E., and Moran, E., 2004. Change detection techniques. *International Journal of Remote Sensing*, vol. 25, no. 12, pp. 2365-2407.

Lu, Y., Trinder, J.C., and Kubik, K., 2006. Automatic building detection using the Dempster-shafer Algorithm. *Photogrammetric Engineering and Remote Sensing*, vol. 72, no. 4, pp. 395-403.

Macleod, R.D., and Congalton, R.G., 1998. A quantitative comparison of change-detection algorithms for monitoring eelgrass from remotely sensed data. *Photogrammetric Engineering and Remote sensing*, vol. 64, no. 3, pp. 207-216.

Makarau, A., Richter, R., Müller, R., and Reinartz, P., 2011, Adaptive shadow detection using a blackbody radiator model. *IEEE Transactions on Geoscience and Remote Sensing*, vol. 49, no. 6, pp. 2049-2059.

Malila, W.A., 1980. Change vector analysis: an approach for detecting forest changes with Landsat. Proceedings of the 6th Annual Symposium on Machine Processing of Remotely Sensed Data, 03-06 June, Purdue University, West Lafayette, Indiana, pp. 326–335.

Marpu, P.R., Gamba, P., and Canty, M.J., 2011. Improving change detection results of IR-MAD by eliminating strong changes. *IEEE Geoscience and Remote Sensing Letters*, vol. 8, no.4, pp. 799-803.

Mas, J.-F. 1999, Monitoring land-cover changes: a comparison of change detection techniques. *International Journal of Remote Sensing*, vol. 20, no.1, pp. 139-152.

Matikainen, L., Hyypä, J., and Hyypä, H., 2003. Automatic detection of buildings from laser scanner data for map updating. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 34, pp. 218-224.

Matikainen, L., Hyypä, J., Ahokas, E., Markelin, L., and Kaartinen, H., 2010. Automatic detection of buildings and changes in buildings for updating of maps. *Remote Sensing*, vol. 2, no. 5, pp. 1217-1248.

Mayer, H., 1999. Automatic object extraction from aerial imagery—a survey focusing on buildings. *Computer vision and image understanding*, vol. 74, no. 2, pp. 138-149.

Melendez, J., Garcia, M., Puig, D., and Petrou, M., 2011. Unsupervised texture-based image segmentation through pattern discovery. *Computer Vision and Image Understanding*, vol. 115, no. 8, pp. 1121-1133.

Melgani, F., Moser, G., and Serpico, S.B., 2002. Unsupervised change-detection methods for remote-sensing images. *Optical Engineering*, vol. 41, no. 12, pp. 3288-3297.

Mena, J.B., 2003. State of the art on automatic road extraction for GIS database update: a novel classification. *Pattern Recognition Letters*, vol. 24, no. 16, pp. 3037-3058.

Meyer, F., and Beucher, S., 1990. Morphological segmentation. *Journal of Visual Communication and Image Representation*, vol. 1, no. 1, pp. 21-46.

## *Bibliography*

Muchoney, D.M., and Haack, B.N., 1994. Change detection for monitoring forest defoliation. *Photogrammetric Engineering and Remote Sensing*, vol. 60, no. 10, pp. 1243-1251.

Müller, R., Krauß, T., Lehner, M., and Reinartz, P., 2007. Automatic production of a European orthoimage coverage within the GMES land fast track service using SPOT 4/5 and IRS-P6 LISS III data. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 36, ISPRS Workshop, Hannover, Germany, 6p.

Murakami, H., Nakagawa, K., Hasegawa, H., Shibata, T., and Iwanami, E., 1999. Change detection of buildings using an airborne laser scanner. *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 54, no. 2, pp. 148-152.

Nadimi, S., and Bhanu, B., 2004. Physical models for moving shadow and object detection in video. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 26, no. 8, pp. 1079-1087.

Nielsen, A.A., Conradsen, K., and Simpson, J.J., 1998. Multivariate alteration detection (MAD) and MAF postprocessing in multispectral, bitemporal image data: New approaches to change detection studies. *Remote Sensing of Environment*, vol. 64, no. 1, pp. 1-19.

Nielsen, A.A., and Canty, M.J., 2008. Kernel principal component analysis for change detection. *Proc. SPIE, Image and Signal Processing for Remote Sensing XIV*, vol. 7109, 10p.

Nielsen, A.A., 2011. Kernel maximum autocorrelation factor and minimum noise fraction transformations. *IEEE Transactions on Image Processing*, vol. 20, no. 3, pp. 612-624.

Ning, J., Zhang, L., Zhang, D., and Wu, C., 2010. Interactive image segmentation by maximal similarity based region merging. *Pattern Recognition*, vo. 43, no. 2, pp. 445-456.

Nock, R., and Nielsen, F., 2004. Statistical region merging. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 26, no. 11, pp. 1452-1458.

Noronha, S., and Nevatia, R., 2001. Detection and Modeling of Buildings from Multiple Aerial Images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, no. 5, pp. 501-518.

Olsen, B.P., 2004. Automatic change detection for validation of digital map databases. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 34, pp. 569-574.

Olsen, B.P, and Knudsen, T., 2005. Automated change detection for validation and update of geodata. *Proceeding of 6th Geomatic Week, Barcelone, Spain*.

Osawa, Y., 2004. Optical and Microwave Sensor on Japanese Mapping Satellite- ALOS, Japan Aerospace Exploratum Agency (JAXA), Japan.

Otsu, N., 1979. A Threshold Selection Method from Gray-Level Histogram. *IEEE Transactions on Systems, Man, and Cybernetics*, vo. 9, no. 1, pp. 62-66.

Paolini, L., Grings, F., Sobrino, J.A., Jiménez Muñoz, J.C., and Karszenbaum, H., 2006. Radiometric correction effects in Landsat multi-date/multi-sensor change detection studies. *International Journal of Remote Sensing*, vol. 27, no.4, pp. 685-704.

Pesaresi, M., and Benediktsson, J.A., 2001. A new approach for the morphological segmentation of high-resolution satellite imagery. *IEEE Transactions on Geoscience and Remote Sensing*, vol. 39, no. 2, pp. 309–320.

Pohl, C., and van Genderen, J.L., 1998. Multisensor image fusion in remote sensing: concepts, methods and applications. *International Journal of Remote Sensing*, vol. 19, no. 5, pp. 823-854.

Poli, D., and Soille, P., 2012. Digital Surface Model Extraction and Refinement through Image Segmentation Application to the ISPRS Benchmark Stereo Dataset, *Photogrammetrie-Fernerkundung-Geoinformation*, E. Schweizerbart'sche Verlagsbuchhandlung, no. 4, pp. 317-329.

Prati, A., Mikic, I., Trivedi, M.M., and Cucchiara, R., 2003. Detecting moving shadows: algorithms and evaluation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 25, no. 7, pp. 918-923.

Qi, J., Chehbouni, A., Huete, A.R., Kerr, Y.H., and Sorooshian, S., 1994. Modified Soil Adjusted Vegetation Index (MSAVI). *Remote Sensing of Environment*, vol. 48, no. 2, pp. 119-126.

## *Bibliography*

Radke, R.J., Andra, S., Al-Kofahi, O., and Roysam, B., 2005. Image Change Detection Algorithms: A Systematic Survey. *IEEE Transactions on Image Processing*, vol. 14, no. 3, pp. 294-307.

Ray, T.W., 1994. A FAQ on Vegetation in Remote Sensing. Available via anonymous FTP at: kepler.gps.caltech.edu/pub/terrell/rsvegfaq.txt.

Reinartz, P., Müller, R., Lehner, M., and Schroeder, M., 2006. Accuracy analysis for DSM and orthoimages derived from SPOT HRS stereo data using direct georeferencing. *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 60, no. 3, pp. 160-169.

Reinartz, P., Müller, R., Schwind, P., Suri, S., and Bamler, R., 2011. Orthorectification of VHR optical satellite data exploiting the geometric accuracy of TerraSAR-X data. *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 66, no. 1, pp. 124-132.

Ridler, T., and Calvard, S., 1978. Picture thresholding using an iterative selection method. *IEEE Transactions on Systems, Man and Cybernetics*, vol. 8, no. 8, pp. 630-632.

Ridley, H.M., Atkinson, P.M., Aplin, P., Muller, J.-P., and Dowman, I., 1997. Evaluating the potential of the forthcoming commercial U.S. high-resolution satellite sensor imagery at the ordnance survey. *Photogrammetric Engineering and Remote Sensing*, vol. 63, no. 2, pp. 997-1005.

Rodriguez, E., Morris, C.S., Belz, J.E., Chapin, E.C., Martin, J.M., Daffer, W., and Hensley, S., 2005. An assessment of the SRTM topographic products, Technical Report JPL D-31639, Jet Propulsion Laboratory, Pasadena, California, USA.

Rosin, P., 2002. Thresholding for change detection. *Computer Vision and Image Understanding*, vol. 86, no. 2, pp. 79-95.

Rosin, P., and Hervás, J., 2005. Remote sensing image thresholding methods for determining landslide activity. *International Journal of Remote Sensing*, vol. 26, no. 6, pp. 1075-1092.

Rottensteiner, F., Trinder, J., Clode, S., and Kubik, K., 2005. Using the Dempster-Shafer method for the fusion of LIDAR data and multi-spectral images for building detection. *Information Fusion*, vol. 6, no. 4, pp. 283-300.



Rottensteiner, F., 2007a. Building change detection from Digital Surface Models and multi-spectral images. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 36, pp. 145-150.

Rottensteiner, F., Trinder, J., Clode, S., and Kubik, K., 2007b. Building detection by fusion of airborne laser scanner data and multi-spectral images: Performance evaluation and sensitivity analysis. *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 62, no. 2, pp. 135-149.

Rouse, J.W., Haas, R.H., Schell, J.A., and Deering, D.W., 1973. Monitoring vegetation systems in the great plains with ERTS," *Third ERTS Symposium, NASA SP-351*, vol. 1, pp.309-317.

Roy, D.P., 2000. The impact of misregistration upon composited wide field of view satellite data and implications for change detection. *IEEE Transactions on Geoscience and Remote Sensing*, vol. 38, no, 4, pp. 2017-2032.

Sanin, A., Sanderson, C., and Lovell, B.C., 2012. Shadow detection: A survey and comparative evaluation of recent methods. *Pattern recognition*, vol. 45, no. 4, pp. 1684-1695.

Sasagawa, H, Watanabe, K., Nakajima, S., Koido, K., Ohno, H., and Fujimura, H., 2008. Automatic change detection based on pixel-change and DSM-change. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Science*, vol. 37, Part B7, pp. 1645-1650.

Satellite Imaging Corporation, 2012a. <http://www.satimagingcorp.com/satellite-sensors/alos.html> (accessed 21th November 2012).

Satellite Imaging Cooperation, 2012b. <http://www.satimagingcorp.com/satellite-sensors/cartosat-1.html> (accessed 21th November 2012).

Satellite Imaging Cooperation, 2012c. <http://www.satimagingcorp.com/satellite-sensors/geoeye-1.html> (accessed 21th November 2012).

Satellite Imaging Cooperation, 2012d. <http://www.satimagingcorp.com/satellite-sensors/ikonos.html> (accessed 21th November 2012).

## *Bibliography*

Satellite Imaging Cooperation, 2012e. <http://www.satimagingcorp.com/satellite-sensors/quickbird.html>(accessed 21th November 2012).

Satellite Imaging Cooperation, 2012f. <http://www.satimagingcorp.com/satellite-sensors/spot-5.html>(accessed 21th November 2012).

Satellite Imaging Cooperation, 2012g. <http://www.satimagingcorp.com/satellite-sensors/worldview-2.html> (accessed 26th November 2012).

Shafer, G., 1976. *A Mathematical Theory of Evidence*, Princeton University Press, Princeton, New Jersey.

Schmid, C., and Zisserman, A., 1997. Automatic Line Matching Across Views, In: *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 666–671.

Schroeder, T.A., Cohen, W.B., Song, C., Canty, M.J., and Yang, Z., 2006. Radiometric correction of multi-temporal Landsat data for characterization of early successional forest patterns in western Oregon. *Remote Sensing of Environment*, vol. 103, no. 1, pp. 16-26.

Shaker, I.F., Abd-Elrahman, A., Abdel-Gawad, A.K., and Sherief, M.A., 2011. Building extraction from high resolution space images in high density residential areas in the great Cairo region, *Remote Sensing*, vol. 3, no. 4, pp. 781-791.

Schindler, K., 2012. An overview and comparison of smooth labeling method for land-cover classification. *IEEE Transactions on Geoscience and Remote Sensing*, vol. 50, no. 11, pp. 4534-4545.

Singh, A., 1986. Change detection in the tropical forest environment of northeastern India using Landsat. *Remote Sensing and Land Management*, edited by M.J. Eden and J.T. Parry (London: John Wiley & Sons), pp. 237-254.

Singh, A., 1989. Digital change detection techniques using remotely-sensed data. *International Journal of Remote Sensing*, vol. 10, no. 6, pp. 989-1003.

Sirmacek, B., and Unsalan, C., 2011. A probabilistic framework to detect buildings in aerial and satellite images. *IEEE Transactions on Geoscience and Remote Sensing*, vol. 49, no. 1, pp. 211-221.

Sohn, G., and Dowman, I., 2007. Data fusion of high-resolution satellite imagery and LiDAR data for automatic building extraction. *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 62, no. 1, pp. 43-63.

Song, C., Woodcock, C.E., Seto, K.C., Lenney, M.P., and Macomber, S.A., 2001. Classification and change detection using Landsat TM data: When and how to correct atmospheric effect? *Remote Sensing of Environment*, vol. 75, no. 2, pp. 230-244.

Straub, C., Tian, J., Seitz, R., and Reinartz, P., 2013. Assessment of Cartosat-1 and WorldView-2 stereo imagery in combination with a LiDAR DTM for timber volume estimation in a highly structured forest in Germany. *Forestry*, vol. 86, pp. 463-473.

Suri, S., and Reinartz, P., 2010. Mutual-information-based registration of TerraSAR-X and Ikonos imagery in urban areas. *IEEE Transactions on Geoscience and Remote Sensing*, vol. 48, no. 2, pp. 939-949.

Tang, X., and Xie, J., 2012. Overview of the key technologies for high-resolution satellite mapping. *International Journal of Digital Earth*, vol. 5, no. 3, pp. 228-240.

Tian, J., Chaabouni-Chouayakh, H., Reinartz, P., Krauß, T., and d'Angelo, P., 2010. Automatic 3D change detection based on optical satellite stereo imagery. *ISPRS TC VII Vienna symposium*, July 2010, Vienna, Austria.

Tian, J., Chaabouni-Chouayakh, H., and Reinartz, P., 2011. 3D building change detection from high resolution spaceborne stereo imagery. *M2RSM*, Jan. 2011, Xiamen, China.

Tian, J., Cui, S., and Reinartz, P., 2013a. Building Change Detection Based on Satellite Stereo Imagery and Digital Surface Models. *IEEE Transactions on Geoscience and Remote Sensing*, in press, pp.1-12.

## *Bibliography*

Tian, J., Reinartz, P., d'Angelo P., and Ehlers, M., 2013b. Region-based automatic building and forest change detection on Cartosat-1 stereo imagery. *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 79, pp. 226-239.

Tian, J., and Reinartz, P., 2013. Fusion of multi-spectral bands and DSM from Worldview-2 Stereo imagery for building extraction, Joint Urban Remote Sensing Event, Sao Paulo, Brazil, pp. 135-138.

Ton, J., and Jain, A.K., 1989. Registering Landsat images by point matching., vol. 27, no. 5, pp. 642-651.

Tong, X., Hong, Z., Liu, S., Zhang, X., Xie, H., Li, Z., Yang, S., Wang, W., and Bao, F., 2012. Building-damage detection using pre-and post-seismic high-resolution satellite stereo imagery: A case study of the May 2008 Wenchuan earthquake. *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 68, pp. 13-27.

Toutin, T., 2003. Block bundle adjustment of IKONOS in-track images. *International Journal of Remote Sensing*, vol. 24, no. 4, pp. 851-857.

Turker, M., and Cetinkaya, B., 2005. Automatic detection of earthquake-damaged buildings using DEMs created from pre- and post-earthquake stereo aerial photographs. *International Journal of Remote Sensing*, vol. 26, no. 4, pp. 823-832.

Vadon, H., 2003. 3D Navigation over merged panchromatic-multispectral high resolution SPOT5 images. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 34, part 5, 5p.

Vincent, L., and Soille, P., 1991. Watersheds in digital spaces: an efficient algorithm based on immersion simulations. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 13, no. 6, pp. 583-598.

Vögtle, T., and Steinle, E., 2003. On the quality of object classification and automated building modelling based on laserscanning data. *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 34, pp. 149-155.

Volpi, M., Tuia, D., Camps-Valls, G., and Kanevski, M., 2012. Unsupervised change detection with kernels. *IEEE Geoscience and Remote Sensing Letters*, vol. 9, no. 6, pp. 1026-1030.

Walter, V., 2004. Object-based classification of remote sensing data for change detection. *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 58, no. 3-4, pp. 225-238.

Waser, L., Baltsavias, E., Ecker, K., Eisenbeiss, H., Feldmeyer-Christe, E., Ginzler, C., Küchler, M., and Zhang, L., 2008. Assessing changes of forest area and shrub encroachment in a mire ecosystem using digital surface models and CIR aerial images. *Remote Sensing of Environment*, vol. 112, no. 5, pp. 1956-1968.

Weidner, U., and Förstner, W., 1995. Towards automatic building extraction from high-resolution digital elevation models. *ISPRS journal of Photogrammetry and Remote Sensing*, vol. 50, no. 4, pp. 38-49.

Wyawahare, M.V., Patil, P.M., and Abhyankar, H.K., 2009. Image registration techniques: an overview. *International Journal of Signal Processing, Image Processing and Pattern Recognition*, vol. 2, no.3, pp. 11-28.

Xiong, Z., and Zhang, Y., 2011. Bundle adjustment with rational polynomial camera models based on generic method. *IEEE Transactions on Geoscience and Remote Sensing*, vol. 49, no. 1, pp. 190-202.

Zadeh, L.A., 1965. Fuzzy sets. *Information and Control*, vol. 8, no. 3, pp. 338–353.

Zebedin, L., Bauer, J., Karner, K.F., and Bischof, H., 2008. Fusion of feature- and area-based information for urban buildings modeling from aerial imagery. *ECCV (4)*, pp. 873-886.

Zhang, L., 2005. Automatic Digital Surface Model (DSM) Generation from Linear Array Images. PhD Thesis, Swiss Federal Institute of Technology Zürich, Switzerland.

Zhu, L., Shimamura, H., Tachibana, K., Li, Y., and Gong, P., 2008. Building change detection based on object extraction in dense urban areas. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 37, Part B7, pp. 905-908.