Toward an Automatic Solution for Updating Building Databases Using Space-borne Stereo Imaging

Gholam Reza Dini

HANNOVER 2014

Toward an Automatic Solution for Updating Building Databases Using Space-borne Stereo Imaging

Von der Fakultät für Bauingenieurwesen und Geodäsie

der Gottfried Wilhelm Leibniz Universität Hannover

zur Erlangung des Grades

Doktor-Ingenieur (Dr.-Ing.)

genehmigte Dissertation

von

Gholam Reza Dini

geboren am 21.09.1979 in Karaj, Iran

HANNOVER 2014

Prüfungskommission:
Vorsitzender: Univ.-Prof. Dr.-Ing. Ingo Neumann
Referenten: Univ.-Prof. Dr.-Ing. habil. Christian Heipke
                          Univ.-Prof. Dr.-Ing. habil. Monika Sester
                          Dr.-Ing. Markus Gerke

Tag der mündlichen Prüfung: 18.07.2014

Dedication

This thesis work is dedicated to my family...

To Hoda who has been a constant source of support and encouragement during research and life. I am truly thankful for having you in my life.

To Dad and Mom who have always supported me unconditionally and whose good examples have taught me to work hard for the things that I aspire to achieve. Your kindness will never be forgotten.

Acknowledgement

"IMAGINation is more important than knowledge"

Albert Einstein

Abstract

Nowadays, along with the development of high resolution spaceborne stereoscopic sensors, Digital Surface Models (DSM) derived from image matching are also being increasingly used particularly for mapping in urban areas. This research addresses the automatic updating of building databases using high resolution spaceborne stereo images assisted by a potentially outdated GIS database. To update building databases, the task is divided into two main parts; (1) 3D building change detection to establish an alarm system, and (2) delineation of building outlines in order to update the footprint of associated building change, accordingly. First, semi-global matching (SGM) is used to derive DSMs and subsequently normalized DSM (nDSM) is generated.

In this study, two scenarios are tested; image-to-image comparison (DSM vs. DSM) and image-to-map comparison (nDSM vs. GIS polygons). In the first scenario, two DSMs are subtracted from each other followed by applying a height difference threshold of 2.5 meters. A larger difference of height is assumed to represent vertical changes. Additionally before DSM subtraction, a least squares adjustment is utilized to eliminate shifting errors stemming from image orientation. To reduce the number of false alarms, it is required to remove matching errors, roads, and vegetation by applying associated removal filters. A final refinement is based upon shape and size thresholds which are carried out using morphological filtering.

In the second scenario, first building blobs are extracted using a threshold for absolute height of 2.5 meters in the nDSM. Then, the building polygons in the GIS database are compared against these blobs. The building hypothesis is verified if a GIS polygon is covered to at least 75% by building blobs. The accepted polygons constitute the outlines for the existing buildings.

When a signal for building change is confirmed, building outlines are extracted based on two different approaches; 3D edge matching and 3D segmentation. For 3D edge matching, first edges in the epipolar images are detected individually in each image using the Canny operator. After removing short edges and extracting straight lines, we find the best corresponding lines using various geometric and proximity constraints. Finally, the topological relationship of 3D edges is employed to reconstruct the rectangular shape of the building footprints using a box-fitting approach. For 3D segmentation, we use active contours based on the level set method to delineate building rooftops. The initialization is carried out using building blobs obtained from the nDSM. The segmentation is performed iteratively until there is no considerable improvement in the building outlines. Finally, these segments are regularized into rectangular primitives to update building footprints of the new construction.

The approaches are tested using IKONOS and GeoEye-1 stereo images in the suburb of Riyadh, the capital city of Saudi Arabia. Note that the proposed method is only applicable for simple rectangular buildings with flat rooftops such as those to be found in our study area. The evaluation of our tests against ground truth information has shown that the elimination of shifting errors and the application of the removal masks decreases the rate of false alarms considerably. In delineation of building outlines using 3D edge matching, our experiments on GeoEye-1 stereo images show promising results only if buildings are large enough, have simple shape and show very good contrast compared to the background. Otherwise there are a large number of miss-detections. In contrast, 3D segmentation of building rooftops delivers results with a better rate of completeness; nevertheless it fails if a building has very poor contrast compared to the background.

Keywords: *building, change detection, stereoscopic, image matching, GIS database, high resolution*

Zusammenfassung

Aufgrund der Entwicklung von hochauflösenden, satellitengetragenen, stereoskopischen Sensoren werden heutzutage auch Digitale Oberflächenmodelle (DOMs), die durch digitale Bildzuordnung generiert wurden, immer häufiger verwendet, besonders für das Kartieren von Stadtgebieten. Diese Arbeit beschäftigt sich mit der automatischen Aktualisierung von Gebäudedatenbanken mittels hochauflösenden, satellitengetragenen Stereobildpaaren unter Einbindung einer möglicherweise veralteten GIS Datenbasis. Die Methode des semi-global matching (SGM) wurde verwendet, um Digitale Oberflächenmodelle (DOM) und daran anschließend normalisierte Oberflächenmodelle (nDOM) abzuleiten. Die Aktualisierung von Gebäudedatenbanken wird hierzu in zwei Hauptteile unterteilt: (1) ein System zur Erkennung von 3D Änderungen der Gebäude und (2) ein System zur Erfassung der Abgrenzung der neuen Gebäudegrundrisse bzw. deren Aktualisierung.

Zwei verschiedene Szenarien werden getestet um das Potenzial von hochauflösenden Stereobilddaten für die Aktualisierung von Gebäuden zu evaluieren: ein Vergleich zweier Oberflächenmodelle (DOM vs. DOM)" und eine Vergleich von Oberflächenmodell zu GIS Daten.

Im ersten Szenario werden zwei DOMs subtrahiert, anschließend werden Gebiete größer als 2,5 Meter betrachtet. Es wird angenommen, dass größere Höhenunterschiede vertikale Änderungen darstellen. Vor der DOM Subtraktion wird eine relative Referenzierung durchgeführt, um Fehler der Bildorientierung auszugleichen. Zusätzlich ist es notwendig Artefakte aus der Bildzuordnung durch Filtermethoden zu entfernen, um die Anzahl der Fehlalarme zu minimieren. Abschließend wird eine morphologische Filterung durchgeführt, mit Schwellwerten für Höhe, Form und Größe, um das Ergebnis zu verbessern.

Das zweite Szenario verwendet alle Pixel oberhalb des Schwellwertes von 2,5 Metern im nDOM für die Generierung von Gebäudehypothesen. Darauffolgend werden alle Gebäudepolygone aus der Datenbank mit den Hypothesen verglichen. Ein Gebäudepolygon wird dabei verifiziert, wenn es zumindest zu 75 % von der zugehörigen Hypothese belegt ist. Die verifizierten Polygone entsprechen dann weiterhin den Gebäudegrenzen.

Im Falle einer Gebäudeveränderung werden die neuen Gebäudegrundrisse auf zwei unterschiedliche Methoden extrahiert: 3D-Kantenzuordnung und 3D-Segmentation. Um eine 3D-Kantenzuordnung durchzuführen, werden zuerst einzelne Kanten in den Epipolarbildern durch Anwendung des Cannyoperators detektiert. Nachdem kurze Kanten entfernt und gerade Linien extrahiert wurden, können die am besten korrespondierenden Linien mittels verschiedener Geometrie- und Lageeinschränkungen gefunden werden. Schließlich werden Topologiebeziehungen der abgeleiteten 3D-Kanten genutzt, um die Gebäudegrundrisse mittels einer Quaderanpassung zu rekonstruieren. In der vorgestellten 3D Segmentationsmethode wird der Startbereich durch die Nutzung eines Schwellwertes von 2,5 Metern für das nDOM definiert. Die Segmentierung ist ein iterativer Prozess und wird wiederholt, bis keine Verbesserung der Gebäudegrenzen mehr erreicht wird. Schlussendlich werden die Segmente in eine rechteckige Form gebracht, um die Gebäudegrundrisse mit den Neubauten zu aktualisieren.

In der Testphase der Methode wurden IKONOS und GeoEye-1 Stereobilder von Riad, der Hauptstadt von Saudi-Arabien, verwendet. Die hier verwendete Methode eignet sich für einfache rechteckige Gebäude mit flachen Dächern, wie im Testgebiet vorhanden. Eine Evaluierung der Ergebnisse mittels Vergleichen zu Referenzdaten hat gute Ergebnisse gezeigt. Die Rauschunterdrückung verringert die Anzahl der Fehlalarme um etwa 32 %, die Anzahl der aufgedeckten Änderungen wurde um ca. 15 % erhöht. Eine quantitative Evaluierung der Ergebnisse hat gezeigt, dass im ersten und zweiten Szenario die geringere GSD der IKONOS Daten einen signifikanten Einfluss auf die Häufigkeit von Fehlalarmen hat. Unsere Tests zur 3D Kantenzuordnung liefern nur gute Ergebnisse, wenn die Gebäude groß genug sind, eine einfache Form haben und große Kontrastunterschiede gegenüber dem Hintergrund besitzen. Ist dies nicht der Fall, sind die Ergebnisse fehlerhaft und unvollständig. Im Vergleich dazu liefert die 3D Segmentierung der Gebäudedächer bessere Ergebnisse im Hinblick auf die Vollständigkeit, allerdings versagt die Methode wenn ein Gebäude einen sehr geringen Kontrast gegenüber dem Hintergrund aufweist.

Keywords: *Gebäude, Aktualisierung, stereoskopisch, Bildzuordnung, GIS-Datenbank, hochauflösend*

Table of Contents

List of Figures

List of Tables

List of Abbreviations

CoG	Centre of Gravity
D-DSMs	Difference of Digital Surface Models
DP	Dynamic Programming
DSM	Digital Surface Model
DEM	Digital Elevation Model
FN	False Negative
FOV	Field of View
FP	False Positive
GCP	Ground Control Point
GSD	Ground Sampling Distance
GT	Ground Truth
HRSI	High Resolution Satellite Imagery
HT	Hough Transform
IFOV	Instantaneous Field of View
LRC	Left-to-Right Consistency
LiDAR	Light Detection and Ranging
LoD	Level of Detail
MI	Mutual Information
MLC	Maximum Likelihood Classification
MAD	Multivariate Alteration Detection
NDVI	Normalized Difference Vegetation Index
nDSM	Normalized Digital Surface Model
PCA	Principal Component Analysis
RANSAC	RANdom SAmple Consensus
ROI	Region of Interest
RPCs	Rational Polynomial Coefficients
SGM	Semi-global Matching
TN	True Negative
TP	True Positive

CHAPTER 1

INTRODUCTION

1.1. Motivation

Along with the development of Earth observation, the image-based updating of geospatial databases is being developed. With respect to man-made objects in urban areas, building detection, extraction and reconstruction are still challenging issues in the field of photogrammetric computer vision. A large number of algorithms have been developed to extract building footprints in dense urban areas at different level of detail (LoD). In most studies, the final goal is to establish an automatic procedure for updating of geospatial databases as well as 3D city modelling.

Meanwhile, the development of very high resolution spaceborne sensors using the capability of stereo imaging provides an improved opportunity for Digital Surface Model (DSM) reconstruction based on image matching. Such a development in sensor technology and algorithms provides a good opportunity to update GIS databases in developing countries where other alternative technologies such as airborne laser scanning are not a viable solution mainly due to economic reasons. For automatic 3D object extraction, several algorithms have been developed, both semi and fully automatic methods. The quality of the obtained results quantified by statistical indices (i.e. completeness, correctness, geometric Root Mean Square Error (RMSE) etc.) has been the focus of research for many years.

Prior to dealing with the problem, it is required to define some terms; *building change detection*, *building extraction* and *building update*. In building change detection, the dataset (vector or raster) in two different epochs are compared against each other in order to detect new construction, demolition or partial change of built-up areas (also called *building monitoring*). In this phase, a test is carried out whether there has been any change in the built-up areas. *Building extraction* is the next phase which in this thesis mainly refers to the extracting of building footprints. The *building update* indicates a processing chain that includes both of the previous phases consecutively; first detects the changes between an old database and a new image/DSM dataset and secondly adapts these changes, demolition, new construction or change of an existing building (e.g. vertical development). This step is known as data capture. Finally the adapted objects should be validated if the new state has a logical consistency (Heipke, *et al.,* 2008).

The updating of building databases through remotely sensing imagery is not only a key point for urban planning, sustainable development and 3D city modelling but can also play an important role in many other applications such as post disaster damage assessments and the estimation of affected population. For example, after a disaster such as an earthquake or tsunami, the government agencies usually need an estimation of building damages to provide related relief and rescue to the affected people as well as planning for reconstruction.

Additionally, in developing countries, such a need has significantly gained importance particularly, in the very densely populated areas in the suburbs of metropolises where, there might be a large number of illegal constructions. It is very important for municipal service offices as well as the infrastructure departments to monitor the rapid changes and control the urban sprawl, particularly illegal construction, consequently the population growth in the suburbs of metropolises to plan for the associated amenities around these developing areas.

1.2. Problem description

As mentioned, the automatic image-based updating of spatial databases is a cost-effective method for different urban objects such as buildings, roads, etc. However, building update is a challenging task in dense urban areas. The automatic updating of buildings in 3D space is the main objective of this study which is carried out by means of satellite stereo imagery. Compared to single images, stereoscopic vision provides a significant advantage to estimate the object height, an appropriate possibility to monitor building changes and is a tool for updating of the building layer within a geospatial databases.

A large number of automatic methods have been developed to extract 2D man-made objects in the urban areas (e.g. roads, roundabouts, fields, etc.) using very high resolution space or aerial images. However, with respect to 3D nature of buildings, it brings with it more challenges compared to the extraction of 2D objects. The generation of height information using HRSI is a cost-effective method for topographic mapping in the urban areas however; it faces with many problems such as limited resolution, shadow, occlusion, etc. The questions arising are: 1) whether spaceborne stereo images deliver satisfying results for building detection, extraction and updating. 2) If so which additional information can be integrated with the stereo images to overcome the associated challenges?

All issues discussed above cause that the development of automatic algorithms for buildings extraction using high resolution satellite images (HRSI) remain as a challenging research topic. While much research has been carried out in the field of building change detection using DSM stemming from laser scanning point cloud, there are only few research works for updating building geospatial databases using space images (Matikainen, *et al.,* 2010), (Chaabouni-Chouayakh, *et al.,* 2010), (Rutzinger, *et al.,* 2010), (Rottensteiner, 2007). It is probably because of the coarse Ground Sampling Distance (GSD) of satellite imageries for building extraction compared to that for very high resolution airborne images. In addition, DSMs derived from image matching are comparatively noisier than those derived from ALS data with the same sampling distance (Gehrke, *et al.,* 2010). Thus, the shape of building footprints in the derived DSM from image matching is not distinct and requires further refinement using additional information. The next problematic point for image matching is areas with low radiometric reflection

(e.g. asphalt roads) and homogeneous texture such as bright or dark rooftops. In these areas, it is difficult to find the corresponding match points. This problem is sometimes called *mismatched areas.* Consequently, the height of such areas can potentially cause a number of false alarms in the change detection phase.

The task is further complicated if trees are adjacent to buildings where building cues in the DSM are mixed with trees, leading to an inaccurate extraction of building footprints. The prevalent solution for this problem could be the separation of vegetation areas using spectral reflection by means of a relevant index (e.g. Normalized Difference Vegetation Index - NDVI). However, if the imagery has no infrared channel, this solution is not possible (Rottensteiner, 2007).

The last challenging point is the level of detail in the verification and updating of building databases, especially when spaceborne image information is used to update the building polygons of a GIS. Due to differences in data source, definition of built-up area, etc., the building outlines in vector and raster dataset may not match to the required accuracy.

Figure 1 shows two samples of building changes in two different epochs (IKONOS 2008 and GeoEye-1 2009) along with an outdated building polygon from the GIS database. First, the difference becomes visually clear when the resolution of two images is compared. Second, it shows how trees adjacent to building may trouble building footprint extraction and the subsequent updating procedure. Third point is the difference of buildings outlines between the original GeoEye-1 image and superimposed GIS polygons. In some cases, the courtyard is categorized as building areas although they are 2D objects (i.e. the difference between building and built-up areas).
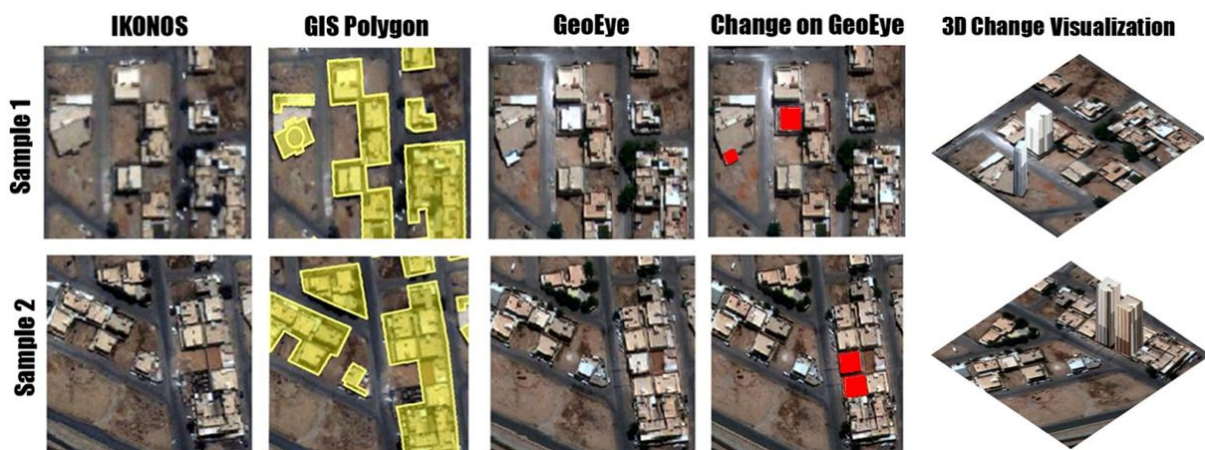
Figure 1- Sample of building changes in two different epochs (IKONOS with GSD 1m acquired in May 2008 and GeoEye-1with GSD 0.5m acquired in Sept. 2009). An outdated building GIS database (yellow polygons) is superimposed to the GeoEye-1 image

1.3. The research objectives

Figure 2 shows the general framework of research to update building databases using satellite stereo images and an outdated GIS database. In fact, buildings and trees are particular structures in urban areas with the specific property of height. Therefore, the height is a unique characteristic to distinguish

these objects from other urban structures. Besides, a visual inspection between the first and second epoch reveals that although change in intensity occurs in many places, it may not be a building or a 3D change. Thus, to establish an alarm system for building changes, the height information is comparatively a more reliable factor than image intensity. However, the smoothing effect in the DSM derived from image matching is a decisive side-effect which should be taken into account for the development of a sound framework to detect building change. To tackle mentioned problems appropriately, it is required to apply a well-developed approach which incorporates height, image and auxiliary information (GIS).

In order to handle the complex issue of building update and obtain reliable signals for building change as well as obtaining the accurate shapes of buildings, a new algorithm is proposed based on height information from image matching and radiometric image information. Image information is used for building change detection and image information is used for the delineation of building outlines and updates the associated building footprints in the GIS. Moreover, to eliminate the side-effects of vegetation and mismatched areas (surface discontinuities e.g. building outlines, asphalt, and tree tops etc.) several noise removal masks are introduced. It is our hypothesis that by applying these masks along with prior knowledge of building properties, the derived DSM from spaceborne stereo matching can be considered a reliable source to detect building changes. Although, removing the entire shadow and occlusion effects in the DSM is not the main objective of this study.
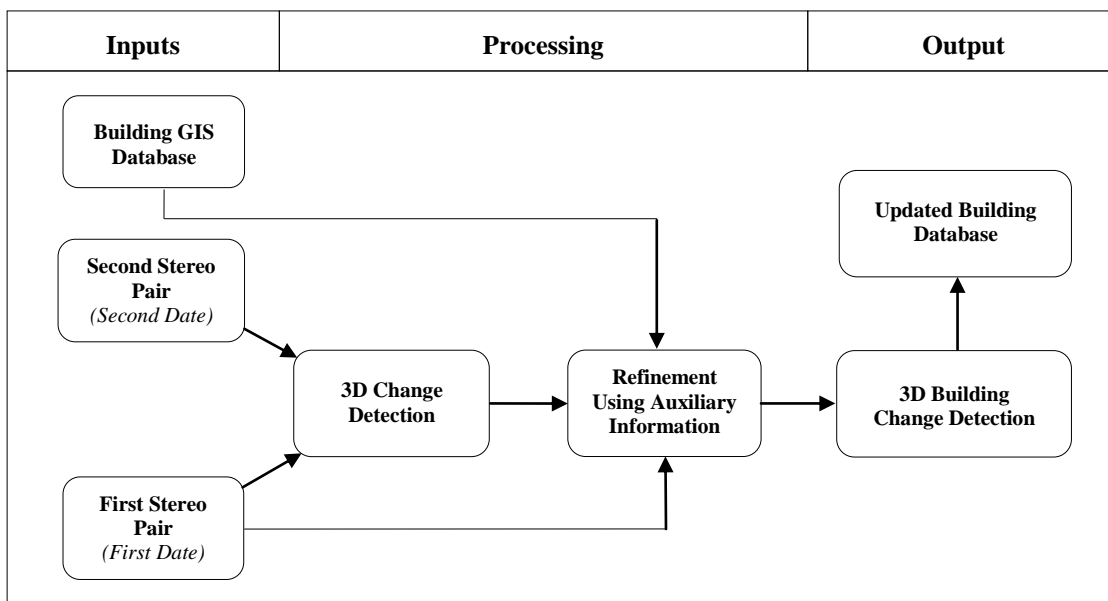
Figure 2- The general framework for the update of building databases using stereo images and a GIS database

After obtaining a change alarm, the next step is to extract the footprint for new buildings. For this task, two different approaches are introduced, namely line-based and region-based. In both approaches, image information is used to delineate building outlines. In the first approach, a 3D edge matching technique is used while for the second approach, we delineate building rooftops using segmentation method based on level set. The derived building outlines are then used to reconstruct building

footprints. The associated height from the nDSM also updates GIS databases. Note that the proposed method is only applicable for simple rectangular buildings with a flat rooftop, like those which can be found in our study area, the suburb of Riyadh the capital city of Saudi Arabia. The main contribution of this research is to develop a building update model by incorporating 2.5D height information (from DSM) with original image information (from stereo images) assisted by auxiliary information.

1.4. Organization of the thesis

This thesis is organized in five chapters. In the second chapter, a short review of literature is reported followed by a discussion and conclusion concerning the pros and cons of the previously developed algorithms. In chapter 3, the theoretical fundaments for image orientation, the geometry of pushbroom space image, image matching, and different approaches for building change detection are discussed. Chapter 4 illustrates the main contribution of this research by proposing a new framework for updating building databases based on the integration of height and image information.

Chapter 5 contains the experimental results of proposed methodology tested on IKONOS and GeoEye-1 stereo pairs. The accuracy assessment is conducted by comparison of the obtained results against reference data which are digitized manually. The quality of the obtained results is evaluated quantitatively using a statistical approach. Final conclusion and recommendations for future research are presented in the chapter 6.

CHAPTER 2

STATE OF THE ART

2.1. Introduction

Nowadays rapid changes are taking place in residential areas, and in particular within developing urban areas, thus updating geospatial databases is an important target for many government agencies in order to develop an automatic solution for this aim. Such an automatic procedure is quite useful when monitoring and updating building changes in a given period of time. The detection of changes using remotely sensed data can act as an alarm system to determine whether or not the updating of geospatial databases is required, in which case the precise outlines of added/removed objects should be delineated.

In this chapter, the task of building update is divided into independent phases; *detecting building changes* and *delineating change outlines* in order to update them on GIS databases. There have been a wide range of algorithms to extract and update building databases over the last 40 years. In this chapter, however, a short review of literature is reported with a discussion at the end to justify the main strengths and weaknesses of each method.

2.2. Related work

This section is divided into two subsections; the first part provides an overview on different approaches in 3D building change detection. The second part focuses on studies concerning the delineation of building footprints using stereoscopic images. For updating building databases using space-born stereo images, in general, the original radiometric image and the associated DSM derived from the image matching are the main input data sources. In this study, these data sources are referred to "*original image*" and "*matching-based DSM*", respectively.

First, previous literature on building change detection and building footprint extraction using spaceborne stereo imageries are introduced. However, in some cases, previous studies using laser scanning points cloud or aerial images are also considered in order to evaluate whether it is possible to extend these algorithms into images with lower resolution (e.g. spaceborne stereo images) by proposing a generalized data model.

In the final discussion, existing methods of building change detection as well as building footprints delineation are discussed along with an assessment of previous studies. Summarising the potential advantages and disadvantages of each method provides a short analytic justification of the literature in addition to demonstrating possible solutions to overcome the challenging points either for change detection phase or for building footprints extraction.

2.2.1. Building change detection review

Monitoring urban sprawl gained importance in urban and regional planning. Remote sensing data, and particularly high resolution images with such characteristics as broad cover in addition to fast and relatively cheap data acquisition, provides a suitable option for tracking changes in urban areas. The captured images using remotely sensed sensors are an appropriate tool to detect changes over dense residential areas (Zhang, *et al.,* 2006). Importantly, regarding the potential of stereo images to possibly generate DSM, the subtraction of these height models from each other within a given period of time can result in building change detection.

A wide range of image processing and computer vision techniques, including spectral indices (e.g. NDVI), geometric (e.g., shape and size) and height information (e.g. DSM, differential DSMs) have been developed for change detection based on remotely sensed images in general (Im, *et al.,* 2005), (Chaabouni-Chouayakh, *et al.,* 2011), (Bouziani, *et al.,* 2010), (Champion, 2007). There are two main strategies which are considered in change detection algorithms using remote sensing data: *change enhancement,* whereby the emphasis is on the image difference without any information addressing the type of change (i.e., changed and un-changed pixels). The second approach is known as *''from–to''* strategy that monitors the land use changes during a period of time by determining the change from land use A to land use B (Im, *et al.,* 2005).

In change detection using remote sensing data and a GIS database, a change map can be computed based on three different types of comparison; (i) map-to-map comparison, which updates an old GIS polygon using an updated one, (ii) image-to-image comparison at two different epochs and finally (iii) image-to-map comparison that investigates the updating of an outdated vector database (i.e. GIS) using an updated image information (raster). For the last case (image-to-map), which is also the most common updating procedure, image segmentation/classification is the most prevalent method. Using segmentation/classification, pixels within a digital image are grouped into separate segments/classes based on their reflections presenting a unique object. Then these segments can be compared against a vector dataset (image-to-map). Malpica, *et al.,* (2010) developed an approach for urban change detection by integrating multi-spectral satellite imagery, LiDAR point clouds and a GIS database. A SVM (support vector machine) was used to classify the image, resulting in a probability layer for buildings. By intersecting the classification result with the GIS building layer, the authors found an increase in the built up area of a few percent.

Durieux, *et al*. (2008) proposed an object-based method for the monitoring of urban sprawl using a region-based segmentation applied to Spot 5 supermode images with 2.5 m GSD. An object-based image analysis (OBIA) approach is employed to extract building footprints and detect building

changes. A hierarchical approach is utilized to classify objects using multi-resolution segmentation by means of physical properties, semantic relationships and neighbourhood relationships of the network's objects. Image objects are then segmented into two classes using multi-resolution segmentation, i.e. from coarse to finer levels. Similar objects are then combined by means of fuzzy-logic operators such as and/or. Tests on Spot-5 images have shown that image resolution impeded accurate detection of building rooftops. Thus, contextual and scale as additional information are employed to support object-based segmentation. Consequently, with Spot-5 images, this building extraction methodology has shown severe limitations at regional scales to be used as a reference for mapping and updating purposes in urban areas. In particular, it exhibits more difficulties representing building shapes.

Vegetation and mainly trees are the main challenges for building change detection as well as footprint extraction. This problem can be more complex if the data acquisition time at two epochs includes a large difference in canopy cover. Rutzinger *et al*. (2010) have introduced an object-based workflow for 3D building change detection in a short time interval. The changes are classified as new construction, demolition and un-changed buildings. They have introduced a few thresholds to measure how unchanged objects can be defined and which level of height difference can indicate real building changes. However, most false alarms are mainly due to different data acquisition times causing different canopy covers between the first epoch (summer) and the second epoch (autumn). In summer, the canopy cover of trees is almost maximal, but, at the second epoch in autumn, it is rather sparse. Consequently, subtraction of DSMs from each other indicates few virtual demolitions which are false alarms because of vegetation. Awrangjeb *et al*. (2012) proposed an effective method for the separation of buildings from trees by employing height, shape, size, color and texture information. They introduced a rule-based approach based on the main orientation of building boundaries in order to recognize building blobs from trees even in some seasons in which leaves' colors are changed or entirely lost.

Zhu *et al*. (2008) also presented a new object-based approach detecting building changes in a suburb of Beijing. This method works based on polygon automatic validation over IKONOS and Quickbird images. The urban objects are generalized using morphological filtering. Then, radiometric information along with texture and neighbourhood analysis is used to propose a new polygon automatic validation method. The variance and average gradient indices demonstrate the difference between texture and neighbourhood information, respectively. Finally, to detect changes between images at two different dates, correlation coefficients of gray value, variance, and average gradient indices have been used. Based on the thresholds defined using training areas, the polygon of each object is validated if its indices exceed the given value. Otherwise, the associated polygon will be automatically eliminated and is referred to as a demolished object.

Knudsen, *et al*. (2003) and Olsen (2004) presented a framework to validate map databases with an emphasis on the building layer using an efficient automatic approach for change detection. They applied classification on aerial high resolution images which was refined by a height filtering in order to validate building polygons over the Danish national map database (TOP10DK). The building layer in the TOP10DK databases is employed as training signatures for the classification to characterize the

different classes and then build up an object model for buildings detection. As the used image had very high spatial resolution, it could demonstrate the complex structures of buildings very well. The height information (LiDAR nDSM) along with the size and shape information supported the extraction of buildings polygons precisely. The classification results were also compared with the existing map to generate the change map using height and size information. Furthermore, noise and misclassification areas are removed from the final change map. The results obviously demonstrated that classification based on spectral signatures alone is not able to extract the complex structure of buildings. This is even more problematic, if the spectral reflections of new buildings are different from the existing buildings in the map database. Therefore, the proposed method is not able to detect building changes if new building shows spectral reflections which are not similar to that of the training sites.

Champion, *et al.* (2010) matched high resolution aerial images to generate a DSM and then updated the building layer of a 2D cadastral database using the derived height information. They subdivided the detection of changes into two main steps, namely the automatic verification of the database and the detection of new buildings. In the first step, all buildings in the database are investigated to check whether they match with extracted objects from images. To quantify the comparison, a similarity measurement is defined using robust geometric criteria. In the second step, a digital elevation model (DEM) is automatically generated from the DSM. The difference of DSM and DEM demonstrated an above-ground mask. From this mask, building blobs are generated and compared with the existing database in order to detect newly constructed buildings. According to the authors, the actual delineation of building outlines is not very accurate, mainly due to shadows.

Le Bris *et al.* (2011) proposed an automatic procedure as an alarm system for 2D change detection. It is used to update building polygons in the GIS topographic database using GeoEye-1 stereo images. The authors used two different approaches, a supervised classification and the subtraction of inter-date DSMs. In the first approach, additional information such as GIS database and DSM-DEM correlation are used to increase the accuracy of classification. It is concluded that for building change detection, this additional information is required particularly, for the detection of complex structure buildings. It is rather difficult to distinguish the urban objects only using radiometric image information. In the second approach, the subtraction of two DSM derived from stereo matching – called *differential DSMs* – referred to the vertical changes. In order to remove false alarms, a vegetation mask is used to eliminate non-buildings areas from the obtained change map. This removal mask significantly increases the correctness of the proposed alarm system. Finally, the small elements in the change map (e.g. sliver polygons) are eliminated and the final alarms are extracted by the incorporation of the results from both approaches.

Doxani, *et al.,* (2010) proposed an object-based classification for the automatic monitoring of buildings changes using multi-temporal images. The Multivariate Alternation Detection (MAD) algorithm and morphological scale space filtering are applied on multi-temporal datasets to overcome the problems observed in the traditional change detection models (e.g. image algebra, image transformation techniques and post-classification analysis). The proposed algorithm was tested on two pan-sharpened Quickbird orthophotos with the ground resolution of 0.6m in two different epochs;

2003 and 2007, respectively. Using MAD components, all pixels are classified into two classes; changed and un-changed pixels however, without addressing the change types. In the next stage, Expectation-Maximization (EM) algorithm is implemented to determine the trend of MAD component between two epochs whether the gray value represents an ascending trend (MAD+) or a descending one (MAD-) and then changed pixels are classified into tiled, bright and dark roofs. In order to simplify the classification procedure in complex urban areas, the scale space filtering was utilized on original multi-temporal images in order to segment urban objects in different resolutions.

Tian *et al*. (2010) used stereoscopic satellite images in order to detect height changes by the computing the difference between DSMs generated in the different epochs. A rectangle was fitted to each extracted blob assumed to be a building footprint. However, most blobs are highly curved, so that the direction of the rectangle edges cannot be computed reliably. Although the proposed method covers the lack of height information in 3D buildings change detection, there is still no solution for the smoothing effects for the above-ground objects (e.g. buildings). The smoothing effects is stemmed from various grounds that cause mismatch (e.g. shadow, occlusion, illumination, etc.). Consequently it causes that the building blobs in the DSM are enlarged (Tian J. , *et al.,* 2010).

2.2.2. Building extraction review

Recently, there has been a significant advancement in new generation high resolution commercial imaging sensors (e.g. IKONOS, GeoEye-1, Quickbird, and Worldview) that are able to provide high quality stereoscopic images. These sensors are well adapted to deal with the 3D imaging in the dense built-up residential areas. Nevertheless, spaceborne images have a relatively low ground sampling distance as well as a small number of overlapping images compared to aerial images.

Building extraction using spaceborne stereo images is usually not studied beyond LoD-1 because detail of buildings cannot be extracted from satellite images. Furthermore, most of proposed algorithms deliver promising results providing that additional rule-based filtering are applied (Lafarge, *et al.,* 2008), (Le Bris, *et al.,* 2011), (Doxani, *et al.,* 2010), (Chaabouni-Chouayakh, *et al.,* 2010), (Champion, *et al.,* 2010), (Matikainen, *et al.,* 2010). A large number of automatic image analysis methods have been developed to extract building footprints from very high resolution stereo images. Most of developed methods utilize the geometric, radiometric, topologic and topographic attributes in the different levels for building change detection as well as building footprints extraction.
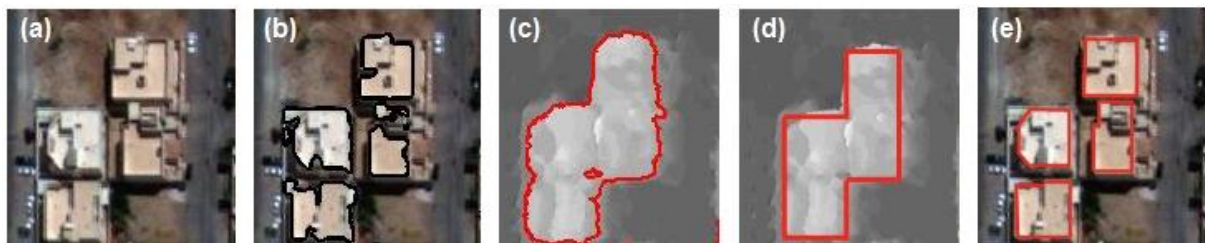
Figure 3- Comparison of different algorithms for the extraction of building footprints (a) original GeoEye-1 image, (b) building extraction using segmentation of building rooftops, (c) the segmentation of DSM derived from image matching, (d) extraction of rectangular shapes using box-fitting over nDSM, (e) ground truth generated manually

Figure 3 shows several prevalent methods for building extraction using image and height information. For building extraction, three different types of algorithms are developed according to the spatial information used. Building extraction may be carried out using *height information* (e.g. DSM) [e.g. (Lafarge, *et al.,* 2008) & (Tian, *et al.,* 2011)], *image radiometric information* [e.g. (Jin, *et al.,* 2005) & (Sirmacek, *et al.,* 2010)] or by the integration of height and image information (hybrid approach) [e.g. (Vallet, *et al.,* 2011)].

As shown by Figure 3, building blobs in the DSM have blurred outlines. This is mainly due to well-known matching problems because of occluded areas and different shadows stemming from the differences in viewing and illumination directions as well as vegetation effects that are accumulated in building blobs. It can be also due to the smoothing effect caused by the interpolation of unmatched pixels. Thus the buildings size in the DSM is enlarged (Alobeid, *et al.,* 2010), (Le Bris, *et al.,* 2011), see also Figure 4. Moreover, the artefacts stemming from image matching cause a number of voids in DSM. As a consequence, it makes several virtual holes inside of building footprints wrongly (see Figure 3c). Hence, these regions should also be filled before fitting a rectangle over building footprints.

The height information delivers promising signal to locate building footprints however, only as a blob. That means in most cases, buildings are substantially recognizable in the DSM but as blurred building blobs. Comparing Figure 3(c) and Figure 3(e) demonstrate the well-known smoothing effects in the DSM derived from image matching (called matching-based DSM) which join individual building blobs by filling out the open space between these buildings resulting larger blobs.

In the following, a brief review of building extraction using spaceborne sensors is reported. In some cases, the algorithms that have been developed for aerial images or for LiDAR points cloud are also explained in order to investigate whether it is possible to generalize these algorithms for spaceborne images. For both types of dataset (image and height) as well as the different methodologies, the pros and cons are discussed.

To extract building footprints precisely two approaches are considered in this study. (i) Approaches that are based on the extraction of building outlines (edge-based approaches). For instance, the extraction of building outlines using 3D edge matching and then building reconstruction according to the polyhedral model of building rooftops. Baillard, *et al.,* (1999) employed a wide range of geometric (epipolar, line orientation, etc.) photometric and topological constraints to find the corresponding lines on very high resolution aerial images. Then planar patches are generated by the grouping coplanar 3D lines. (ii) Approaches that extract building footprints based on the segmentation/classification of building rooftops (region-based approaches). The segmentation can be carried out on the image information or associated DSM to detect individual building footprints. Figure 3(b) and 4(c) show two examples from the segmentation of planar rooftops. Chehata, *et al.,* (2005) utilized a graph cut approach to extract 3D primitives of the roofs on spaceborne images with a submetric resolution (45-70cm). The obtained 3D segments provide a hybrid approach to generate a digital elevation model with a higher level of detail particularly, with a better demonstration of building outlines in the DSM.
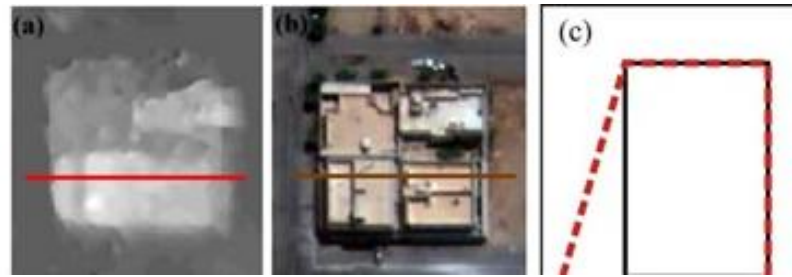
Figure 4- (a) Sample matching-based DSM, (b) corresponding image along with a cross section, (c) the schematic cross section from the side (red: DSM, black: original object)

Different approaches are tested to decrease the smoothing effects as shown by Figure 4. A solution for this problem is the extraction of individual building footprints based upon segmentation or delineating building footprints using building outlines. The segmentation can be carried out using the original points cloud or the radiometric image information. This approach extracts building parts with a higher level of detail (LoD). However, if the building extraction is carried out using segmentation of points cloud or DSM, a good sampling distance is also required.

Vosselman (1999) tackled the building reconstruction based on the clustering of laser scanning points cloud. Building boundaries are extracted based on the Delaunay triangulation using original points cloud and then a threshold is set to detect the connected components based upon planar size. The proposed method can preserve the building shapes very well because there is no smoothing effect due to interpolation. However, like other data-driven approach; it fails to reconstruct the small annexes of building rooftops such as dormers and chimneys. In addition, it delivers promising results on very high resolution airborne data however, utilising the same methodology over spaceborne matching-based DSM with a lower GSD may not be expected.

Poli *et al*. (2012) proposed a hybrid method in order to enhance the DSM derived from stereo Cartosat-1 images. A hierarchical segmentation on WorldView-1 image is employed to enhance the object outlines in the DSM. Results are evaluated against a DSM extracted from ALS points cloud as reference data. Using hierarchical segmentation, the adjacent pixels are partitioned based on the connectivity constrains. If the dissimilarity of a pixel with respect to the surrounding pixels exceeds a given threshold, it belongs to a new segment. The dissimilarity is defined based on the pixel intensity. By increasing the threshold value, the hierarchical connected components are arranged from a course to fine level. After segmentation of WorldView image, all segments are intersected with the DSM and then the mean height value for each segment is assigned to associated pixels in the DSM. In fact, by this approach, the DSM is segmented using radiometric image information while the height value is obtained from original DSM. Although the proposed algorithm enhanced the object outlines in the DSM, it fails to detect an object if there is a good contrast between the target object and the background.

To extract rectangular buildings footprints from DSM, Lafarge, et al. (2008) investigated a marked point process to obtain the general shapes of building footprints in the DSM derived from PLEIADES sub-metric satellite images. This object-based method fits a rectangle on buildings in order to simplify

complex buildings structures into the primitive components. To do so, by using the global energy minimization, the best rectangle is fitted to the building footprints based on the Markov Chain Monte Carlo. The basic concept is that there is usually a sharp discontinuity - with a considerable height difference - between building outlines and the neighbourhood pixels (known as point of interest). The optimization of internal and external energy resulted in fitting the best rectangle over building footprints. Finally, the topological relationships between the connected components are utilized to generate the final regularized buildings footprints.

Fraser, *et al*. (2002) used 1-m IKONOS panchromatic satellite imagery to reconstruct rectangular shapes of building footprint in LoD-1. The image matching is performed by using multi-image datasets from three different viewing angles (backward, nadir, forward). Image matching carried out using least squares adjustment. The results have shown a sub-meter accuracy, the planimetric and height accuracy are reported between 0.3-0.6m and 0.5-0.9, respectively. The experimental results have shown that, using IKONOS stereo images and under the ideal conditions for radiometric quality, image orientation with ground control/checkpoints, coupled with favourable imaging geometry (e.g. B/H> 0.8) and also provision of sensor calibration, the proposed method can be used for building reconstruction. However, the authors did not discuss the quality of extracted building footprints particularly the quality of the building shapes. This is of course an important issue particularly, when the DSM is generated based on the least squares matching because it demonstrated a high level of objects smoothing in the DSM and subsequently, building shapes are deformed.

Jin, *et al*. (2005) also demonstrated a method based on various structural, contextual, and spectral information to extract buildings footprints using IKONOS image (1-meter panchromatic and 4-meters multispectral). The differential morphological profile (DMP) is used to generate building hypotheses assisted by the shape information. Shadows and vegetation are extracted using the DMP and NDVI, respectively. Shadow as complementary information indicated the position and the size of corresponding buildings. The watershed segmentation is used to extract building cues based on the spectral information; those that were not detected using structural and contextual information. Finally, structural, contextual, and spectral information are integrated to extract the final building footprints. The detection rate and quality percentage are reported 72.7% and 58.8%, respectively.

Sirmacek, *et al*. (2010) presented an automatic method to enhance the DSM derived from image matching by utilizing a box-fitting approach. The general idea of the proposed method is energy minimization in order to find buildings outlines and then to reconstruct rectangular shapes for the building footprints. The authors used the Canny operator to extract building edges on the IKONOS panchromatic image. The box-fitting begins from the centre of each cue toward building outlines. This phase is carried out by testing the all probabilities for the orientation angle as well as the size of rectangle, iteratively. The best fit is found with respect to the cost function based on the energy minimization. Finally, the boundary of the fitted box is used to enhance the DSM. The main shortcoming for the proposed method is that the outcome depends on the Canny operator. Note that this method may face with another serious difficulty because in this method, there is no difference

between 2D and 3D edges. That means neither a 3D edge matching algorithm nor a DSM reduce the search space to find the real building outlines.

2.3. Discussion

A literature review for 3D building change detection and extraction of building footprints is reported. As stated, the change detection acts like an alarm system. However, it is also an important task to refine building outlines and update building footprints of the GIS database. Hence, while the first step detects the building changes, the second step functions as a complementary phase to delineate building outlines for new constructions. The subtraction of DSMs of different epochs delivers a reliable signal for building change detection so that the matching-based height information is utilized to address building changes. However, a DSM is not a sufficient source to delineate building outlines precisely because in such a DSM building footprints are usually enlarged, and building shapes are deformed in the matching-based DSM (Champion, *et al.,* 2010), (Alobeid, *et al.,* 2010), (Doxani, *et al.,* 2010).

After reviewing the literature for building udpates, the advantages and limitations of each method is summarized. In term of building change detection, it can be concluded that due to complex building structures in dense urban areas as well as several shortcomings of matching-based DSMs, the hybrid approaches are more capable than other methods to detect building change. These approaches usually incorporate image and height information as well as auxiliary information such as texture, shape, size, and topological information. It is also reported by (Le Bris, *et al.,* 2011), (Doxani, *et al.,* 2010) and (Lafarge, *et al.,* 2008) that the image context without additional height information is normally not able to detect buildings within a complex scene. Therefore, in this research, we propose a novel algorithm for 3D building change detection by incorporating image and height information, assisted by additional information (GIS database, building property, etc.) to improve the quality of the change detection.

For the extraction of building footprints, an overview of related works indicated that associated approaches are classified into two main groups; (i) reconstruction of polyhedral building model using extraction of building outlines (ii) segmentation of building rooftops. If the source of a DSM is from ALS points cloud, by the segmentation of DSM or points cloud, individual building footprints can be obtained. However, due to the smoothing effect of matching-based DSMs, this strategy does not deliver promising results. The first approach based on 3D edge matching could also successfully reconstruct bulding footprints using very high resolution aerial images. However, it is required to test the efficency of 3D edge matching using stereo space image because the image resolution is coarse than the resolution of aerial images. To detect building footprints at LoD-1 using spaceborne images, several studies were also carried out based on the fitting a rectangle on building outlines.

CHAPTER 3

THEORETICAL BACKGROUND

In the past decades, following with the development of remote sensing sensors, there have been significant developments in earth observation methods including the monitoring of natural, residential and industrial areas. Hence, the change detection based on remotely sensed data is being widespread used in order to update geospatial databases. In addition, change detection is also used at numerous applications such as video surveillance, traffic management, human tracking, tracking of moving objects, assessment of ecological resources, disaster management and particularly damage assessment, medical diagnosis and urban objects monitoring (Porter, *et al.,* 2009), (Radke, *et al.,* 2005).

Monitoring of changes using remote sensing data in general have been studied for few decades however, as discussed in the previous chapter, buildings have a specific property, the height which assists building change detection. Height reconstruction using stereoscopic imaging provides a reliable tool compared to radiometric image information because, change of intensity in the image domain cannot necessarily indicate building changes (as shown by Figure 3). However, as discussed the image information can act complementary for the height information to extract building footprints.

On the one hand, height characteristics facilitate change detection for 3D urban object (buildings, trees, etc.). That means the height is the main difference between buildings and the other two dimensional urban objects. Therefore, new generation spaceborne sensors which provide stereoscopic images with a suitable angle of convergence (height-to-base ratio) are useful to update building databases. On the other hand, this characteristic of building can act as a limitation for the detection and extraction phases. Shadows and occlusions are examples of limitations caused by three dimensional nature of buildings.

In this chapter, the main theoretical concepts of our processing chain including image matching, DSM generation and building change detection are explained. It contains the theoretical background of related techniques, methodologies, strategies as well as the challenging points to generate a DSM and subsequently to detect building changes by the DSMs subtraction.

3.1. Sensor-oriented rational polynomial coefficients (RPCs)

For pushbroom sensors, the image orientation is usually carried out through bias corrected rational polynomial coefficient (RPC). The image orientation using RPC is possible with or without ground control points (GCP). The geometric relationship between image and ground coordinates can be approximated by a 3D interpolation in object space with polynomials based on the sensor geometry and direct sensor orientation (Dowman, *et al.,* 2012). RPCs consist of a ratio of two polynomials of object space coordinates. As shown by equation 3.1, two separate rational functions are used to transfer the position from the object space into the image space.

$$x_{i,j} = \frac{P_{i1}(X,Y,Z)j}{P_{i2}(X,Y,Z)j} \qquad y_{i,j} = \frac{P_{i3}(X,Y,Z)j}{P_{i4}(X,Y,Z)j} \tag{3.1}$$

Where $x_{i,j}$ and $y_{i,j}$ are the normalized image coordinates; X, Y, Z are the normalized ground coordinates. Equation (3.2) demonstrates the polynomial function

$$P(X,Y,Z) = \sum_{i=0}^{n1} \sum_{i=0}^{n2} \sum_{i=0}^{n3} a_{i,j,k} X^i Y^j Z^k \tag{3.2}$$

In normal case, the order of the polynomials (n1, n2, n3) should be between 0 and 3. Moreover, the RPC equation's order - the summation $(n1 + n2 + n3)$ - should be equal or less than 3. Each P(X, Y, Z) is a third-order equation expressed by the 20-term polynomial as per below

$$P_{i1}(X,Y,Z)_j = a_1 + a_2.Y + a_3.X + a_4.Z + a_5.Y.X + a_6.Y.Z + a_7.X.Z + a_8.Y^2 + a_9.X^2 +$$
$$a_{10}.Z^2 + a_{11}.Y.X.Z + a_{12}.Y^3 + a_{13}.Y.X^2 + a_{14}.Y.Z^2 + a_{15}.Y^2.X + a_{16}.X3 + a_{17}.X.Z^2 +$$
$$a_{18}.Y^2.Z + a_{19}.X^2.Z + a_{20}.Z^3 \tag{3.3}$$

Where (X,Y,Z) denotes the longitude, latitude and the height in ground coordinates, respectively. The a_n are the coefficients $(a_1$-$a_{20})$ which are available in the metadata. The polynomials P_{i2}, P_{i3} and P_{i4} have the similar form as P_{i1}, making a total number of 80 coefficients (Kim, 2000), (Grodecki, *et al.,* 2001), (Dowman, *et al.,* 2012). The RPCs values are computed based on the direct sensor orientation using the sensor position which is captured by a global navigation system (i.e. GPS) and attitude information by the star sensors and gyros, together with a calibration system. Similar to other image orientation methods, the use of RPCs requires a large number of 3D corresponding points in object and image coordinates, being well distributed throughout the scene (Dowman, *et al.,* 2012).

3.2. Epipolar geometry of linear array scanners

To perform image matching using global image matching methods (e.g. dynamic programming and semi-global matching), it is required to reduce the search space to one dimension. This step is usually carried out by the generation of epipolar images. Firstly, it is an essential step for each global image matching methods. Secondly, the epipolar images reduce the search space into the x-direction speeding up image matching procedure. By generation of epipolar images, the number of incorrect

matches is also reduced, considerably. In our scheme, the generation of epipolar images is required for the building outlines delineation through 3D matching and 3D segmentation.

In epipolar images, the corresponding pixels have identical y-coordinates so that the parallax in the x-coordinates is directly related to the object height, the larger the parallax, the higher the object is. Unlike frame cameras with perspective geometry, high resolution spaceborne sensors use mainly pushbroom technology supported by the CCD line sensors, scanning the earth line by line.

The fundamental characteristic of pushbroom sensors is that there are separate projection centres for each scan line whereas the frame sensors have only one central perspective for each frames (see Figure 5). As a consequence, the epipolar lines of these sensors are not straight as in the traditional frame cameras, but are in form of a curve (Oh, 2011). However, in space images, these curves can be approximated to a straight line due to the smooth satellite trajectory.

Topographical variation of landscape and image acquisition out of nadir location also cause associated geometry distortions. In Geo or Ortho-ready Standard images, it is assumed that using a simple linear transformation, the object space can be transformed into the image space providing that the study area is flat and the image is acquired almost from nadir. In such a case, the transformation of images into the epipolar geometry can be considered as a 2D problem. Therefore, the rotation of images around the base direction can reconstruct the epipolar geometry resulting in the *quasi-epipolar* images (Jacobsen, 2011). Terrain undulation may cause a shift in the image space compared to the correct X-Y position. These displacements depend on the viewing angle as well as the amount of height difference within the scene.
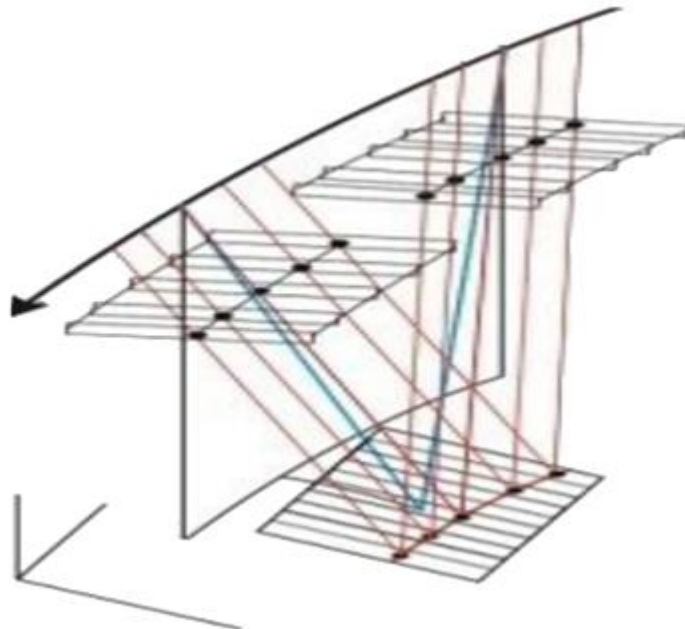
Figure 5- Geometric reconstruction of epipolar lines in the pushbroom
sensors (Dowman, *et al.,* 2012)

3.3. Challenges for image matching

Image matching is a procedure to find the corresponding points within a stereo pair automatically. It is a key issue for the topographic mapping using remote sensing data. Disparity map refers to the apparent pixel difference or the parallax between corresponding pixels within a stereo pair being in the epipolar direction. They encode the relative height of a pixel in relation with the other pixels by representing the parallax which is caused by object height so that the larger the parallax, the higher the object is. The disparity map can be translated into a DSM using the image orientation parameters.

There are two main methodologies for image matching; window-based methods which find the corresponding pixels between stereo pairs using a template known as *local methods*. The second approach is the pixel-based approach which is also called *global methods*. The drawback of the first group is that, usually they smoothen the sharp edges (i.e. building outlines) in the height model. This problem is further intensified in case of the occlusion, shadow and poor textured areas such as asphalt roads. In contrast, global methods minimize the energy function to find the best corresponding pixels. These methods are much more robust against illumination directions as well as shadow and occlusion.

There are a number of error sources influencing the results of image matching. Occlusion, shadow, trees, homogeneous areas, viewing and illumination directions are the most important factors causing matching errors. Figure 6 shows shadow and occluded areas within a stereo pair GeoEye-1. As shown the occluded areas depend upon the viewing angle, A and B in Figure 6 show the associated occlusion caused by two different viewing angles. The extent of these areas has a reverse relationship with the incidence angle.
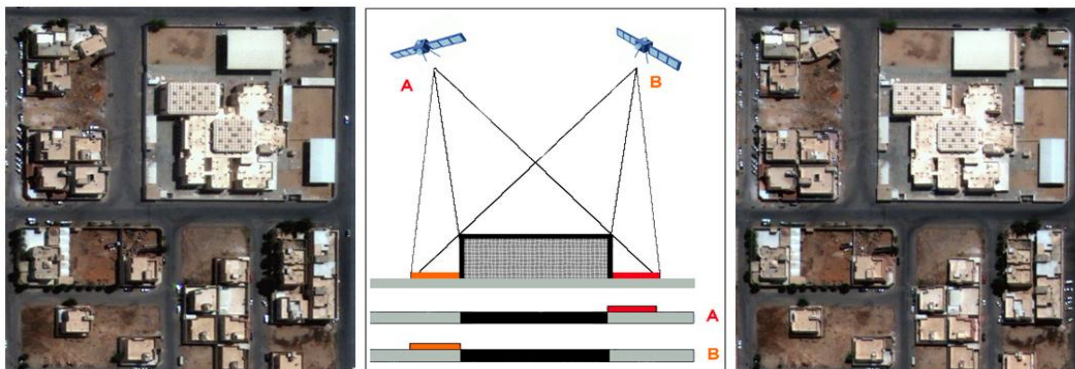
Figure 6- Location of occluded areas in one stereo pair with respect to the imaging directions

A number of problems are reported for the DSM generation using image matching techniques. Narrow streets between buildings - which are invisible in one of the stereo images - are an example for mismatched areas due to occlusion. The accuracy of the derived DSM from image matching depends upon two factors; GSD of the sensor and the geometric imaging configuration, dominated by the height-to-base ratio (Dowman, *et al.,* 2012), (Alobeid, 2011).

In automatic image matching, the precision of the height determination is a key issue. Equation (3.4) demonstrates the height precession derived from the image matching.

$$\sigma_Z = \frac{H}{f} \cdot \frac{H}{B} \sigma_{px} \tag{3.4}$$

Where σ_Z is the standard deviation of the object height, H is the height above ground, *B* indicates the image base (distance of projection centres), *f* is the focal length of camera and σ_{px} is the standard deviation of the x-parallax. As ratio (*H/f*) is the scale factor, the equation (3.4) can be expressed as per below

$$\sigma_Z = \frac{H}{B} \cdot S \cdot \sigma_{px} \tag{3.5}$$

For digital images ($S. \sigma_{px}$) can be expressed as a fraction of the GSD, leading to the equation 3.6

$$\sigma_Z = \frac{H}{B} \cdot a \cdot GSD \tag{3.6}$$

The factor *a* depends on various parameters; the land cover of the area (i.e. open area, forest, urban etc.), the contrast of the image, the matching method as well as the height-to-based ratio (Dowman, *et al.,* 2012).

Image matching is a process to evaluate the similarity between a stereo pair; hence a small angle of convergence causes smaller occlusions and subsequently results in more similarity between stereo images. That means with small convergence angle of the space images, the associated x-parallax has a smaller standard deviation. However, the base of the stereo pair is also smaller, decreasing the accuracy of height estimation (see equation 3.4). Hence, the optimal height to base relation should be determined according to the topography variation of the study area (Büyüksalih, *et al.,* 2007).

The difference in the viewing angle as well as the illumination angle may cause some changes in the object reflection resulting in un-matched pixels. As discussed in the section 2.2.1, vegetation and particularly trees usually pose problems for image matching and subsequently for the building extraction if they are close to the buildings causing enlarged building footprints in the DSM. Moreover, interpolation of unmatched areas can result in the similar problem. Note that all above mentioned problems might be accumulated together yielding in larger false alarms.

3.4. Depth reconstruction using semi-global matching

Alobeid (2011) evaluated three different methods for stereo image matching in densely built-up urban areas: least squares matching based on the region growing, dynamic programming (DP) and semi-global matching (SGM) in order to assess the efficiency of each method for the DSM generation. Among all, SGM has shown the best results by the quantitative assessments. Furthermore, a visual inspection revealed that SGM represents building shapes in the DSM better than the other methods. Therefore, in this study, the image matching is carried out using SGM.

A short review of semi-global matching is introduced. Hirschmueller (2008) has developed the SGM algorithm that uses an approximation of the global model to find corresponding pixels in the stereo

pairs. In this method, the final disparity is generated by using a global smoothness constraint with a cost function along multiple linear paths. SGM requires epipolar geometry and its processing chain consists of four consecutive phases as per below. For more detail see Hirschmueller (2008).

- *Calculating matching costs:* the matching cost is computed based on the Mutual Information (MI). MI uses the entropy and joint entropy of the images instead of direct intensity value so that it can robustly compensate the radiometric differences of input images caused by different reflection or illumination (Hirschmueller, 2008). Equation 3.7 - 3.9 demonstrate the calculation of the matching cost function.

$$MI_{I_1,I_2} = H_{I_1} + H_{I_2} - H_{I_1,I_2} \tag{3.7}$$

Where H_{I_1,I_2} is the joint entropy of stereo pairs and H_{I_1} and H_{I_2} denote the entropy of the first and second images, respectively. They are defined using probability distributions P of intensities of the associated images, which are approximated by the histogram.

$$H_{I_1} = -\int_0^1 P_{I_1}(i) \, log \, P_{I_1}(i) di \tag{3.8}$$

$$H_{I_1,I_2} = -\int_0^1 \int_0^1 P_{I_1,I_2}(i_1,i_2) \, log \, P_{I_1,I_2}(i_1,i_2) \, di_1 \, di_2 \tag{3.9}$$

- *Cost aggregation:* minimising the matching cost is used to derive the corresponding matching candidates but generally pixel-wise cost calculation contains ambiguity. This is mostly due to noise from various sources. Thus, in order to overcome the ambiguity case, additional constraints are enforced during the cost aggregation. It supports smoothness by penalizing changes of the neighbouring disparities. The pixel-wise cost and the smoothness constraints are expressed by defining the energy E_D that depends on the disparity image D as shown by the equation 3.10

$$E_D = \sum_P C(p, D_p) + \sum_{q \in N_p} P_1 \, T[|D_p - D_q| = 1] + \sum_{q \in N_p} P_2 \, T[|D_p - D_q| > 1] \tag{3.10}$$

Where

$\sum_P C(p, D_p)$: the sum of all pixel matching costs for the disparities of D

P: image coordinate of current pixel

P1: a value that penalizes disparity changes between a pixel and its neighboring pixels

P2: a value that penalizes the disparity changes of more than one pixel between neighboring pixels

The first term in the equation above shows the sum of all matching costs for the disparities of *D*. The second term adds a constant penalty of P1 for the all pixels q in the neighborhood N_p of p, for which the disparity changes slightly (1 pixel). The third term adds a larger constant penalty P2 if there is a

considerable change in disparity (more than one pixel). The second and third terms are associated with the small and large disparity changes, respectively. Consequently, by setting a lower penalty for the small changes, this type of dissimilarity is easier accepted for adaptation, however, in case of large changes, the third term adds a larger constant penalty (*P2*). Nevertheless, in case of a sharp disparity difference between the neighbouring pixels, it preserves the potential height discontinuity (e.g. sharp building outlines).

- *Disparity computation:* Generally, SGM works on the basis of dynamic programming but the disparity for a given pixel is computed in the various directions (multi-baseline matching). A left-right consistency check is carried out to reduce the number of false matches, for instance in occluded areas; a disparity value is accepted only if the difference of the values in both cases does not exceed a given value.

- *Disparity refinement:* few percentages of matched pixels generated from previous stages have outliers that should be eliminated. These peaks are mainly due to the differences in illumination or viewing angle between stereo pairs, textureless foreground regions (e.g. homogeneous building roofs), homogeneous background areas (e.g. asphalt, bare soil), noise etc. To eliminate these peaks, the disparity is segmented, and then very small segments are removed by setting a threshold on the segments size. The next point in the disparity refinement concerns the recognition of discontinuity between the foreground and background areas, particularly when one of them contains outliers because of the homogenous texture.

After generation of the disparity map, it is converted to DSM through georeferencing by rational polynomial coefficient (RPC) as described in the section 3.1. Finally, the DSM filtering is carried out to generate a Digital Elevation Model (DEM), and subsequently a normalized Digital Surface Model (nDSM) is computed, demonstrating the building height. This step is described in the chapter 4.

3.5. Co-registration of DSMs

Since the building change map is obtained from the difference of two DSMs, it is required that the corresponding objects in both epochs refer to the same coordinates system. This is provided by an accurate image orientation utilizing the ground control points (GCP). In case the orientation is carried out without GCP, it is required to co-register DSMs against each other. In this study, this step is carried out using the least squares adjustment (LSA) in order to eliminate a potential systematic shift between two DSMs. Note that only by performing this step, the subsequent background subtraction in the height domain demonstrates actual building changes. This section describes the fundamentals of DSMs co-registration by utilizing the least squares adjustment.

In schematic Figure 7, it is assumed that the first epoch is well registered to the ground coordinate but the second one has a georeferencing error. Consequently a co-registration is required to adjust the coordinates of the corresponding objects against each other before performing the height subtraction otherwise it causes object displacement for the new buildings as well as false alarms for the existing buildings.

Prior to performing the least squares adjustment, it is required to select one of the DSMs - which is accurately geo-referenced - as reference DSM and then calculate the shifting of the next DSM with respect to the reference DSM. To eliminate the shift within three dimensionality, a least squares adjustment is utilized to calculate the residual errors between two DSMs using an iterative approach as explained by Heipke, *et al.*(2002).
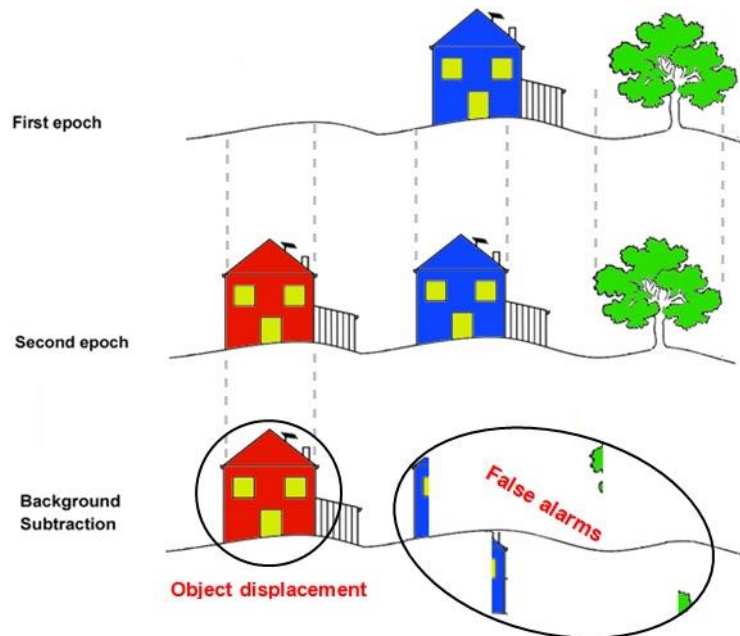
Figure 7- Schematic presentation of the georeferencing errors in the 3D building change detection based on the height subtraction

3.6. Approaches for building change detection

The approach for solving the complex issue such as 3D building change detection can be categorized with respect to a wide range of perspectives. As stated previously, the height is a unique characteristic that can separate 3D objects (e.g. buildings, trees, electricity pylons and etc.) from 2D objects in urban areas so the height is known as a reliable element for the building change detection. However, a novel approach is introduced to overcome the challenging points in the building change detection by the incorporation of height and image information. From another point of view, since the change detection is constituted by the comparison of similarity and dissimilarity between two datasets, all objects within a given scene at two different epochs can be classified into foreground and background pixels (see Figure 8). According to this approach, an object during a given period of the time is classified as background pixels (if static, showing no change) or foreground pixels (if dynamic, showing a change). Consequently, there are two main strategies to find the similarities and dissimilarities (changed and un-changed pixels) by comparison of two DSMs in the pixel and object level.

3.6.1. Background subtraction approach (pixel-based)

Nowadays background subtraction is often used in change detection (e.g. in the image sequencing, monitoring of pedestrian and vehicle traffic and particularly using video surveillance). In fact, change detection using multi-temporal datasets is similar to monitoring using video datasets.
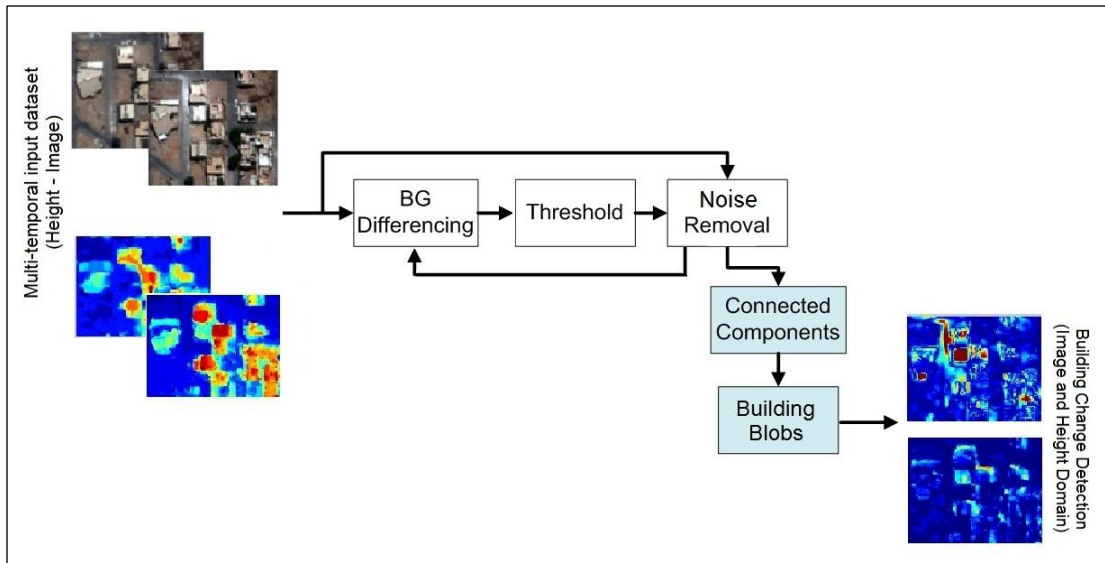
Figure 8- Building change detection based on the background subtraction in the image
and height domains

Change detection using remotely sensed data is a process to detect the dissimilarity between two datasets in the specific period of time. Building change detection can be processed in the height domain or image domain. Height as specific characteristic of building is used to detect building changes more reliable than image information. For example, if there is a change in the intensity due to the different illumination conditions, the change detection based on radiometric information may cause a number of false alarms because, in this case, the radiometric difference between two epochs does not refer to the new construction or demolition (see Figure 8).

Pixel-wise comparison in the height domain is carried out through DSM subtraction on a per-pixel basis. Equations 3.11 and 3.12 demonstrate the general form of the pixel-based approach using background subtraction in the height domain.

$$VC(i,j) = H_{t2}(i,j) - H_{t1}(i,j) \tag{3.11}$$

$$CM\,(i,j) \begin{cases} demolition & if \ VC(i,j) < -threshold \\ construction & if \ VC(i,j) > +threshold \\ no\ change & if \ |VC(i,j)| < threshold \end{cases} \tag{3.12}$$

Where, $H_{t1}(i,j)$ and $H_{t2}(i,j)$ are the height values of the first and second epochs, respectively. In case, the vertical change $VC\,(i,j)$ is greater than a given threshold, it demonstrates a significant height change between time $t1$ and $t2$. This indicates the value of height change, quantitatively. The pixels are classified as unchanged pixels if they demonstrate a height difference less than minimal building height (e.g. 2.5 meters). These types of vertical changes may be caused by cars, trucks, shrubs, etc. The changed pixels are also divided into two groups; positive changes that indicate *new construction* and negative changes which show *demolished* buildings.

Building change detection can be obtained using the image radiometric information in two different epochs. It is introduced mainly based on the image differencing, classification and the color space transformation of multispectral channels. Multivariate alteration detection (MAD), Principal Components Analysis (PCA) and morphological scale space filtering are the most prevalent methods for change detection using radiometric information (Doxani, *et al.,* 2010). As stated, the main drawback of building change detection using image information is that they may not detect the vertical change if there is no intensity change and vice versa. The analysis of pixel-based approach demonstrates that the most important disadvantages of these methods are the large number of false alarms (Niemeyer, *et al.,* 2003), (Tzotsos, *et al.,* 2011).

3.6.2. Foreground validation approach (object-based)

As previously stated, background subtraction using height information refers to the particular property of building; the height. In urban areas with the different types of objects, the height can simply separate 2D and 3D objects (above the ground e.g. buildings, trees and electricity pylons). In such areas, buildings are assumed as foreground objects while the other 2D objects constitute the background of the scene. Consequently, building change detection can be obtained directly by the identification of foreground markers demonstrating buildings or vice versa by the subtraction of background pixels representing no vertical change.

In building change detection based on foreground validation, first the initial building blobs are extracted using a threshold in the nDSM. These building candidates are then compared with the building polygons in the next epochs namely *"verification of the hypotheses"* (Vallet, *et al.,* 2011). Building polygons can be reconstructed by 3D edge matching, 3D segmentation of building rooftops or from an existing building GIS database. Foreground validation is a suitable approach for change detection within noisy datasets in order to detect the changes by identifying the strong signals. This change alarm is then verified if it is verified by the existing building polygons; otherwise, it is classified as new building. It is the main drawback of this approach that it misses the small parts of buildings such as lofts.

CHAPTER 4

IMPLEMENTATION OF BUILDING UPDATING APPROACHES

4.1. Introduction

The updating of building topographic databases is an important task because it provides a basis for many applications, such as urban planning, damage assessment and 3D city modelling. As discussed in chapter 2, the updating of geospatial databases may be carried out in three different manners, namely *map-to-map*, *image-to-map* and *image-to-image*. The aim of this research is to update the spatial databases of an outdated building GIS by using updated satellite stereo images. The updating procedure includes the updating of topographic database as well as building polygons on GIS.

The main goal of this research is to develop a generalized algorithm, allowing the updating of building databases under various circumstances (i.e. *image-to-map and image-to-image*) for updating building databases. Figure 9(a) shows an outdated building GIS database overlaid on a recently captured GeoEye-1 image. The intersection of DSM derived from a GeoEye-1 stereo pair with old GIS polygons exhibits a necessity for *image-to-map* updating. Figure 9(b) shows an *image-to-image* comparison used to detect 3D building changes. In this study, 3D building change detection is acquired through not only comparing two DSMs at different epochs but also by comparing GIS polygons against DSM.

The processing chain starts from the generation of DSM using SGM. The derived DSM is compared against an older DSM based on background subtraction or against outdated GIS polygons (foreground validation). Several removal masks and shape and size metrics are then introduced to remove false alarms. The processes up to this stage result in establishing an alarm system for building change. 3D edge matching and 3D segmentation of building rooftops (using active contours based on a level set approach) are utilized to fulfil building footprint extraction. Note that our study areas contains only buildings with flat rooftops so in this research, our methods developed for building footprint extraction (3D edge matching and 3D segmentation) are valid for flat building rooftops only. A final justification of method reveals under which circumstances 3D edge matching or 3D segmentation is suitable for building footprint extraction.

4.2. Organisation of the framework

4.2.1. Analysis of input data

In order to propose a general method delivering promising results in various circumstances, a novel framework is required with an emphasis on the advantages of height and radiometric image information as well as GIS polygons. For optimised utilisation of imagery and height information, the advantages and disadvantage of each type of dataset used for updating building databases are first discussed. After identification of the limitations of image and height information, we can utilize each type of dataset (image or height information) within the proposed algorithm appropriately.
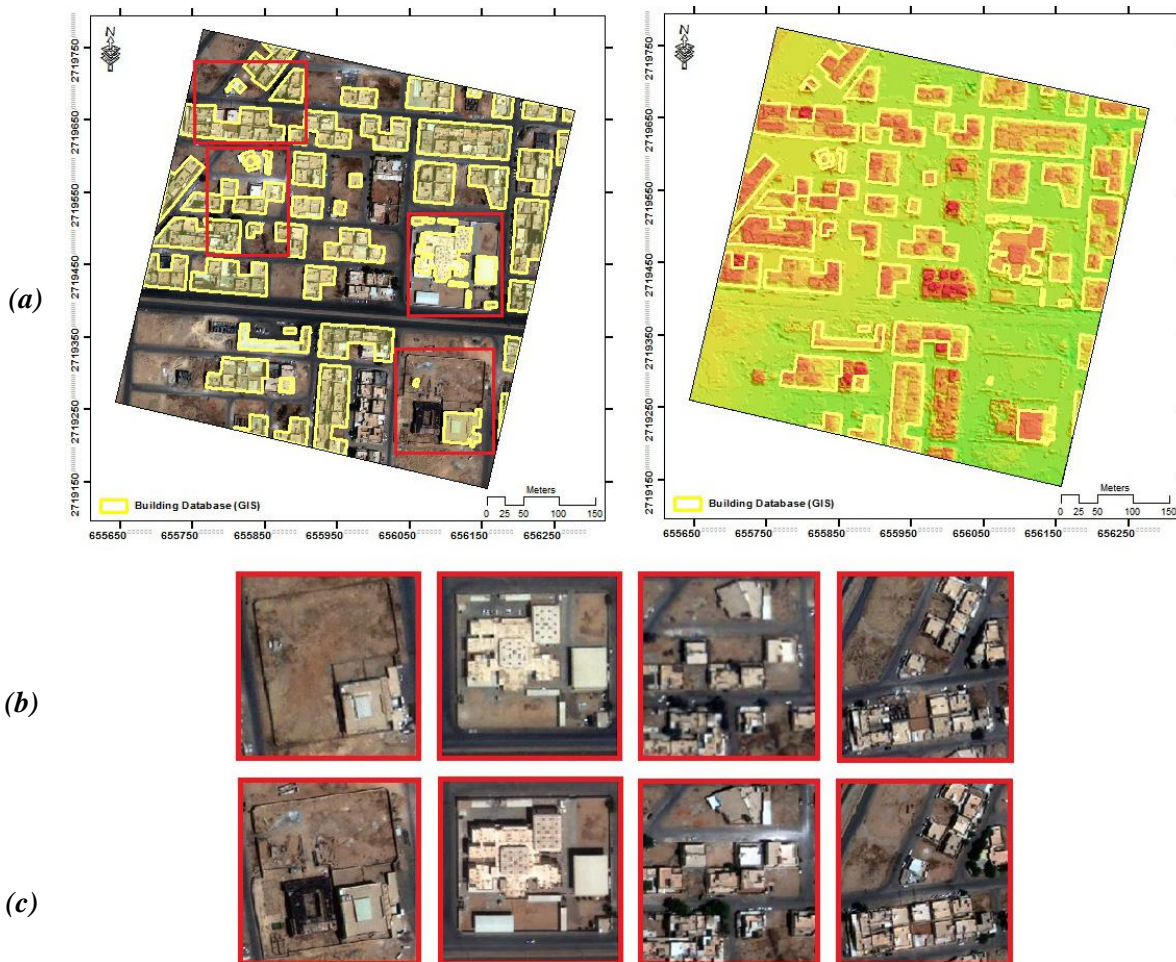
Figure 9- (a) image-to-map comparison. An outdated building GIS with polygons (yellow lines) overlaid on a recent GeoEye-1 image (left) and a DSM derived from GeoEye-1 stereo images (right), (b&c) image-to-image comparison, zoom to building changes shown on IKONOS (b) and GeoEye-1(c) images captured in 2008 and 2009, respectively

The inputs of proposed method are at the least a recent stereo pair and an outdated GIS database. Note that the images used for building update should be high resolution images usually with a ground resolution around 0.5-1m (e.g. IKONOS, GeoEye-1, Quickbird, WorldView, etc.) in order to be able to detect building shapes as well as building details within the associated DSM. As described in section 2.2 satellite stereo images are used in two forms in this thesis: (i) The original radiometric image information, known as *"original image"*. They are fused images (pan-sharpened RGB images)

that are captured from two different imaging angles. (ii) The height information derived from image matching, namely "*matching-based height information*". Note that the meaning of height information in this thesis refers to matching-based height information.

The height information can assist to detect building change with a higher success rate compared to radiometric/intensity information. However, it cannot deliver high quality results for building extraction mainly due to several reasons including matching errors, vegetation influence and particularly trees adjacent to buildings which cause overestimated building sizes in the DSM whereas original image information can be used for more accurate footprints extraction (Dini, *et al.,* 2013).
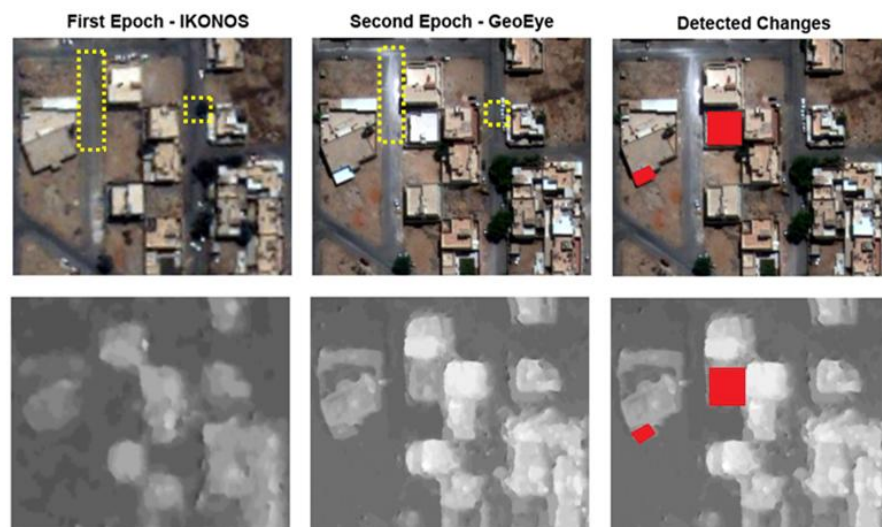
Figure 10- Showing the influence of matching errors and the associated overestimation of building size in DSM (brightness is proportional to elevation), yellow hatches show two sample of mis-detection of building changes using image information

Despite the overestimation of building size on DSM/nDSM, the extracted building footprints from nDSM are identified in the form of blobs which do not maintain the original building shape. The level of shape deformation depends on image resolution, contrast of building rooftops with surrounding objects, building details, shadows, occlusion and the existence of trees adjacent to buildings. Hence, extracted building footprints from matching-based DSM are normally not only enlarged, but also the building outlines are deformed so that they are not sharp and usually show as jagged lines instead of straight outlines (Figure 10). However, these problems are observed less in SGM compared with other image matching techniques (e.g. least squares matching) but it is still a drawback for SGM (Alobeid, 2011). Figure 10 shows that individual buildings (blobs) extracted from SGM do not show similar size/shape as actual ones in original images.

In the building change detection phase, most false alarms after DSMs subtraction suffer from matching errors, vegetation influence and the different geometric accuracy of image orientation (Dini, *et al.,* 2012). Therefore, before DSMs subtraction, it is required to co-register DSMs against each other so that only after the elimination of shifting errors, matching-based height information can detect building change in 3D space. In addition, after DSMs subtraction, further auxiliary information is also

utilized to refine the initial building change map based on structural properties of building (shape, size, etc.) and to improve the quality of the change map.

Although, image information can delineate building footprints more precisely it cannot robustly detect building changes as well as height information. For instance, the intensity of the object might have changed significantly between the first and second epochs, but this change does not necessarily indicate any building changes in 3D space. Such a problem is marked by yellow color in Figure 10. As shown, the radiometric reflection within the yellow rectangle has changed remarkably. However, there is no height (building) change between the first and second epochs. Therefore, original image information alone shows more limitations than the height information does for the building change detection phase.

Proposing a sound framework for updating building databases with the integration of height/image information first requires the identification of the advantages and disadvantages of height, image and GIS databases. Table 1 summarizes the pros and cons of each data type for building updating. Height information supports the proposed algorithm as an alarm system for building changes (change finder), although it is essential to take into account its smoothing effect as well as noise removal masks. Besides, image information acts as a complementary for the height information to assist the delineation of building footprints more precisely but it also needs the support of height information which locates building changes as blobs.

Data Type	Advantage	Disadvantage
Height Information	reliable to identify 3D changes (alarm system)	unreliable to delineate outline of buildings (blobs)
Image Information	reliable to delineate objects outline	unreliable in separating between 2D and 3D changes
GIS Feature	reliable to delineate objects outline & verify building hypothesis over nDSM	not include new constructed or demolished buildings

Table 1- Presenting the advantages and disadvantages of image/height information and GIS database for the update of building databases

The above-mentioned challenges are the main limitations for height/image information and GIS databases to update building databases. In order to decrease the influence of the limitations of height/image information within the proposed framework, the first phase (building change detection) is based on height information. However, further post-processing is required to eliminate the influence of several shortcomings. Image information is also suitable to use for delineating building footprints.

A visual inspection of Figure 10 revealed that the left yellow rectangle shows changes on the asphalt road between the two epochs without any vertical change (height values). The right rectangle shows a small tree close to the building which is absent in the second epoch. The red rectangles show why height information is suitable for building change detection. Furthermore, it shows the smoothing

effect over DSM, causing enlarged building blobs. As previously stated, improving the rate of correct alarms for building changes is one point while another point is to improve the quality of building shapes in the DSM.

4.2.2. The general procedure for updating building databases

As stated in chapter 2, the rationale for the challenging task of building updates is twofold: "*building change detection*" and "*delineation of building footprint*". Performing these two phases consecutively within a processing chain, results in the updating of building databases. The epipolar geometry is used to compute the height at pixel-level based on the associated parallax of a given object using stereo pair resulting in DSM.

As explained in chapter 3, *matching-based DSM*, as the final product of image matching, has several specific characteristics compared with DSM derived from LiDAR points cloud. The main point is the overestimation of building size on DSM (building blobs) compared with their actual size. This overestimation is due to unmatched pixels which cannot find a corresponding pixel and consequently have a disparity value so that their height value is derived from the interpolation of neighbouring pixels. This problem is observed mainly around building outlines that cause smoothing effects as described in section 2.2.2 (shown by Figure 4). Occlusion, shadows and trees close to buildings are the main reason for mismatched pixels, often around building outlines.

The proposed framework for updating building databases should consider this special characteristic of matching-based DSM and image information, allowing that which type of information is suitable for change detection and footprint extraction phases, respectively. The outdated building GIS database is also used as an input. There are three main functions (see section 4.5) for the existing building database (Bouziani, *et al.,* 2010), (Jin, *et al.,* 2005). In this study, the outdated GIS database has two main functions, firstly; to validate building blobs hypothesis (over nDSM) and secondly; to delineate the building outline by using building polygons of GIS. We will describe these functions in section 4.5 in order to see how the outdated GIS can be used as an input for updating building databases.

Height Domain	Image Domain
Stereo Pairs	**Stereo Pairs**
Left Images / Right Images	Left Images / Right Images
Image Matching Techmique	**ROI (Construction-Demolition)**
DSM 1 nDSM1 DSM2 nDSM2	Straight Lines / Initial Plotting
Background Sub. Foreground Val.	**Line-based Region-based**
Applying Noise Removal Masks	Edge Matching / 3D Segmentation
Shape/Size Metrics Thresholding	**Box Fitting - Boundary Regularization**
Alarm System for Building Change	Building Footprints

Figure 11- Functions of height and image information for update of building databases

Figure 11 shows the proposed processing of image/height information within our framework, showing the function of each data type. Further illustration and the linkage between the following processes within a processing chain are illustrated in section 4.4.

The final outcome is to update building spatial databases on GIS at LoD1. It includes updating the topographic database as well as building polygons of GIS. When DSMs subtraction with additional noise removal masks is used as a *"pixel-based"* approach to detect building change, the 3D edge matching and 3D segmentation are then illustrated to delineate the building outlines as *"edge-based"* and *"region-based"* approaches, respectively. The prime strategy in this study consists of using height information for the building change detection phase and image information for the second phase; the precise delineation of building footprints.

Inputs	Processing		Output
	First Phase	Second Phase	
GIS Database **First Stereo Images** **Second Stereo Images** →	**Building Change Detection** →	**Extraction of Building Footprints** →	**Updating Building Databases (Topographic Attributes & Polygons)**
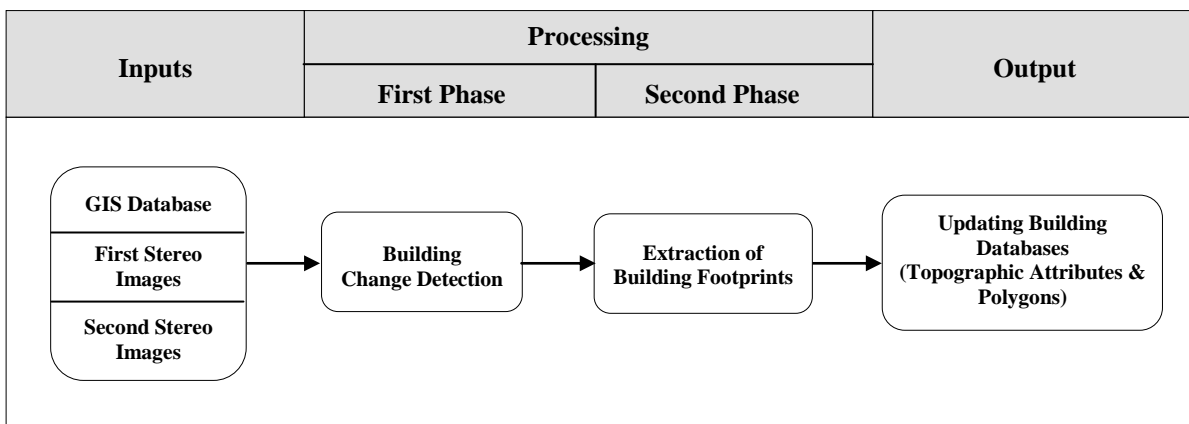
Figure 12- General procedure to update buildings databases

4.3. Approaches for building update using stereoscopic images

As stated in previous sections, the updating of building databases is divided into two main phases; *building change detection* and *the delineation of outlines for new buildings*. As shown by Figure 13, different approaches are employed for each phase.
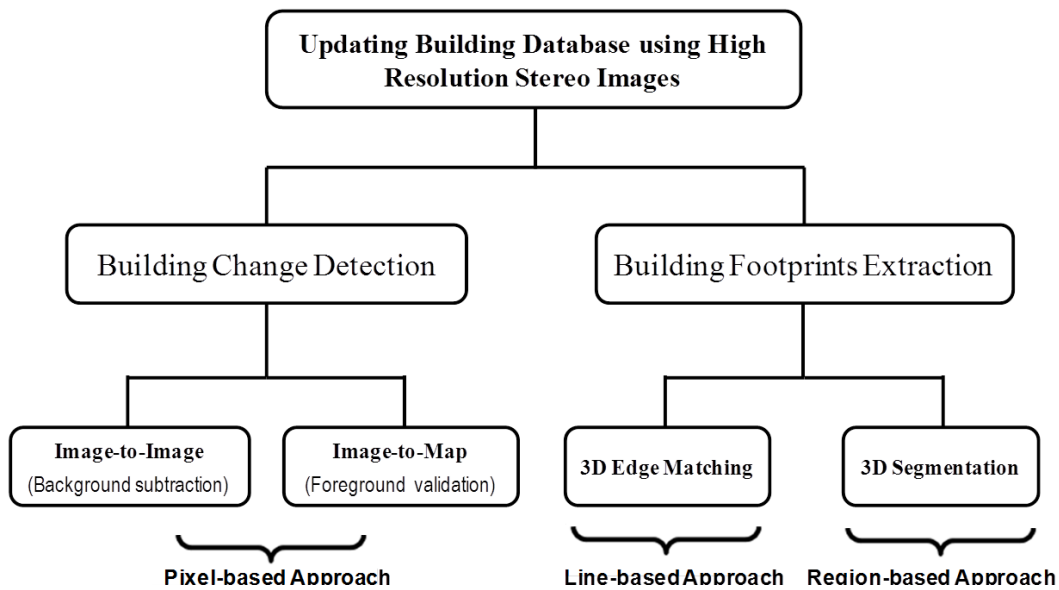
Figure 13- General framework to update building databases based on different approaches

For change detection, a *pixel-based* approach using SGM is utilized in order to generate DSMs and then detect vertical changes. Building footprints extraction is carried out using *line-based* and *region-based* approaches. However, using satellite stereo images with limited ground resolution, extraction of building footprints remains a challenging task.

4.3.1. Pixel-based approach

This section gives an overview of the three approaches used in this thesis. As stated in section 3.4 with respect to data type, the use of the *pixel-based* approach for building change detection as well as footprint extraction can be carried out using image or height information. In *image-based* models, building detection is usually based on the classification of multispectral bands. However, the associated auxiliary information such as shape and/or texture information can be introduced to assist the accurate classification of spectral information into target objects. A wide range of building extraction methods using classification of gray value (radiometric image information) are proposed by (Grigillo, *et al.,* 2012), (Scott Lee, *et al.,* 2003) and (Benarchid, *et al.,* 2013).

Height-based models typically concern the clustering of pixels based upon their height value, usually to *ground points* and *above-ground points*. This task is typically carried out by setting a threshold on height value. The detection of building change in height domain may be estimated by comparison of two height models. In this study, the pixel-based comparison of DSMs is used in order to indicate vertical changes (subtracting two DSMs from each other which contain potential building change). To distinguish actual building changes from false alarms three removal masks as well as filtering based on size and shape are employed. As an alarm system, the final blobs demonstrate building changes as well as addressing the change type.

4.3.2. Line-based approach

In this thesis, line-based and region-based approaches are used to extract building outlines when the pixel-based approach using DSMs subtraction indicates building changes. 3D edge matching technique is employed over epipolar images in order to find buildings outlines. It separates 3D edges which belong to building outlines from 2D edges that cannot be a part of building footprints (i.e. edges representing road boundaries, occlusion and shadow).

For 3D edge matching, the most significant challenge is the lack of strong constraints. Ideally, for each line there should be a unique corresponding line in the next epipolar image (pair-wise matching) but in urban areas with complex structures, and particularly in dense built-up areas, enforcing these constraints might not eventuate in finding a unique corresponding line. These are known as *ambiguity cases.* The mismatch cases stemming from occlusion or shadow are typical for ambiguous cases. Thus usually only with the integration of various geometric (including line orientation, mid-point, epipolar and proximity), radiometric and topologic constraints eventuate in a unique match (Baillard, *et al.,* 1999). Another solution for this problem is increasing the number of stereo images (additional views) in order to recognize the best corresponding line (Ok, *et al.,* 2012) however, it is more applicable with aerial images. Moreover, additional height information derived from SGM can also support identifying the best match. Finally, after pair-wise matching, the rectangular shape of building footprints must be reconstructed by fitting a box over the detected building outline.

4.3.3. Region-based approach

In a region-based approach for the segmentation of building rooftops, the homogenous areas on original images are segmented to delineate building outlines by grouping those of similar intensity value. It follows a bottom-up, region-growing segmentation procedure starting from centre of building as region of interest (ROI). The smoothing effect of matching-based DSM – as discussed in chapter 2 and shown by Figure 4 – motivated the author to perform segmentation over original images because in matching-based DSM from dense built-up areas, the gap between individual buildings are filled so that the individual building footprints are merged together, presenting a blob. However, note that the matching-based height information can efficiently recognize the prime region of interest (ROI) locating building footprints as initialization for the segmentation.

In many region-based methods, the segmentation procedure starts from a number of initial pixels and then merges the similar pixels around them into larger segments through intensity or pattern similarity. The selection of initial ROI is, therefore, a fundamental point for many region-based methods because it has a significant influence on the final outcome. In this study, the building blobs on nDSM are assumed as the initial ROI. It includes a rough location of building footprints and segmentation and is then carried out using region-based active contours based on level set method (Chan, *et al.,* 2001). Determining an optimum number of iterations is also a very important point which can balance the processing time and the accuracy of segmentation effectively. Similar to the extraction of building footprints using edge matching, after the segmentation of building rooftops, a box fitting algorithm is required in order to regularize the unshaped segments into rectangular parts.

4.4. A framework for building change detection

In order to update building databases, at first it is required to investigate whether there has been any building changes. Building changes include demolition, new construction, extension of building height (caused by adding new a floor), reduction of building height (caused by demolition of a floor) and adding a new part (annex) to an existing building. The comparison can be performed either between an old GIS database and an updated stereo imagery or by comparison of an old stereo image vs. an updated stereo image. However, if both an old GIS and two stereo pairs are available (as in this study), building blobs on nDSM are employed to verify building polygons of GIS. Figure 14 shows the general framework for updating building databases.

4.4.1. Blob detection in matching-based DSMs

For image-to-image comparison, image matching was carried out using SGM (Hirschmueller, 2008) resulting in DSMs for both epochs. The grid spacing should be set to the same GSD if they are different in order to equalize the comparing conditions for both epochs. As shown by Figure 14, an nDSM is also required. To generate an nDSM, the DEMs were generated for each epoch by filtering the DSMs according to (Niemeyer, *et al.,* 2010), and normalised digital surface models (nDSM) were subsequently computed as point-by-point differences between DSM and DEM.

After pre-processing (e.g. DSM co-registering, nDSM generation, etc.), the DSM of the first epoch is subtracted from the second one resulting in difference of DSMs (namely D-DSMs) that demonstrates

vertical changes containing both building and non-building vertical changes (e.g. tracks, new trees, temporary shelters etc.). This stage is based on background subtraction at pixel level representing height difference between the first and second epochs.

As stated, if there is a shifting error between DSMs due to geo-referencing, it should be eliminated before performing subtraction. The shift errors mainly stemmed from inaccurate of one of the DSMs to the earth coordinate. To eliminate this, a 3D least squares adjustment is applied (as illustrated in section 3.5) to co-register DSMs against each other. Most often such a problem occurs if image or DSM orientation is carried out without utilizing ground control points (GCP). The shift elimination should be performed in all three coordinate components.
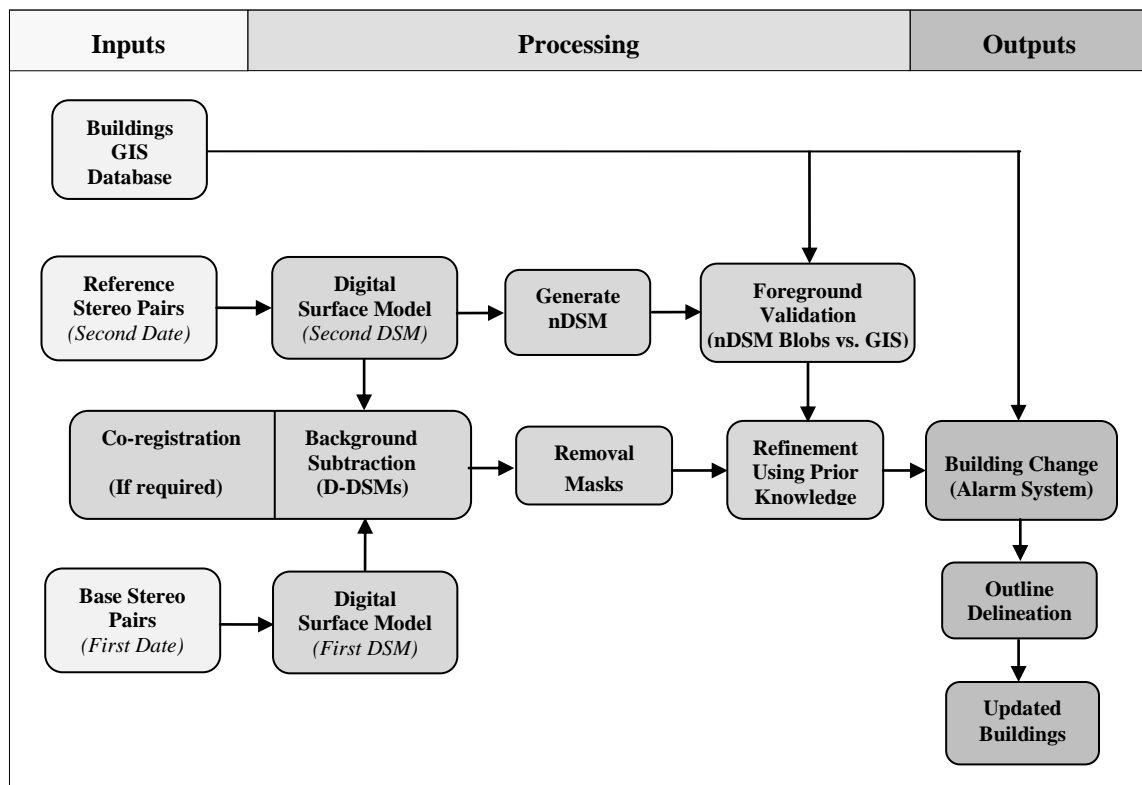
Figure 14- A general framework for building update using stereo images and existing GIS databases

After shift elimination, the difference of height is computed pixel-by-pixel for each position in the object space as shown in the equation below.

$$\Delta H_{(i,j)} = H_{t2(i,j)} - H_{t1(i,j)} \qquad (4.1)$$

Where $\Delta H_{(i,j)}$, $H_{t2(i,j)}$ and $H_{t1(i,j)}$ represent the height change and the DSM values for position (*i, j*) in the first and the second epoch, respectively. In order to find building change pixels within all vertical changes, it is required to introduce a height threshold by keeping only the pixels with absolute height values larger than 2.5m; the vertical changes with height values less than this are considered as noise which should be discarded from building changes. By applying this height threshold, the D-DSMs are converted into a binary map demonstrating vertical changes. All pixels with an absolute value of 2.5m in D-DSMs are discarded and are known as un-changed pixels.

The above processing procedure describes image-to-image comparison based on background subtraction. For the image-to-map updating procedure, an nDSM is compared against an outdated GIS database based on foreground validation. If the existing buildings in a GIS databases are confirmed by nDSM, then the remaining blobs demonstrate the changes (see 4.6). It can be concluded that both processing based on background subtraction and foreground validation acts as an alarm system in order to find changes (see Figure 15), although it is required to refine the D-DSMs using noise removal filtering, at least the peaks are those generated by image matching. It improves the quality of change alarms considerably by eliminating false alarms.
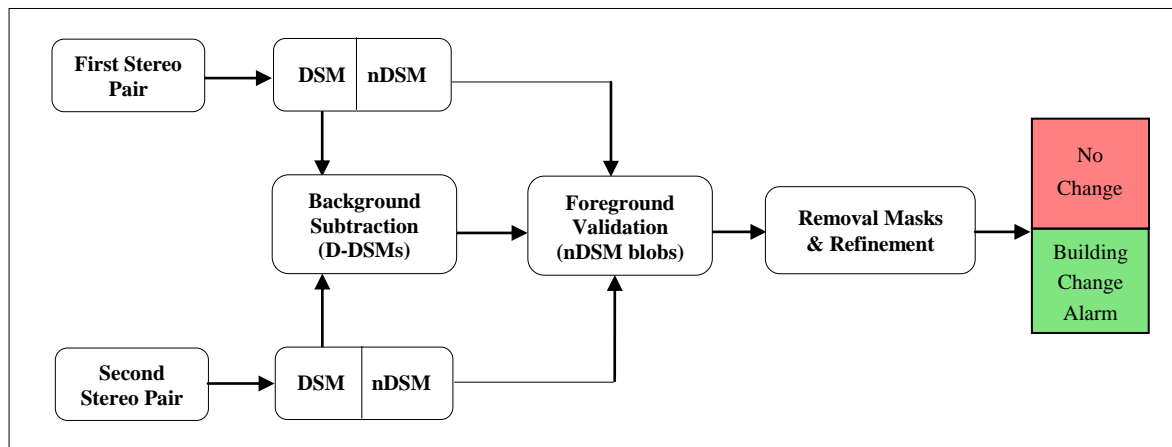
Figure 15- Building change detection using height information based on the background subtraction and foreground validation concepts

4.4.2. Removal masks

After the subtraction of DSMs from each other, despite the fact that there is no vertical change, the corresponding pixels of the DSMs may not show a similar height value. It is mainly due to a number of drawbacks stemming from matching-based DSM – as explained in section 3.3 – causing false alarms for building change as well as for footprint extraction from DSM. In order to reduce the number of false alarms and obtain reliable alarms for building changes, it is required to eliminate these undesirable alarms from the change map. Three removal masks are intersected with the change map so that these non-buildings pixels are removed from actual building changes.

Problems can stem from image matching errors (because of vegetation, asphalt reflection, shadow and occlusion effects). This section introduces a number of noise removal masks to eliminate these areas – those that do not belong to built-up areas – from real buildings. Note that although these masks are not able to remove the systematic influences of matching errors, vegetation and roads from DSM, they can impede the accumulation of these errors on our results by decreasing the number of false alarms. In addition, after removing the influences of these peaks, further refinement is required to eliminate elongated shapes and very small blobs which may not represent a building change.

- *Matching errors*

In dense built-up areas, occlusion, shadows, and un-textured regions are the most well-known reasons for outliers in disparity map (Krauß, *et al.,* 2011). SGM is relatively robust for these image matching

challenges, although a large difference between intensity of corresponding pixels causes an inevitable mismatching in the associated disparity map (Hirschmueller, 2008).

As shown by Figure 16, the DSM derived from image matching from left-to-right and right-to-left consistencies delivers different DSM, particularly in building outlines which are probably caused by shadow or occlusion.
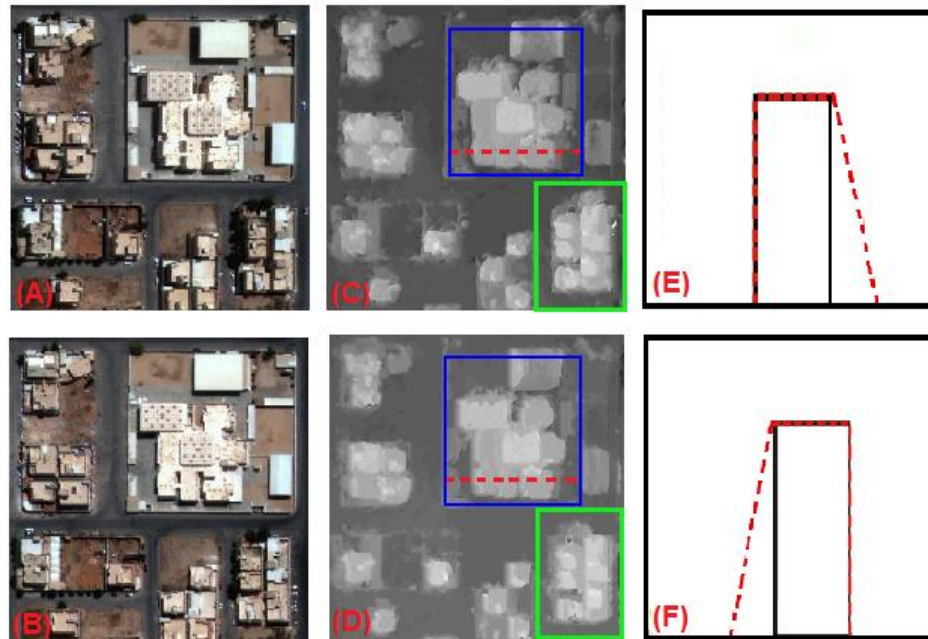
Figure 16- Occlusion, shadow and tress causing mismatched areas (A),(B) show the left and right epipolar images (pan-sharpened),(C) DSM derived by the left-to-right image matching, (D) DSM derived by the right-to-left image matching, (E),(F) schematic cross sections of highlighted buildings (red: DSM and black: original object), blue and green box represent smoothing effect caused by shadow and vegetation, respectively

Ideally, a match should deliver a unique disparity for both left-to-right and right-to-left matching, but error sources can pose a major obstacle towards the successful generation of disparity maps. A pixel is considered as *mismatched pixel,* if matching from the left to the right images, and subsequently back from the right to the left image yields a position difference in the left image exceeding a given threshold. The mismatched areas are detected by employing left-to-right consistency (LRC) checks (Hannah, 1989). In fact, enforcing a left/right or symmetric consistency between epipolar images result in the detection of mismatched areas.

- *Vegetation mask*

This section deals with the elimination of high vegetation (trees) artefacts that cause false alarms for building change detection. Trees may also enlarge the size of building footprints on matching-based DSM by smoothing building outlines. As previously discussed one main reason for this problem is vegetation influence in image matching in general and mismatched stemming from trees in particular. Vegetation and, mainly trees usually pose problems when the height values of two DSMs are subtracted from each other. This problem might be even more complicated if (i) it is not possible to

calculate the normalized difference vegetation index (NDVI) which allows for the separation of vegetated from sealed areas or (ii) the resolution of multi-temporal stereo images and subsequently derived DSMs are different, therefore the stereo pair that has lower resolution causes extensive smoothing effects, thus causing false alarms. Additional challenges due to occlusion of buildings by trees have to be dealt with separately.

These negative influences caused by trees should be removed using related mask after DSMs subtraction. As previously stated, the separation of buildings blobs from trees are challenging tasks for building updates, however, if the images are captured in winter, the problem is somewhat reduced, at least for deciduous vegetation. As the reflection of vegetation is changed during the different seasons of a year (phenological changes) vegetation masks based on color information might exhibit a poor detection quality. If an infrared channel is available, the NDVI is a more efficient index than masks obtained from RGB channels so that it is of course a better vegetation mask.

If there is no access to an infrared image channel (e.g. this study) then, it is not possible to detect vegetation using NDVI. In this case, an alternative is to generate a vegetation mask based on a classification assisted by height information. Hence, a maximum likelihood classification (MLC) of pan-sharpened RGB image along with height information from an nDSM as an additional channel is employed to generate the vegetation mask.

In fact, height information is used along with spectral information in order to compensate for the lack of an infrared channel and to improve the quality of classification, because in classification using pan-sharpened images, vegetation might be mixed with other classes (generate inter-classes). The reason is that in pan-sharpened images vegetation usually represents a low reflectance, similar to dark areas such as asphalt roads and areas in shadow (see Figure 16). Furthermore, by computing NDVI, vegetation can be detected. However, this is not able to separate the difference between low and high vegetation (i.e. grass, shrubs and trees) (Le Bris, *et al.,* 2011). This is why height variation is introduced as a new channel within classification because it can simply assist the separation of on-ground (e.g. grass and shrubs) and above-ground (e.g. trees) vegetation.

- *Road mask*

In addition to the above-mentioned masks that eliminate the influences of peaks stemming from above ground objects (e.g. trees, shadows caused by building and occlusion), there are a number of error sources from 2D urban objects (e.g. roads) that may complicate building change detection. Roads - particularly asphalts - are one of the main error sources that suffer in the matching process because of homogenous texture.

The masking of asphalt roads becomes an important task if the roads are partially covered by shadows or higher objects, such as trees, posing mismatched areas. The outcome of image matching in such areas results in outliers in the disparity map, as a consequence leading to a false height value in the DSM. This problem impedes the accurate extraction of building footprints. Therefore, the subtraction of DSMs may show the real building changes. In the same way, the outline of building footprints are probably extended over road networks, causing false detections.

A large number of methods have been developed in order to extract road networks from spaceborne images. However, if there is a vector layer containing road networks (as in our study area); it is recommended to use that as a road mask. The assumption here is that building footprints do not overlap with road networks. Also, the roads are assumed to be contained in the GIS database with correct geometric positions. A buffer of 1 meter around the roads was generated, representing the road network mask. These three masks were then intersected with the blob obtained from the previous stage (binary image of D-DSMs), therefore mismatched areas, roads and trees are discarded from the building changes.

4.4.3. Refinement of building blob using prior knowledge

Up to this step, vertical changes are detected between two dates, followed by the use of removal masks. It is expected that D-DSMs resulting from previous stages do not contain building changes only. As previously discussed, this is mainly due to well-known image matching artefacts stemming from different occluded areas, different shadow lengths due to discrepancies in view and illumination direction [e.g. (Alobeid, *et al.,* 2010); (Le Bris, *et al.,* 2011)]. Furthermore, the different resolution of stereo images (at the first and second epochs) as well as the displacement of object due to georeferencing accuracy issues causes such false alarms. Some vertical changes may also be caused by dumps, land excavation or land filling and the presence of petrol tanks. Most of these false alarms have an elongated shape; however some of them are small blobs like salt-and-pepper noise. These artefacts on D-DSMs may not exhibit a building change, therefore they should be removed.

This section deals with the refinement of D-DSMs using prior knowledge from buildings properties and in particular the geometric characteristics of buildings. Height, size and shape information are the most important building properties used to separate building changes within D-DSMs. The criteria based on building characteristics are introduced to separate building changes from all non-building changes (vertical changes that may not indicate building changes). A rule-based approach is applied to refine vertical changes and verify building blobs with respect to buildings height, shape and size metrics. After this stage, the final change map demonstrates alarms for building changes (Dini, *et al.,* 2012).

Thresholds based upon *height information* are performed on D-DSMs and nDSM of the first and second epochs. As described in the section 4.5.1, the first height threshold (A) is 2.5m and is performed on D-DSMs. This term demonstrates all vertical changes with an absolute height difference greater than 2.5m (representing new construction or demolition), thus obtaining a binary change map.

$$Potential\ building\ changes \begin{cases} (A):\ if\ |D-DSMs| > 2.5m \\[6pt] and\ at\ least \\[6pt] (B):\ nDSM_1 > 2.5m\ \ or\ \ nDSM_2 > 2.5m \end{cases} \qquad (4.2)$$

The second set of conditions (B) is thresholding on nDSM. This is also a threshold of 2.5m used to separate building changes from amongst all vertical changes, demonstrating building changes by discarding other vertical changes that are caused by trucks, low shrubs, and other vertical changes.

These change alarms do not correspond to building changes (e.g. the accumulation of noise in D-DSMs). The second set of conditions corresponds to the floor height of buildings. These two sets of conditions include the main characteristics of candidate pixels for a building change. Such a threshold also can be carried out using DSM-DEM correlation that locates the potential building candidates (Beumier, *et al.,* 2012), (Le Bris, *et al.,* 2011). With respect to the first and second sets of conditions, building changes are classified into the following classes as shown by Table 2.

Change Type	D-DSMs	nDSM$_1$	nDSM$_2$
(1): No Building Change	|D-DSMs| < 2.5m	-	-
(2): New Construction	D-DSMs ≥ 2.5m	nDSM$_1$ < 2.5m	nDSM$_2$ > 2.5m
(3): Height Extension (New Floor)	D-DSMs ≥ 2.5m	nDSM$_1$ > 2.5m	nDSM$_2$ > 2.5m
(4): Demolition	D-DSMs ≤ -2.5m	nDSM$_1$ > 2.5m	nDSM$_2$ < 2.5m
(5): Height Reduction (Demolished Floor)	D-DSMs ≤ -2.5m	nDSM$_1$ > 2.5m	nDSM$_2$ > 2.5m
(6): Noise (False Alarms)	|D-DSMs| ≥ 2.5m	nDSM$_1$ < 2.5m	nDSM$_2$ < 2.5m

Table 2- Classification of building changes with respect to changes in the D-DSMs and nDSM

As shown, applying the first condition is an essential criterion which is required for all types of building changes. If a pixel does not have the absolute height value of D-DSMs greater than 2.5m, then it is not possible to be a building change (1), so other sets of conditions on nDSMs are not tested. The second set of conditions (B) indicates that at least in one of the nDSMs, the absolute height value should be greater than 2.5m, otherwise the change is not a building change [i.e. (6) in Table 2]; probably indicating a vertical change caused by noise accumulation or it is caused by land excavation and land filling at the first and second epochs, respectively. In this manner, the accumulation of matching errors in the D-DSMs is reduced. Hence, all building changes are classified into four classes: (2), (3), (4) and (5).

The next important thresholding is the masking of D-DSMs using *size* and *shape information*. Metrics used to characterize the shape and size of building blobs are solidity and elongation. That means that size of shape filtering not only removes the elongated structures but also small blobs on D-DSMS known as salt-and-pepper noise, although those may not demonstrate a building blob. Theoretically, a building blob should have a minimal size (area) and also demonstrate a solid shape (sufficiently a massive blob on D-DSMs), otherwise, a detected blob does not represent a building. A morphological opening is employed to eliminate such noise from real building changes.

Let $\gamma_{\lambda i}^*$ be a morphological opening operator which is reconstructed using a structural element and $SE = \lambda_i . \lambda_0$ is the SE with only one element, and the size of λi increases with increasing $i \in [0, n]$, where n is the total number of iterations. The opening profile $\prod_\gamma(x)$ at the point x of the image I is defined as a vector. The morphological opening is constituted of morphological erosion and dilation, respectively (Jin, *et al.,* 2005).

$$\prod_\gamma(x) = \left\{ \prod \gamma_i : \prod \gamma_i = \gamma_{\lambda i}^*(x), \qquad \forall i \in [0, n] \right\} \qquad (4.3)$$

With respect to building size in urban areas, the filter mask is chosen based on the usual minimal size and width of a building. A size of $50m^2$ and 4m is applied as a minimum building area and width, respectively (Dini, *et al.,* 2012).

4.5. Rooftop hypothesis verification using GIS polygons

The subtraction of two DSMs is sufficient to detect building changes in three dimensions. However, if an old GIS database is available, it can be used as auxiliary information in order to verify the changes as well as building outlines. Furthermore, if an image-to-map comparison with the aim of map updating is desired, the GIS database should be verified against building blobs on nDSM. There are three main functions for an old GIS database for updating building polygons using high resolution imagery. First, it can be used as training sites, if the detection of building footprints is carried out using supervised classification techniques, then building polygons can provide hypotheses for the positions and sizes of the building training class (Walter, 1999). In this study, building footprints are not extracted by supervised classification, therefore that this function is not used and is shown by the dotted line in Figure 17.
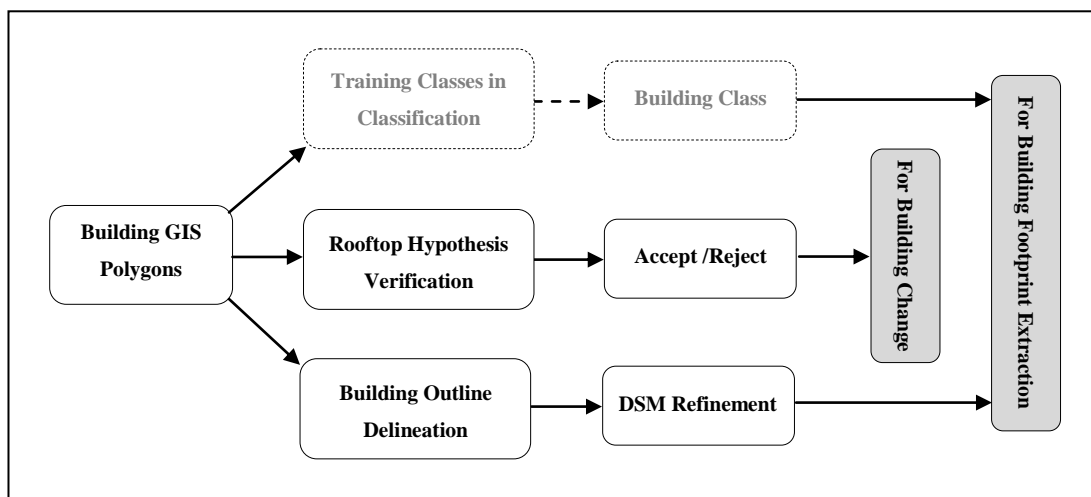
Figure 17- The functions of GIS polygons for the update of building databases

Second, GIS polygons can be used as reference data for the checking of rooftop hypotheses through comparing these polygons against building blobs derived from an updated source (e.g. satellite imagery or DSM) (Jin, *et al.,* 2005), (Bouziani, *et al.,* 2010). If those building blobs are accepted, they confirm the existence of old buildings by an updated image or DSM. Otherwise, they represent an alarm for demolition or new construction.

Finally, if a building GIS polygon is verified by comparing it against building blobs in the nDSM, GIS polygons can also be utilized for the refinement of the associated DSM. However, for newly constructed buildings, there are no polygons in the outdated GIS database. Thus, two solutions for the delineation of these building footprints are proposed in the next sections. As shown by Figure 17, the first and last functions of GIS databases deal with the delineation of building footprints, but the second one concerns the building change detection phase.

For the reasons explained above, candidates for building change from the D-DSMs have to pass two more tests as described in the previous section: first, at least in one epoch a blob must be detectable in the nDSM, and second, in epoch 1 detected building blobs need confirmation from the GIS building layer. Based on experience, a blob must cover at least 75% overlap of the GIS building object, otherwise that blob is either salt-and-pepper noise which should be removed or it is a new building.

Overlapping index ($overlap$) is an area-based similarity index which is defined by the ratio of intersection of two sets to their union

$$Overlap = \frac{A_{t1} \cap A_{t2}}{A_{t1} \cup A_{t2}} \tag{4.4}$$

The range of the overlapping index varies between 0 (0% overlap) and 1 (100% overlap). It measures the similarity of two overlapped objects with respect to size (area) and centre of gravity (relative location) quantitatively. However, this index does not represent the shape similarity as well as the level of detail.

After checking each blob against GIS polygons and removing salt-and-pepper noise, all resulting building blobs are classified into existing and new building candidates. Blobs either correspond to buildings in the GIS layer in which case the GIS information is confirmed and kept or they do not in which case we assume to have found a candidate for a new building, if a number of additional constraints are fulfilled.

For existing buildings, the updating of building databases refers to adding the building height from the nDSM to the GIS polygons. To do so, building polygons are intersected with the nDSM, and then the mean of height value from the nDSM is assigned to the GIS polygons as an updated height value.

4.6. A framework for delineation of new building footprints

After the detection of building changes using the proposed alarm system, the next stage is building boundary delineation for new constructions (detected changes). To delineate the outline of building changes, two different approaches (line-based and region-based) are tested to observe the quality of each method and then choose a method which shows higher quality: (i) edge-based approach based on 3D edge matching and (ii) 3D segmentation of building rooftops.

4.6.1. Delineation of building outlines using 3D edge matching

This section deals with the delineation of building footprints for newly constructed buildings by using 3D edge matching. New constructions also include new parts that are added to existing buildings. However, these added annexes should first be large enough in order to be detectable in the building change detection phase. After checking nDSM blobs against GIS polygons, the remaining blobs are considered to be candidates (ROI) for new buildings.

In order to accurately delineate the building outlines, the image information is employed to detect building edges in both epipolar images and then match corresponding lines. Here the rationale is twofold: (a) image information can deliver a higher level of detail of buildings than matching-based DSM, particularly demonstrating finer building outlines and (b) building outlines are clearly

detectable in original images, otherwise edge matching is not able to extract associated footprints. Therefore, extracting straight lines in the stereo images is an essential stage which allows the matching procedure to find the corresponding 3D edges and subsequently to delineate the footprints.
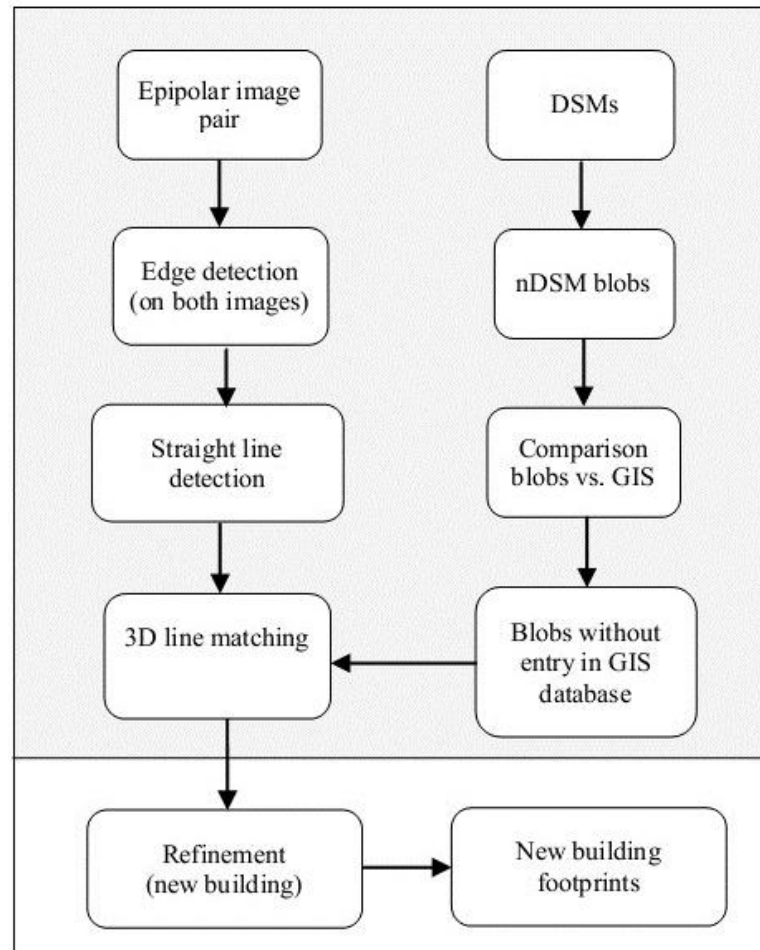
Figure 18- A workflow for footprints extraction of new buildings using 3D edge matching

After line matching, the resulting 3D lines are further processed to find the topological relationships composing rectangular building footprints (Dini, *et al.,* 2013). Building footprints have either a simple rectangular shape which can be reconstructed by a box fitting algorithm, or a complex structure (e.g. L or U shape) which can be decomposed into rectangular primitives. Similar to previous sections, the height values of newly constructed buildings are updated by nDSM. The general workflow for building footprints extraction is shown by Figure 18.

- *Edge detection and extraction of straight line*

In order to match building edges in two epipolar images, first all edges in the stereo pair are detected individually using the well-known Canny operator because among the various edge detection operators, it exhibited better performance (Shrivakshan, 2012). Then these edges are pre-processed by removing short lines (< 3m) because these lines probably do not belong to building outlines due to their very short length. Note that if a given line is detected as several individual small parts, it is automatically discarded from the further line matching process. Fitting a unique line over such small

parts could be an alternative but it causes another problem; individual outlines of row-house building blocks may join together, impeding the elimination of the smoothing effect between building blocks which is our goal in this section.

In the next phase, the straight part of an edge should be detected. It is very important because in case the corresponding lines are not detected appropriately or even detected at all and there is a deviation between the orientation of the fitted straight line and the original edge derived by Canny detector, the matching procedure may fail.

There are different methods to detect straight lines (e.g. Hough transformation, RANSAC, etc.). An overall investigation of the various methods revealed that the PCA-based method is more robust than other approaches (Lee, *et al.,* 2006), (Ok, *et al.,* 2012), (Werner, *et al.,* 2002). However, this method is rather sensitive to the template size, noisy images or very short edges. In contrast, Hough transformation (HT) has a high misdetection rate if the scene includes many straight lines near or parallel to each other (e.g. dense build up areas). In such areas, it is rather likely that HT assumed all parallel lines as a unique line and approximates a line with the best fit across them.

Furthermore, HT acts as a parametric model to estimate a line across all edge pixels. Hence, the fitted straight line may have a slight difference of orientation vs. original lines while the PCA-based method keeps the original orientation of line segments, which is important in the 3D line matching phase. The main advantages of HT are that it is flexible in the detection of straight lines even with the presence of gaps in an edge map. Moreover, it is relatively unaffected by image noise or occlusion, however, as HT acts as a mandatory method for grouping collinear lines and joins the outlines of different buildings if they are in the same direction (Lee, *et al.,* 2006), such a systematic false alarm may pose a problem in the separation of building footprints, especially in dense built up areas which contain a lot of systematic man-made structures with repetitive linear patterns. Hence, the PCA-based method is used to detect straight segments among extracted edges from the Canny operator.

Inspired by (Lee, *et al.,* 2006), first for all edge pixels, two eigenvalues are calculated in a given neighbourhood by applying principal component analysis (PCA). The eigenvector demonstrates the main direction of the distribution of the pixels of a given line and the eigenvalue shows the length of this line. Ideally, in a straight line, the second eigenvalue should be zero. In practical terms, it should be smaller than a given value. Therefore, pixels with one large and one small eigenvalue are accepted as part of a straight line. Equation 4.5 shows the general form of the scatter matrix of edge segments.

$$S = \begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix} \tag{4.5}$$

If n is the number of pixels in a line and (x_i, y_i) are the coordinates of the ith pixel of the line

$$S_{11} = \frac{1}{n} \sum_{i=1,\dots,n} (x_i - x_m)^2 \tag{4.6}$$

$$S_{12} = S_{21} = \frac{1}{n} \sum_{i=1,\dots,n} (x_i - x_m)(y_i - y_m) \tag{4.7}$$

$$S_{22} = \frac{1}{n} \sum_{i=1,\dots,n} (y_i - y_m)^2 \qquad (4.8)$$

Where

$$x_m = \frac{1}{n} \sum_{i=1,\dots,n} x_i \qquad (4.9)$$

$$y_m = \frac{1}{n} \sum_{i=1,\dots,n} y_i \qquad (4.10)$$

The large eigenvalue λ_1 and the small eigenvalue λ_2 of the scatter matrix are

$$\lambda_1 = \frac{1}{2}\left\{S_{11} + S_{22} + \sqrt{(S_{11} - S_{22})^2 + (4S_{12})^2}\right\} \qquad (4.11)$$

$$\lambda_2 = \frac{1}{2}\left\{S_{11} + S_{22} - \sqrt{(S_{11} - S_{22})^2 + (4S_{12})^2}\right\} \qquad (4.12)$$

Only the proportion of the elements of the eigenvector is known, so the angle of the line (θ) is calculated as below

$$\theta = \tan^{-1}\frac{(\lambda_1 - S_{11})}{S_{12}} \qquad or \qquad \theta = \tan^{-1}\frac{S_{21}}{(\lambda_1 - S_{22})} \qquad (4.13)$$

The above two equations are identical because $det.(\lambda_1 I - S) = 0$. In this way, the straightness of a line and its angle are obtained (Lee, *et al.,* 2006). Finally, after the detection of straight segments, the "*Douglas–Peucker algorithm*" is applied to simplify these segments into straight lines.
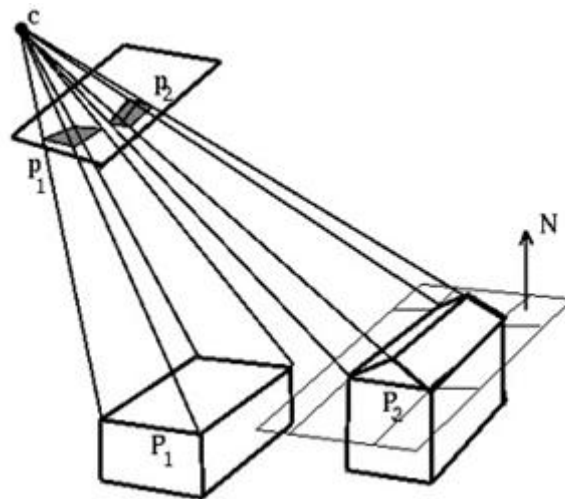
Figure 19- Geometry of flat and gable rooftop and their back-projection into image space (Jaynes, *et al.,* 2003)

For different roof types, the back-projection of object space into image space results in different 2D building outlines. A simple comparison is shown by Figure 19. It demonstrates the difference between

gable and flat roofs when rooftop edges are back-projected into image space. As shown, the difference refers to the fact that a flat roof is one plane, whereas a gable roof is composed of two planes. Hence, a flat roof has a simple back-projection into image space (as in our study area). For our case, a perspective transformation simply transforms building rooftops into image space (Hartley, *et al.,* 2004). Note that for oblique views, the back-projection of building outlines into image space results in different building footprints. In this study, it is assumed that the epipolar images are vertical or somewhat close to the nadir position so the deviation from this case is not important.

In Figure 20, the flowchart of 3D edge matching, starting from epipolar images and ending at the extraction of building footprints is described. In order to find the corresponding edge within epipolar images, several constraints are introduced in this study. First, epipolar constraints reduce the search space within a parallelogram. The geometric constraints are the next constraints to further reduce search space. After enforcing these constraints, if a pair-wise corresponding is found, the matching could successfully detect 3D edges otherwise, it is ambiguity case. To solve this challenging task, the proximity constraint and height information from DSM are used to find the best matching case.
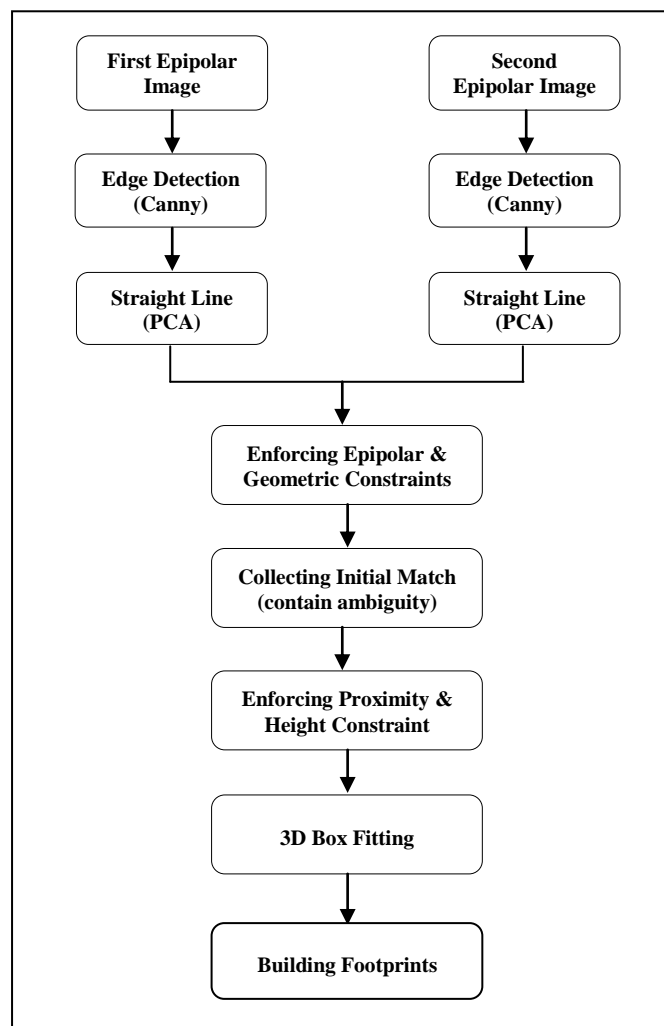
Figure 20- Extraction of building footprints using 3D edge matching

- *Epipolar constraint*

To match the corresponding lines of epipolar images, several constraints are employed. Epipolar constraint is used to reduce the search space for collecting possible correspondence cases. First, by the back-projection of object space into image space, an area around the blob under investigation is determined in image space. To do so, image orientation is required. This is approximated by RPCs block adjustment determining the relative position of two images with respect to one another.

The epipolar constraint is often used in line matching (see Ok *et al*. 2012). It yields an area in which the corresponding line should lie (quadrilateral constraint). The extent of this area varies according to the maximal allowed parallax. The higher the building results in the larger the parallelogram. Using epipolar constraints, for each straight line within image 1, a parallelogram is reconstructed collecting possible correspondences (called initial correspondences). Note that additional constraints are required in order to reduce the search space for finding the best match for each individual line. The end-points of homologous straight lines are usually not identical in both images. This is due to differences in illumination and the viewing direction of the stereo pair. In order to compensate the displacement of line end-points, a buffer of 2 meters is applied to broaden the search space similar to (Baillard, *et al.,* 1999). It increases the chance of matching so as not to miss homologous lines due to the displacement of end-points.

- *Geometric constraints*

In order to reduce the number of possible matches as well as the dimension of ambiguity, similar to (Ok, *et al.,* 2012), further geometric constraints are introduced as below.

- Orientation angle
- Line length
- Line mid point

The orientation is the most stable property of straight lines. On the other hand, as shown by Figure 19, the line orientation in image space depends on the line orientation in object space (e.g. roof shape). In our study area, most roofs are horizontal planes. It is assumed that the homologous lines have the same orientation in the two epipolar images. Therefore, a perspective transformation can transform object space into image space. We use a threshold of 5° for the maximum difference between the orientations of homologous straight lines.

Furthermore, we assume the lengths and the midpoint of the two lines to be similar, again using user-determined thresholds. If the length or mid-point of corresponding lines has a displacement more than 4 meters, this pair cannot be corresponding edges.

- *Proximity constraint*

For simple and clear buildings, applying epipolar and geometric constraints may result in a one-to-one matching, however, for complex buildings, enforcing the described constraints probably does not result in one-to-one corresponding pairs. Additional constraints (e.g. radiometric constraint) based on flanking regions around corresponding lines (Ok, *et al.,* 2012) is also not applicable in our case, mainly due to shadow, occlusion and, importantly, the similarity of the limited ground resolution of

the satellite images compared to aerial images. Therefore, further refinement is carried out by a new constraint which is named *proximity constraint*.

Assuming the roofs do not have any detectable roof structures and therefore show a homogeneous texture, we argue that lines depicting the building outlines are those which have the shortest distance from the building center - any lines representing shadows or other objects such as roads lie further away from that point. Starting from the centre of gravity of the blob under investigation (which we take to be the building centre) we thus resolve ambiguities by selecting the lines nearest to the building centre as homologous. Note that due to the other geometric constraints, these lines lie on the same side of the building centre (left or right in the epipolar image) and have the same orientation. Distances are computed as Euclidean distances between the building centre and the mid-point of the line.

In addition to these constraints, there is a possibility to test the height difference (from nDSM) at two sides of a given edge. If an edge belongs to a building outline being a 3D edge, there must be a considerable elevation difference at two sides of the building outline; otherwise it is a 2D edge. For a given edge, if the difference of height of both sides is greater than 2.5m that edge is assumed to outline. However, as the nDSM comes from matching, the smoothing effect does not always allow us to use it for separating 2D and 3D edges. Therefore, this constraint may be used as an additional constraint for very high resolution DMS (e.g. LiDAR), while in our case, it is not applied.

- ***Topological reconstruction of buildings outlines using box-fitting approach***

Up to this stage, the building outlines are extracted using 3D edge matching. They will be used later in order to reconstruct building footprints. However, due to different illumination and imaging direction, the length of building edges may differ between image 1 and image 2. This problem causes the extracted building outlines to be shorter than the real ones. A simple solution can be reached, if only buildings with a rectangular footprint are considered. In this case, for each building a box fitting algorithm can be utilized e.g. (Sirmacek, *et al.,* 2010).
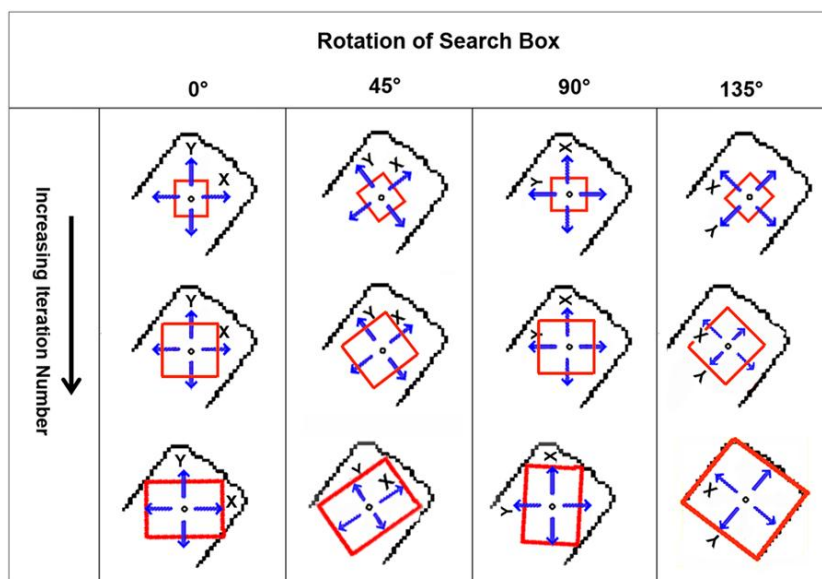
Figure 21- Box-fitting algorithm by the expansion of the search box and the rotation of orientation angle (black lines are straight parts of building outlines and the red boxes show search boxes)

This method can reconstruct building footprints if at least three sides of a rectangle are extracted. However, in case one of the rectangle sides is extracted partially, as explained above, the fitted box will be accordingly smaller than the real footprints. The proposed method fits the best box on 3D edges based on an energy minimization approach. Obviously, as this method depends on edge detection and fitting straight lines into edges, if an edge is not detected in both epipolar images, it is missed in the edge matching stage and consequently the used box fitting is not able to reconstruct the associated box.

Figure 21 demonstrates a schematic mechanism of box fitting. As shown, the search box (red rectangle) begins from the centre of each building cue expanding itself outward, iteratively. First, using building blobs in nDSM, the centre of gravity is calculated for each individual building blob, with each considered as an initial seed point for box-fitting. The search box is expanded in two orthogonal directions. The search box is rotated in Cartesian coordinates with a step of $\pi/30$ ($\theta = [0, \pi/30, \pi/15 \dots 2\pi]$) iteratively where, θ is the orientation of the search box. In each iteration, the size of the search box is increased by one pixel in the X and Y directions until it reaches the building outline. The energy $E(\theta)$ for each iteration is calculated with respect to the overlap of the search box (red box) and the binary straight edge (black lines). The more the overlap between the search box and the edge, the less the energy term is (energy minimization).

$$E(\theta) = \sum_{i=1}^{n} \min \sqrt{(x_s(i) - x_e(j))^2 + (y_s(i) - y_e(j))^2} \qquad (4.14)$$

Where $(x_s(i), y_s(i))$ represent coordinates for ith pixel on the edges of the virtual box (shown by red color), $(x_e(j), y_e(j))$ are the coordinates of jth pixel on the building edge (shown by black color).
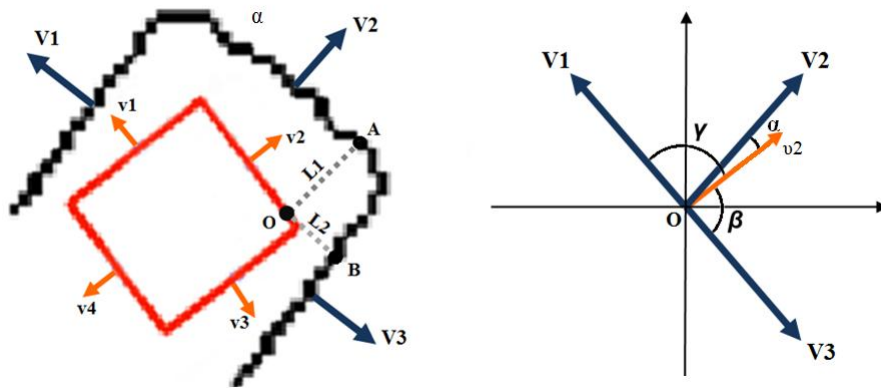
Figure 22- Showing the selection of corresponding pixels with respect to the line orientation in the box fitting algorithm

For each pixel in the search box, one pixel in the building edge is chosen to calculate the distance. The chosen pixel is the one which has the minimum distance to the pixel of the virtual box, providing that it has a similar normal vector. That means the corresponding pixels between the search box and building outline are chosen first based upon the line orientation (normal vector), if the normal vector of both lines are rather similar then the building edge pixel which has the shortest distance to the virtual box pixel is chosen as the corresponding pixel. As shown by Figure 22 (left) for point O of the

red box, although the distance to the point *B* (*L2*) is smaller than point *A* (*L1*), the point *A* is chosen as the corresponding point for the point *O* because the normal vector *v2* is much closer to the normal vector *V2*. The difference of orientation between possible corresponding lines is shown in the Figure 22 (right). As the angle α is smaller than β and γ, point *A* is the corresponding point for point *O*. The energy is computed for all rotation angles and the box with minimal energy is selected as the best fitted box. As shown by Figure 21, after the required iteration, the box-fitting algorithm is able to reconstruct the best rectangle over building edges if at least three sides of a given building are completely detected by 3D edge matching.

4.6.2. Delineation of building outline using 3D segmentation

The delineation of building footprints based on 3D edge matching may show some degree of success when used to reconstruct rectangular building footprints. However, as stated extracted 3D edges can reconstruct building footprints only if the Canny operator is able to detect these edges, otherwise this method fails to extract building footprints. 3D edge matching finds building outlines through discarding 2D edges and keeping 3D ones but complex building structures as well as the low resolution of satellite imagery pose many problems for the proposed approach. More problematically, if the buildings have jagged outlines then we are unable to fit a straight line to the edge, and subsequently line matching procedure fails to find 3D edges.

A novel solution to handle the extraction of complex building footprints is proposed based on the 3D segmentation of building rooftops. Using this method, not only are buildings with sharp edges as well as straightforward ones detected but also buildings with medium contrast and jagged outlines are also delineated. The segmentation is carried out by employing a region-based segmentation using active contours as proposed by (Chan, *et al.,* 2001).

In this study, the main reason for the utilization of rooftop segmentation using active contours without edges is that, unlike other region-based methods (e.g. region growing), this method not only extracts the building rooftop as a blob, but it preserves the object edges very well. That means that segmentation using active contours without edges considers the building outlines more than other methods so that for most building rooftops it is likely that the curve evolution is terminated in the building outlines if they have a poor contrast compared to their backgrounds. This is a very important characteristic for preserving the object edges, particularly around the corners and cusps of complex buildings. Therefore, it can be used to delineate rooftop outlines, thus propagating smoothing effects on DSMs derived from SGM.

The segmentation is performed on epipolar images. For the initialization, the height information from nDSM is used to assist the approximate locating of building rooftops as seed points for segmentation. The overall framework of the proposed method is illustrated by Figure 23. To initialize the active contour, the building blobs in the nDSM are taken into account as initial ROIs. They contain the pixels with height values larger than 2.5m and without an overlap with the removal masks. The initialization of segmentation using nDSM delivers a rough approximation of buildings blobs so the segmentation is started from a correct point (i.e. the core of building footprints). It is the main aim of initialization using height information as it increases the chance of curve evolution within building rooftops only.
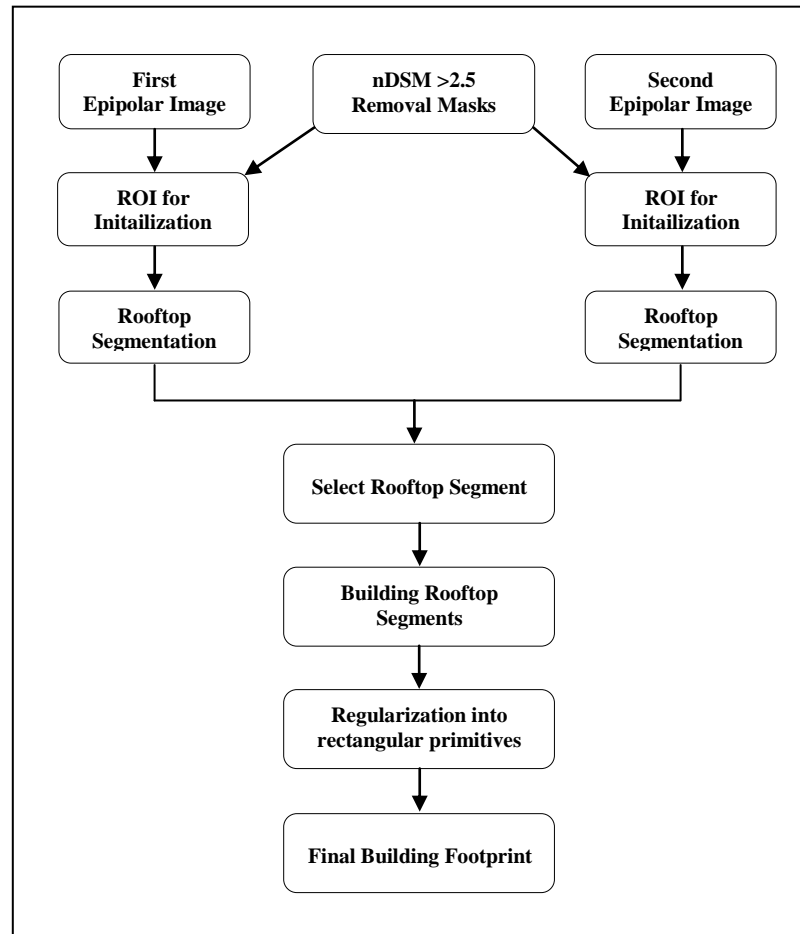
Figure 23- Extraction of building footprints using 3D segmentation of rooftops

Similar to other active contour methods, segmentation using a level set function is based on energy minimization. It is a numerical technique which detects homogenous parts of a given object by tracking the object boundary. In this thesis, we use a variational level set method introduced by (Chan, *et al.,* 2001). The curve within a building rooftop is evolved from the initial region of interest to find the building outline (known as the zero-level). Similar to parametric active contour models, this variational energy function (level set) consists of two terms; internal and external energy terms. The internal energy controls the smoothness of the contour. The external energy pushes the zero level curves toward the object boundaries - for more details of this approach see (Chan, *et al.,* 2001).

Note that by minimizing the energy function, the curve is pushed to the point of maxima, acting as an edge-detector. However, it preserves the smoothness in the corner of the object boundary (Chan, *et al.,* 2001). This characteristic of level set functions causes the building rooftops derived from 3D segmentation to be in the form of blobs instead of rectangular building footprints. To overcome this problem, a box fitting algorithm is employed in order to decompose building footprints into rectangular primitives. The regularization of segments into rectangular shapes means that building footprints have straight outlines and also, sharp corners and cusps, instead of being round.

In active contours, initialization plays an important role. This becomes clear when comparing the segmentations with different initializations - the initial contours of building outlines - which cause

topological changes in building rooftops. Hence, in the 3D segmentation of building rooftops with an initialization by height information, as in this study, it is important to generate the initial building ROI through careful height thresholding. In this study, a height threshold of 2.5m in the nDSM is selected while a slight change of the height threshold may cause topological changes in building footprints and result in different segments. For example, as shown by Figure 3(c) applying the threshold of 2.5m generates small islands within building footprints while selection of 3m as the height threshold may not make such an island for the initialization ROI. Small gaps between buildings or the courtyards of buildings surrounded by walls also cause the same problem by causing topological changes.

The initial buildings ROI from nDSM can locate building blobs, however, in matching-based DSM, due to the smoothing effect; several individual buildings are joined together generating a building block instead of individual footprints. To decrease this side effect of segmentation, it is required to conduct curve evolution into building rooftops through correct initialization because, if the segmentation is started from non-building objects, it may segment other objects (e.g. shadow, trees etc.) and wrongly take them into account as part of building footprints.

To do so, after setting a threshold of 2.5m, the associated ROI is intersected with removal masks that take into account building areas without any intersection by trees, roads and shadow, as well as occluded areas. Thus these regions are removed from the initialization ROI. As the entire smoothing effect on DSM is probably not eliminated by these masks, the morphological erosion is additionally applied to shrink the ROI by 1meter further. It increases the chance of curve evolution within building rooftops impeding the segmentation of other objects instead of buildings. The filtering size is chosen based on the usual minimal size and width of a shadow or occlusion in the study area. This buffer removes the influence of shadow from building ROIs that probably have not been already removed however, as stated above; these masks only discard false alarms from the initialization.
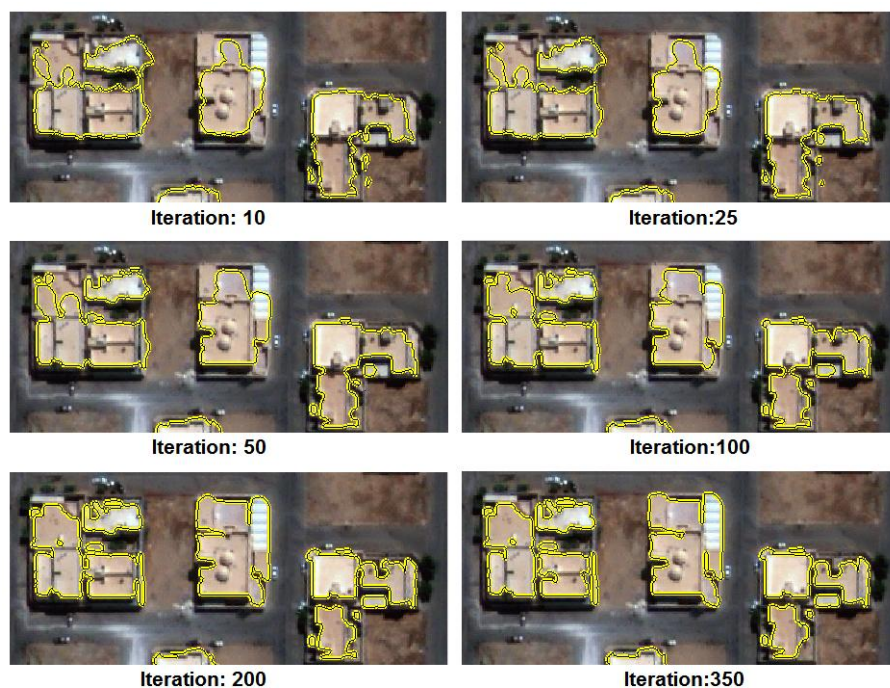
Figure 24- Level set based segmentation of building rooftops under different number of iterations

The next point is the estimation of an optimized number of iterations representing the best balance between the accuracy of segmentation and processing time. The number of iterations depends on building size as well as the complexity of building structures. As shown by Figure 24, segmentation starts from initial ROIs. It is iteratively expanded over homogenous building rooftops. In case the curve evolution reaches a building outline with a sufficient sharp edge, the segment is not expanded further, terminating the iteration on the building outlines.

As demonstrated by Figure 24 and also illustrated by the associated graph in Figure 25, from the 1st to the 100th iteration, the overlap between the segmentation and ground truth information has sharply increased. At the 254th iteration, the overlap rate reaches the optimal value of 91.15% where there is no further improvement afterwards. As the ROI for initialization is chosen from building blobs on nDSM, as shown even at the first iteration there is an overlap of more than 70% between segmented rooftops and ground truth. As previously stated, such an initialization conducts the initialization into target objects (building rooftops) decreasing the probability of curve evolution outside of building ROIs.

The above statistic is driven from the segmentation of rooftops in a small subset with 11 sample buildings. These rooftops have different shapes, sizes, and reflections, as well as consisting of different types of buildings (with various levels of complexity). The obtained segments are compared with manually generated ground truths in order to estimate the iteration number for the segmentation of all building rooftops in the study area.
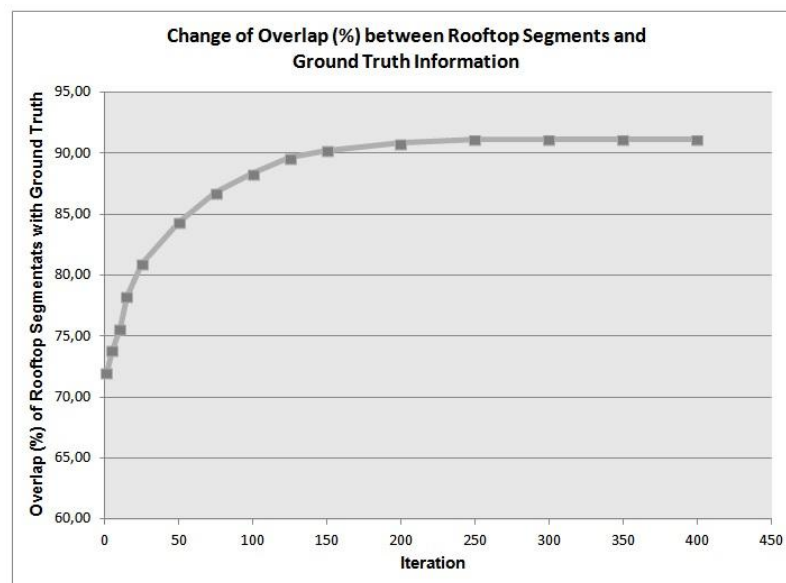
Figure 25- Overlap (%) between rooftop segments and ground truth under
the different number of iterations

After the segmentation of rooftops individually in the stereo images, similar to 3D edge matching, the epipolar constraint is used to approximate the locations of the corresponding segments (Chehata, *et al.,* 2002). In epipolar images, the corresponding segments should be usually found in the same coordinates (with $\pm$ maximal parallax). If a given segment is visible in both epipolar images, this ensures us that it is a building, not a segment showing shadow or occlusion which is visible only in

one image. The final building rooftop is obtained from one of the epipolar image which has the smallest incidence angle close to the nadir location.

The reason to choose the epipolar image with the smallest incidence angle is that as the segmentation is not carried out on orthophotos images, the image with the smallest incidence angle has less shape deformation compared with the nadir point. The collected segments generate the *initial building outline*. These obtained polygons of building rooftops are then assumed as new ROIs for re-initialization so the segmentation is performed iteratively until there is no significant improvement in the delineated building outlines. Finally, the obtained segments are regularized into rectangular primitives, generating the final building footprints (see next section).

4.6.3. Regularization of segments into rectangular primitives

As shown by Figure 24, the rooftop segments extracted using 3D segmentation have irregular outlines. Therefore, in order to use the segments to update building databases at LoD1, it is required to regularize the building footprints in the form of rectangular shapes. Guercke, et al. (2011) applied Hough transformation (HT) as a parametric model to decompose a complex building polygon into straight primitives. The proposed model is employed to re-sample the jagged building outline. A Hough buffer based on the main orientation and distance to the origin is then performed to estimate the straight lines of building outlines among all possible lines in Hough space. Finally, a least squares adjustment is applied for the generalization of footprints.

Because of the limited ground resolution of satellite images, Hough transformation is not able to regularize segments derived from 3D segmentation. Moreover, it is also the main drawback of Hough transformation that it extracts a large number of wrong lines, particularly in very dense built-up areas, therefore in this study an enhanced version of a minimal bounding box (called an object-oriented bounding box) is used to fit the best box over rooftop segments. In Figure 26, the blue blob is a schematic building rooftop which should be regularized into rectangular primitives.
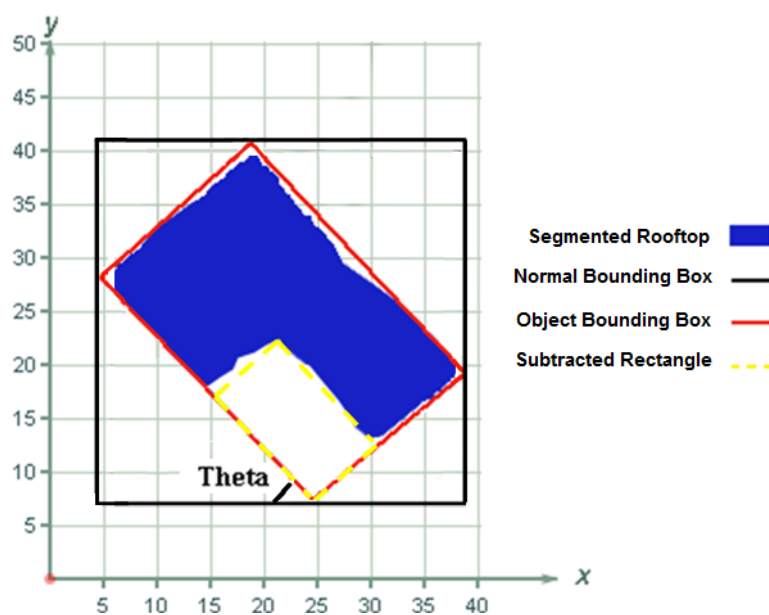
Figure 26- Minimal bounding box and object-oriented bounding box fitted into a segment

In the minimal bounding box (shown by black color in Figure 26), the fitted box is aligned with the axes of the coordinate system. It is also known as an axis-aligned bounding box which surrounds an object by fully enclosing it within the minimal rectangle. Hence, the major axes of the segment are parallel with the X and Y axis of Cartesian coordinates, so the actual dimensions of the axis-aligned bounding box are based on the maximum and minimum position values of the segment along each of the X and Y coordinates. In the object-oriented bounding box, the X and Y axes are rotated based on the main orientation of the segment (Theta in Figure 26).

The image skeleton technique (Matlab, 2011) is utilized in order to detect the main direction of the segment. In shape analysis, the image skeleton shows the general shape of an object as a thin version of that shape. It demonstrates the geometrical and topological properties of the shape e.g. connectivity, topology, length as well as the main direction which is important for us here. The object-oriented bounding box is arbitrarily oriented with respect to the coordinate system so the X and Y coordinates are rotated to be fitted on the major axes of the segment. The red box in Figure 26 demonstrates an object-oriented bounding box for the blue blob.

In case building rooftops have a simple rectangular shape, the object-oriented bounding box simply fits a rectangle representing the building footprint. However, for the extraction of complex building shapes (e.g. L shape or U shape buildings), the segment derived from 3D segmentation is subtracted from an associated object-oriented bounding box, then a morphological opening is performed.

(1) Original rooftop segment	(2) Fit object-oriented bounding box	(3) Subtract original segment from object bounding box
(6) Final L shape building footprint	(5) Subtract object-oriented bounding box from (4)	(4) Morphological opening on remaining blob
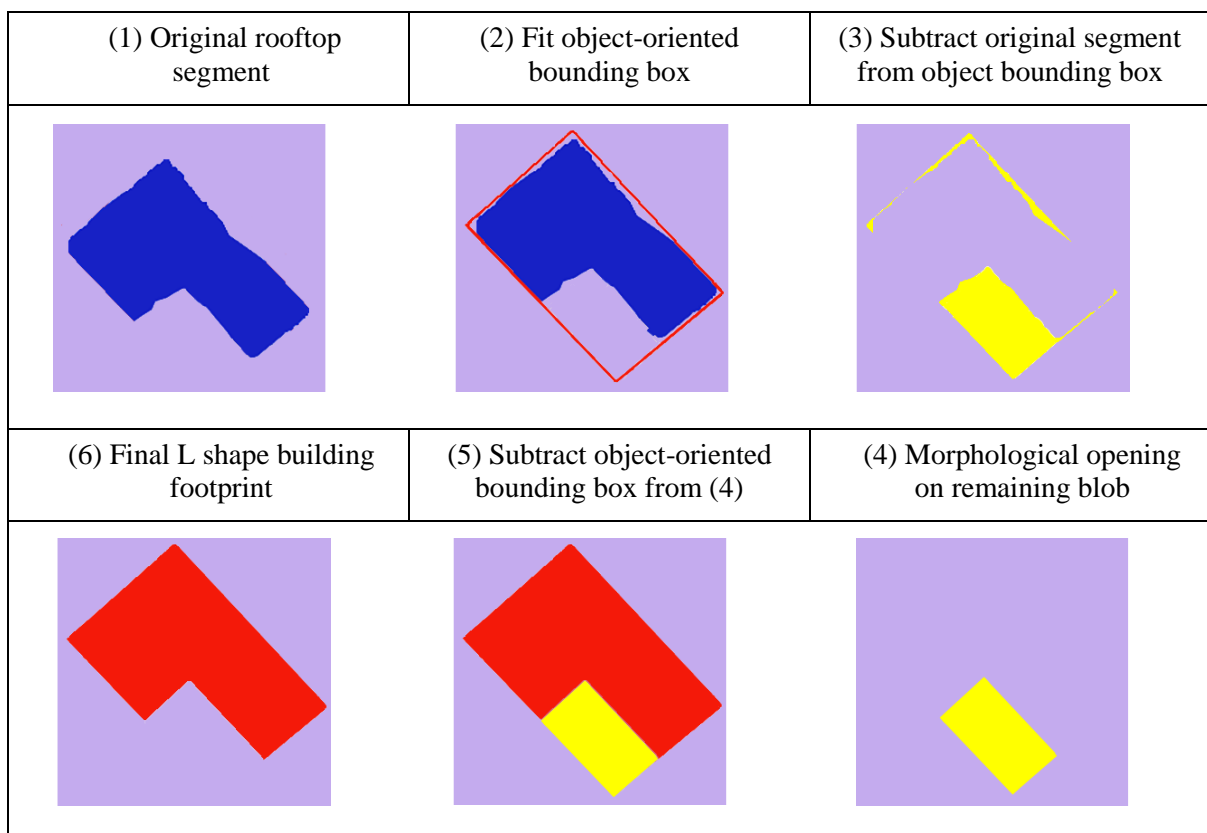
Figure 27- Schematic procedure to extract L and U shapes of building footprints by regularization of segments (rooftops) into rectangular primitives

The structural element of morphological filtering is empirically found as 1m. If a building footprint has a simple rectangular shape after morphological filtering, the obtained blob (shown by the yellow hatched line in Figure 26) should have vanished; otherwise it indicates that the footprint has a complex shape so it is required to subtract the yellow box from the object-oriented bounding box representing the final shape of the building footprints. Figure 27 demonstrate the extraction of the L shape building footprint in six steps.

4.7. Concluding remarks

For updating building databases using spaceborne stereo images, the task is divided into two main categories; *building change detection* phase and the extraction of building footprints through precise *delineation of new constructed buildings*. For the first stage, background subtraction (DSMs) along with removal masks is employed to detect potential building change areas. The updated building blobs derived from nDSM are used to validate building polygons in the GIS database, and then confirmed polygons are used as 2D building footprints.

Although the pixel-based subtraction of two DSMs derived from SGM delivers promising signals for the detection of building changes, it is not able to delineate building polygons mainly due to the smoothing effect causing an overestimation of building size. This problem becomes complicated for buildings with poor radiometric contrast. In such cases, not only the building size is overestimated in the DSM but also the building shape may be deformed due to the potential mismatch, particularly on building outlines. Therefore the matching-based DSM does not deliver a precise building shape.

Information fusion (image information with DSM derived from SGM) is the main idea proposed in this thesis to detect building changes and then update building polygons through original stereo images and associated matching-based DSM. To delineate buildings outlines, two different methods, "edge-based" and "region-based" approaches, are employed. However, a comprehensive review of previous works indicates the success of region-based approaches more than edge-based ones, because building footprint extraction based on 3D edge matching usually delivers promising results using very high resolution aerial images (Baillard, *et al.,* 1999), (Suveg, *et al.,* 2004) and (Habib, *et al.,* 2010). The experimental results in the next chapter will demonstrate the success of each method for the delineation of building footprints.

In our region-based approach, the fusion of height and image information is evaluated using 3D segmentation of building rooftops. Nevertheless, this method is referenced more in the literature for building footprint extraction using satellite images; the main deficiency of this approach is over and under-segmentation that may necessitate an additional step (e.g. split-and-merge method) in order to extract building footprints as in reality (Khoshelham, *et al.,* 2005).

It can be concluded that each method has its own advantages and disadvantages. Furthermore, note that ground resolution has a key effect on the quality of results in both phases of building updates, so when an approach might be theoretically applicable, the limited ground resolution probably impedes the gain in expected experimental results. As a consequence, a unique algorithm is not recommended to obtain promising results for different ground resolutions (Lu, *et al.,* 2004).

CHAPTER 5

EXPERIMENTAL RESULTS AND ANALYSIS

In this chapter, the test results are presented. The results are analysed for further justification and conclusion in chapter 6. First, the experimental results are displayed to show the efficiency of the proposed approach, and then a quantitative assessment is carried out to evaluate the accuracy and quality of each step in relation to the ground truth. Finally, these results are interpreted to figure out the main contributions of this thesis.

5.1. Datasets and pre-processing

5.1.1. Datasets and ground truth

The experiments used pan-sharpened RGB stereo images from IKONOS-2 (epoch 1) and GeoEye-1(epoch 2), acquired on May-24, 2008 and Sept.-15, 2009 with ground sampling distances (GSD) of 1m and 50cm, respectively (see Figure 28). The study area is a suburb of Riyadh, the capital of Saudi Arabia, including individual buildings as well as row-house building blocks.

The slant angle is 11° toward West for both images and the height-to-base ratio is in the same range with 1:1.75 for IKONOS and 1:1.51 for GeoEye-1stereo pair. This ratio has an important role in image matching. With respect to the angle of convergence, view angle, shadow length as well as image quality, our dataset is geometrically suitable for DSM generation and building detection. Table 3 shows details of the acquisition parameters of the dataset. Further general technical characteristics of IKONOS-2 and GeoEye-1are illustrated in the appendix 1.

In addition, we have at our disposal a somewhat outdated building and road GIS database including buildings polygons as well as their attributes. These GIS polygons are generated based on aerial images captured in 2007 with a scale around 1:1000. The map projection is UTM (Zone N-38). The building database does not contain all the buildings which can be seen on IKONOS or GeoEye-1but building blobs in the nDSM (building hypothesis) are used to verify the GIS polygons. The GIS road layer is employed to generate the road mask. Reference data for the building change detection as well as for boundary delineation were generated manually by comparing and digitizing the buildings visible at both epochs.

Parameter \ Stereo Pair	IKONOS-2 (first epoch)	GeoEye-1 1 (second epoch)
Acquisition Date/Time	(1): 2008-05-24 07:46 GMT (2): 2008-05-24 07:46 GMT	(1): 2009-09-15 07:42 GMT (2): 2009-09-15 07:43 GMT
Processing Level	Standard Geometrically Corrected	Standard Geometrically Corrected
Interpolation Method	Cubic Convolution	Cubic Convolution
Map Projection - Datum	UTM – Zone N-38 (WGS84)	UTM – Zone N-38 (WGS84)
Spectral Channel - GSD	Data fusion (Pan-RGB) – 1 meter	Data fusion (Pan-RGB) – 0.5 meter
Radiometric Resolution	8 bits per pixel	8 bits per pixel
Scan Direction	Reverse	Reverse
Nominal Collection Azimuth	(L): 217.66° (R): 319.08°	(L): 348.77° (R): 236.44°
Nominal Collection Elevation	(L): 64.32° (R): 76.10°	(L): 63.40° (R): 73.65°
Sun Angle Azimuth	(L): 100.85° (R): 100.62°	(L): 140.63° (R): 141.06°
Sun Angle Elevation	(L): 74.72° (R): 74.53°	(L): 62.98° (R): 63.12°
Reference Height	742.85 meters	742.78 meters

Table 3- Technical acquisition summary of stereo pairs for the first and second epoch, (L) and (R) refer to left and right images, respectively

An investigation reveals that GeoEye-1stereo pair has a good image quality while IKONOS images show a lower quality due to a lower ground sampling. This low resolution might be accumulated with smoothing effects in the DSM. Consequently, it probably causes a problem for building change detection based on background subtraction, resulting in false alarms. Furthermore, as the ground sampling of GeoEye-1 and IKONOS images is not equal, to generate a comparable condition for height subtraction, we have re-sampled the GeoEye-1 DSM into 1 meter as IKONOS DSM.
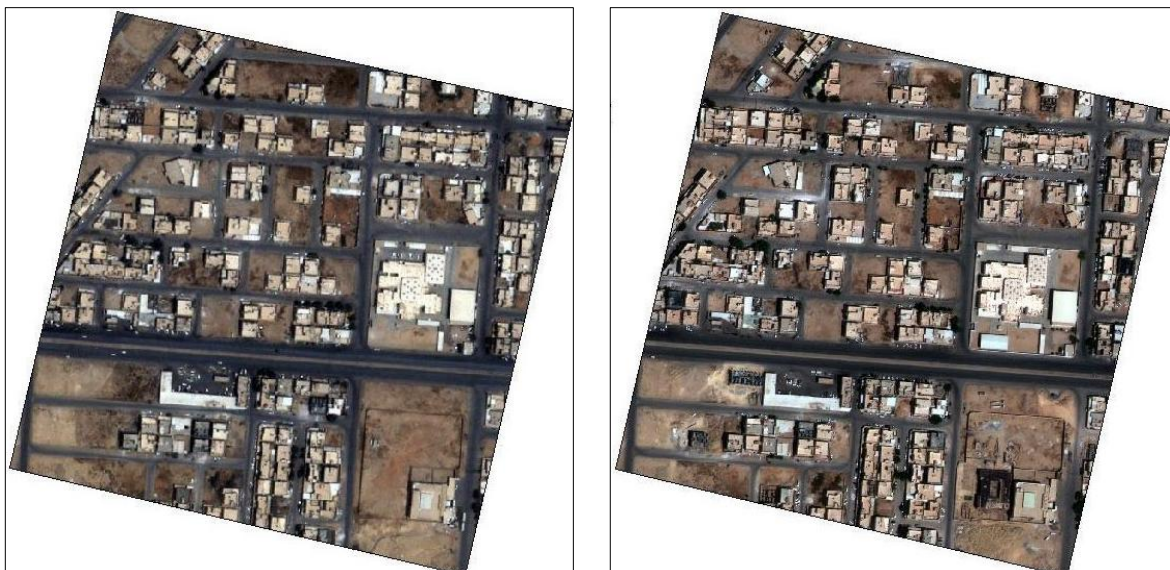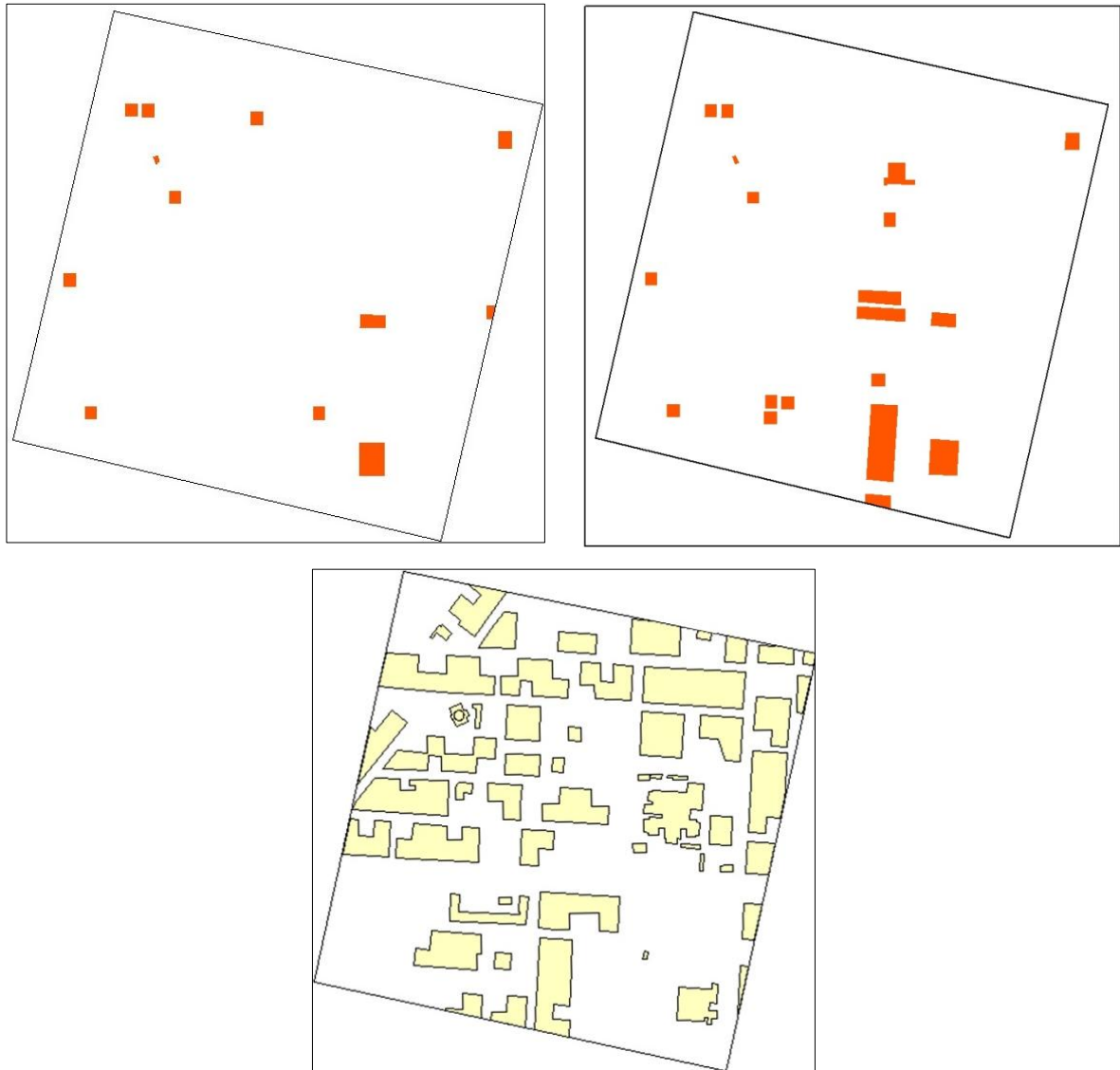
Figure 28- (first row): IKONOS image, 2008 (left), GeoEye-1 images, 2009 (right), (second row-left): Ground truth information for the first scenario (IKONOS vs. GeoEye-1), (second row-right): Ground truth information for the second scenario (GIS polygons vs. GeoEye-1), (third row): building polygons in the outdated GIS database

A visual survey reveals that the prevalent building's type in the study area is mainly small separate buildings or row-house building blocks that are very close together. The images are cloud-free and as shown in table 3, there is a 16 month delay between the acquisition of the first and second epochs. In such a rapidly developing urban area, this time interval is enough for a number of building changes to occur, including new construction, demolition etc. Note that the vertical development of existing buildings is only detectable for the first scenario as the buildings do not have a reliable height attribute in the GIS. The ground truth information is generated manually for both scenarios (IKONOS vs. GeoEye-1 and GIS polygons vs. GeoEye-1). Moreover, a visual inspection of IKONOS and GeoEye-1 demonstrates a very slight change in tree canopy cover, although our study area is an arid urban area with little vegetation cover. Some trees have been cut entirely and some slight change is because of the difference in image acquisition time; spring for IKONOS and the end of summer for GeoEye-1.

5.1.2. Generation of epipolar images and image matching

As previously stated, for a pixel-based matching approach (e.g. semi-global matching and dynamic programming), it is required to generate epipolar images, though not necessarily to generate rectified images. The corresponding pixels in stereo pairs are recognized using SGM and results in a disparity map. Hence before image matching, a transformation of the images into epipolar geometry is required.

As discussed in section 3.2, in pushbroom sensors, the epipolar lines are curves (hyperbolas) which are relatively difficult to model. However, the solution is to approximate an epipolar curve to a line through linearization of the equation. In our case, images are projected into a plane in object space so it is sufficient to rotate the images around the viewing direction resulting in the x-axes of both image coordinate systems being parallel to the base, followed by a shift of 2.5 pixels in the y-direction. By rotating stereo pairs around the base direction, quasi-epipolar images have been generated (see Figure 29). Note that this method delivers satisfying results if the scene is not located in an extreme mountainous area (same as our study area).
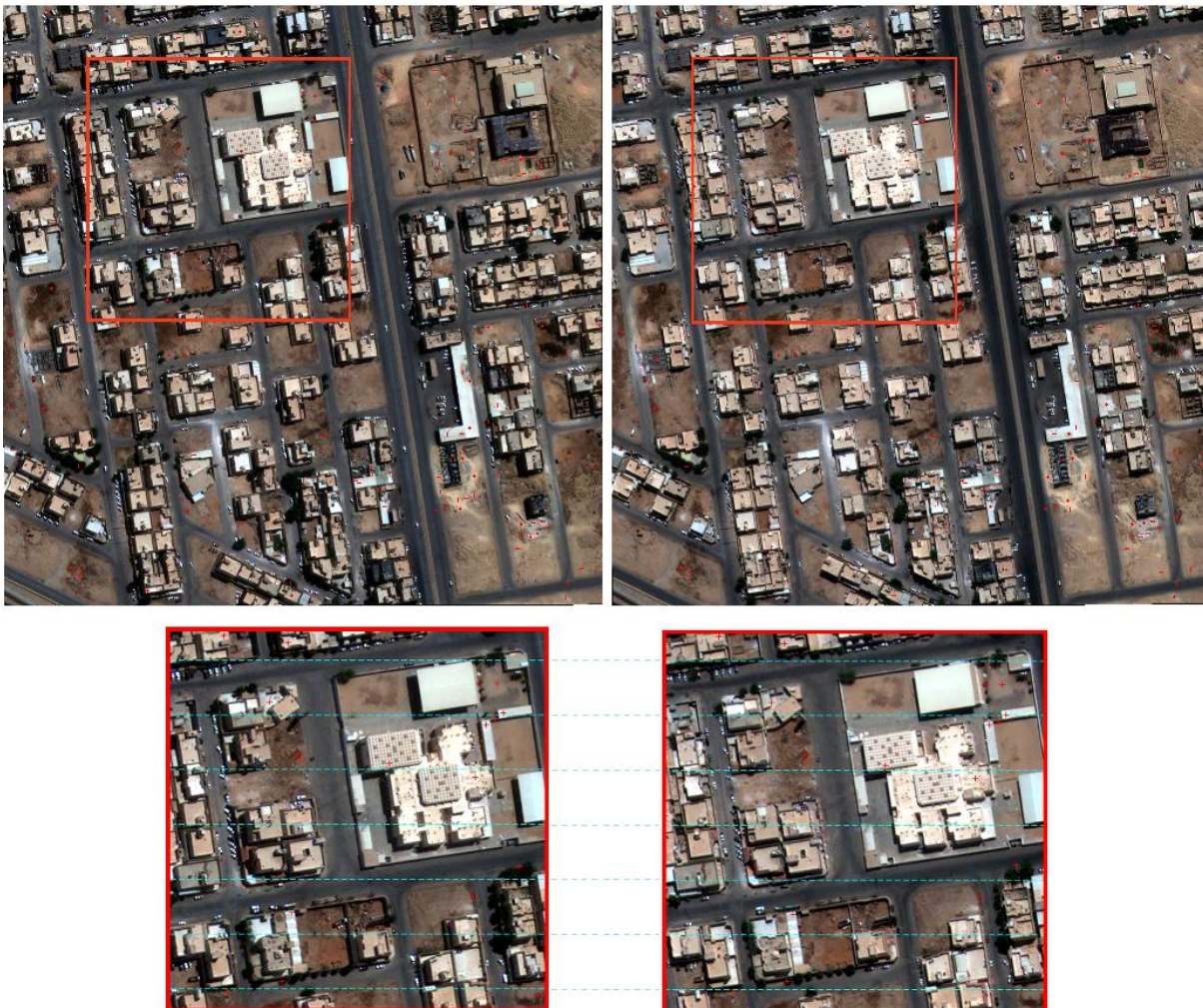
Figure 29- (left): Quasi epipolar lines superimposed to the GeoEye-1 stereo pair (corresponding pixels have the same y coordinate in the image space)

5.1.3. DSM generation and pre-processing

For the subsequent DSM generation, semi-global matching (SGM) as per (Hirschmueller, 2008) has been used resulting in the disparity maps for both epochs being transformed into DSM. To compute the optimal disparity, user-defined parameters are chosen empirically. To find corresponding pixels, the searching is carried out in the maximum of accumulated directions (2^4=16 paths). The penalty for changes in disparity (one pixel and more than one pixel) is set to 6 and 12, respectively.

This section deals with the transformation of the obtained disparity map into the associated DSM at each epoch individually (registration of the disparity map into the object coordinates). Image orientation was provided by means of Rational Polynomial Coefficients (RPC). Figure 30(a,b) demonstrates the obtained DSMs from IKONOS and GeoEye-1 stereo pairs.

A visual inspection noticeably shows the influence of GSD on the quality of DSM because, as shown in table 3, all the image acquisition parameters in general and the *b/h ratio* (the angle of convergence) in particular are relatively similar for IKONOS and GeoEye-1 stereo pairs except for the ground resolution. As illustrated, in the GeoEye-1 DSM, the building shapes in general and building outlines in particular contain more detail than in IKONOS DSM. It will create more problems for the detection of small annexes which are added to existing buildings. Moreover, in our study most buildings have lofts. It is an upper story or attic over the buildings, directly under the roof. Alternatively, it can be used as a second story area for storage or a large adaptable open space, often converted for residential use. With respect to the small size of such building annexes and also their low height, they cannot be detected as their height is less than the proposed threshold so they are probably missed within the noise of DSM.
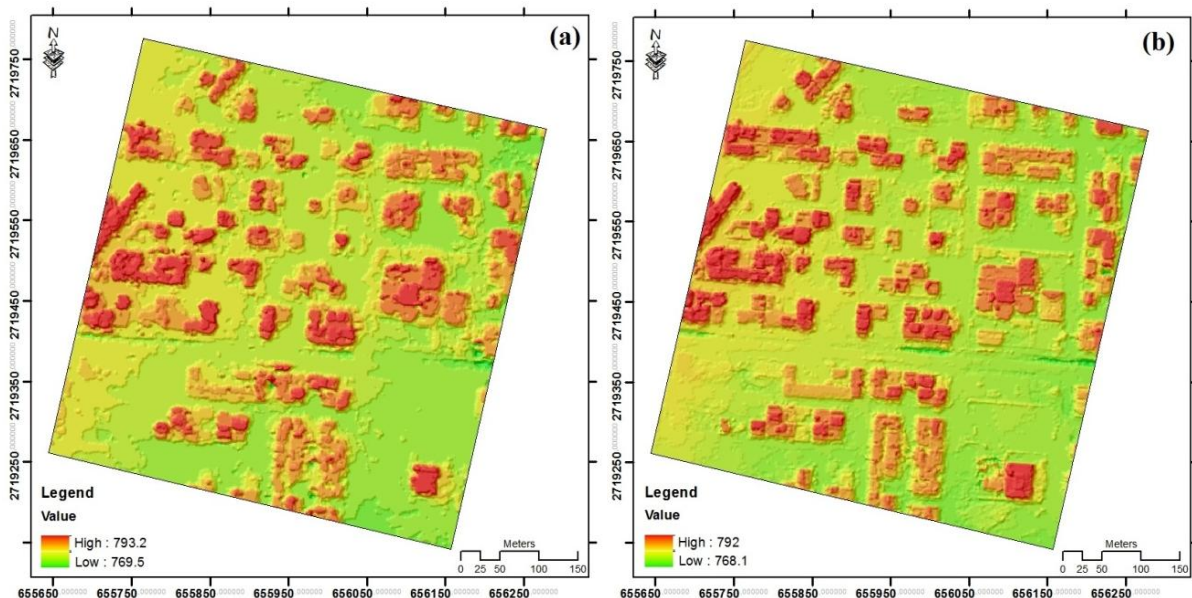
Figure 30- Derived DSM using SGM from (a) IKONOS stereo pair, (b) GeoEye-1 stereo pair

Since the orientation is carried out without any ground control points (GCP), it is required to eliminate shift errors in the DSMs before performing the background subtraction. To do so, we thus applied a shift in all three coordinates (*x, y, z*) to the second DSM with respect to the first. It is computed

automatically using three parameters based on the 3D least squares adjustment (LSA) model as per (Heipke *et al.,* 2002). The shifting is amounted to 7.2 m in *x*, 1.7 m in *y* and 1.3 m in *z* directions.

5.1.4. Generation of normalized digital surface model (nDSM)

In both cases, either for existing buildings or for new constructed buildings, a normalized digital surface model (nDSM) is required. It is generated through DSM filtering in order to compute a DEM first, and then followed by a pixel-based subtraction of DEM from DSM that results in an nDSM.

DEM is generated based on a local roughness index, and then each pixel is classified as a "*rough*" or "*smooth*" object. Subsequently, a region growing approach is carried out for the grouping of DSM pixels to on-terrain (flat) and off-terrain (above-ground) segments.

Finally, the morphological opening is applied to keep the main structural elements and eliminate very small objects showing salt-and-pepper noise (Niemeyer, *et al.,* 2010). We employ a binary mask to keep only those pixels on the nDSM which have absolute height values larger than 2.5m. Such regions of interest (ROIs) contain potential building blobs at the first and second epochs.
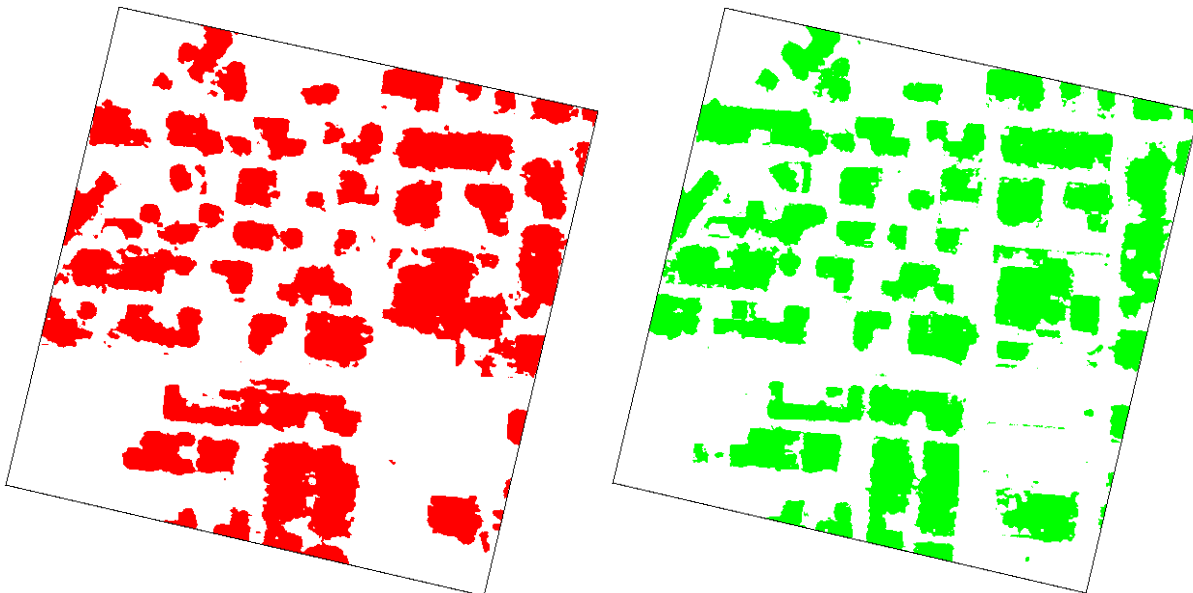
Figure 31- nDSM of epoch 1(left, red) and epoch 2 (right, green) as binary maps after applying a height threshold of 2.5m

5.2. Building change detection

5.2.1. Vertical change detection

This section describes the detection of building changes by comparing the DSMs of the two different epochs. After applying 3D least squares adjustment some vertical changes - those that stemmed from shifting due to false coordinates- are eliminated, however, it is required to refine these vertical changes to obtain the building changes.

As described in chapter 4, the initial D-DSMs are computed by the direct subtraction of DSMs. A threshold of 2.5m for the absolute difference was then used for each pixel individually, which as expected significantly reduced the amount of potential false changes. This binary thresholding

discards all non-building changes over the vertical change map. The resulting binary D-DSMs are depicted by Figure 32(b). A first visual inspection reveals that while most building changes are contained in the D-DSMs, there are also a number of false alarms. Some refer to potential building activity and represent excavations; others refer to matching errors, probably due to poor contrast.
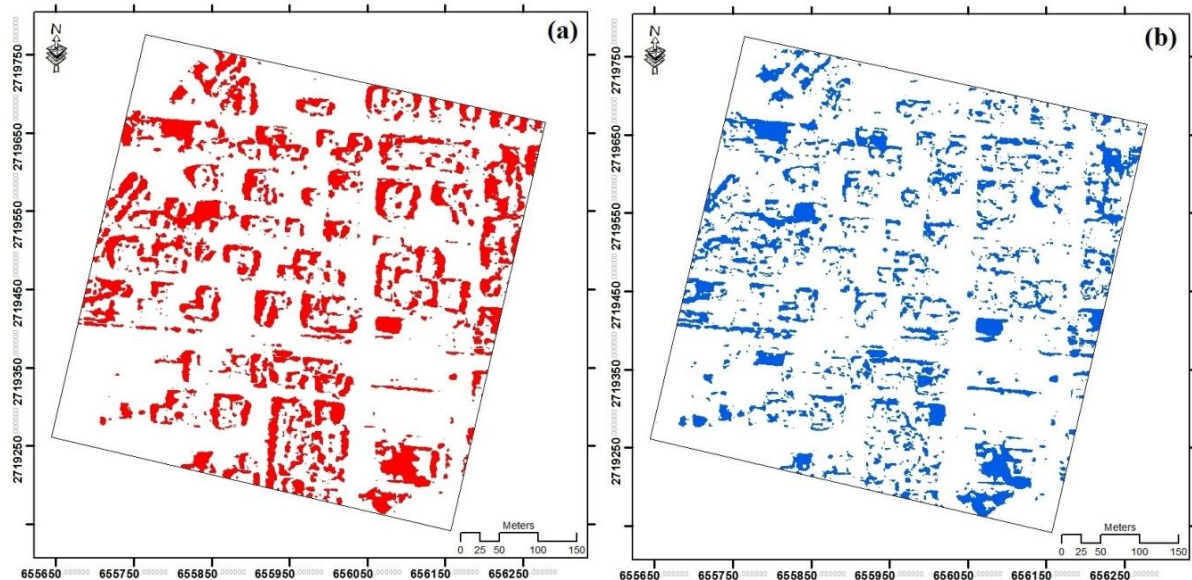
Figure 32- Initial binary change maps with the D-DSMs larger than 2.5m: (a) before (red), and (b) after (blue) shift elimination

An example of these errors within D-DSMs is shown by Figure 33. The green boxes superimposed on the GeoEye-1 images show correct building changes, but the red boxes are false alarms: the upper one is due to a group of parked trucks as well as poor texture over an asphalt road and the lower one belongs to a dump caused by a building under construction. As described in chapter 4, to eliminate such vertical changes from the real building changes another threshold of 2.5m is applied to nDSMs that could successfully separate building changes within all vertical changes.
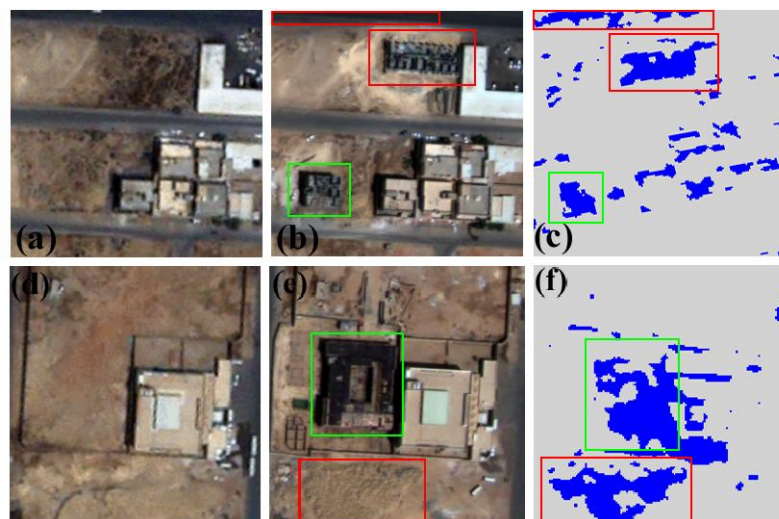
Figure 33- Two samples of the D-DSMs filtering based on height thresholding -
left: IKONOS image (epoch 1), centre: GeoEye-1 image (epoch 2), right: D-DSMs.

Up to now, an initial building change map is derived but as previous work has shown (Dini, *et al.,* 2012) there are still a considerable number of false positives which indicate building changes incorrectly. To eliminate these false alarms, three removal masks and morphological filtering (based upon size and shape) are applied as described in chapter 4.

5.2.2. Refinement of vertical change using removal masks

To refine the initial change map and eliminate false alarms, three removal masks are introduced, followed by a morphological opening in order to detect building change alarms. Figure 34 demonstrates two DSMs from left-to-right and right-to-left matching using GeoEye-1 stereo pairs.
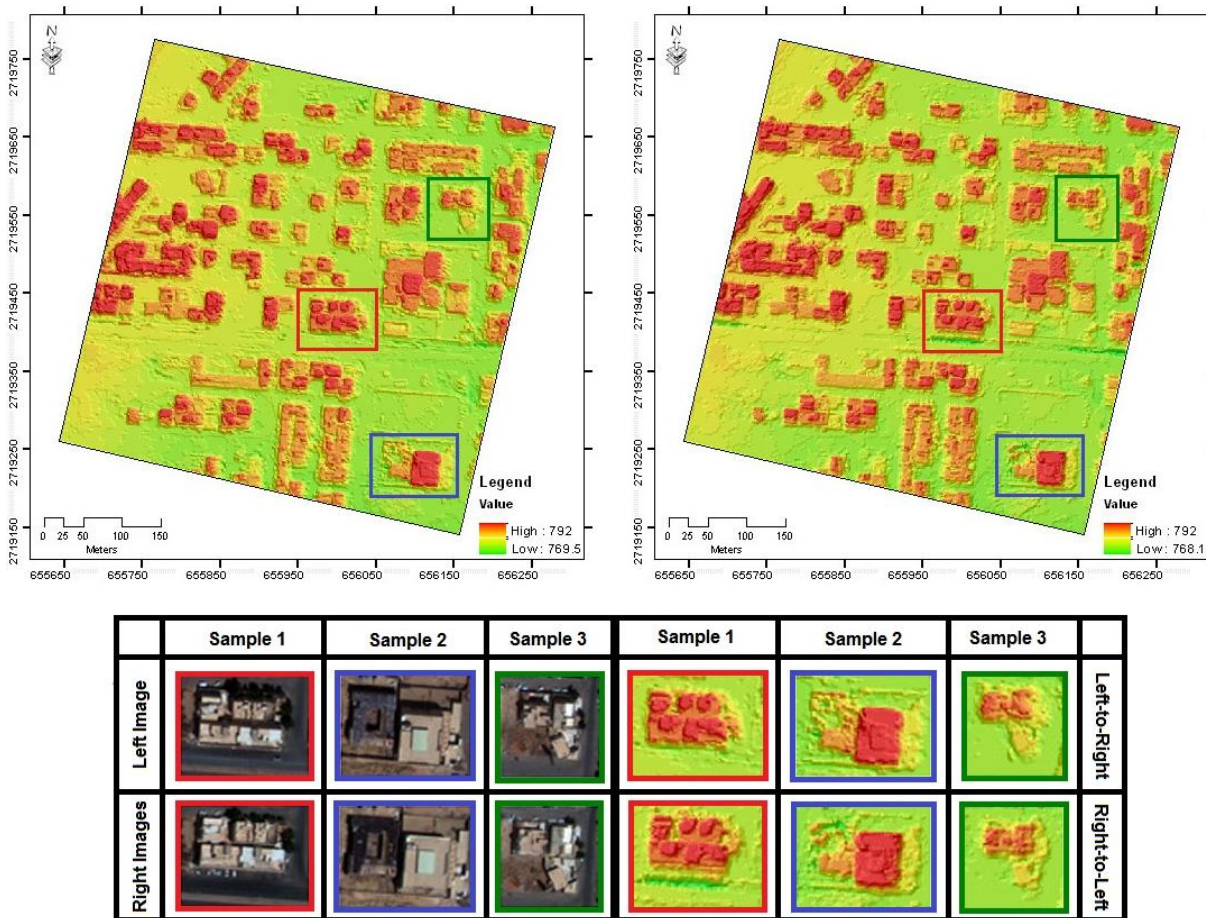
Figure 34- Matching-based DSM (left): left-to-right matching, (right): right-to-left matching, (down): showing buildings shapes for three different samples (shown by red, blue and green colors) overlaid on stereo pairs of GeoEye-1 and their associated DSMs

Matching errors are detected based on the subtraction of two DSMs from left-to-right and right-to-left matching (see Figure 34). A visual inspection reveals that large buildings and also buildings with good contrast show the same footprints in both DSMs, while small buildings or those that have very poor contrast demonstrate a slight change in the disparity and subsequently in the associated DSM. This becomes clear when sample 2 of Figure 34 is compared against sample 1 or 3. Sample 1 of Figure 34 also demonstrates the influence of trees that enlarge building sizes over the associated DSM.
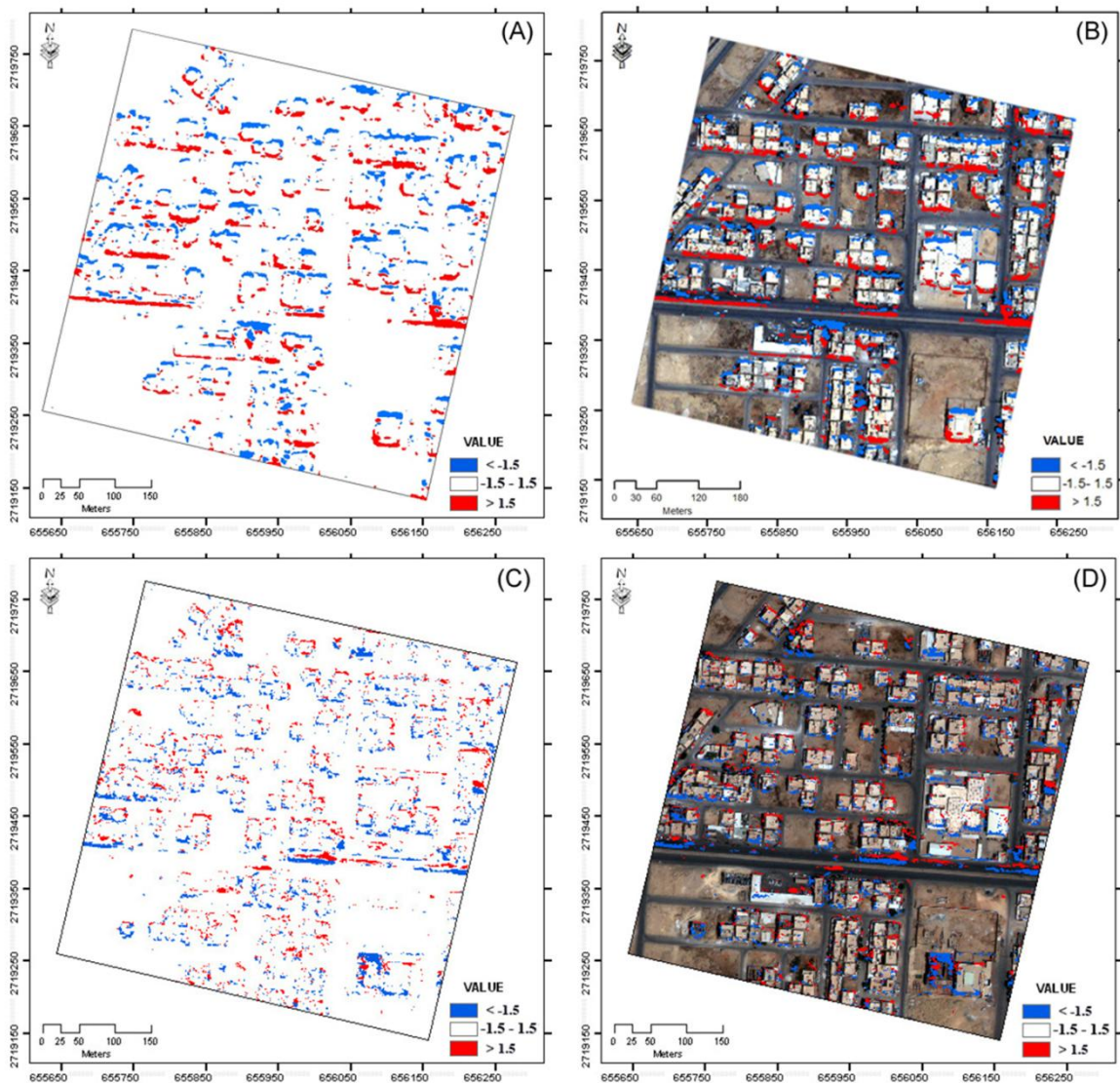
Figure 35- Mismatch masks derived from the difference between the left-to-right and the right-to-left matching (A): from IKONOS stereo pair, (B) superimposed to one IKONOS image, (C): from GeoEye-1 stereo pair, (D) superimposed to one GeoEye-1 image

Figure 35 shows the mismatch masks (difference of left-to-right and right-to-left) for GeoEye-1 and IKONOS. Only those pixels with an absolute value greater than 1.5 meters are assumed as mismatched pixels. Those pixels with a minus value are shown by blue and those with a plus value by the red color. As expected, analysis of the mismatch masks of IKONOS and GeoEye-1 reveals that the major matching errors are located on asphalt roads (particularly in the IKONOS mask) while trees and occluded areas around building outlines also cause problems. Hence, by discarding errors stemming from asphalt roads (road mask) and vegetation, the mismatch masks for IKONOS and GeoEye-1 are almost the same, containing only elongated errors around buildings caused by occlusion. It is probably because of the similar image acquisition conditions for both IKONOS and GeoEye-1.

Figure 36 compares the amount of mismatched pixels for both the left-to-right and right-to-left matching of GeoEye-1 stereo pairs. The mismatched pixels are classified into three classes according to the standard deviation of each pixel within a 1×1 meter grid. In order to subtract the left-to-right

DSM against the right-to-left one, all matched pixels (with the ground sampling of 0.5 meters) are intersected with an empty 1×1 meter grid. Therefore, each cell contains approximately four pixels. The standard deviation of these pixels within each cell demonstrates the consistency of matched pixels within a cell for left-to-right and right-to-left matching. As shown, the amount of mismatched pixels in the DSM derived from the right-to-left matching is considerably more than the left-to-right one. It indicates that for the generation of D-DSMs, the DSM from left-to-right matching delivers a better result. It shows that matching is not a symmetric process when searching for conjugate pairs, as only the visible pixels in one image are matched. If the role of left and right images is reversed, new conjugate pairs are found (Fusiello, *et al.,* 1997).
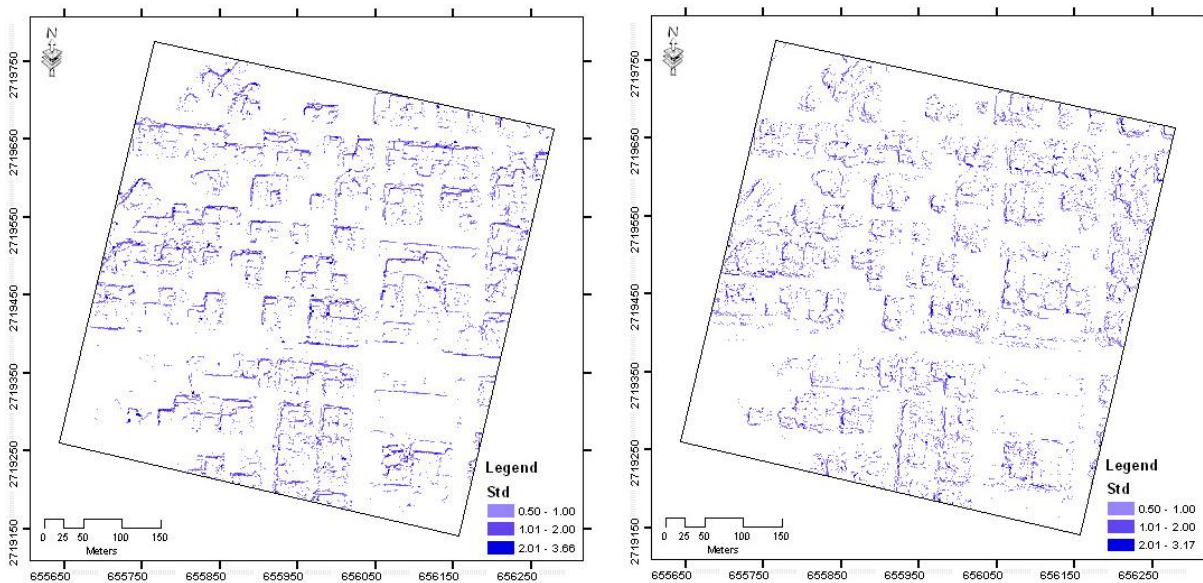
Figure 36- Standard deviation of matched points for the GeoEye-1 stereo pair within a 1×1 meter grid (left): on left-to-right image matching, (right): on right-to-left image matching

Figure 36 reveals two points (i) the number of mismatched points with left-to-right consistency is considerably less than the right-to-left ones. That means it is better to use left-to-right image matching to generate DSM and subsequently D-DSMs (ii) the mismatched areas are usually around building outlines, however additional mismatches because of asphalt reflection and vegetation have to be dealt with. Mismatched areas are mainly caused by occlusion (often along buildings outlines) and homogenous areas. To eliminate matching errors from D-DSMs, un-matched areas are intersected with D-DSMs and then overlapped areas are removed from the potential building changes.

In addition to the matching errors stemming from occlusion, un-textured regions (e.g. asphalt roads, bright roofs, etc.) also cause some problems. As stated in chapter 4, a buffer of 1 meter around road polygons in GIS results in a road mask which should be removed from D-DSMs. The road mask superimposed on one of the GeoEye-1 images is shown by Figure 37.
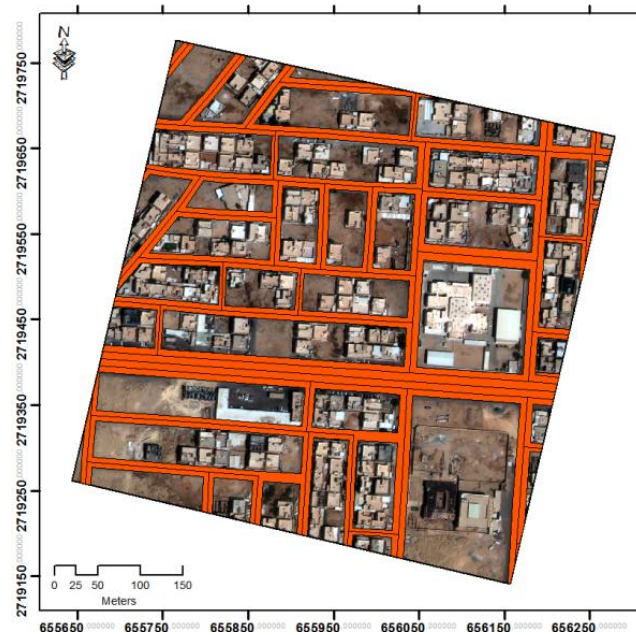
Figure 37- Road mask, superimposed to one GeoEye-1 image

The third mask is a vegetation mask which is generated by supervised maximum likelihood classification (MLC) using pan-sharpened image and the nDSM as an additional channel. Two classes, namely vegetation and non-vegetation, are defined via carefully selected training sites (see Figure 38). After applying these three removal masks to the D-DSMs, the remaining blobs still contain salt-and-pepper noises. In fact, there are a number of very small changes that cannot be considered candidates for building changes because of their shape (elongated) or size (small blob) which cause false alarms.

Figure 38- Vegetation mask derived from MLC classification using (left) IKONOS image, (right) GeoEye-1 image

These errors are the main drawbacks of background subtraction. Therefore, fundamentally a further refinement is required in order to eliminate these false alarms from new constructed buildings. These

blunders (small and elongated shape blobs on D-DSMs) are removed using a morphological opening filter as described in chapter 4.

However, morphological filtering deforms the buildings shape somewhat with respect to the filter size but it also effectively eliminates some salt-and-pepper noise caused by limited ground sampling (at the first epoch) as well as the influence of moving objects (e.g. cars and trucks). In our dataset, the filter size is set to 4m. This value was determined empirically. Figure 39 shows the final building changes as red blobs.
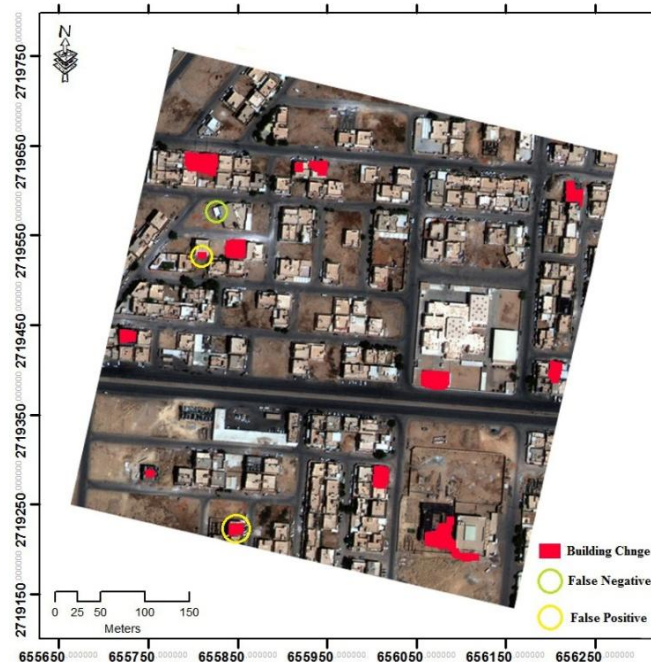
Figure 39- Final change map after refinement by morphologic opening (size and shape refinement)

We have only two misdetections and one missed building out of 12 building changes. Note that box fitting or any other means used to consider the fact that buildings are usually rectangular have not yet been applied. This alarm system is used to detect building change in the region of interest, then in the next sections 3D line matching and 3D segmentation are employed to delineate building outlines.

5.2.3. Verification of existing buildings on GIS database

As shown in the previous section, image-to-image comparison (IKONOS DSM vs. GeoEye-1 DSM) followed by removal masks demonstrated building changes, however, two false alarms and one misdetection were also observed. In addition to image-to-image comparison as explained in the previous section, this section describes the experimental results of the verification of GIS polygons using building blobs in the GeoEye-1 nDSM.

To perform image-to-map comparison, first the existing GIS polygons are compared against the GeoEye-1 nDSM. If a polygon is verified, the verified building polygons show the outlines for the associated building blobs in the nDSM. For image-to-map comparison, at first pixels with an absolute height greater than 2.5m on the GeoEye-1 nDSM generate the initial building blobs and then removal masks are applied to eliminate false alarms. After applying the three removal masks to the blobs, the

remaining blobs are refined using morphological filtering then they are compared against the GIS building polygons. Polygons which are covered at least 75% by of one or more blobs are considered as verified polygons and so are classified as *existing buildings*. The remaining blobs represent candidates for new buildings. In Figure 40, the yellow polygons represent building outlines of the GIS database, the superimposed green areas show accepted buildings and blobs shown by the red color are classified as *new buildings*.
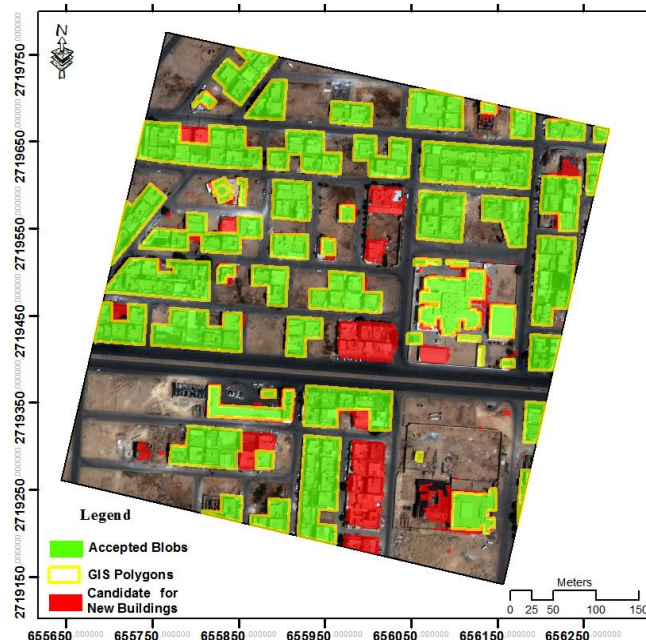
Figure 40- GIS polygons (yellow), verified buildings (green), blobs representing new building (red), superimposed to one GeoEye-1 image

After the identification of existing buildings, GIS polygons are assumed as the footprints for the existing buildings. Thus the verified polygons on GIS are used to be replaced instead of the 2D outlines of building blobs in the nDSM because as shown by Figure 41, in the matching-based DSM the buildings shape are deformed.

Figure 41- Buildings 3D visualization based on the DSM derived from SGM
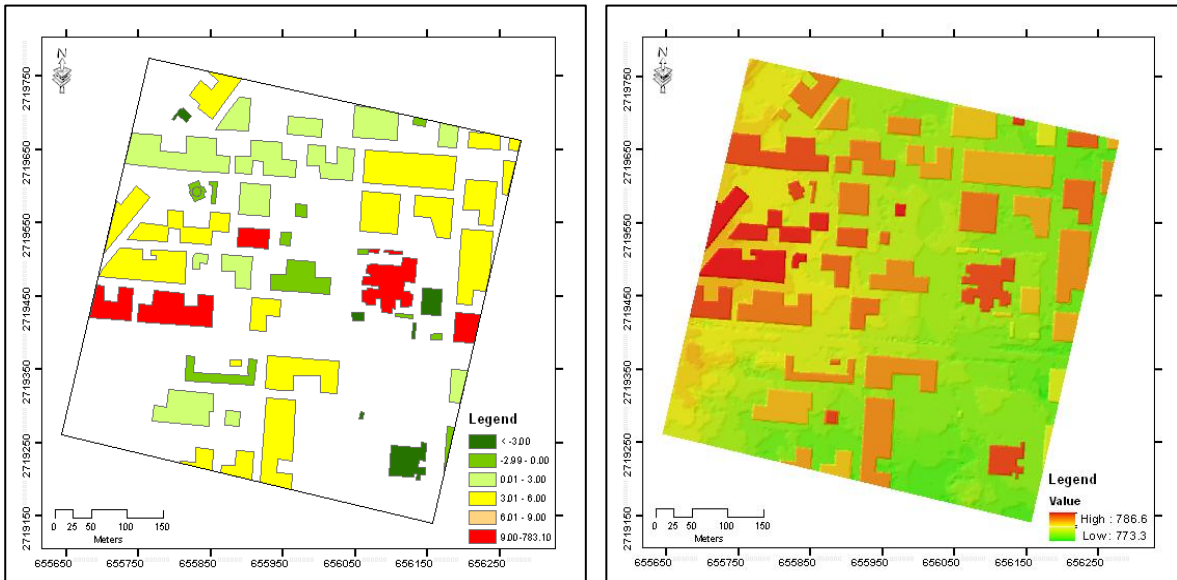
Figure 42- (left): Difference of building height between GIS database and DSM extracted from SGM, (right): enhancement of DSM derived from SGM using buildings polygons of GIS database

In addition to the building footprints, the height value of the GeoEye-1 nDSM is compared against associated attributes of the GIS database. Unfortunately, the height value for a number of buildings in the GIS database is null and while the other buildings show the height value as somewhat similar to the GeoEye-1 nDSM. Hence, the derived height value from SGM is more updated than the GIS database. For the updating of height attributes in the GIS polygons, the mean of height value from nDSM is assigned into the GIS polygons as the updated building height (see Figure 42, left). However, this approach smoothens the height value for different parts of rooftops (e.g. small lofts) representing flat rooftops.
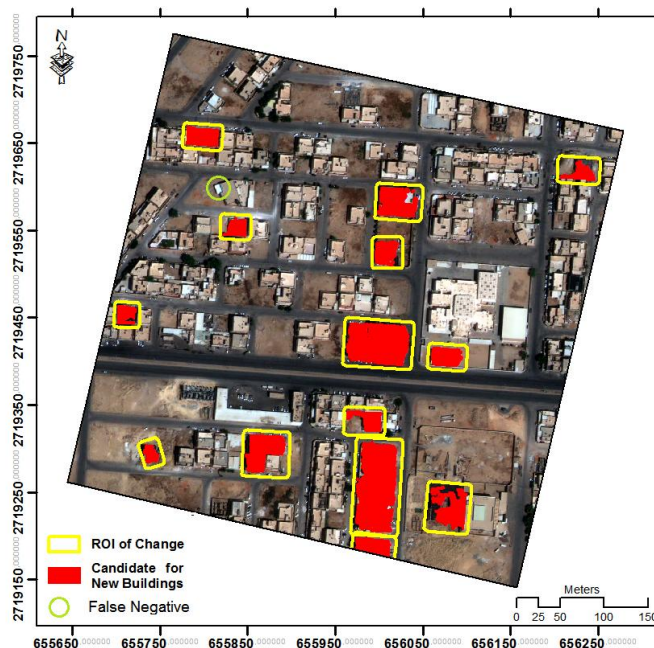
Figure 43- Final change map (comparison of GIS polygons against GeoEye-1 nDSM); areas of change potentially containing new buildings (red), superimposed to one GeoEye-1 image

Figure 43 demonstrates the final building changes that are obtained by comparison of the GeoEye-1 nDSM against GIS polygons. The yellow polygons are fitted as bounding boxes over building changes, being ROIs for building outline delineation. As shown here there is no false alarm. Similar to the IKONOS-GeoEye-1 comparison, there is only one false negative case which is a very small annex added to an existing building.

Figure 44- Building 3D visualization based on the DSM enhanced with GIS polygons

Figure 44 visualizes the study area using enhanced DSM three dimensionally (with buildings outlines from the GIS and height value from the nDSM). It obviously shows (also by Figure 43) that the enhanced building footprints have obtained very sharp outlines compared to the original DSM derived from SGM. Note that the new constructed buildings are not applied.

5.3. Updating of buildings outlines

Up to this stage, the region of building changes has been detected, as shown in Figure 43, and existing buildings are also updated (both 2D building footprints and height value). In order to delineate the outlines of new constructed buildings a buffer with a size of 0.5 meters is performed on the region of change. This buffer ensures that the boundary of new constructed building is within the given bounding box (see Figure 43). Sections 5.3.1 and 5.3.2 show several experimental results of 3D edge matching and 3D segmentation, respectively.

5.3.1. Delineation of new buildings using edge-based approach

In order to find corresponding matches within a pair-wise 3D line matching approach using various constraints, it is required to extract straight lines in both the epipolar images. First, using Canny operator, all the edges are detected on each epipolar image individually (Figure 45 left) then the edges with very small lengths are eliminated (less than 3 meters). Using a PCA-based method the straight parts of the remaining edges are extracted. Finally, as described in section 4, by applying the Douglas-Peucker algorithm, a straight line is fitted to the associated edges (Figure 45 right).

Figure 45- Straight line detection: (left) edge detection using the Canny algorithm and then finding straight edges by a PCA-based method; (right) fitting a straight lines on edges based on the Douglas-Peucker line simplification

After the detection of straight lines in the epipolar images, the first epipolar constraint is applied in order to collect the initial correspondence. This constraint is used to reduce the search space with respect to the epipolar geometry and the associated height of a building causing a parallax. The yellow box in Figure 46 shows how the epipolar constraint constructs a parallelogram to reduce the search space. As shown, the straight line in the left image is simply deleted from the further matching process, because there is no corresponding line in the right image within the associated parallelogram.
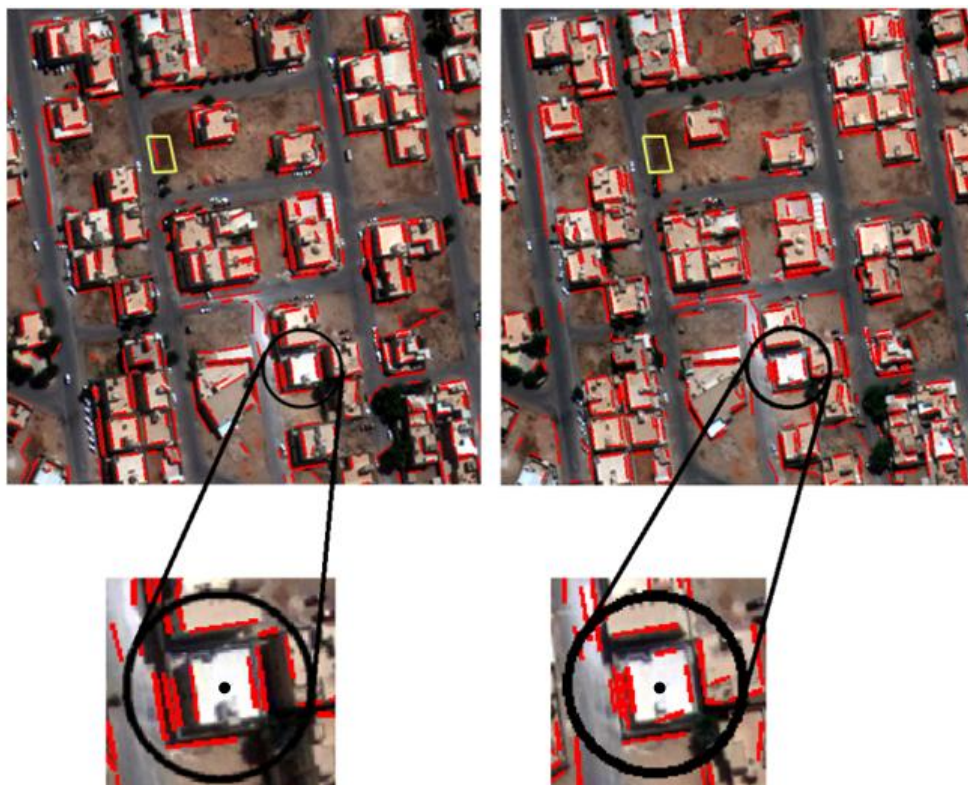
Figure 46- Matching of straight lines in epipolar images; the yellow parallelograms show how the epipolar constrain reduces search space; the black circles show ambiguous cases and how proximity constraint detects the corresponding edges (for further explanation see text)

The orientation angle, line length and line mid-point are used to detect the best corresponding match resulting in initial corresponding matches (between the epipolar images). After applying epipolar and geometric constraints, if the matching could successfully establish a one-to-one relationship between

corresponding cases, the matching is finished for aforesaid corresponding cases therefore a further matching process is not required for these lines, otherwise there are ambiguity cases which require further constraints.

To solve the ambiguity cases, the proximity constraint is applied to find the best corresponding case for a given ambiguous case. Figure 46 shows an ambiguous case. It clearly shows the mechanism of proximity constraint and how it finds real building outlines among numerous candidates (ambiguity cases) by choosing an edge with the shortest Euclidian distance to the centre of building. As shown, again most false matches are caused by occlusion or shadow.

Figure 47 shows the final results of 3D edge matching. The green lines successfully found a proper match; whereas lines which could not be matched are shown by the red color. As shown, only building footprints in the first and second columns of Figure 47 are successfully extracted and the others are missed, because at least three sides of the building outlines are not found (a condition required for footprint extraction using 3D edge matching). The last row demonstrates the superimposement of matched and unmatched edges over the nDSM. It clearly shows that, due to the smoothing effect of matching-based DSM, the height information cannot provide further useful information as an additional constraint.

Our results indicate that if buildings have more complex structures or there is not a good contrast between building rooftops and their backgrounds, no homologous lines can be found. However, for most buildings with simple structures and a good contrast, 3D edge matching is able to delineate building outlines. Additional checking of matched lines against DSM also confirmed this point (see Figure 47). Furthermore, the experimental results have demonstrated that in very complex building structures, only the proximity constraint is able to suppress the ambiguity cases effectively, as the other constraints (including height) are not effective here.
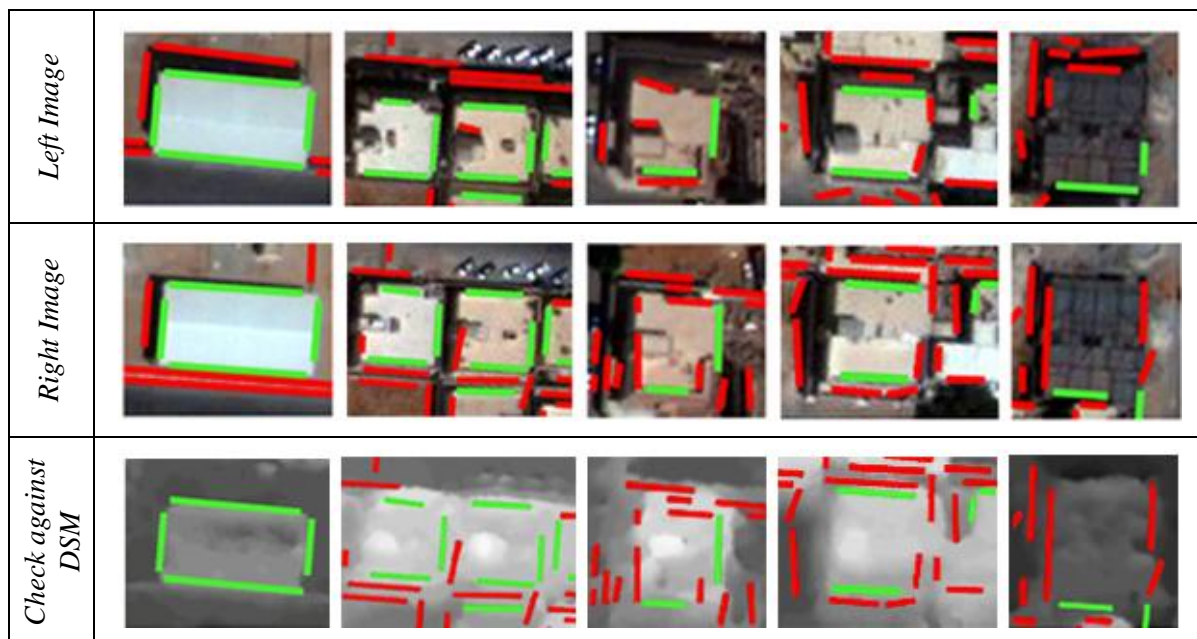
Figure 47- Samples from matched (green) and un-matched (red) straight lines. Only building footprints in the first and second columns could be reconstructed successfully

The 3D edge matching was successfully able to reconstruct building shapes using very high resolution aerial images (Baillard, *et al.,* 1999). However, using our dataset it gained little success. This is probably due to the low ground sampling of our dataset which is comparatively less than that of very high resolution airborne imagery. Figure 47 shows the building outlines that are detected using 3D edge matching (red lines). If at least three sides of a rectangular building footprint are detected completely, the box fitting algorithm can simply reconstruct the associated rectangle.
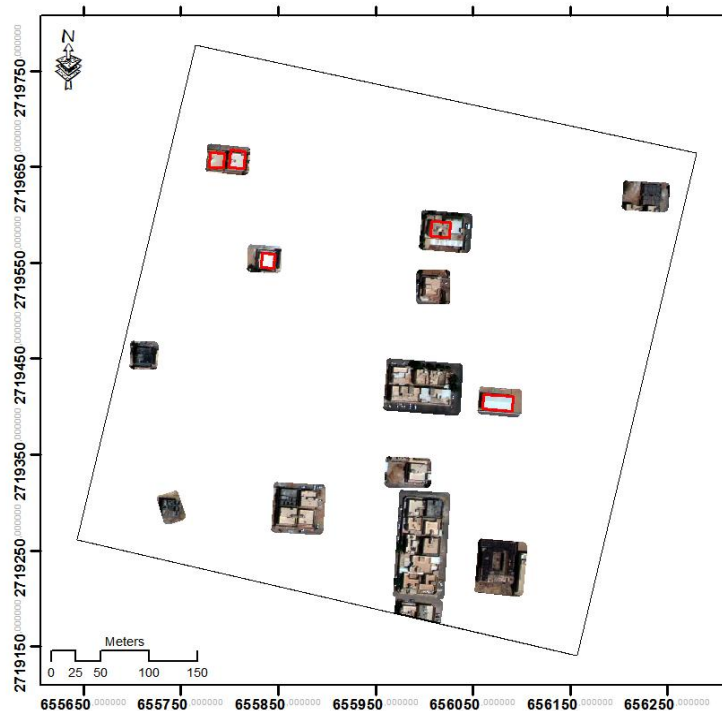
Figure 48- Outline delineation for new buildings based on 3D edge matching

Among all new constructions, only four building footprints could be reconstructed based on 3D edge matching and the rest were missed. Also, one of them is also reconstructed using three sides, one of which is too short; therefore the final box is smaller than the real size of the building footprint.

5.3.2. Delineation of new buildings using region-based approach

This section demonstrates some experimental results of the 3D segmentation of building rooftops. As previously explained the height information is employed in order to locate building areas and to delimit segmentation into these ROIs (initialization). This strategy locates building areas so segmentation delivers a near to real building outline even after ten iterations (approximately over 70% overlap between evolving segments and the ground truth information – see Figure 25). It supports the segmentation procedure by conducting the segmentation into built-up areas and impedes the segmentation of non-building objects by taking into account building footprints.

This innovative solution adopts the region-based segmentation for building footprint extraction according to (Chan, *et al.,* 2001); hence we perform the initialization using height information (nDSM) so the segmentation is conducted into building footprints only.

Figure 49- Delineation of building footprints using 3D segmentation (left): segmentation of building rooftops, (centre): building outlines superimposed to original image, (right): building outlines superimposed to DSM derived from SGM

The segmentation is carried out over epipolar images and if a given rooftop segment is observed within both epipolar images, it confirms the building footprint. However, if the final segment is chosen from the image with the smaller incidence angle (close to the nadir view - in our dataset it is the right image) then the derived building rooftops are regularized into rectangular building footprints. Figure 49 shows rooftop segmentation over the GeoEye-1 image followed by the regularization of footprints into rectangular shapes.
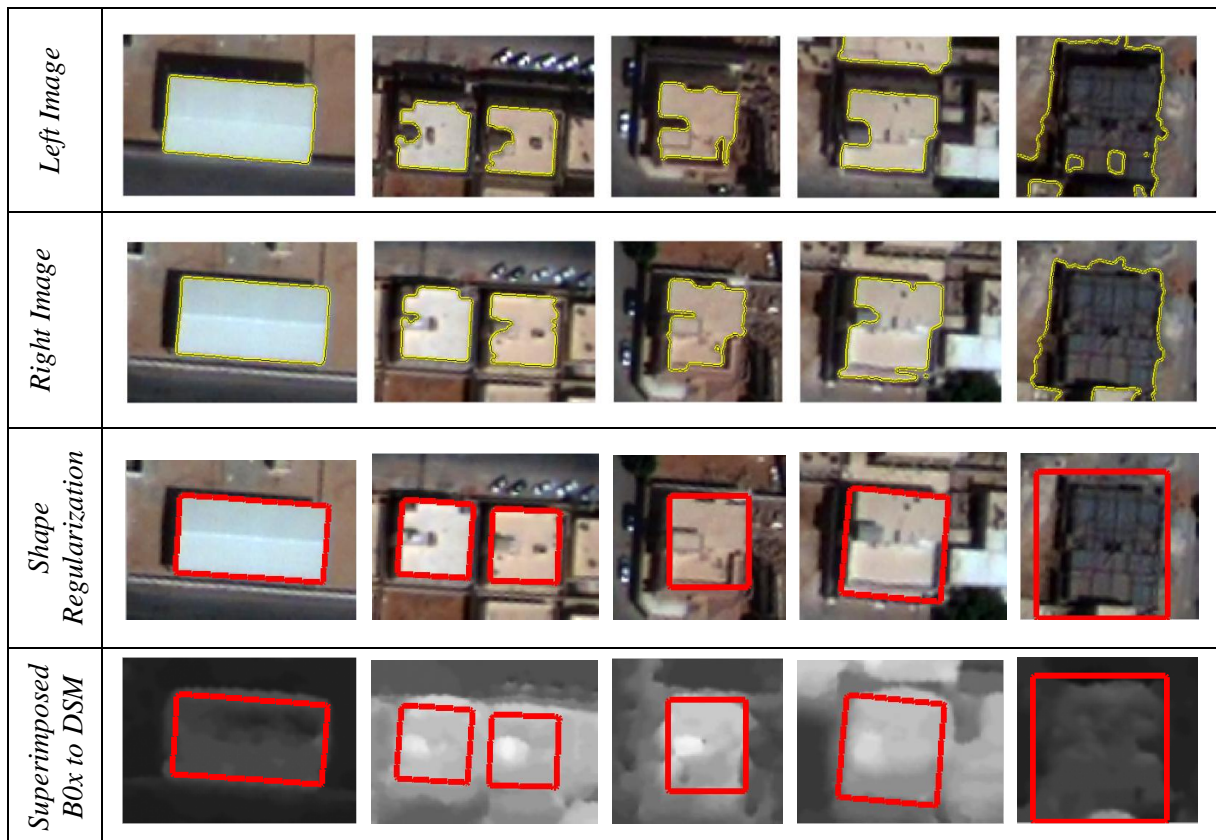
Figure 50- 3D segmentation of building rooftops and regularization into rectangular shapes superimposed to one of epipolar images and associated DSM, respectively

One of the main challenges for the delineation of building outlines using segmentation is the determination of a unique iteration number for the segmentation of all rooftops within a scene. With respect to the different structures of building rooftops as well as the level of building complexity, it is

relatively difficult to determine a unique iteration number (Khoshelham, *et al.,* 2005). Overestimation and underestimation of iteration numbers cause over-segmentation and under-segmentation, respectively. Besides, as the main goal is to develop an automatic solution, determining the iteration number for each single building rooftop is not a suitable solution. Hence, as stated in chapter 4, the iteration number is computed for 11 sample rooftops with different shapes, sizes, and radiometric reflections as well as for different types of building. As estimated (also shown by Figure 25), the 254th iteration provides the best overlapping between rooftops segments and manually delineated building footprints. Figure 50 also demonstrates the efficiency of the chosen value.

In the 3D segmentation of rooftops, the image information is used (instead of height information) to delineate building outlines. Thus, the delineation is not dealt with the smoothing effect as matching-based DSM. It becomes clear when Figure 49 (middle) is compared with Figure 49 (right). As demonstrated, the enlargement of building footprints on the DSM is due to trees that are close to buildings (green arrows), and mismatched areas stemming from occlusion (yellow arrow) and shadow (black arrow). Figure 49 shows these influences superimposed on to the original image and the associated DSM, respectively.

Figure 50 shows several samples of building footprint extraction based on 3D segmentation over the same area as the 3D edge matching shown by Figure 47. The 3D segmentation successfully delineated the building outlines for all samples except the last column of Figure 50 while the 3D edge matching failed to find corresponding edges and consequently could not deliver a promising result for the third, fourth and fifth columns of Figure 47. Nevertheless, similar to edge matching, the 3D segmentation also has not successfully delineated building outlines for the last column because it shows a very poor contrast against the background. The last row also demonstrates the building footprints superimposed on to the associated DSM. As shown, because of smoothing effects, the height constraints are not also able to suppress an unsuccessful result (as in the last column) to fit an appropriate rectangle.
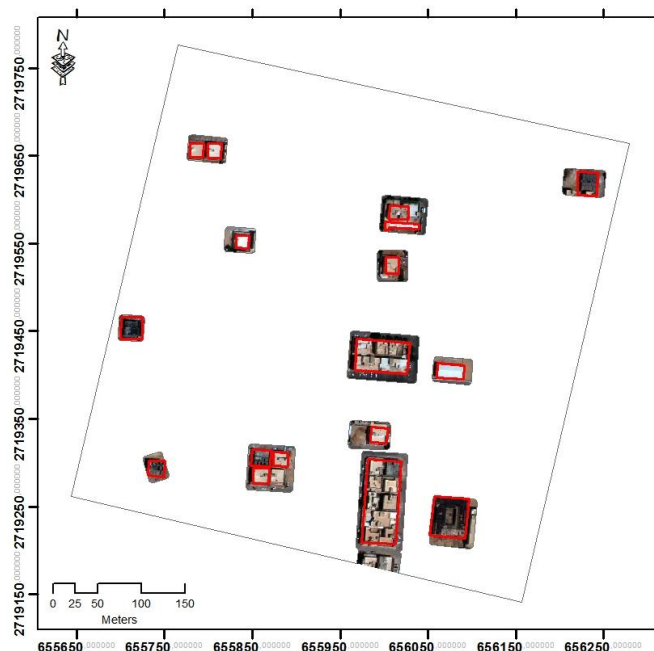
Figure 51- Outline delineation for new buildings using 3D segmentation

The next point is that 3D edge matching usually fits a smaller rectangle than the real size of a building, if one side of the building footprint is detected partially (shorter than its real size), but with the 3D segmentation of rooftops an extracted building footprint is somewhat bigger than the real size of a building, because if in case a building does not have a sharp contrast compared to its background (i.e. the building rooftop is not satisfyingly homogenous, e.g. the last sample in Figure 50), segmentation might continue into the background area. However, the building size cannot be bigger than the nDSM blobs derived from SGM as the segmentation is already restricted within change alarms (see Figure 43& Figure 51).
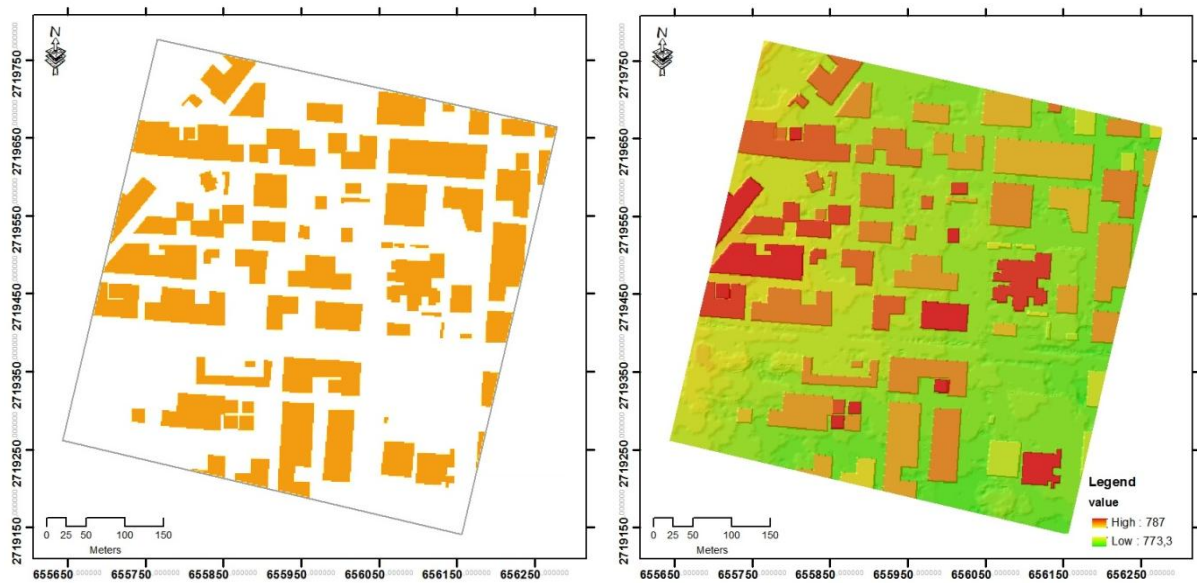
Figure 52- (left): Final building footprints (right): Final DSM which is enhanced using building polygons in the GIS and 3D segmentation of building rooftops

Figure 52 shows the final up-dated building footprints and the associated DSM which is refined by GIS polygons and the 3D segementation of building rooftops over GeoEye-1 images. The 3D visualization of building footprints shown by Figure 53 also demonstrates that, compared to Figure 41, the building shapes are considerably regularized showing clear building outlines.

Figure 53- Buildings 3D visualization using final enhanced DSM

Furthermore, the mean the of height value from the nDSM is assigned to each polygon (either GIS polygons or new building footprints derived from 3D segmentation). Hence, not only building outlines in 2D space but also the height value of buildings are clear for each building footprint, representing a promising result for building updates in LoD 1.

5.4. Quantitative evaluation of the results

In order to evaluate the efficiency of our proposed approach independently at each stage, we captured the vegetation masks as well building changes for two scenarios (IKONOS vs. GeoEye-1 and GIS polygons vs. GeoEye-1) manually. The following parameters are used to compute quantitative evaluations of building change detection for two different scenarios (IKONOS vs. GeoEye-1 and GIS polygons vs. GeoEye-1).

True Positive (TP): Both obtained and ground truth data demonstrate a building change.

True Negative (TN): Both obtained and ground truth data demonstrate a non-building change.

False Positive (FP): Only the obtained method demonstrates a building change.

False Negative (FN): Only ground truth data demonstrates a building change.

IKONOS vs. GeoEye-1: In order to evaluate the role of removal masks for building change detection, these indices are calculated for two scenarios (i) without applying removal masks and (ii) with applying removal masks. Table 4 demonstrates the quantitative evaluation for the building change detection phase. As shown by applying these removal masks, the number of false alarms is decreased significantly.

Parameter	TP	FP	FN	Sum
Building Change Detection without Removal Masks	11	5	1	17
Building Change Detection with Removal Masks	11	2	1	14

Table 4- Quantitative evaluation of building change detection (IKONOS vs. GeoEye-1)

The ground truth contains 12 newly constructed buildings but without any demolition. As demonstrated, we could automatically detect 11 out of the 12 new constructions. Note that there is only one FN (missed change) in our results which is a small annex added to an existing building. Because of its very small size, this annex is absorbed into the background pixels. Importantly, applying removal masks could successfully decrease the number of false detections (FP). Therefore, comparison of (i) and (ii) confirms that applying these masks increases the efficiency of the proposed approach for it to be a promising alarm system for building changes.

GeoEye-1 vs. GIS Polygons: In the second scenario building blobs derived from the GeoEye-1 nDSM are refined using removal masks and subsequently morphological filtering. The obtained blobs are then compared against building polygons in the GIS database. A TP was declared if 75% of the area of

the new building was covered by pixels indicating change. The small annex which is added to an existing building is here missed (FN) but the most important point is that there is no false detection (FP) as observed in the first scenario while there are no false alarms for the second scenario.

Parameter	*TP*	*TN*	*FP*	*FN*	*Sum*
GeoEye-1 DSM vs. Building Polygons in GIS	14	56	0	1	71

Table 5- Quantitative evaluation of building change detection (GIS Polygons vs. GeoEye-1)

14 building changes are detected correctly (TP). 56 building polygons in the GIS database are verified using the GeoEye-1 DSM (i.e. they did not change and are called "True Negative" according to the introduced terminology). As shown by Table 5, there is no false change detection. This was the main difference between the first and second scenarios.

In terms of the evaluation of 3D line matching and 3D segmentation, as shown in previous stages, 3D segmentation can successfully deliver promising results for 13 out of 17 building footprints except for 4 buildings which have very poor contrast. Because of the very poor contrast, segmentation is not terminated at the rooftop outlines, hence the building footprints are enlarged for these buildings (as well as SGM), while 3D edge matching was only able to delineate buildings outlines for 4 buildings. This is mainly because of the limited ground resolution which is not sufficient to detect building edges therefore building edges are not detected in both epipolar images, whereas 3D segmentation does not have such a dependency.

5.5. Summary

The results of 3D building change detection (as an alarm system) as well as for building footprint extraction for new buildings are presented. As the study area is a suburb of a developing city, there were no demolition cases and all the building changes indicate new constructions. The experimental results demonstrated that the subtraction of matching-based DSMs acts fruitfully as a reliable alarm system to detect building changes providing that the removal masks eliminate the associated matching errors.

For the delineation of building footprints using space-born stereo images, 3D edge matching has been shown to have a fundamental difficulty with the extraction of building outlines. It is mainly due to the very limited ground resolution of satellite images which is not sufficient for building extraction. However, in a previous study (Baillard, *et al.,* 1999), 3D edge matching successfully reconstructed building footprints using very high resolution aerial images (e.g. using polyhedral models). As shown in our tests, compared to the 3D edge matching, the 3D segmentation delivers comparatively a good result for building footprints extraction.

Furthermore, as our results have shown that robust geometric criteria such as elevation have priority to the radiometric information in the detection phase, because change detection based on image information is considerably dependent upon illumination conditions (Champion, 2007). On the other hand, for building footprints extraction image information is relatively more useful than matching-

based DSM. Table 6 compares the pros and cons of each approach for building change detection as well as for building footprints extraction using high resolution spaceborne images.

Subject	Pixel-Based	Edge-Based	Region-Based
Methodology	DSM generation using SGM DSMs subtraction refined with noise removal filters Thresholding on nDSM	3D edge matching of epipolar images to extract building outlines Box fitting	Locating buildings ROIs from nDSM Roof segmentation to delineate building outlines Regularization of obtained segments into rectangular shapes
Advantage	Reliable method compared with building detection using image information Useful as an alarm system for building changes	Building outlines are clear Delineate building outlines precisely	Building outlines are clear Delineate building outlines precisely Its efficiency is independent from buildings size and shape
Disadvantage	Needs a fine GSD Building outlines are jagged Building footprints are enlarged Needs removal masks to deliver promising results	Needs a fine GSD Miss small buildings those have poor radiometric contrast compared to surrounding objects (background) Inapplicable for complex structures with few straight outlines	Inapplicable for buildings those have poor radiometric contrast compared to background In fitting box over segmentation, the main direction of building may deviate slightly

Table 6- A synopsis comparison of proposed approaches to update building databases using high resolution stereo images

CHAPTER 6

CONCLUSION AND FUTURE WORKS

6.1. Conclusion

In this thesis, we have described approaches for the automatic updating of building databases from stereo high resolution satellite images. The updating of building databases is divided into two main stages; *building change detection* and *delineation of building outlines*. Our experimental results have shown that the proposed method can be used as an alarm system for updating building geospatial databases. Comparison of GIS building polygons against building blobs derived from nDSM can efficiently establish a system to verify building changes. Meanwhile, 3D line matching and 3D segmentation are investigated to delineate the outlines of new constructed buildings in order to fulfil the second stage of updating procedure.

For the first stage, our approach is based upon background subtraction and foreground validation in order to detect building changes; however, we have also utilized *a priori* information derived from an existing building topographic database as well as building size, height, and shape information to refine building changes by discarding false alarms. As experimental results have shown, while direct background subtraction in height domain delivers a large number of false alarms, the removal masks could effectively eliminate most of these blunders and improve the quality of building change detection considerably. The previous work (Dini, *et al.,* 2012) with the same study area has shown five false alarms. However, after applying removal filters, the number of false detections is reduced to two cases only. Auxiliary information (e.g. GIS polygons) is utilized to refine building footprints in the DSM. In addition, if absolute orientation is not carried out using ground control points (GCPs), it is recommended to co-register DSMs against each other before background subtraction in height domain, otherwise the number of false alarms is increased considerably.

The results have demonstrated that the height information detects building changes with a better correctness where the radiometric image information is suitable to delineate building outlines. This was the main reason that in the proposed flowchart, the height information is used for change detection phase and image-based approach is applied for building footprints extraction. It is shown that although the matching-based DSMs contain more noise when compared with the DSM derived from laser scanning DSM, although by employing removal masks, it is capable of establishing a

promising alarm system locating building changes. That means image radiometric information along with the height information derived from image matching provides an updating system for building databases. Nevertheless each source alone -image or height information- has some difficulties to provide a high quality result.

On the one hand, in a noise-affected DSM derived from image matching, there is a considerable amount of height fluctuation even at flat surfaces or textureless areas such as asphalt roads (see Figure 41). Due to these artefacts stemming from image matching, the matching-based DSMs are not ideal for automatic extraction of building models if the level of detail is beyond LoD1 (Macay Moreira, *et al.,* 2013) as the building outlines within such DSMs are usually curved instead of straight outlines. On the other hand, using radiometric image information, it is possible to detect building footprints in complex urban areas as far as image information is not influenced by smoothing effects such as DSM. In contrast, it is not an easy task to locate building rooftops using image information alone because buildings have a wide range of reflectance so it is rather difficult to separate building class from non-building objects using image information only.

The assessment of our experiments shows a significant influence of trees in the building change detection. The same problem is also reported by (Rutzinger, *et al.,* 2010) when most misclassifications are observed around buildings that are adjacent to the vegetation. Parked cars or trucks pose fewer problems as they are relatively small and are not as high as trees. Moreover, if trucks are not close to the buildings, in the DSM, they are usually not mixed with building footprints.

Two scenarios (IKONOS DSM vs. GeoEye-1 DSM and GIS polygons vs. GeoEye-1 DSM) are tested to investigate the influence of image ground resolution in the building change detection. The results show that even using removal filtering there are still two false alarms for the first scenario while the second scenario demonstrated no false alarms.

The misdetection and false alarms in the tests are mainly buildings with very poor radiometric or contextual information as well as small buildings or building annexes. The matching-based DSM alone is not capable of delineating building outlines, particularly in the half-occluded areas around building outlines. In addition to the quantitative assessment, a visual inspection of building changes reveals that a wide range of noise removal filters are applied to refine the result of background subtraction but building outlines are not still robust to the occlusion. It is mainly because of several reasons: high level of noise within matching-based DSM, low radiometric image quality (pan-sharpened image) as well as low spatial resolution, particularly IKONOS images with 1 meter ground resolutions as reported by (Rottensteiner, 2008) and (Champion, *et al.,* 2010).

In image matching, the object boundaries and particularly building outlines are deformed (especially in IKONOS DSM). This problem is accumulated by background subtraction causing more deformation in the building shapes. This effect appears frequently in the building outlines and reduces the correct detection of building change. Consequently, the overall shape of buildings is further deformed in direct subtraction of DSMs (see Figure 32). It indicates that it is necessary to delineate the building outlines using approaches that are based on image information because they are not influenced by smoothing effects of matching-based DSM. It can be concluded that in building change

detection, the height information provides more stable signals if there has been a vertical change in the built-up areas, while the radiometric image information and GIS polygons provide more accurate information in relation with building outlines.

For the second stage, 3D edge matching and 3D segmentation of building rooftops are employed in order to delineate outlines for building changes. For delineation of building outlines using 3D edge matching, the hypotheses are generated using several epipolar, geometric and topologic criteria. Using 3D edge matching, our experiments deliver promising results if buildings are large enough, have a simple rectangular shape, and a good radiometric contrast compared to surrounding objects. On the other hand, for complex buildings, our approach has more problems, particularly when building outlines cannot be clearly detected in both epipolar images.

Compared to investigations using very high resolution aerial images (Baillard, *et al.,* 1999), satellite images have more problems with extracting building outlines based on 3D edge matching. For instance, the main problem in 3D edge matching using GeoEye-1 stereo images is missing a large number of edges due to the low ground sampling, low image quality as well as the misdetection of some building edges because of different view direction or illumination. Therefore if a given edge is not detected in one of the epipolar images (or even if it is detected partially), it is impossible to find a corresponding line in the next image (refer to the ambiguity case). Such a problem arises in very dense built-up areas where buildings are very small with respect to the ground resolution or have very poor contrast compared to the background. As stated, different viewing or illumination angles probably cause the same problem.

Most of these problems are not highlighted in the 3D segmentation of building rooftops because delineation of small buildings is independent from fitting a straight line. This means in 3D edge matching many building outlines are partially or completely missed but in 3D segmentation of building rooftops, building footprints are detected nevertheless if it contains straight or a jagged outlines. Our tests also show that for the detection of complex structured building footprints using 3D edge matching, a better ground resolution is required while 3D segmentation is a relatively more flexible approach which is applicable to a wide range of high resolution sensors.

In can be concluded that while in our previous investigation (Dini, *et al.,* 2012) the subtraction of two DSMs derived from SGM can be used as an alarm system for building change detection, it is not an efficient tool for the delineation of building outlines. In contrast, line matching and particularly 3D segmentation can delineate building outlines provided that they have a good radiometric contrast with an acceptable size and shape. In this case 3D segmentation delivers better results with a higher rate of detection. On the contrary, 3D edge matching using high resolution satellite images is not an effective solution if buildings have poor contrast or complex structures containing curved outlines.

6.2. Future works

According to the obtained results which were summarized in the previous section, the following directions are proposed as a development of this research in future. Although a more refined building extraction approach will improve the results to some degree, as a first priority a finer ground sampling

distance than 0.5m is required so as to obtain better results. Hence, the first point is to test the proposed algorithm over finer datasets (such as images with higher ground resolution and quality). Such a test checks the influence of image resolution in the quality of acquired results (either for building change detection or building footprints extraction) particularly using 3D line matching which has shown poor results using GeoEye-1 stereo pair.

Furthermore, with such a dataset we would be able to handle complex building footprints and changes of building parts (e. g. new annexes of an existing building). In addition to the spatial (ground resolution) and radiometric (image quality), a better spectral resolution is required to provide a possibility for the generation of vegetation mask using infrared channels. As mentioned before, our dataset does not only contain the infrared channel but it is also a false color composite of pan-sharpening so it does not represent the true reflectance of objects as an RGB channel. This may cause a problem for classification in order to distinguish between vegetation and other objects like shadow or asphalt. Therefore, it is also required to increase the accuracy of vegetation mask by using an infrared channel.

Another direction for future work is the application of a sophisticated classification method (e.g. conditional random field) to generate vegetation mask. It probably detects vegetation mask better as it works based on a stochastic framework which places emphasis on the texture and spatial context. Hence, such a classification method can detect vegetation particularly; where trees partially cover building rooftops and also it considers the neighbouring information in the classification process.

The next point is the improvement of the accuracy of image orientation. Accurate georeferencing of DSMs at different epochs using ground control points (biased corrected RPC) is another issue while a more detailed comparison of extracted blobs and the building ground plan is another separate topic of investigation. It is also recommended to test the efficiency of proposed algorithms using tri-stereoscopic satellite data such as Pleiades images providing observations at three viewing angles (backward, nadir, forward). Of course, a DSM generated from tri-stereoscopic images can partially compensate the occlusion effects and delivers a DSM with finer level of detail (building blobs), providing that it has the same ground sampling as GeoEye-1 or a better one.

Furthermore, we also consider the possibility of optimising the segmentation procedure by adding further terms, mainly by introducing more shape information into the existing method. This could improve the quality of 3D segmentation of building rooftops, in particular those with poor contrast. Moreover, detection of pairs of parallel lines in 3D line matching probably supports the extraction of rectangular building footprints.

Bibliography

Alobeid, A., Jacobsen, K., & Heipke, C. (2010). Comparison of Matching Algorithms for DSM Generation in Urban Areas from IKONOS Imagery. *Photogrammetric Engineering and Remote Sensing , 76(9)* , pp. 1041-1050.

Alobeid, A. (2011). *Assessment of Matching Algorithms for Urban DSM Generation from Very High Resolution Satellite Stereo Images.* Ph.D. dissertation, Gottfried Wilhelm Leibniz Universität Hannover, ISSN 0174-1454, Nr. 293.

Awrangjeb, M., Zhang, C., & Fraser, C. (2012). Building Detection in Complex Scenes Thorough Effective Separation of Buildings from Trees. *Photogrammetric Engineering and Remote Sensing , Vol. 78, No. 7* , pp. 729–745.

Baillard, C., Schmid, C., Zisserman, A., & Fitzgibbon, A. (1999). Automatic Line Matching and 3D Reconstruction of Buildings from Multiple Views. *In: IntArchPhRS, Vol. 32 (Part 3-2W5), pp. 69–80.*

Benarchid, O., Raissouni, N., Adib, S. E., Abbous, A., Azyat, A., Achhab, N. B., Lahraoua, M., & Chahboun, A. (2013). Building Extraction using Object-Based Classification and Shadow Information in Very High Resolution Multispectral Images, a Case Study: Tetuan, Morocco. *Canadian Journal on Image Processing and Computer Vision , 4* (1), 1-8.

Beumier, C., & M., I. (2012). Building Change Detection from Uniform Regions. *Lecture Notes in Computer Science , 7441* , pp. 648-655.

Bouziani, M., Goita, K., & He, D. (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 , 65* , pp. 143-153.

Büyüksalih, G., & Jacobsen, K. (2007). Comparison of DEM Generation by Very High Resolution Optical Satellites. *EARSeL. Band "New Developments and Challenges in Remote Sensing", Rotterdam: Millpress, ISBN 978-90-5966-053-3, pp. 627-637.*

Chaabouni-Chouayakh, H., & Reinartz, P. (2011). Towards Automatic 3D Change Detection inside Urban Areas by Combining Height and Shape Information. *Photogrammetric Fernerkundung Geoinformation , 4* , pp. 205-217.

Chaabouni-Chouayakh, H., Krauss, T., dAngelo, P., & Reinartz, P. (2010). 3D Change Detection Inside Urban Areas Using Different Digital Surface Models. *: IntArchPhRS, Vol. XXXVIII, Part 3B, pp 86-91.*

Champion, N., Boldo, D., Pierrot-Deseilligny, M., & Stamon, G. (2010). 2D Building Change Detection from High Resolution Satellite Imagery: A Two-step Hierarchical Method Based on 3D Invariant Primitives. *Pattern Recognition Letters , 31* , pp. 1138 - 1147.

Champion, N. (2007). 2D Building Change Detection from High Resolution Aerial Images and Correlation Digital Surface Models. *In: IntArchPhRS, Vol. 36, 3/W49A., pp. 197-202.*

Chan, T. F., & Vese, L. A. (2001). Active Contours Without Edges. *Transactions on Image Processing , Vol. 10, No. 2 ,* pp. 266-277.

Chehata, N., M., P.-D., & Stamon, G. (2005). Hybrid Digital Elevation Model Production Guided by 3D-Primitives: A global Optimization Algorithm using Graph Cuts. *IEEE International Conference on Image Processing ICIP'05, pp. 117-120.*

Chehata, N., Pierrot-Desseilligny, M., Jung, F., & Stamon, G. (2002). Extraction of 3D Primitives from Stereopairs of Aatellite Images for Automatic Reconstruction of Buildings. *IAPR Workshop on Machine Vision Applications (MVA 2002), Nara, Japan, pp. 636-639.*

Dini, G. R., Jacobsen, K., & Heipke, C. (2013). Delineation of Building Footprints from High Resolution Satellite Stereo Imagery using Image Matching and a GIS Database. *In: IntArchPhRS vol. XL-1-W1-81, pp. 81-85.*

Dini, G. R., Jacobsen, K., Rottensteiner, F., Alrajhi, M., & Heipke, C. (2012). 3D Building Change Detection Using High Resolution Stereo Images and a GIS Database. *: IntArchPhRS, Vol. XXIX-B7, S. 299-304.*

Dowman, I., Jacobsen, K., Konecny, G., & Sandau, R. (2012). *High Resolution Optical Satellite Imagery.* Dunbeath, Scotland, UK: Whittles Publishing.

Doxani, G., Karantzalos, K., & Tsakiri-Strati, M. (2010). Automatic Change Detection in Urban Areas under a Scale-Space, Object-Oriented Classification Framework. In E. A. Coillie, & F. Van (Ed.), *In: IntArchPhRS, Vol. XXXVIII-4/C7.*

Durieux, L., Lagabrielle, E., & 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 , 63 ,* pp. 399–408.

Fraser, C., E., B., & Gruen, A. (2002). Processing of Ikonos Imagery for Submetre 3D Positioning and Building Extraction. *ISPRS Journal of Photogrammetry and Remote Sensing , 56 ,* pp. 177–194.

Fusiello, A., Roberto, V., & Trucco, E. (1997). Experiments with a new Area-Based Stereo Algorithm. *International Conference on Image Anlysis and Proceedings.* Florence, Italy.

Gehrke, S., Morin, K., Downey, M., Boehrer, N., & Fuchs, T. (2010). Semi-Global Matching: an Alternative to LiDAR for DSM Generation. *In: IntArchPhRS, Vol. XXXVIII, part1.*

Grigillo, D., Fras, M. K., & Petrovič, D. (2012). Automated Building Extraction from IKONOS Images in Suburban Areas. *International Journal of Remote Sensing , 33* (16), 5149–5170.

Grodecki, J., & Dial, G. (2001). IKONOS Geometric Accuracy. *Joint ISPRS Workshop on HRM from Space*, (pp. 77-86).

Guercke, R., & Sester, M. (2011). Building Footprint Simplification Based on Hough Transform and Least squares Adjustment. *Proceedings of the Joint Workshop Geographic Information on Demand, p. CD/Web.*

Habib, A., Zhai, R., & Kim, C. (2010). Generation of Complex Polyhedral Building Models by Integrating Stereo-Aerial Imagery and LiDAR Data. *Photogrammetric Engineering \& Remote Sensing , 76 (5)* , pp. 609–623.

Hannah, M. (1989). A System for Digitial Stereo Image Matching. *Photogrammetric Engineering and Remote Sensing , 55(12)* , pp. 1765-1770.

Hartley, R., & Zisserman, A. (2004). *Multiple View Geometry.* Oxford: Cambridge University Press.

Heipke, C., Koch, A., & Lohmann, P. (2002). Analysis of SRTM DTM - Methodology and Practical Results. *Journal of the Swedish Society for Photogrammetry and Remote Sensing , 1* (Photogrammetry meets geoinformatics, Anders Boberg (Ed.)), pp. 69-80.

Heipke, C., Woodsford, P. A., & Gerke, M. (2008). *Updating Geospatial Databases from Images.* In: Li, Z.; Chen,J.; Baltsavias E. (Hrsg.): Advances in Photogrammetry, Remote Sensing and Spatial Information Sciences. London: Taylor and Francis, pp.355-362.

Hirschmueller, H. (2008). Stereo Processing by Semiglobal Matching and Mutual Information. *IEEE Transactions on Pattern Analysis and Machine Intelligence , 30(2)* , 328-341.

Im, J., & Jensen, J. R. (2005). A Change Detection Model Based on Neighborhood Correlation Image Analysis and Decision Tree Classification. *Remote Sensing of Environment , 99* , pp. 326-340.

Jacobsen, K. (2011). Characteristics of very High Resolution Optical Satellites for Topographic Mapping. *In: IntArchPhRS vol. XXXVIII-4/W19, Hannover, 2011, 6 S. CD.*

Jaynes, C., Riseman, E., & Hanson, A. (2003). Recognition and Reconstruction of Buildings from Multiple Aerial Images. *Computer Vision and Image Understanding , Vol. 90* , pp. 68-98.

Jin, X., & Davis, C. H. (2005). Automated Building Extraction from High-Resolution Satellite Imagery in Urban Areas Using Structural, Contextual, and Spectral Information. *EURASIP Journal on Applied Signal Processing , 14* , pp. 2196–2206.

Khoshelham, K., Li, Z., & King, B. (2005). A Split-and-Merge Technique for Automated Reconstruction of Roof Planes. *Photogrammetric Engineering \& Remote Sensing , Vol. 71, No. 7* , pp. 855–862.

Kim, T. (2000). A study on the Epipolarity of Linear Pushbroom Images. *Journal of Photogrammetric Engineering & Remote Sensing* , 66(8): pp. 961-966.

Knudsen, T., & Olsen, B. P. (2003). Automated Change Detection for Updates of Digital Map Databases. *Photogrammetric Engineering and Remote Sensing , 69* , pp. 1289-1296.

Krauß, T., & d'Angelo, P. (2011). Morphological Filling of Digital Elevation Models. *IntArchPhRS Vol. XXXVIII-4/W19*, (pp. 165-172).

Lafarge, F., Descombes, X., Zerubia, J., & Pierrot-Deseilligny, M. (2008). Automatic Building Extraction from DEMs Using an Object Approach and Application to the 3D-City Modeling. *ISPRS Journal of Photogrammetry and Remote Sensing , 63* , pp. 365-381.

Le Bris, A., & Chehata, N. (2011). Change Detection in a Topographic Building Database Using Submetric Satellite Images. In S. U. al. (Ed.), *PIA11, In: IntArchPhRS, Vol. 38 (3/W22), pp. 25-30.*

Lee, Y.-S., Koo, H.-S., & Jeong, C.-S. (2006). A Straight Line Detection Using Principal Component Analysis. *Pattern Recognition Letters , 27, no. 14 ,* 1744-1754.

Lu, D., Mausel, P., Brondizio, E., & Moran, E. (2004). Change Detection Techniques. *International Journal of Remote Sensing , 25* (12), pp. 2365-2401.

Macay Moreira, J., Nex, F., Agugiaro, G., Remondino, F., & Lim, N. (2013). From DSM To 3D Building Models: A Quantitative Evaluation. *In: IntArchPhRS, Vol. XL-1-W1, pp. 213-219.*

Malpica, J. A., & Alonso, M. C. (2010). Urban Changes with Satellite Imagery and LiDAR Data. *In: IntArchPhRS, Vol. XXXVIII, Part 8, pp. 853-858.*

Matikainen, L., Hyyppae, J., Ahokas, E., Markelin, L., & Kaartinen, H. (2010). Automatic Detection of Buildings and Changes in Buildings for Updating of Maps. *Remote Sensing , 2(5) ,* pp. 1217-1248.

Niemeyer, I., & Canty, M. (2003). Pixel-Based and Object-Oriented Change Detection Analysis Using High Resolution Imagery. *25th Symp. Safeguards and Nuclear Material Management, online: http://www.definiens-imaging.com/documents/publications/esarda03_nieca.pdf.*

Niemeyer, J., Rottensteiner, F., Kühn, F., & Soergel, U. (2010). Extraktion Geologisch Relevanter Strukturen auf Rügen in Laserscanner-Daten. *In: DGPF Jahrestagung, pp. 298-307.*

Oh, J. (2011). *Novel Approach to Epipolar Resampling of HRSI and Satellite Stereo Imagery-based Georeferencing of Aerial Images.* PhD Dissertation, The Ohio State University.

Ok, A., Wegner, J., Heipke, C., Rottensteiner, F., Soergel, U., & Toprak, V. (2012). Matching of Straight Line Segments from Aerial Stereo Images Of Urban Areas. *ISPRS Journal of Photogrammetry and Remote Sensing , 74 ,* pp. 133-152.

Olsen, B. P. (2004). Automatic Change Detection for Validation of Digital Map Databases. *In: IntArchPhRS, Vol. XXX IV, Part B2, pp. 569-574.*

Poli, D., & Pierre, S. (2012). Digital Surface Model Extraction and Refinement through Image Segmentation - Application to the ISPRS Benchmark Stereo Dataset. *PFG Photogrammetrie, Fernerkundung, Geoinformation , 4 ,* pp. 317-329.

Porter, R., Harvey, N., & Theiler, J. (2009). A Change Detection Approach to Moving Object Detection in Low Fame-Rate Video. *In: Proc. SPIE 7341, Visual Information Processing XVIII, 73410S; doi:10.1117/12.818622.*

Radke, R. J., Andra, S., Al-Kofahi, O., & Roysam, B. (2005). Image Change Detection Algorithms: A Systematic Survey. *IEEE Transactions on Image Processing , 14* (3), pp. 294-307.

Rottensteiner, F. (2007). Building Change Detection from Digital Surface Models and Multi-Spectral Images. In S. U. . (Ed.), *PIA07, In: IntArchPhRS, Vol. 36 (3/W49B), pp. 145-150.*

Rottensteiner, F. (2008). Automated Updating of Building Databases From Digital Surface Models and Multi-spectral Images. *In: IntArchPhRS 37 (B3A), pp. 265-270.*

Rutzinger, M., Ruef, B., Hoefle, B., & Vetter, M. (2010). Change Detection of Building Footprints from Airborne Laser Scanning Acquired in Short Time Intervals. *In: IntArchPhRS, Vol. XXXVIII, Part 7B, pp. 475-480.*

Scott Lee, D., Shan, J., & Bethel, J. S. (2003). Class-Guided Building Extraction from Ikonos Imagery. *Photogrammetric Engineering & Remote Sensing , 69* (2), 143–150.

Shrivakshan, G. (2012). A Comparison of various Edge Detection Techniques used in Image Processing. *IJCSI International Journal of Computer Science Issues, Vol. 9, Issue 5, No 1* , 269-276.

Sirmacek, B., dAngelo, P., & Reinartz, P. (2010). Detecting Complex Building Shapes in Panchromatic Satellite Images for Digital Elevation Model Enhancement. *ISPRS Istanbul Workshop 2010 on Modeling of optical airborne and spaceborne Sensors, IntArchPhRS Vol. XXXVIII, part 1/W17.*

Suveg, I., & Vosselman, G. (2004). Reconstruction of 3D Building Models from Aerial Images and Maps. *ISPRS Journal of Photogrammetry and Remote Sensing , 58 (3-4)* , pp. 202–224.

Tian, J., Chaabouni-Chouayakh, H., & Reinartz, P. (2011). 3D Building Change Detection from High Resolution Spaceborne Stereo Imagery. *Multi-Platform/Multi-Sensor Remote Sensing and Mapping (M2RSM), Print ISBN: 978-1-4244-9402-6.*

Tian, J., Chaabouni-Chouayakh, H., Reinartz, P., Krauss, T., & dAngelo, P. (2010). Automatic 3D Change Detection Based On Optical Satellite Stereo Imagery. *In: IntArchPhRS, Vol. XXXVIII, Part 7B, pp. 586-591.*

Tzotsos, A., Karantzalos, K., & Argialas, D. (2011). Object Based Image Analysis Through Nonlinear Scale Space Filtering. *ISPRS Journal of Photogrammetry and Remote Sensing , 66* , pp. 2-16.

Vallet, B., Pierrot-Deseilligny, M., Boldo, D., & Brédif, M. (2011). Building Footprint Database Improvement for 3D Reconstruction: A Split and Merge Approach and its Evaluation. *ISPRS Journal of Photogrammetry and Remote Sensing , 66 (5)* , pp. 732–742.

Vosselman, G. (1999). Building Reconstruction Using Planar Faces in Very High Density Height Data. *In: IntArchPhRS XXXII, Part3/2W5, pp. 87-92.*

Walter, V. (1999). Automated GIS Data Collection and Update. In i. D. Spiller (Ed.), *'Photogrammetric Week '99', Wichmann Verlag* , (pp. 267 – 280). Heidelberg, Germany.

Werner, T., & Zisserman, A. (2002). New Techniques for Automated Architectural Reconstruction from Photographs. *ECCV '02 Proceedings of the 7th European Conference on Computer Vision-Part II, pp. 541-555.*

Zhang, L., & Gruen, A. (2006). Multi-Image Matching for DSM Generation from IKONOS Imagery. *ISPRS Journal of Photogrammetry and Remote Sensing , 60* , pp. 195-211.

Zhu, X., Zhang, H., & Pang, G. (2008). A Change Detection Method with High Resolution Images Based On Polygon Automatic Validating. *In: IntArchPhRS, Vol. XXXVII. Part B7., pp. 841-845.*

Appendix A: Technical Characteristics of IKONOS-2 and GeoEye-1-1

Performance Parameter	IKONOS-2	GeoEye-1–1
Orbital Information		
Altitude	681 km	681 km
Inclination	98.1° / sun sync	98.1° / sun sync
Orbit Duration	98 minutes	98 minutes
Equatorial Crossing Time	10:30 am	10:30 am
Revisit at best NIIRS	3 days	3 days
Off Nadir (Terrestrial) Imaging	60 degrees	60 degrees
Image Quality		
Spatial Resolution	Pan = 0.82 m @ nadir MSI = 3.2 m @ nadir	Pan = 0.41 m @ nadir MSI = 1.65 m @ nadir
Imaging Bands	1 Pan / 4 MSI	1 Pan / 4 MSI
Spectral Range	Pan = 526–929 nm MSI 1 (B) = 445–516nm MSI 2 (G) = 506–595 nm MSI 3 (R) = 632–698 nm MSI 4 (NIR) = 757–853 nm	Pan = 450–800 nm MSI 1 (B) = 450–510 nm MSI 2 (G) = 510–580 nm MSI 3 (R) = 655–690 nm MSI 4 (NIR) = 780–920 nm
Bits/Pixel	11-bits/pixel	11-bits/pixel
Minimum Image Size	121 sq.km	231 sq.km
Time Delay Integration	32 stages (selectable)	64 stages (selectable)
Best GSD @ nadir	0.82 meters	0.41 meters
NIIRS	4.5	5.5
Collection Capacity		
Daily Pan Area	300,000 sq.km	700,000 sq.km
Daily MSI Area	300,000 sq.km	350,000 sq.km
Swath Width @ nadir	11.3 km	15.2 km
On-Board HSSU Capacity (store & dump)	80-gigabits	1-terabit
Geolocation		
Mono: CE90 w/out external GCPs	$\leq$15 meters	<5 meters
Stereo: CE90 / LE90 w/out external GCPs	10 meters / 10 meters	4 meters / 6 meters
Duration		
Design Life	5 years	7 years
Expected In-service Life	10 years	10 years

reference: http://www.GeoEye-1.com

Curriculum Vitae (C.V)

- **Personal Information**

Given Name	Gholam Reza
Surname	Dini
Date and Birth	21.09.1979
Place of Birth	Karaj, Iran
E-mail	dini@daad-alumni.de
Personal Homepage	www.rezadini.webs.com

- **Education**

1985-1991	Primary School, Karaj, Iran
1991-1994	Guidance School, Karaj, Iran
1994-1998	High School, Karaj, Iran
1998-2002	**B.Eng.** Golestan University, Department of Watershed Engineering, Gorgan, Iran.
2002-2005	**M.Sc.** Tarbiat Modares University, Department of Remote Sensing and Geographic Information Systems (GIS), Tehran, Iran.
2010-2014	**Ph.D.** candidate and DAAD scholar at Gottfried Wilhelm Leibniz Universität Hannover, Institute of Photogrammetry and Geoinformation (IPI), Hannover, Germany.

- **Work Experience**

2003 - 2005	Geomantic Expert, Basir Remote Sensing Institute, Tehran, Iran
2006 - 2007	Teaching Assistant (Part time), Azad University, Ray Campus, Department of Surveying and Cartography, Tehran, Iran.
2006 - 2008	Teaching Assistant (Part time), Tehran University, Environment Engineering Department, Tehran, Iran.
2005 - 2009	GIS and Geodatabase Expert, Tehran Regional Water Authority (TRWA), Tehran, Iran.
2010 - 2014	Ph.D candidate and scholar of German Academic Exchange Service (DAAD) at the Leibniz University of Hannover, Institute of Photogrammetry and Geoinformation (IPI), Hannover, Germany.
Sept.-Dec. 2013	Visiting Researcher at the Telecommunications and Remote Sensing Laboratory - Department of Electronics and Computer Engineering, University of Pavia, Italy.

Wissenschaftliche Arbeiten der Fachrichtung Geodäsie und Geoinformatik der Leibniz Universität Hannover

(Eine vollständige Liste der Wiss. Arb. ist beim Geodätischen Institut, Nienburger Str. 1, 30167 Hannover erhältlich.)

Nr. 287	Schramm, Matthias	Unüberwachte lineare spektrale Entmischung mit Anwendungen für Baumdichteabschätzungen in semiariden Regionen Afrikas (Diss. 2010)
Nr. 288	Alsayel, Mohammed Y.	Towards Standardizing International Boundary Determination and Quality Consequences on surveying and Mapping (Diss. 2010)
Nr. 289	Tavakkoli Sabour, Seyed Mohammad	Multi-Temporal Classification of Crops Using ENVISAT ASAR Data (Diss. 2011)
Nr. 290	Vennegeerts, Harald	Objektraumgestützte kinematische Georeferenzierung für Mobile-Mapping Systeme (Diss. 2011)
Nr. 291	Bulatov, Dimitri	Textured 3D reconstruction of urban terrain from UAV-borne video sequences (Diss. 2011)
Nr. 292	Grote, Anne	Automatic road network extraction in suburban areas from aerial images (Diss. 2011)
Nr. 293	Alobeid, Abdalla	Assessment of Matching Algorithms for Urban DSM Generation from Very High Resolution Satellite Stereo Images (Diss. 2011)
Nr. 294	Fang, Xing	Weighted Total Least squares Solutions for Applications in Geodesy (Diss. 2011)
Nr. 295	Wegner, Jan Dirk	Detection and height estimation of buildings from SAR and optical images using conditional random fields (Diss. 2011)
Nr. 296	Göpfert, Jens	Snakes for Adapting GIS Road and River Objects to Airborne Laser Scanning Data (Diss. 2011)
Nr. 297	Muhle, Daniel	Gegenseitige Orientierung von Mehrkamerasystemen mit nicht überlappendem Sichtfeld (Diss. 2011)
Nr. 298	Siriba, David Nyangau	Conflation of Provisional Cadastral and Topographical Datasets (Diss. 2012)
Nr. 299	Helmholz, Petra	Verifikation von Ackerland- und Grünlandobjekten eines topographischen Datensatzes mit monotemporalen Bildern (Diss. 2012)
Nr. 300	Deng, Zhiguo	GPS Meteorology with Single Frequency Receivers (Diss. 2012)
Nr. 301	Zeißler, Maik	Zur Ermittlung von Bodenrichtwerten bei fehlenden Kaufpreisen unbebauter Grundstücke (Diss. 2012)
Nr. 302	Paffenholz, Jens-André	Direct geo-referencing of 3D point clouds with 3D positioning sensors (Diss. 2012)
Nr. 303	Weinbach, Ulrich	Feasibility and impact of receiver clock modeling in precise GPS data analysis (Diss. 2012)
Nr. 304	Heiker, Andrea	Mutual validation of Earth orientation parameters, geophysical excitation functions and second degree gravity field coefficients (Diss. 2012)
Nr. 305	Voigt, Christian.	Astrogeodätische Lotabweichungen zur Validierung von Schwerefeldmodellen (Diss. 2013)
Nr. 306	Thiele, Antje	3D Building Reconstruction From High Resolution Multi-Aspect Interferometric Synthetic Aperture RADAR Data (Diss. 2013)
Nr. 307	Alrajhi, Muhamad	Semi-Automated Procedure for Orthophoto Generation from High Resolution Satellite Imagery (Diss. 2013)
Nr. 308	Wang, Jin	Towards deformation monitoring with terrestrial laser scanning based on external calibration and feature matching methods (Diss. 2013)
Nr. 309	Naeimi, Majid	Inversion of satellite gravity data using spherical radial base functions (Diss. 2013)
Nr. 310	Seatovic, Dejan	Methods for Real-Time Plant Detection in 3-D Point Clouds (Diss. 2013)
Nr. 311	Schunert, Alexander	Assignment of Persistent Scatterers to Buildings (Diss. 2014)
Nr. 312	Guercke, Richard	Optimization Aspects in the Generalization of 3D Building Models (Diss. 2014)
Nr. 313	Ziems, Marcel	Automatic verification of road databases using multiple road models (Diss. 2014)
Nr. 314	Dini, Gholam Reza	Toward an automatic solution for updating building databases using high resolution space-borne stereo imaging (Diss. 2014)