GIMA

Geographical Information Management and Applications

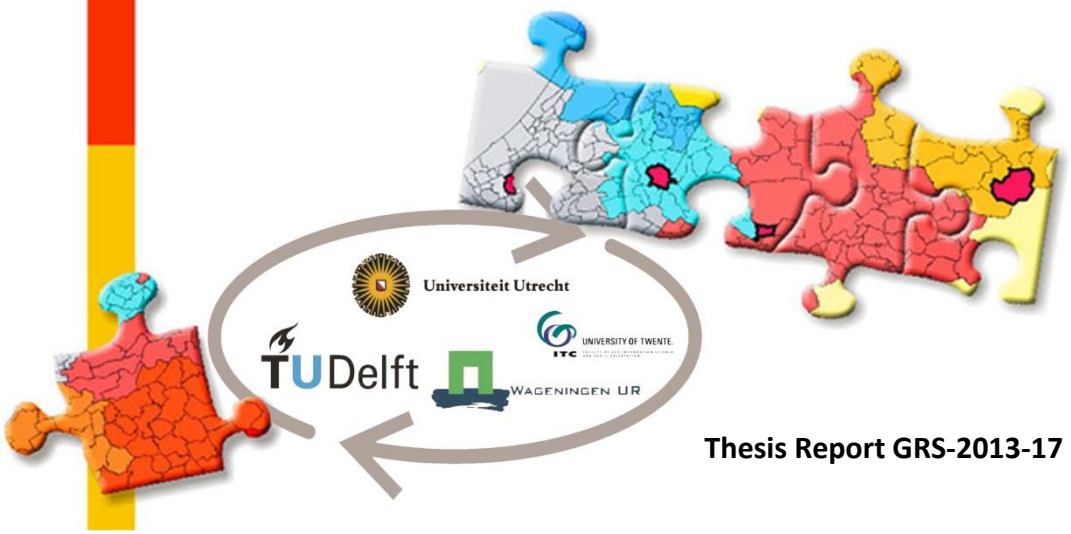
Geo-specific terrain modeling for military simulations

A study focusing on the potential benefit of spectral analysis based feature extraction for DSM to DTM filtering in the geo-specific terrain production process, using very high resolution multispectral stereo satellite imagery.

MSc Thesis
August 2013

Joost Michael
jcp.michael@mindef.nl

Professor: Prof. dr. ir. Arnold Bregt (WUR)
Supervisor(s): Dr. ir. Jan Clevers (WUR)
Ir. Frido Kuijper (TNO)



Thesis Report GRS-2013-17

Geo-specific terrain modeling for military simulations

A study focusing on the potential benefit of spectral analysis based feature extraction for DSM to DTM filtering in the geo-specific terrain production process, using very high resolution multispectral stereo satellite imagery.

Thesis submitted in partial fulfillment of the degree – Master of Science in Geographical Information Management and Applications (GIMA)

August 2013

Author: Joost Michael
Professor: Prof. dr. ir. Arnold Bregt (Wageningen University and Research centre)
Supervisor(s): Dr. ir. Jan Clevers (Wageningen University and Research centre)
Ir. Frido Kuijper (Netherlands Organisation for Applied Scientific Research)

Disclaimer

I hereby declare that I am the sole author of this thesis. This thesis has been prepared by the author for academic purposes only and does not necessarily reflect the view of the universities organizing the GIMA program or the Dutch Ministry of Defense on the subject of geo-specific terrain modeling for military simulations. This is a true copy of the thesis that has been submitted after assessment by the GIMA supervisors.

ABSTRACT

Geo-specific terrain models are the geospatial source of information used for the creation of military simulations of existing real world environments. These simulations are used to gain situation awareness without having set foot on the ground and for mission rehearsal. Geo-specific terrain models consist of a digital terrain model (DTM) representing bare earth elevation only, an orthorectified imagery layer and basic feature layers such as building footprints, forests, water and infrastructure. Given the temporal nature of missions and operations it is important to shorten the production cycle of geo-specific terrain models as much as possible through the development of semi-automatic geo-specific terrain modeling processes. Creating geo-specific terrain models from existing data with various sources is challenging considering the correlation of varying data accuracies, currencies and scales. Furthermore, the derivation of a DTM from the more common digital surface model (DSM) that still includes all above ground objects such as buildings and forests is a subject recurrently studied and for which no panacea exists. Satellite imagery is the only guaranteed source of accurate and current information when creating geo-specific terrain models over foreign territories since it is commercially available with near global coverage. Therefore the use of the state of the art stereo satellite imagery provided by the WorldView-2 satellite is investigated as a single source of input for the derivation of all required geo-specific terrain model layers.

This research focusses on the improvement of terrain modeling efficiency by answering four questions, namely: What suitable DSM to DTM filtering techniques exist? Can spectral based features be combined with geometric information for the purpose of DSM to DTM filtering? Which DSM to DTM filtering strategy yields the best result? Are features extracted significantly better when using eight spectral bands instead of the typical four imagery bands? Photogrammetric techniques are implemented to generate DSMs and orthorectified imagery. From the orthorectified imagery features are extracted automatically using the DSM as a secondary source for the identification of building features. Both four and eight spectral bands imagery have been used for automatic feature extraction to determine whether the use of more non-typical spectral bands yields better feature extraction results. Finally the suitability of using either the spectral features, the geometric information present in the surfaces or a combination of both in the process of DSM to DTM filtering is investigated by comparison of results.

It is found that the accuracies of automatically extracted road and building features do not suffice for the direct removal of above ground objects from the photogrammetrically derived surfaces, although the forest and water features can be used in the combined spectral feature based and geometry based DSM to DTM filtering process. The use of these features in the process of combined DSM to DTM filtering yielded the best results in both study sites. Eight spectral bands showed a general improvement in the feature extraction results. The biggest improvements were made in the complex road and building feature extraction.

Keywords: Military simulation, geo-specific terrain modeling, stereo satellite imagery, WorldView-2, photogrammetry, surface, terrain, feature extraction, OBIA, DSM to DTM filtering.

ACKNOWLEDGEMENTS

First of all I would like to thank my wife who supported me throughout the last three years of study and kept me going when the going was tough on the both of us. Without her support, understanding and endless supplies of coffee I would not have been able to complete this study. I would also like to thank my parents who have always believed and supported me in the choices I made, even when I wanted to become a paleontologist or a tropical forester! They taught me the value of enjoying the work you do since you will be doing it for quite some time.

I would like to thank the Royal Netherlands Army Geographic Agency for helping me achieve my goals. They believed I could do it, helped finance my study and provided me with the data necessary for my research. I would also like to thank my colleagues from the RNLAGA. I always enjoy our endless discussions on geography from which I believe one can learn so much and it has often been a source of inspiration or insight during my study.

I would also like to thank Jan Clevers and Arnold Bregt from the Wageningen University and Research Centre (WUR), whom have been willing to supervise my research even though my thesis was a bit off topic from a GIMA perspective. I have very much appreciated your comments and suggestions from which I have learned a lot.

I would like to express my gratitude to Frido Kuiper from the Netherlands Organization for Applied Scientific Research (TNO) who has offered to supervise my thesis from his content matter expert point of view. Without your additional supervision, that you had offered without hesitation and that had to be done outside regular working hours, I could not have studied this topic.

I would like to thank the people from IMAGEM. As distributor of the Erdas software you have been subjected to endless questioning from my side and I have very much appreciated your thoughts on the matter and your quick response time.

Last but certainly not least I would like to thank the entire GIMA staff. I am proud to have taken part in the GIMA program and found it a truly meaningful experience from which I have learned so much. I would also like to thank all my fellow part- and full time students for this great experience. I have very much enjoyed the collaborations, discussions, presentations and off course the dinners and beers we enjoyed after those long days during the contact weeks.

Joost Michael
Arnhem 2013

TABLE OF CONTENT

LIST OF ABBREVIATIONS

AA	Accuracy Assessment
AGO	Above Ground Object
ATE	Automatic Terrain Extraction
AHN2	Actueel Hoogtebestand Nederland 2
BAG	Basisregistratie Adressen en Gebouwen
BAGLV	BAG Landelijke Voorziening
CE90	Circular Error with 90% probability
CityGML	City Geography Markup Language
CL	Confidence Level
COTS	Commercial Off-The-Shelf
DEM	Digital Elevation Model
DSM	Digital Surface Model
DTM	Digital Terrain Model
eATE	Enhanced Automatic Terrain Extraction
EF	Extracted Features
EM	Electromagnetic
FE	Feature Extraction
GCP	Ground Control Point
GPS	Global Positioning System
GEOBIA	Geographic Object Based Imagery Analysis
GIScience	Geographic Information Science
GSD	Ground Sample Distance
HCS	Hyperspherical Color Space
IDW	Inverse Distance Weighting
IR	Infrared
LE90	Linear Error with 90% probability
LIDAR	Light Detection and Ranging
LOD	Levels of Detail
LPS	Leica Photogrammetry Suite
MS	Multispectral
nDSM	Normalized DSM
OBIA	Object Based Imagery Analysis
OGC	Open Geospatial Consortium
POS	Passive Optical Sensor
QA	Quality Assurance
RMSE	Root Mean Square Error
RNLAGA	Royal Netherlands Army Geographic Agency
SFP	Single Feature Probability
SRTM	Shuttle Radar Topography Mission
TIN	Triangulated Irregular Network
UTM	Universal Transverse Mercator
VHR	Very High Resolution
WBS	Work Breakdown Structure
WGS84	World Geodetic System 1984

1 INTRODUCTION AND BACKGROUND

1.1 Military simulations

Military simulations are perhaps as old as war itself and are implemented within the army for good reasons. Through experimentation and manipulation in simulated (representative) environments, using various scenarios, military professionals are able to analyze and evaluate what-if scenarios to determine the optimal solutions for a given objective and can experiment with the consequences of decision making through trial and error (Niemeyer, 2003). The decision making faced by military professionals during tactical combat is a difficult task that will even be challenging for the most experienced given its possible complexity and consequences (Naseer et al., 2008). Since one of the best ways of learning is to make mistakes, simulations are especially suitable for military training and learning processes in order to prevent mistakes made in the field, with potential catastrophic outcomes.

Where simulations have traditionally been in analogue form, computer technologies and the game industry have given a major boost to military simulation development and as a result realistic commercial game like training simulations are developed and successfully implemented in various armies to motivate and engage the “game-savvy” soldier of this generation into tactical decision making (Fong, 2004). The application of simulation within armed forces has both easily quantifiable benefits (e.g. costs of ammunition, logistics and damage) as well as very heavy weight intangible benefits (e.g. casualties). Naseer et al. (2008) have conducted research to investigate the potential applications of simulation modeling techniques for healthcare and in doing so monitored the successful application of simulation modeling present in the military. Their research showed that within the military many successful applications can be found. They showed that the most heavily used application was planning (combat, operation, strategic, trajectory, mission etc.), followed by training and warfare.

In order for a simulation to be applied effectively it is important that the proper types of terrain models are available and of sufficient quality. Terrain models used for military simulation can be divided into two categories, the geo-typical and geo-specific models. Geo-typical terrain models represent fictional terrains that can be tailored for each training scenario, whereas with geo-specific terrain models it is attempted to represent an existing terrain as accurate as possible. Through training with geo-specific terrain models, trainees will gain situational awareness and a better understanding of how decisions made in the simulation will manifest during real live missions (Smelik et al., 2010). Simulation based training purposes vary among the different user groups, allowing for these different terrain modeling strategies. For example, the training of basic infantry tactics or the training of vehicle operation does not necessarily require a geo-specific terrain to practice in as long as representative terrain is modeled. This allows for the flexible usage of terrains, making it easier to create various scenarios to train in. However, mission rehearsal in a geo-specific terrain has the main advantage that a military professional can gain situational awareness of a mission area before setting foot on the ground, and the soldier has the possibility to experience how the results of his decisions could possibly manifest in a specific area.

For both geo-typical and geo-specific terrain modeling the costs of production are high and laborious, and therefore design and production processes need to be automated as much as possible (Smelik et al., 2008). This research will focus on the automated production of geo-specific terrain databases.

1.2 Geo-specific terrain models

Geo-specific terrain modeling requires the modeling of elevation, imagery and objects on its true geospatial location and with the right properties (e.g. height, object type). Geo-specific terrains are typically modeled completely automatic based on existing data sources, semi-automatic from existing imagery and elevation and manual vector extraction, or completely manual (Kuijper et al., 2010). These typical strategies are often undesirable for the geo-specific simulations needed by the military. Manual labor will result in long production cycles. Another problem is that the terrains to be modeled are often remotely located with very limited existing sources of good quality information. Because existing sources have various scales, resolutions and inaccuracies, the correlation of these different data sources is assumed to be troublesome and a cause of substantial error propagation, evident in the resulting terrain.

Building a geo-specific terrain requires a layer stratification similar to the one shown in Fig. 1. These could be considered the base layers of the geo-specific terrain model, but higher levels of detail (LOD) can be modeled depending on the use and desired degree of realism of the simulation. CityGML, a semantic information model created by the open geospatial consortium (OGC) for the cross platform visualization of urban environments includes up to 5 levels of detail (OGC, 2007). The layer stratification as shown in Fig. 1, is the equivalent of the CityGML LOD1. First of all a digital terrain model (DTM) is required as an elevation layer. A DTM is an elevation model that contains elevation information of the bare earth, meaning vegetation such as forests are removed. A DTM could include man-made objects for some analysis such as inundation modeling. However, in this research the term DTM is used solely for an elevation model from which all above ground objects are removed. A DTM can be created from a digital surface model (DSM) which, in contrast to a DTM, includes the elevation information of all objects and vegetation. The difference between the DSM and the DTM is an elevation layer which contains the elevations of all above ground objects only and no bare earth elevation values and is referred to as normalized DSM (nDSM) in this research. It can be created by subtracting the DTM from the DSM. The term digital elevation model (DEM) is considered a generic term for both types of elevation models in this research. The different DEMs are illustrated in Fig. 2. The process of creating a DTM out of a DSM is referred to as DSM to DTM filtering in this research.

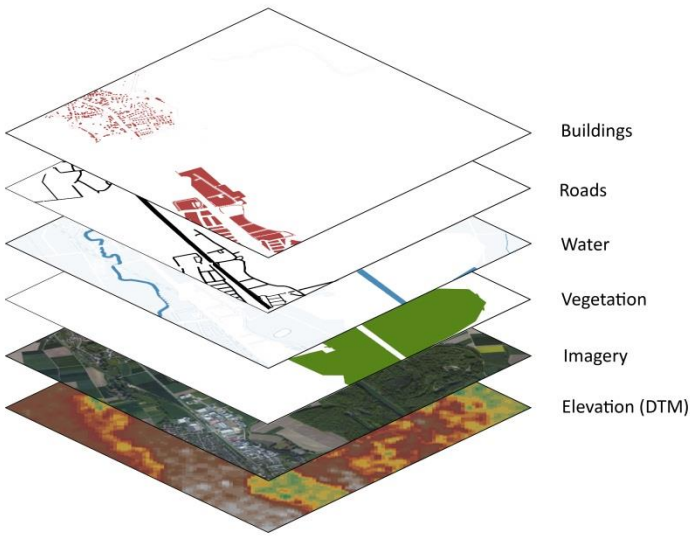


Fig. 1 Required geo-specific terrain model layer stratification.

A DSM can be created using well established photogrammetric techniques and using stereo imagery recorded by passive optical sensors (POS), which measure the reflected electromagnetic (EM) energy of the sun on the earth surface. Alternatively active sensors can be used for DSM generation, which actively emit an energy signal and measure the returning energy. This is the case with laser scanning. Both laser scanning as well as photogrammetry can be used to create DSMs. Laser scanning has as an advantage in this process that an emitted beam can have various returns. For instance a beam directed at the canopy of a forest can bounce back on top of the canopy but the beam or a part of it can also penetrate the canopy and bounce back on the ground surface. If an active sensor is capable of recording multiple returns of an emitted beam, bare earth elevation points can be obtained in forested areas, whereas the latter cannot be obtained through photogrammetric processes. A disadvantage of the laser scanning is that it doesn't have the imagery which is essential for simulation and the processes of detection, classification, identification and measurement of objects (Baltsavias, 1999).

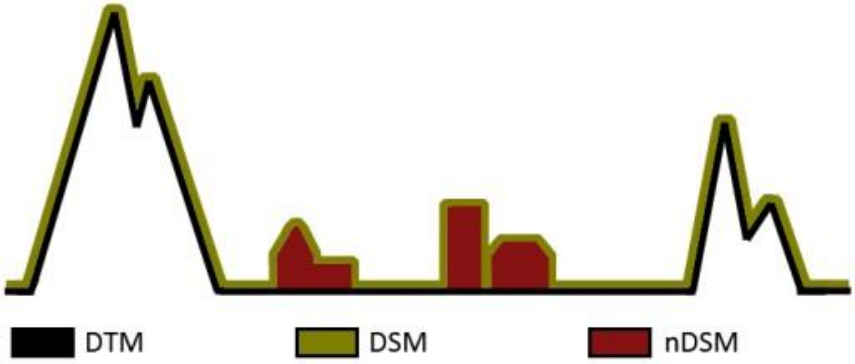


Fig. 2 Different digital elevation models (DEMs).

The second layer consists of high resolution orthorectified imagery, meaning the imagery is orthographically re-projected in order to eliminate scale variation and displacement caused by the natural relief of the terrain and tilting of a sensor (Lillesand et al., 2008). The orthorectification process of imagery requires a digital surface model. Both elevation as well as imagery are raster data (regular grids), and the remaining layers are superimposed vector object layers. In order to create these objects layers, manual labor is required or the objects are extracted from the imagery in an automated fashion. The latter is done using object oriented feature extraction techniques using both elevation data and imagery allowing for the identification and classification of features based on both spectral characteristics as well as geometric characteristics.

1.3 Problem statement

Given the strengths of using geo-specific terrains in simulations (e.g. situational awareness) it is especially important to shorten the production chain as much as possible considering the temporal nature of a mission and the large area sizes of operation. The sooner mission rehearsal and training can begin the better, however typical productions face problems with correlating various scales and qualities of source information when using existing data, laborious manual editing and the absence of suitable geospatial information of remote areas. For some areas only satellite imagery is available, in combination with a low resolution surface model and old topographic maps (scale 200K/250K). These sources have various scales, resolutions, accuracies and currencies. The topographic maps are

likely to be outdated, are generalized representations, and are possibly purposefully inaccurate (Monmonier, 1996), making it a bad source of information for feature extraction. Given the low resolution of the surface model, its geometric information cannot be effectively used in the DSM to DTM filtering process and the extraction of features. From the figure below (Fig. 3) it becomes clear that it is impossible to identify above ground objects from low resolution elevation data because pixels represent a mixture of elevation values. When stereo imagery is not available it is not possible to generate the required higher resolution DSM using photogrammetric processes.

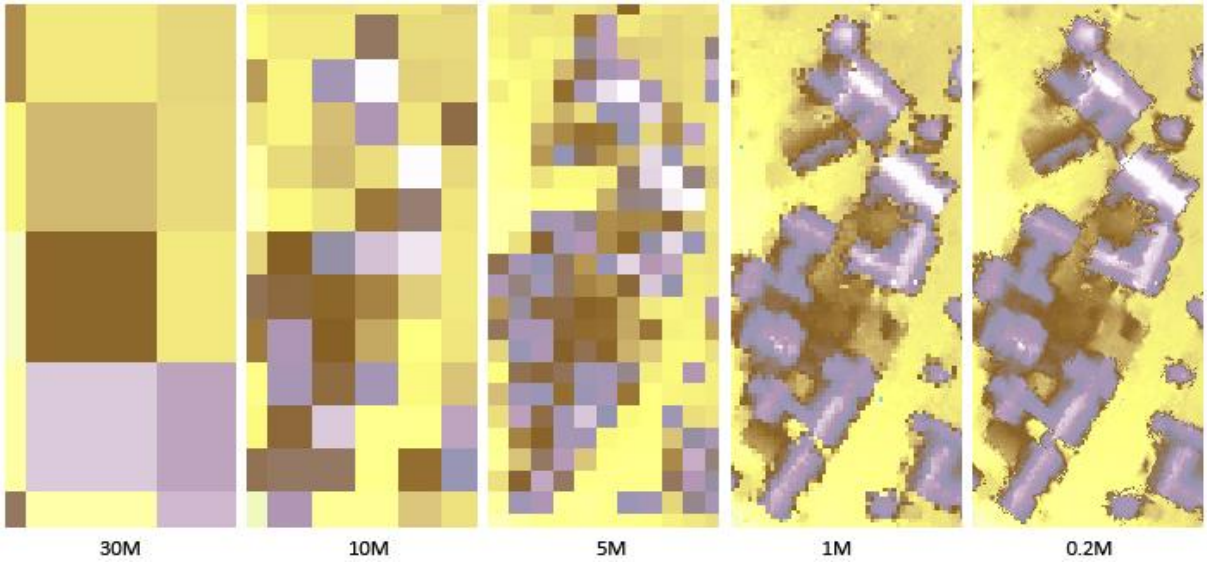


Fig. 3 Example of the relation between surface resolution and building object recognition

Another problem in geo-specific terrain modeling is that both the process of DSM to DTM filtering and feature extraction are not without error. Given the complex nature of our environment, extraction of bare earth elevations and objects from imagery is often considered both an art and a science, for which no 100 percent solution exists. During the DSM to DTM filtering process bare earth terrain features can mistakenly be removed because of the filtering strategies and parameters applied, or the application of certain strategies and parameters cannot prevent that some above ground objects that should be removed remain in the terrain. During the feature extraction process a target object can be found to have a significantly different spectral signature, e.g. different types of roof tops, making it hard to extract based on one spectral class. Furthermore, there is no all-inclusive methodology for the extraction of all layers required for geo-specific terrain modeling.

1.4 Innovation aimed at

Given the need for fast and automated processes, the unavailability of high quality data of remote areas, the problems of correlation that would occur when sources of data are present, and recent technological advancements, a strategy is investigated which uses very high resolution (VHR) multispectral stereo imagery as a single source of data for geo-specific terrain modeling. Through the use of photogrammetric processes a DSM is created, that is used for the orthorectification process of the imagery. The orthorectified imagery in combination with the surface elevation information is used for the feature extraction of objects. In this research the potential benefit of the feature extraction output on the process of DSM to DTM filtering is emphasized. A comparison is made

between various DSM to DTM filtering techniques that include DSM to DTM filtering using only geometric criteria, only spectral criteria and a combination of the latter. So far no example of such a comparison is found in literature, and it is believed to be of additional value in the geo-specific terrain modeling process. A better understanding of the value of the spectral bands in the automated extraction of features and DSM to DTM filtering will lead to more effective terrain modeling for military simulations, but will also prove to be of value for topographic mapping agencies and other fields of research such as computer vision or artificial intelligence.

When using stereo imagery as a single source of information in the terrain modeling process various fields of research are addressed such as photogrammetry, DSM to DTM filtering, automatic feature extraction, object based image analysis (OBIA), machine learning and computer vision. As a result geo-specific terrain modeling can be considered a multi-disciplinary process. There is a substantial amount of examples of DSM to DTM filtering in literature (Baillard et al., 1998; Passini et al., 2002; Vosselman, 2000) and although proposed techniques could be considered successfully implemented for specific study sites, the process of DSM to DTM filtering is considered a challenging task (Jung et al., 2004) that leaves room for improvement. Most recent researches focus on the DSM to DTM filtering of LIDAR derived surfaces which are normally of superior quality when compared with photogrammetrically derived surfaces. Commonly applied techniques for DSM to DTM filtering are mathematical morphology and the use of surface segmentation in the filtering process. In order to overcome the existing problems of DSM to DTM filtering, the use of a combination of geometric and radiometric information is advised (Vosselman, 2000). In this research both the geometric and the radiometric information are used for feature extraction, and in turn the obtained results are used to improve the photogrammetrically derived DSM to DTM filtering process.

It is shown that objects from digital surface models can be extracted solely based on geometric segmentation criteria in order to improve the DSM (Gamba et al., 2000), or making use of very high resolution imagery as shown by Joshi (2010). The above researches focus on building extraction, making their methods hard to apply for the more remote and rural areas. Another research conducted by Haala et al. (1998) shows that classified multispectral imagery can be used to create a DTM in the DSM to DTM filtering process. Their strategy is to create an initial DTM based on mathematical morphology and use this DTM to create an nDSM containing all above ground objects and their elevation. This nDSM is used in the imagery classification process and the resulting classification is used in turn to filter the initial DTM even more based on classified objects such as forests and buildings. At present date higher resolutions and more spectral information is available to research the above mentioned approach and is assumed to produce better results.

In this study the aim is to improve the process of DSM to DTM filtering by using VHR multispectral stereo imagery obtained by the WorldView-2 satellite similarly to Haala et al. (1998), although in this research the goal is to integrate the geometry based and spectral feature based information in the DSM to DTM filtering process and not to perform two consecutive steps (first geometric filtering followed by spectral filtering of a DSM to DTM). WorldView-2 produces multispectral imagery using 8 different spectral sensors that focus on a specific strip within the electromagnetic spectrum. All bands record in the visible to near-infrared range, and as stated by Digitalglobe (2009a) should enhance segmentation and classification like no other information carrier. Novack et al. (2011) concluded from their research that the WorldView-2 imagery resulted in more simple and effective

feature extraction models compared to Quickbird-2 imagery, and also got better results in the extraction of certain roof and asphalt types using WorldView-2.

1.5 Research definition

1.5.1 Overall objective

The overall objective of this research is to improve both the effectiveness and the efficiency of geo-specific terrain modeling from a single source of input (high resolution 8-bands multispectral stereo imagery produced by the Worldview2), by improving DSM to DTM filtering of photogrammetrically derived DSMs and by investigating the potential benefit of additional non-typical spectral information as provided by the WorldView-2 satellite for the feature extraction and DSM to DTM filtering processes. This will result in higher quality geo-specific simulations with minimum production effort and time. Training in simulations using state of the art geo-specific terrains will increase the situational awareness and experience required for real missions in military simulations.

1.5.2 Research questions

In order to reach the overall objective 4 research questions are formulated that need to be answered.

1. What are suitable methods & techniques for geometry based DSM to DTM filtering as described in literature for photogrammetrically derived DSMs?
2. Can spectral feature based and geometry based DSM to DTM filtering be combined in order to improve the photogrammetrically derived DSM to DTM filtering?
3. Which DTM generation strategy yields the best results in terms of accuracy?
   - Geometry based DSM to DTM filtering.
   - Spectral features based DSM to DTM filtering.
   - A combination of spectral features based and geometry based DSM to DTM filtering.
4. Is there a significant difference in feature extraction output quality when using 8 or 4 bands WorldView-2 imagery?

1.6 Research scope

1.6.1 Scope

The scope is to create a workflow from existing and new techniques that is suitable for the creation of geo-specific terrain databases in possibly remote areas, using a single source of data in the process. A high quality DTM is believed to be a key element in the production and use of these geo-specific terrain models, therefore 3 methods of DSM to DTM filtering are researched in order to find the best method for geo-specific terrain modeling that can be applied for a wide variety of landscapes. The potential benefit of spectral information in the DSM to DTM filtering process is investigated using the 8 bands multispectral spectral imagery provided by the WorldView-2 satellite

as well as a reduced set of spectral bands which constitutes of red, green, blue and near infrared bands in order to determine the added value of additional non typical spectral bands for feature extraction.

1.6.2 Out of scope

It is not within the scope of this research to obtain the highest possible horizontal accuracy, which seems a paradox considering previous statements on quality improvement. Ground control points (GCPs) are measured in the field and will be derived from aerial imagery in order to gain lower circular errors, however the main reason for using GCPs is the comparison with reference data. This gives good insights into the quality of investigated methods. Typical geo-specific terrain modeling is done in remote areas without the availability of GCPs. The reason for using WorldView-2 in those areas is that it already has a reasonable spatial accuracy without the GCPs as described in the data section. After having defined the optimal production workflow for single source data, spatial accuracy is no issue in the correlation between various layers since they all suffer from the same (acceptable) spatial inaccuracy.

In this research both 4 and 8 band multispectral imagery is used in the feature extraction process. Although it is assumed that 8 bands will perform better in the feature extraction process, depending on its significance and complexity, it will be considered as suitable input for future geo-specific terrain modeling. In this research it is not attempted to compare the feature extraction results using 8 or 4 bands with other typical VHR imagery sources with the typical 4 bands. It is possible that a subset of 4 bands from WorldView-2 will not yield the best results when compared to other VHR MS imagery with 4 corresponding bands, because the 4 bands of WorldView-2 only cover a narrow strip of the electromagnetic spectrum (Fig. 6), but WorldView-2 is chosen as a potential single source of input given its empirically evident horizontal accuracy without the use of GCPs, as shown in the data section. If similar geo-location accuracies can be obtained by other satellites, and there is no significant difference in extraction between the 4 bands these satellites are suitable alternatives as well.

In this research it is attempted to improve the DTM quality by combining the spectrally based features resulting from feature extraction and geometry information in the DSM to DTM filtering process. No modifications or improvements are attempted to be made on existing geometry based DSM to DTM filtering techniques described by others.

1.7 Structure of this document

In general the structure in the project report coincides with the consecutive project processes undertaken to reach the project objective and to find answers to the specific research questions. A thorough literature study has been conducted but will be addressed separately for every project process to keep a clear and chronological research structure. Answering specific research questions is a complex task which in this research is always depending on the results of several project processes. Therefore not one chapter provides an answer to a specific research question, but answers are found when combining the results from the different project processes. In Fig. 4, the complex relationship between the project processes (chapters) and the research questions is given. It clearly shows that in

order to answer 3 out of the 4 research questions a combination of project processes is required. Answering the second research question, whether spectral and geometry based DSM to DTM filtering can be combined, for example depends on the obtained results from the feature extraction, the theoretic support found for the method in literature, and the obtained results when performing this type of DSM to DTM filtering. Below the research structure is explained in more detail.

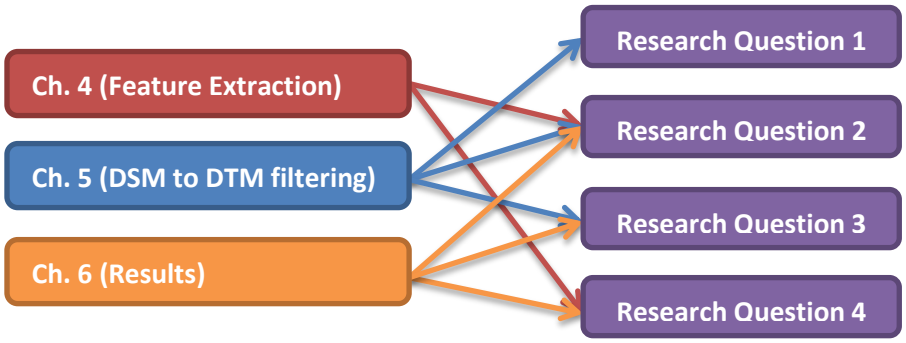


Fig. 4 Research question / research chapter relationship.

Chapter two describes the study sites, the data used, and finally the software used in more detail. Next the photogrammetry chapter (chapter 3) describes the background and applied methodology of all implemented photogrammetry processes including the applied methods for the output accuracy assessment (AA). The photogrammetry processes do not directly provide answers to the stated research questions, but is described separately because of its complexity and influence on all other processes. The photogrammetric output is believed to be the foundation for successful feature extraction and DSM to DTM filtering and as such it is crucial to have a better understanding of these photogrammetrically produced foundation layers (DSM and orthorectified imagery). The following chapter is dedicated to the process of feature extraction (chapter 4). Here the background information, applied methodologies and applied accuracy assessment methodology are described similarly to the photogrammetry chapter. Next the process of DSM to DTM filtering (chapter 5) is described in terms of background, methodology and accuracy assessment. This is the only chapter providing a direct answer to a research question. The background literature study conducted provides the answer to what suitable photogrammetrically derived DSM to DTM filtering techniques exist. In the results chapter (chapter 6) all the results from the applied methodologies described in chapter 3, 4, and 5 are described in detail. All the results combined provide answers to the remaining three research questions. All results are discussed thoroughly in chapter 7 and finally conclusions are drawn in regards of the individual research questions and the overall project in chapter 8.

2 RESEARCH SITE, DATA AND SOFTWARE

2.1 Study sites

Two study sites are used in this research. The first study site is located in the South-West of the Netherlands, near the Belgium border and includes the cities of Vlissingen and Middelburg of the Zeeland province. The study site can coarsely be divided into the land cover classes: water, residential, industry and cultivated land. The terrain is relatively flat and the area does not contain any natural forests and contains only small patches of planted forests. Although the terrain poses fewer challenges in terms of relief effects and forested areas, the study site is optimal since a reference DTM dataset is available for an accuracy assessment on the created terrain. The water body depicted in Fig. 5 is the Westerschelde estuarine which connects the Antwerp harbor to the Northsea. For that reason it is not a part of the Dutch delta works, which limits the influence of the sea in other parts of the region. The Westerschelde is characterized by a complex set of fairways and various sandbanks.

The second study site is located in the center of the Netherlands and includes Amersfoort city. The study site can coarsely be divided into the land cover classes: residential, cultivated land, forest and bits of industry. In contrast to the first study site this site has more variation in elevation. The increase of elevation is caused by the moraine named the Utrechtse Heuvelrug which was formed in the second last glacial period. This study site also has relatively large amounts of forests, making it a good addition to the first study site.

The study sites are not used entirely. Because of considerable computation intensity and time, smaller sites (5 * 5 km) are selected. A criterion for selection is diversity in both land cover and elevation.

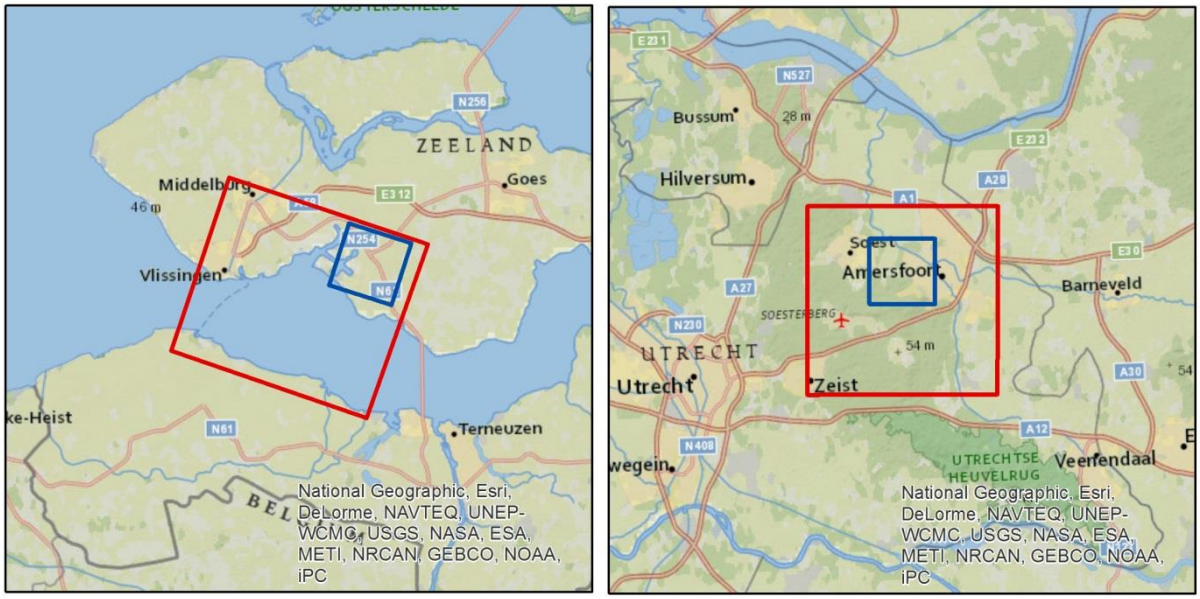


Fig. 5 Study sites (left: Vlissingen, right: Amersfoort). The red rectangle shows the footprint of the satellite imagery. The blue rectangle represents the study site used for processing.

In Fig. 5 the bounding box of the source imagery is shown with the red polygons, and the smaller sites, used for computation, with the blue polygons.

2.2 Source data

In October 2009 the satellite WorldView-2 was launched, and is now successfully part of the Digitalglobe constellation together with Worldview-1 and Quickbird. The WorldView-2 satellite is a bi-directional pushbroom scanner, meaning it will record a single strip in one overpass, without the need of a sweeping motion of a mirror to distribute the energy on the recorder. The main advantage of a pushbroom satellite is that each pixel within the strip is recorded with a single detector on a so-called charged-coupled device (CCD). Because no motion is required to capture an entire strip, more time is available for imagery capture, resulting in less noise and a more stable geometry (Kerle et al., 2004). Being bi-directional enables the satellite to capture stereo imagery.

Worldview-2 imagery has a geo-location accuracy specification of 6.5m circular error with 90% confidence (CE90), and a predicted performance between 4.6 and 10.7 meters (CE90), excluding effects caused by the terrain and off-nadir effects (Digitalglobe, 2009b). A study conducted by the National Geospatial-Intelligence Agency in which 64 panchromatic level 1B stereo pairs are processed showed an estimated Circular error with 90% confidence of 3.1m and a linear error with 90% confidence (LE90) of 4.5 (Bresnahan, 2011). Other studies such as the research of Cheng et al. (2010) had similar results. Even better results are obtained when making use of Ground Control Points (GCPs) for location accuracy improvements. Studies have showed horizontal accuracies of <3.5m CE90 with examples of accuracies of <1m (Hobi et al., 2012). More technical specifications of the WorldView-2 satellite are given in the table and figure below (Table 1 & Fig. 6).

Table 1 WorldView-2 satellite optical sensor specifications (Digitalglobe, 2009b).

General Digital globe specifications	
Sensor bands	Panchromatic + 8 MS bands (red, blue, green, near-IR1, red edge, coastal, yellow and near-IR2
Resolution	Panchromatic: 0.46 (resampled to 0.5) meters (GSD) at nadir MS: 1.84 meters (GSD) at nadir
Dynamic range	11-bits per pixel
Swath Width	16.4 kilometers at nadir
Geo-location accuracy (CE90)	6.5m CE90, excluding terrain or off-nadir effects, with registration of GCPs in the image 2m CE90
Retargeting agility	Time to slew 300 Kilometers: 9 seconds

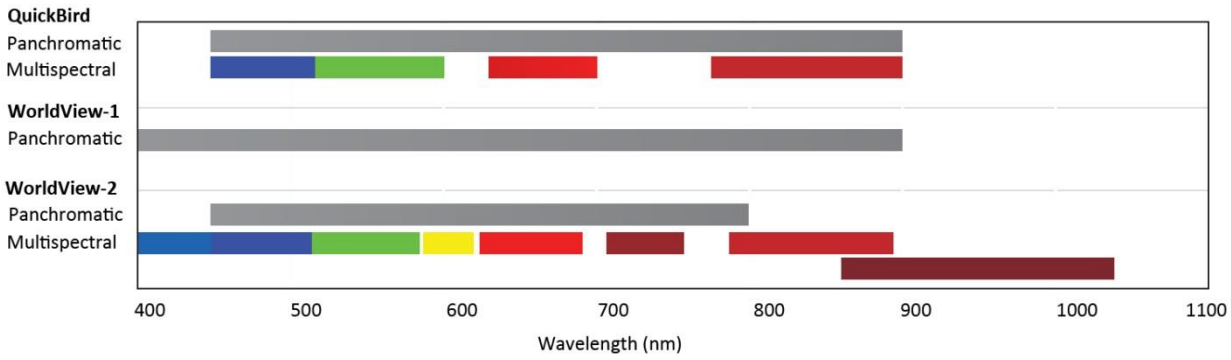


Fig. 6 WorldView-2 imagery band wavelengths (Digitalglobe, 2009a).

The specifications of the stereo imagery obtained over the study sites are given in the tables below (Table 2 & Table 3).

Table 2 Stereo imagery specifications of study site 1 (Vlissingen).

Specific imagery specs.	First image (stereo)	Second image (stereo)
Acquisition date	4 September 2012	4 September 2012
Time	11:26:26	11:28:23
Scan direction	Forward	Reverse
Projection	UTM 31 N (WGS84)	UTM 31 N (WGS84)
Mean in-track view angle	27.1 ^{o}	-24.8 ^{o}
Mean cross-track view angle	-8.40 ^{o}	-10.7 ^{o}
Mean off nadir view angle	28.3 ^{o}	26.9 ^{o}

Table 3 Stereo imagery specifications of study site 2 (Amersfoort).

Specific imagery specs.	First image (stereo)	Second image (mono)	Third image (stereo)
Acquisition date	19 August 2012	19 August 2012	19 August 2012
Time	12:15:04	12:11:01	12:15:41
Scan direction	Forward	Forward	Forward
Projection	UTM 31 N (WGS84)	UTM 31 N (WGS84)	UTM 31 N (WGS84)
Mean in-track view angle	9.8 ^{o}	-0.2 ^{o}	-12.7 ^{o}
Mean cross-track view angle	-3.7 ^{o}	-4.1 ^{o}	-4.6 ^{o}
Mean off nadir view angle	10.4 ^{o}	4.1 ^{o}	13.6 ^{o}

The WorldView-2 source images are single bands. The 8 spectral bands are combined (layer stacked) into two 8 bands multispectral images for the Vlissingen study area and into three 8 bands multispectral images for the Amersfoort study area. After the bands are combined the resulting 5 multiband images were subjected to hyperspherical color space (HCS) resolution merge (pan-sharpened). When pan-sharpening an image, ideally you want the resolution found in the panchromatic bands in combination with the exact colors found in the multispectral bands, however in reality there is always a loss in one of the two characteristics. The HCS resolution merge has proven to outperform other available pan-sharpening techniques on WorldView-2 imagery (Padwick et al., 2010).

The second important source of data is a set of ground control points, which are essential for the improvement of the satellite image geo-location accuracy. 19 GPS measured ground control points[1] are available well distributed over the Amersfoort study site. These ground control points include the horizontal and vertical coordinates as measured in the field using a Leica GPS system 500, with accuracies up to 3 centimeters. For the Vlissingen study area, these points measured in the field are not available and therefore aerial orthorectified imagery[2] from 2011 with a resolution of 0.4 meters and an average horizontal error of 0.5 meters is used in the process of deriving ground control points. These points are locations on relatively easily identifiable ground locations in the aerial orthorectified imagery as well as the satellite imagery. This way the horizontal coordinates can be derived from the reference dataset. In total 7 horizontal ground control points have been extracted.

[1] Courtesy of TNO
[2] Courtesy of the Royal Netherlands Army Geographic Agency (RNLAGA)

2.3 Reference data

In general three types of reference data are used. These are reference imagery, reference elevation models (DSM and DTM) and reference large scale topographic feature objects. Reference imagery and elevation models are used for the production and accuracy assessment of the photogrammetric output and the accuracy assessment of the produced DTMs. The large scale topographic data is used for the accuracy assessment of the automatic feature extraction output.

2.3.1 Reference imagery and elevation models

Three datasets are used as reference data in the photogrammetric phase for production and accuracy assessment purposes. The first reference dataset is the 0.4 meter spatial resolution orthorectified aerial imagery mentioned in the previous section, which is available for both study sites. Apart from being a source of control points, the same reference imagery is used after the photogrammetric processes to determine the horizontal accuracy of the produced imagery, surfaces and terrains.

The second dataset is a 90m resolution surface model from the Shuttle Radar Topography Mission (SRTM version2). SRTM version2 is also referred to as the "finished SRTM" version which has undergone substantial editing in order to improve its quality (Farr et al., 2007). This surface model is used in the process of DSM extraction because it will potentially improve the extraction process. In this process it is used as an initial elevation seed source. SRTM version2 is used in this process, because the goal is to optimize the workflow for remote areas where products with higher spatial accuracies than SRTM are likely not (freely) available. SRTM is freely available for most parts of the world. In Table 4, the determined errors found in the SRTM data are given, divided in global regions.

Table 4 SRTM errors in meters with 90% CL (Rodríguez et al., 2005).

	Africa	Australia	Eurasia	Islands	N. America	S. America
Absolute Geo-location Error (m)	11.9	7.2	8.8	9.0	12.6	9.0
Absolute Height Error (m)	5.6	6.0	6.2	8.0	9.0	6.2
Relative Height Error (m)	9.8	4.7	8.7	6.2	7.0	5.5
Long Wavelength Height Error (m)	3.1	6.0	2.6	3.7	4.0	4.9

The third reference dataset used in the photogrammetry phase is the Actueel Hoogtebestand Nederland 2 (AHN2). This data is also used in the final quality assessment of the DTM outputs. It is used as a reference elevation dataset to estimate the vertical accuracies of the produced surface and terrain models. The AHN2 is a laser altimetry generated nationwide elevation model, produced for the executive body of the Dutch Ministry of Infrastructure and the Environment. It has a spatial resolution of 0.5 meters. Two versions are available that both cover the entire study site. The first one is a non-interpolated DSM and the second a non-interpolated, but filtered DTM as shown in Fig. 7. DSM to DTM filtering of the AHN2 reference elevation has partly been done automatically and for a large part manually. The height accuracy is stated to be no more than 5 cm standard deviation and no more than 5 cm systematic deviation (van der Zon, 2012).

The AHN2 data has to be preprocessed in order to be used effectively as a reference dataset, e.g. the vertical datum of the AHN2 elevation model has to be transformed from the national vertical datum

to the global vertical datum as output by the photogrammetric processes. This preprocessing is described in more detail in appendix V.

Fig. 7 Reference elevation (left: AHN2 unfiltered, right: AHN2 filtered).

2.3.2 Large scale topographic reference data

The results obtained from the automatic feature extraction processes are compared to the feature objects from existing large scale topographic datasets that are similar in terms of scale with the extracted features from the imagery. Two datasets[3] are used, the object oriented national topographic product TOP10NL, which is produced by the Dutch national mapping agency (Kadaster), and a national dataset containing municipal information in regards to buildings and addresses called the Basisregistratie Adressen en Gebouwen (BAG). The BAG is a Dutch key register containing municipal information from al buildings and addresses in the Netherlands. All municipalities are custodians of the data, and the Dutch national mapping agency funnels all this information into one dataset through the so-called BAG national facility (BAGLV) of which they are custodian and distribute the BAG as open data to various parties in the Netherlands.

The TOP10NL is a 1:10000 scale topographic product and is well suited as reference data for the targeted feature object classes, forest vegetation, roads and water in the accuracy assessment of the feature extraction process. For the targeted building feature object class the TOP10NL is less suited for accuracy assessment since it has been subjected to various forms of generalization. Instead the BAG is used. The BAG dataset contains buildings geometries with a 1:1000 scale, and is therefore better suited as reference data for the automatic feature extraction accuracy assessment of buildings in comparison to the TOP10NL.

2.4 Software

The software used in this research for the photogrammetric processes and the feature extraction is ERDAS IMAGINE (2011), with the IMAGINE Objective add-on for the feature extraction process (object-based imagery analysis), and the ERDAS Leica Photogrammetry Suite (LPS) with the enhanced Automatic Terrain Extraction (eATE) module for all photogrammetric processes including terrain extraction. The desktop GIS application ArcGIS 10.1 is used for the more generic processes of preprocessing, visual inspections and accuracy assessment. Another application used for visual

[3] Courtesy of the RNLAGA

inspections is the FugroViewer. FugroViewer is freeware produced by Fugro Geospatial Services, and is effective in the visualization of large point clouds that result from DSM extraction from stereo images. Other than visualization no analyses have been performed with the software.

Yet another freeware product used is PyScripter, which is used in conjunction with the programming language python 2.7. Using PyScripter it is possible to import the ArcGIS geoprocessing modules and automate processes. This is particularly done in the DSM to DTM filtering process.

The choice for commercial off-the-shelf software products (COTS) and freeware is related to availability. It is assumed that when using different software packages the results obtained will be similar to the results in this research since the aim is to utilize a set of existing and well documented techniques.

3 PHOTOGRAMMETRY

The first major phase in the production of a geo-specific terrain model is the so-called photogrammetry phase. Photogrammetry is the art, science and technology to make measurements from (stereo) imagery. This imagery could be aerial, from satellites, from ground cameras and more recently also from video stills. In this photogrammetry phase the terrain will be extracted from the stereo imagery using photogrammetric processes and with the resulting DSM, orthorectified imagery is created that is later used in the automatic feature extraction phase. The required photogrammetry processing steps are streamlined within the available software package (ERDAS LPS). These consecutive steps are:

Step 1: Interior orientation
Step 2: Exterior orientation
Step 3: DSM extraction
Step 4: Orthorectification

These photogrammetric processes rely heavily on the colliniarity equations, that describe the relationship between image coordinate, ground coordinates and the exposure station (Lillesand et al., 2008). Using the colliniarity equations the X Y Z ground coordinates of every pixel can be determined. This phase does not address any research question, but is considered a phase nevertheless since a lot of processing steps are included, which need fine-tuning in order to get the best results. In the following subsection background information is given on the standard photogrammetric process. Furthermore, the applied methodology is described for the processes, as well as a detailed description on the accuracy assessment of the output DSM and orthorectified imagery.

3.1 Orientation and triangulation background

The interior orientation defines the internal geometry (principal point, focal length, fiducial marks and lens distortion) of the camera or sensor at the time an image is captured (ERDAS, 2010c). These properties can be determined manually from analogue aerial images, from camera calibration reports created in laboratories and are provided with modern day digital aerial imagery or in various schemas accompanying digital satellite imagery. In the case of WorldView-2 imagery the properties are provided with the satellite imagery and are automatically extracted by the software.

After the interior orientation is successfully established for the camera the next step is solving the exterior orientation. The exterior orientation defines the position and angular rotation of the camera or sensor at the moment of image capture, through the mathematical relationship that is established between the image coordinates and Ground Control Points (Jensen, 2007). Ground control points (GCPs) are points of which X, Y and optimally Z coordinates are known. Ground control points can be measured in the field using GPS, or can be obtained from other sources, such as maps or other imagery sources. The better the quality of ground control points, the more accurate the relation between points in the captured imaged and its true geospatial location.

In Erdas LPS various types of points can be inserted into the project in order to help solve the exterior orientation e.g. tie points, horizontal control points, full control points and check points. Tie points are points that are manually or automatically identified in overlapping images, but of which no coordinates are known. Horizontal control points are points found in overlapping images of which only the horizontal coordinates are known. Full control points are points that are identified in the overlapping imagery and of which the horizontal and vertical coordinates are known. Check points are horizontal or full control points that are used only as an input for quality reporting and are not used to solve the exterior orientation.

After inserting full control or horizontal control points into an LPS project a mathematical relation is established between the ground coordinates, the image coordinates in the captured satellite image and the position and rotation of the camera. After this is established a desired set of extra tie points can be extracted automatically. The final step in solving the exterior orientation is the process of block triangulation. In this process the horizontal and vertical coordinates are determined for all tie points, horizontal control points, full control points and check points. Block bundle adjustment is used by the software to minimize and distribute the errors in the triangulation process.

3.2 Applied orientation and triangulation methodology

For both study areas the multispectral images (8 bands), panchromatic bands and Pan-sharpened images were loaded into an LPS project. The interior orientation was solved by referring to the available information provided with the images. The Pan-sharpened images are not used in solving the interior or exterior orientation, but will later on be used in the orthorectification process. The interior-, exterior orientation and triangulation is done for the entire images and not on a smaller portion of the images. Only while extracting a DSM and during the orthorectification the image, a smaller portion of the imagery is used.

For Amersfoort the exterior orientation is resolved using 19 high accuracy GCPs (full control points) that have been captured inside the study area[4]. In the Vlissingen study area no field measured GCPs were available and instead existing aerial orthorectified imagery is used as a source of GCPs (horizontal control point). In total 7 horizontal control points were used to solve the exterior orientation. After the insertion of full- or horizontal control points extra tie points are generated distributed throughout the study area. All the inserted and generated points are input into the triangulation process and horizontal and vertical coordinates are determined. The triangulation report is reviewed to identify tie or control point that have a relatively large negative effect on the RMSE of the triangulation process and these points are removed from the set of points. After the removal of erroneous points the triangulation is re-run. This process of point removal is repeated until the RMSE value is stabilized and does not decrease any further. The completion of the triangulation process with an acceptable RMSE means that the exterior orientation is solved and the next step, DSM extraction, can be initiated.

[4] Courtesy of TNO

3.3 DSM extraction background

DSMs are created using the parallax principle. Parallax is the relative change in object position at different viewing angles (Lillesand et al., 2008). Objects closer to the lens (at a higher elevation) have a relative larger displacement than objects farther away from the camera. By measuring the relative displacement from overlapping images the surface height can be calculated. When using ERDAS LPS, two approaches are available, namely Automatic Terrain Extraction (ATE) and enhanced ATE (eATE). ATE identifies a set of feature points (interest points) based on its grey level variation and contrast. The conjugate feature points are sought in the overlapping images. When conjugate points are located, correlation coefficients are determined to eliminate so-called low certainty point matches. The correlation coefficient is a value between 0 and 1 in which the value 1 reflects a 100 percent certainty point match. After some final statistical cleanup strategies the 3D coordinates are computed for the remaining points (ERDAS, 2010b). Using the eATE module in LPS, allows you to search conjugate pixels in the overlapping images for every pixel instead of a set of interest points. As a result a larger set of mass points is used in the image matching, and a larger set of points is output from the process. Although the quality of the 3D coordinates is not affected by the use of eATE over ATE, resulting point clouds are eventually interpolated into a closed raster surface. When fewer points are put into an interpolation process, more values are estimated, and when a denser point cloud is available, fewer values have to be estimated. This results in a potentially higher accuracy DSM when using a denser point cloud (eATE), however the quality of the points used in the image matching are the same.

The process of image matching takes place at specified pyramid levels when using ATE or eATE. The process is started at a low resolution pyramid level and iterates through the pyramid levels until the original resolution or selected pyramid level is reached. At every pyramid level the approximate point location, correlation coefficient determination and selection of points based on its correlation coefficient takes place. Typically the threshold for the correlation coefficient becomes higher towards the last iterations, which results in a higher amount of (low certainty) points being removed.

A large variety of parameters and strategies are available for the DSM extraction process. The most influential are the search windows size, correlation window size and correlation coefficient threshold. The search window size is the window that is used to identify the conjugate feature point after its approximate location is found. The correlation window size is the window that is used to determine the correlation coefficient of the matched points. It is typically larger in areas with less relief and topographical complexity and smaller in areas with high topographic relief such as urban areas. With the correlation coefficient threshold the operator can choose to lower the coefficient threshold at will. By doing so more points will be output, since less certain (lower correlation coefficient) point matches are allowed. When using ATE the above parameters and various others are predefined in a set of available strategies that can be applied on the entire area or on specified areas of interest. It is also possible to create a custom image matching strategy by adjusting the parameters at will.

Both ATE and eATE have the option to use seed sources in the DSM extraction process. Possible seed sources are breaklines, a relatively more accurate DSM, or point elevation data. The influence of the seed sources can be set as an initial elevation value after which consecutive process iterations do not take the seed source into account, or the seed sources are written to the output file as is. By default

a global DSM with 1km resolution is used by the software to approximately align the first pyramid level used in the iterations with a vertical datum.

3.4 Applied DSM extraction methodology

In this research the entire process of DSM extraction is done in two steps, using both ATE and eATE. The method chosen is considered the optimum after various trials and visual inspections. In the first stage the multispectral images are used in the ATE image matching process with the projects tie and control points and a 90 m resolution DSM (SRTM version2) as seed sources. In the second step the panchromatic bands are used in the eATE image matching process with the projects tie and control points and ATE output surface as seed sources. All DSM extraction processes are performed on a representative 25 square kilometer subset of the original study area to shorten the production time. Pan-sharpened imagery has not shown to generate better results in the DSM extraction and is therefore not used. Below a more detailed description of the separate DSM extraction steps is given.

3.4.1 ATE DSM extraction

Seed sources are an important aspect in the generation of DSMs, and therefore the SRTM surface is chosen as the seed surface in the first step. The SRTM dataset has a vertical error smaller than 9m with a 90 percent confidence level, and a horizontal error smaller than 13 meters with a 90 percent confidence level (Rodríguez et al., 2005), and is most likely the best available elevation model for the areas of interest of military geo-specific terrain modeling. This is also the main reason why the AHN-2 surface model is not used in the production of the DSM, but only as a reference for accuracy assessment. The projects tie and control points are also used as a seed source in the ATE DSM extraction. These points, as well as the seed surface, are used solely in the first iterations during the image matching process. In this way the seed sources help with the initial vertical datum alignment, but are limited in influence because the consecutive iterations no longer take these seed sources into account.

The multispectral images were used in the first DSM extraction step using ATE. The relatively lower resolution of the multispectral imagery is sufficient because the goal of the first step is only to create a high quality seed surface for the next DSM extraction step. The multispectral imagery has a 2m spatial resolution and is used to create a 2.5m spatial resolution DSM. Alternatively the panchromatic imagery could also have been used up until the second to last iteration, considering that the pyramid level used at that point is equal to the resolution of the multispectral imagery.

A search window of 11 pixels is used in order to locate the conjugate points in overlapping imagery. After a point match was made the correlation coefficient was determined using a window of 7 pixels. A correlation coefficient threshold of 0.8 was imposed on the output points.

3.4.2 eATE DSM extraction

In the second step the higher resolution panchromatic imagery is used for the image matching. Image matching is performed using eATE in order to create a dense set of matched points that is eventually interpolated into a 0.5m spatial resolution DSM. The output surface from the ATE step is

input as a seed source along with the same tie and control points used in the first step, in order to align the elevation values to the desired vertical datum. The seed surface is set to influence only the first iteration and the seed points are set to influence all but the final iteration.

During the second step the correlation window size is increased to 9 pixels and the correlation coefficient threshold is lowered to 0.6 since more output points are desired to be output by the DSM extraction process. A consequence of lowering the coefficient threshold could be the increased presence of spikes in the interpolated surface. However, these can be removed in the DSM to DTM filtering process by applying e.g. low pass convolution filtering.

3.5 Satellite imagery orthorectification

Orthorectified images can be created from aerial imagery and satellite imagery. Orthorectification refers to the photogrammetric process in which relief displacement and altitude variations (scale variations) are removed from the image by using the DSM as a mean to alter the conical / perspective projection of the image into and orthogonal projection (Lillesand et al., 2008). In Fig. 8 an example of typical relief displacement is given.

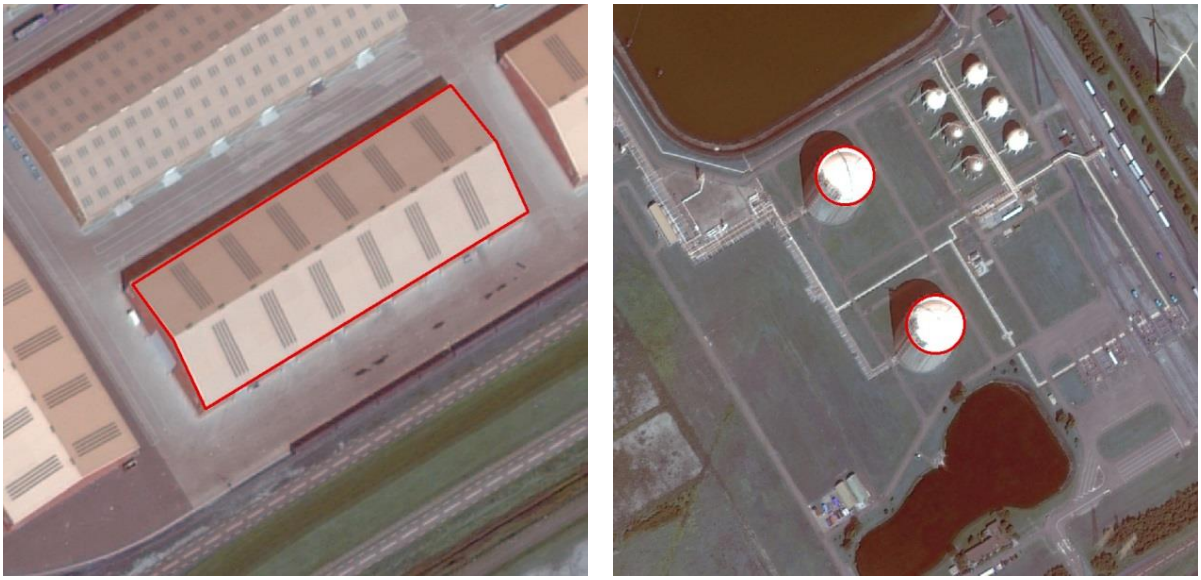


Fig. 8 Relief displacement present in orthorectified imagery. The object footprints do not share the same coordinates as the rooftop footprints indicated in red.

Severe distortion, which is particularly evident at building edges in urban areas, is a known phenomenon in large scale orthorectified imagery and is caused by the DSM (Jensen, 2007). This issue can be resolved by manually correcting the DSM to create flatter rooftops and crisper building edges. There is a trade-off between the reduction of relief displacement allowed in the orthorectified imagery and quality of the surface model. The precise positioning of an orthorectified image depends for the most part on the quality of the extracted surface model. Especially in urban areas the quality of orthorectified imagery depends on the representation of objects as geometric shapes in the surface (Palà et al., 2002). The better the surface is modeled the better local relief displacement (particularly present in buildings) can be reduced. The density of points matched in the DSM extraction is therefore of great influence on the orthorectification and often does not suffice for urban areas when using a photogrammetrically derived DSM. The reduction of spatial resolution in

the DSM reduces visual anomalies in the orthorectified image, but as a result some relief displacement is still visible in the orthorectified imagery.

The relief displacement or lean of buildings are causing the occlusion of other parts of the image and therefore methodologies have been created based on the automatic identification of occluded image parts to create so-called true orthorectified imagery (Habib et al., 2007). The top of the buildings are moved to their true planimetric location causing empty bits of imagery that can be filled up by other overlapping images taken from different camera locations and angles. In this research true orthorectified imagery production is out of scope. As a result there is a potential loss in horizontal accuracy when automatically extracting features since the building rooftops are not necessarily covering their true planimetric location, but the amounts of horizontal displacement are considered acceptable for the production of geo-specific terrain models.

3.6 Applied orthorectification methodology

The WorldView-2 imagery bands are combined into 8 bands multispectral image (Red, Green, Blue, Near IR1, Near IR2, Red edge, Coastal & Yellow) with a 2m spatial resolution, and are pan-sharpened into 0.5m spatial resolution multispectral imagery using the available panchromatic band. The pan-sharpened multispectral imagery is orthorectified with a 5m spatial resolution DSM. Using a DSM with a reduced spatial resolution will reduce anomalies in the orthorectified imagery that could otherwise only have been avoided by laborious manual improvements of the DSM. However, some relief displacement will remain and has to be taken into account.

3.7 Photogrammetry accuracy assessment

The accuracy of a DSM or DTM is the single most important aspect in terms of quality or success (Li et al., 2005). The accuracy of a DSM or DTM can be assessed in various ways depending on the available resources and the desired level of accuracy determination. It is the set of requirements that determine the suitability of a generated DSM or DTM. Two types of quality determination are described by Kasser et al. (2002), namely the internal and external validation. The internal validation is performed by looking at the properties of features found in the image with generic knowledge of landscape features. The visual inspection of the direction of flow in a river, or the confirmation of expected building properties in the results can already give valuable information on the quality. When comparing the results from two methodologies it is possible to identify the level of error found when comparing the two and is clear that one is more accurate than the other. However, it is only possible to properly determine the absolute error in a DSM extraction result when it is compared to reference data which is assumed to be error free (Fisher et al., 2006). The external quality validation is performed using high quality reference data for comparison, such as existing surface models, or measurements from the field, such as GCPs made using GPS.

Visual inspections are performed (internal validation) and reference surfaces and orthorectified imagery is used in this research for external quality validation. Although the determination of quality is best made through measurements, visual inspections of output data are believed to provide valuable insights in the output quality and can direct further, more detailed research on the quality of the output data. Even though the military geo-specific terrain models are most likely required for

areas where high quality large scale geospatial information is absent, using reference data in this research provides insights in the potential quality of geo-specific terrain modeling in other areas. A reference DSM and reference orthorectified imagery is used in order to assess the horizontal and vertical deviation from these reference datasets. The initial inaccuracy of the reference data is considered insignificant for the determination of output quality in this research and is therefore not taken into account here.

3.7.1 Applied visual accuracy assessment methodology

Visual inspection will be performed on the output DSM and the orthorectified imagery. Two products are output by the software when creating a surface model, namely a point cloud, which is the direct result of the image matching and an interpolated rasterized DSM. The point cloud will be inspected through the visualization of the point cloud in a Triangular irregular network (TIN), and a raster representing the slope of the interpolated and rasterized point cloud is created and inspected. As shown in Fig. 9 any systematic errors made in the image matching become clearly evident when visualizing the slope of a surface, since spikes introduced by bad image alignment or incorrect image matching cause high slope values.

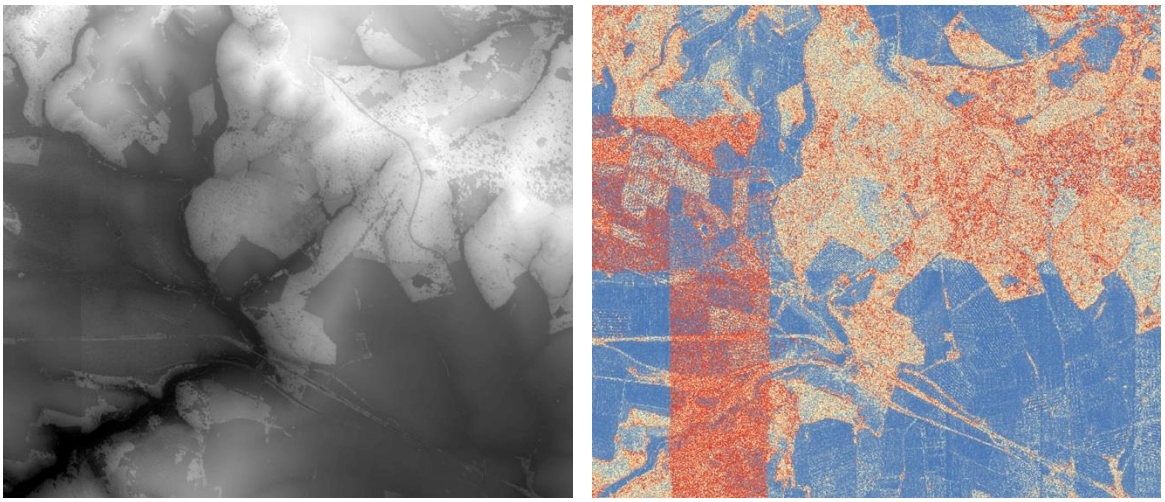


Fig. 9 Visual DSM error inspection (images courtesy of RNLAGA). Left: surface shown in greyscale, right slopes visualized using color ramp.

The orthorectified imagery will be inspected in terms of abnormalities such as blurred or deformed objects in the imagery that could be the result from outliers (spikes) in the DSM used in the orthorectification process, or the nature of the DSM itself as described in the section on orthorectification.

Another inspection, referred to as discrepancy mapping, is a visual as well as a statistical vertical accuracy assessment methodology. The reference DSM is subtracted from the generated DSM, and from the resulting raster layer large discrepancies between the two surfaces become immediately evident, which will be subjected to a more detailed inspection as to see whether there are logical explanations for these discrepancies. It is also possible to see the general discrepancy between the layers and whether the discrepancy is local or present in the entire surface. After the visual inspection the discrepancy layer is converted to include only positive values i.e. the negative values are turned into positive values.

3.7.2 Applied horizontal accuracy assessment methodology

The horizontal accuracies of the DSM and orthorectified imagery are assessed by a set of distributed points over the two study areas. For both study areas around ten points are distributed over the area. These points are placed at specific locations using the orthorectified satellite imagery as guidance. All points were placed on bare earth terrain, at relatively easily identifiable locations that are not suspected to change within the time difference of a year. Points are for instance located on the marking of a sport field or the markings on a road. In Fig. 10 the distribution of these check points is shown.

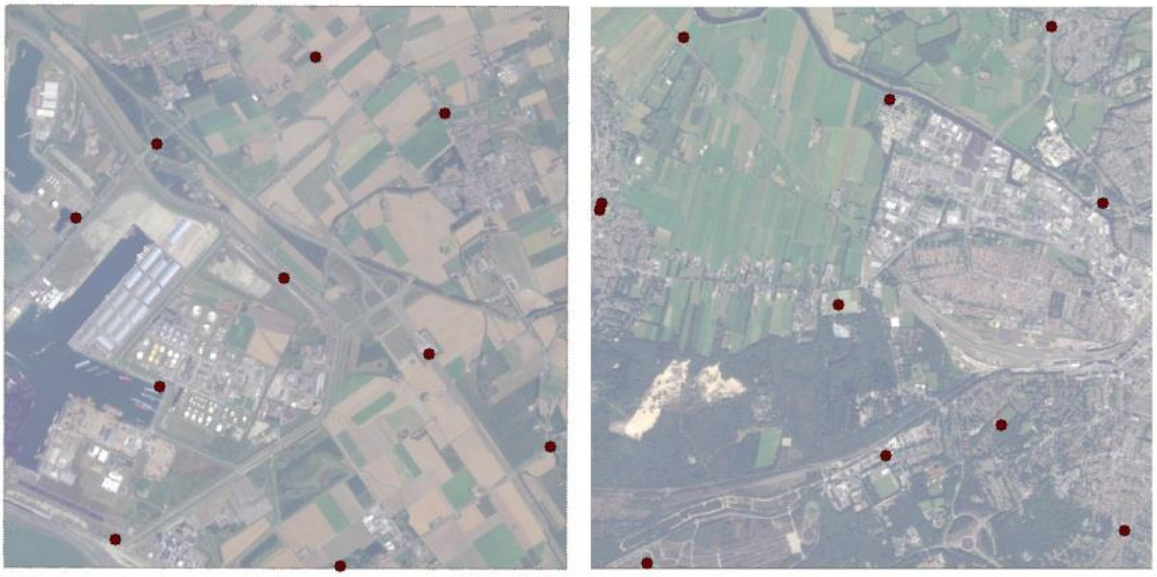


Fig. 10 Distribution of accuracy assessment check points (left: Vlissingen, right: Amersfoort).

Since the check points are located at relatively easily identifiable locations, the same points are located in the reference aerial orthorectified imagery. From these two sets of points the horizontal deviation between the two point sets is determined and used in the quality assessment. The distance in meters between the corresponding points is determined as well as the angle of deviation. Considering the fact that the DSM is produced from the imagery and later on used as input for the orthorectification of the satellite imagery, it is assumed that the horizontal accuracy of the satellite imagery is the same as the horizontal accuracy of the DSM.

3.7.3 Applied vertical accuracy assessment methodology

The vertical accuracy assessment is performed on the DSM using the same set of points that was created for the horizontal accuracy assessment. However, when comparing the points located using the generated orthorectified imagery with point locations in the reference data, the horizontal accuracy becomes an important factor. When the horizontal deviation between generated and reference imagery is too high this would impede the analysis of the vertical quality. Whenever the horizontal deviation between the generated imagery and the reference data is linear in nature, a correction in horizontal position is made on the reference data in order to perform the quality analysis. This correction is based on the displacement of conjugate points in the reference data and therefore does not take into account any remaining relief displacement present in the generated

orthorectified imagery. Note that some horizontal inaccuracy is accepted in the output data, considering the locations of the geo-specific terrain models, and the absence of GCP for these locations. Therefore, the reference data can be shifted horizontally in order to align optimally with the photogrammetric output. The elevation values at the point locations in the generated surface and the reference surface are statistically compared to give insights in the vertical accuracy.

4 AUTOMATIC FEATURE EXTRACTION

The process of feature extraction is considered the second major production phase, and together with the process of DSM to DTM filtering will help answering the second research question, whether spectral and geometric DSM to DTM filtering can be combined considering the output quality of extracted features. The process of feature extraction, however, does provide a comprehensive answer to the fourth research question, whether there is a significant difference in feature extraction output quality when using 8 or 4 band WorldView-2 imagery. If there are no significant differences this will lead to the conclusion that if other VHR resolution imagery sensors have at least the same geo-location accuracy as WorldView-2, and a near-IR band, they are potential candidates to be a single source of input as well. In this chapter background information is given on automated feature extraction in general, the specific feature extraction workflow implemented in Imagine Objective, and the accuracy assessment of the extracted features. The application of the feature extraction methodology and feature extraction accuracy assessment is described subsequently.

4.1 Feature extraction background

Traditionally image analysis has been pixel based, and so was the extraction of features. However, a shift in the paradigm of image analysis has taken place from the pixel based image analysis towards the object based image analysis (OBIA) in the last two decades (Hay et al., 2008). OBIA emulates the way humans organize the images they see in order to comprehend them (Hay et al., 2006). This means that in contrast to pixel based analysis, other visual cues, e.g. textures, shapes and context are also taken into account. With the advent of computer technologies and VHR satellite imagery, the desire for large scale feature extraction in an automated fashion emerged, and the object based image analysis approach was found to have significant benefits over pixel based analysis. Many examples of successful OBIA implementation can be found in (Blaschke, 2010).

OBIA is built upon the well-established techniques of image segmentation and classification. The classification is conducted supervised or unsupervised. However, for the direct extraction of the specific features required for geo-specific terrain models a supervised classification is required. Segmentation divides an image into regions or so-called geometric primitives, based on the applied classification, training samples and input parameters. These segments can have additional attributes such as a pixel probability value, spectral mean or variance. For this very reason a segment can hold more spectral information than individual pixels. This is a major benefit for the process of OBIA and therefore the process of image segmentation is considered the cornerstone of OBIA. However, in order to have a successful segmentation the objects of interest should be bigger than the individual pixels. With the increase in spatial resolution found in satellite imagery products, larger scale objects such as individual buildings can be found with OBIA. The scale at which segmentation is most suitable differs per object and possibly even per image, and although most frequently the segmentation scale and parameters are determined on a visual trial and error basis, new methods are created to overcome this biased form of determination such as the method of using local variance as a measure to determine optimal segmentation scales (Drağuţ et al., 2010).

Information sources can be fused together in the process of OBIA in order to add both spectral and geometric information. When combining multiple sources the data needs to be co-registered (Kiema, 2002), and correlation issues need to be taken into account. It is also possible to add multiple sources of information for feature extraction from the same satellite or aerial imagery source. Because the imagery and surface are produced within the same project they both suffer from the same errors in geo-location. However, they have no correlation problems and don't require co-registering the vertical and horizontal datum and projections. The inclusion of elevation information in the OBIA process, whether derived from another source or the same source, will provide the required additional value for image interpretation (Sowmya et al., 2000). Optimally the added elevation information is either a DSM or an nDSM. An nDSM can be used in the building extraction process as shown by (Mason et al., 1997) in their study on informal settlement building extraction. By applying a threshold on the nDSM low objects, such as cars, are filtered from the nDSM. However, this method does not take into account individual trees and forests, which are present in the study sites in this research. The DSM can be used to estimate a delta height between the extracted objects and its surroundings, and a threshold can be applied to eliminate unwanted objects.

In contrast to the WorldView-2 product documentation and other researches such as Novack et al. (2011), that conclude that the extra non typical bands of the WorldView-2 imagery do provide additional value in terms of feature extraction, there are also examples in literature stating that there is a large amount of information redundancy found in multispectral feature extraction making it harder to extract key features. Various methods for dimensionality reduction exist and have been subjected to comparative studies (Tian et al., 2005). However, a comparison between a lower dimension feature extraction and a full VHR multispectral imagery based feature extraction has not been found in literature and therefore it is assumed that indeed the multispectral nature of the WorldView-2 imagery will have an advantage over the imagery with reduced dimensionality. It remains interesting to see whether dimensionality reduction would outperform multispectral imagery in OBIA, and whether the input of 8 bands multispectral imagery would yield better results in OBIA when reduced in dimension in comparison to dimensionality reduced 4 bands multispectral imagery.

The term Geographic Object-Based Image Analysis (GEOBIA) was introduced by Hay et al. (2008) to differentiate the latter from the more generic term OBIA, which is also applied in the fields of computer vision and biomedical imaging, and to clearly emphasize that the application of OBIA on remotely sensed earth surface imagery is a sub-discipline of Geographic Information Science (GIScience) and inherits all its benefits and challenges (e.g. atmospheric distortions and horizontal positional accuracies). Throughout this report the terms OBIA is used.

4.2 Imagine Objective workflow

IMAGINE Objective, much like OBIA, attempts to extract features based on the same principles of human visual image interpretation. The software uses training samples (supervised classification) as a method for machine learning in combination with the quantification of visual image interpretation cues such as color/tone, texture, size, shape, shadow, site, pattern and association on both the pixel and object level (ERDAS, 2010a). The strategy for human vision emulation applied by Imagine Objective has also been described as the combination of inferential learning and expert knowledge

(Lane, 2009), meaning that the input of the training samples in terms of shape and spectral characteristics, and the optimal workflow as determined by the expert are stored and can be used at a later date for other locations. In Fig. 11 the generic Imagine Objective feature extraction process flow is shown. Different feature objects require different extraction strategies, and the specific use of the processes can be tailored for every object type.

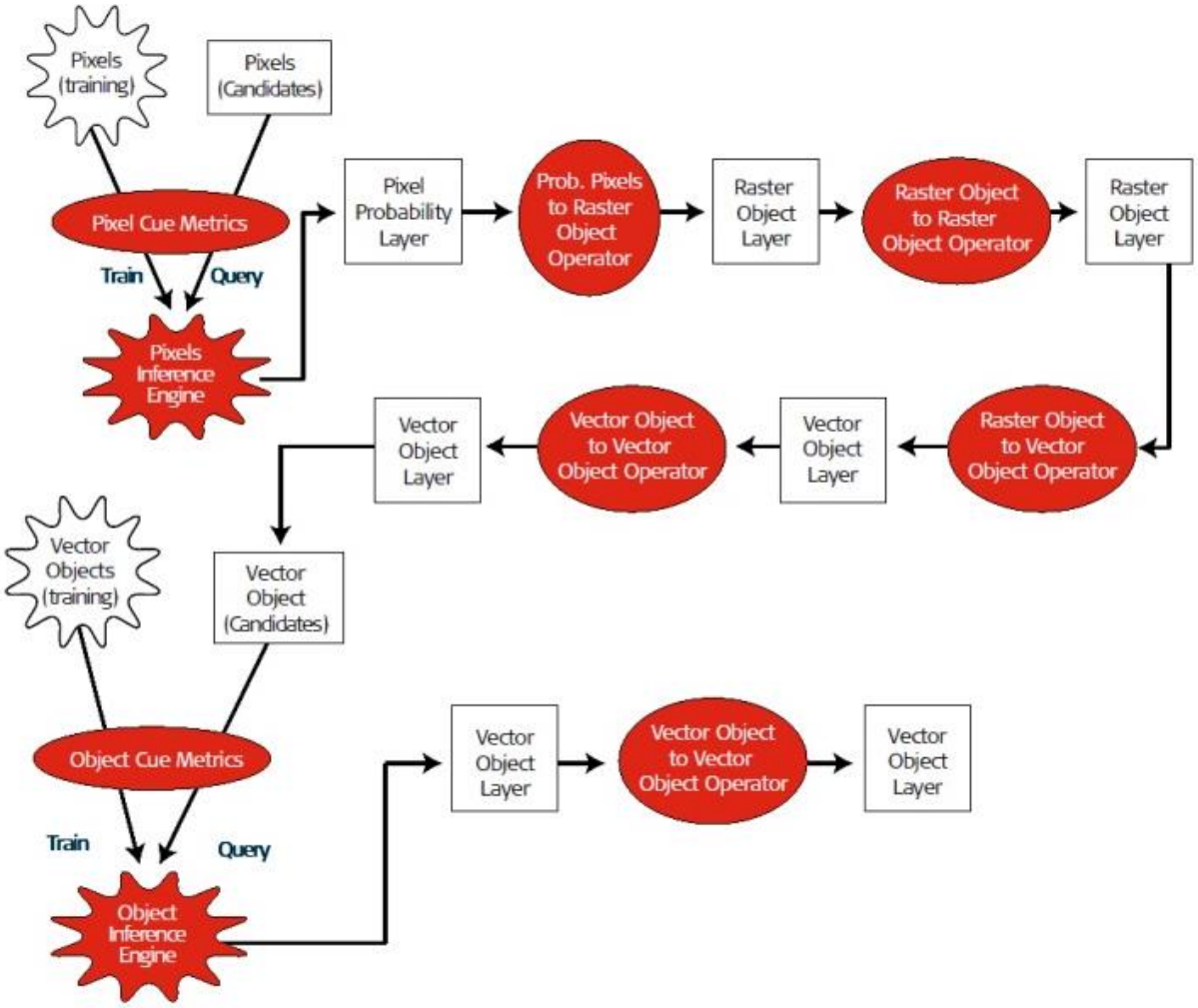


Fig. 11 Imagine Objective workflow (Intergraph, 2013).

In the first process a pixel classifier is trained through the manual insertion of object pixel candidates and background pixel candidates. After a representative set of training samples is created for the targeted feature object, a probability layer is computed using metric pixel cues such as spectral values and texture. In this layer every pixel is assigned a value between 0 and 1, in which higher values indicate a higher probability for the targeted feature object. An example of the pixel probability layers is shown in Fig. 12.

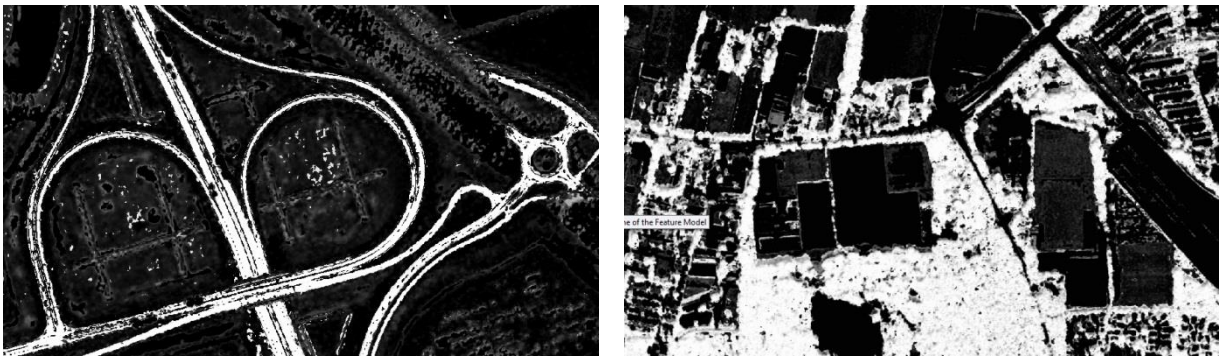


Fig. 12 Pixel probability raster (left: Vlissingen roads, right: Amersfoort forest).

In the probability pixels to raster object operator, the individual pixels are converted into raster objects i.e. a continuous cluster of pixels with a shared ID and attributes (zonal mean pixel probability). This is the process where segmentation can be applied, although other methods to turn raster pixels to raster objects also exist and can successfully be applied in the feature extraction process. One such example is the “threshold and clump” operation in which a threshold is applied to the probability raster layer and the remaining pixels are clumped (clustered) into raster objects. An example of the application of the segmentation and threshold and clump operators is shown in Fig. 13.

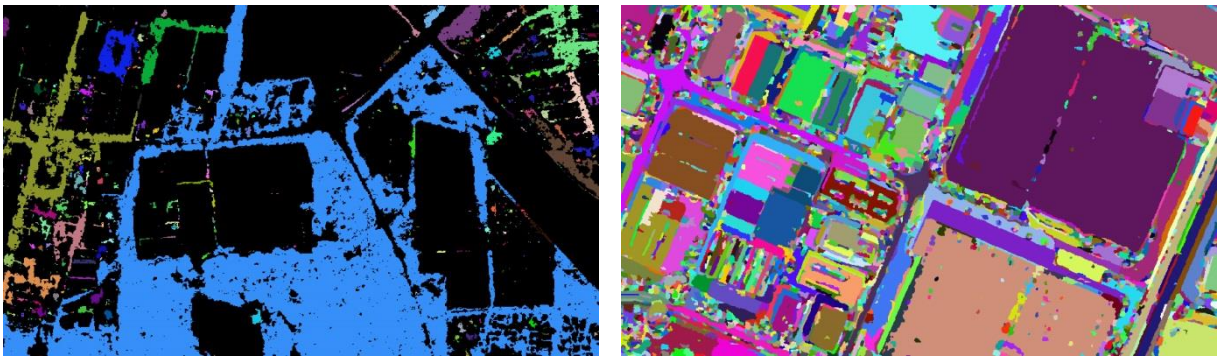


Fig. 13 Raster object layers (left: threshold & clump, right: segmentation).

In the next process various operators are available that alter the raster object data while remaining in the raster object domain. Examples of operators are probability filters, size filters, centerline conversion, mathematic morphological operators such as dilate and erode. After these operators, the conversion from raster objects to vector objects (lines or polygons) takes place. Entering the vector domain opens up many new possibilities to alter the extracted features so far, such as the filling of islands within polygons, generalization and the creation of skeletons from polygons. This process takes place in a vector object to vector object operator. Next training takes place again, using the initial set of training samples or a new set of samples or even template shapes. The resemblance to the training samples is attributed to the vector objects and can be used in further processing. Also new cue metrics can be quantified for the vector objects, which was not possible in the raster object domain such as orthogonality, orientation or the ratio between an object boundary and area. In the final stage a vector to vector operator is used again, but now the new set of cue metrics can be applied in a final cleanup operation of the extracted feature objects. Two examples of the extracted vector objects after the final cleanup operation are shown in Fig. 14.

Fig. 14 Vector objects (left: Vlissingen roads, right: Amersfoort buildings).

4.3 Applied feature extraction methodology

In order to provide an answer to whether the use of more spectral bands produces better feature extraction results, a 4 bands multispectral subset is made from the orthorectified 8 bands multispectral imagery. Spatially the images are equal and only differ in the amount of bands. The two sets of imagery have been subjected to the same feature extraction strategies (that can differ for each study area) in which all bands are taken into account in order to make comparisons between the results. Not all images are used in the feature extraction process, but the images that had the smallest off-nadir viewing angle were used for both study sites limiting the effect of the remaining relief displacement in the orthorectified images.

Since the focus is on the comparison of 8 bands and 4 bands imagery as input to OBIA, and not the creation of one extraction strategy that can be applied at different locations, the same targeted objects can have a different strategy for each study site. In appendix I these strategies and applied operators are shown. Example strategies are provided in the software documentation. However, these were not found fit for use for the two study areas. The used strategies were based on software documentation and through a thorough process of trial and error.

The features that were targeted for extraction were water, forests, roads and buildings. Within these features no distinctions were made between different types, for the sake of simplicity. However, the first results with the building extraction showed that the use of building subclasses improves the extraction results. These subclasses do not reflect logical building classes from a topographic point of view, but rather they are a pragmatic implementation to improve the extraction results. Visually the most common roof types were identified as greyish roofs, red roofs, and large rectangular industrial buildings. The greyish and red roofs were extracted using the same strategy, only different training samples were used. The large rectangular industrial buildings were extracted with a different strategy and different training samples. The large and rectangular nature of these buildings allow for the use of more cues for extraction. For instance the filtering of small objects or the orthogonality metric simplified the extraction of these buildings. The three different types of buildings were merged into one feature set in which the building distinction was disregarded. One of the most influential processes in the extraction of buildings was the use of elevation information. The DSM was used to determine delta heights between extracted objects and their near vicinity and this delta height value is used to filter unwanted objects given a certain threshold. The delta height implementation as offered within Imagine Objective causes memory failures and therefore the delta heights and subsequent threshold filtering were conducted using a workflow in ArcGIS, which is the

equivalent of the Imagine Objective implementation. A more detailed description on this alternative approach can be found in appendix IV.

In all strategies around a hundred training samples are collected for each object of interest, and a single feature probability operator was used to determine the probability values for each pixel. Afterwards the pixels are transformed into raster objects through segmentation or threshold and clumping. In the segmentation all pixels are present in the output as part of a raster object, whereas the threshold and clump operator applies a threshold on the input probability layer. The segmentation size was found to play a critical role in these two types of raster object creation. For instance, when searching for large rectangular industrial buildings, a large segment size proved to be of additional value.

Not all the extraction process steps as defined in the Erdas Imagine Objective workflow (section 4.2) are implemented in the extraction of all the features as can be seen in appendix I. In the last two process steps of the predefined workflow object cue metrics are computed and the final object cleanup takes place. However, by implementing these last two steps to create e.g. buildings that are more orthogonal than the originally extracted building feature, the results look representative given the occasional flaw in shape or direction, but are not necessarily better from a DSM to DTM filtering point of view. Because it is the goal to create better DTMs using spectral information and not to create the best looking object features, the roads have not been converted to line or ribbons (A ribbon shaped polygon feature) in the extraction process, since it will compromise the effect it has on the filtering of the DTM. In Appendix II a more detailed description is given on the effect of various process options on the extraction results.

4.4 Feature extraction accuracy assessment background

Much research on the accuracy assessment of land classification has been conducted, from which OBIA and automatic feature extraction benefit equally. A common approach for the accuracy assessment of feature extraction is to take (stratified) random samples of the generated data and compare it to the real earth or high quality reference imagery. The results from these comparisons are summarized in an error matrix (confusion matrix / contingency matrix) that shows the overall quality, and the type I and type II errors, also referred to as commission and omission, respectively (Kerle et al., 2004). Another method is the so-called Kappa analysis, which basically returns measures that indicate significant deviations between a determined error matrix and other error matrices that are created randomly or result from other classification methodologies (Smits et al., 1999). This provides answers to whether the right image or method has been applied in terms of quality.

4.5 Applied feature extraction accuracy assessment

The assessment of the feature extraction accuracy is done very similarly to the common error matrix approach. Instead of random (stratified) sampling, all data is compared to large scale high quality reference data that is available for both study sites. Three measures are determined. The amount of overlap between the extracted features and the features found in the reference data (overlap), the amount of features extracted, but not found in the reference data (type I error) and finally the

amount of features that should have been extracted according to the reference data, but were not (type II error).

The most accurate way to determine feature extraction accuracy is believed to be the extraction of targeted features by hand first and to compare those results with the automatically extracted features. Using this method you would extract only features optically and spectrally from the imagery in contrast to the objects present in the reference data that have partly been created by fieldwork. For example roads located under dense tree canopy cover would not be extracted by hand, although they will be present in the large scale topographic reference data. However, this method is labor intensive and the comparison of extracted features with reference features is sufficient for the level of accuracy assessment in this research context.

Two different reference datasets are used for the accuracy assessment of the feature extraction. The large scale topographic dataset TOP10NL, was used as a reference for the water, forest and road features. An example of the reference dataset is given in Fig. 15. Object classes and types were combined in order to better overlap the extracted feature objects, e.g. various types of roads and forest classes are combined into a road and forest class.

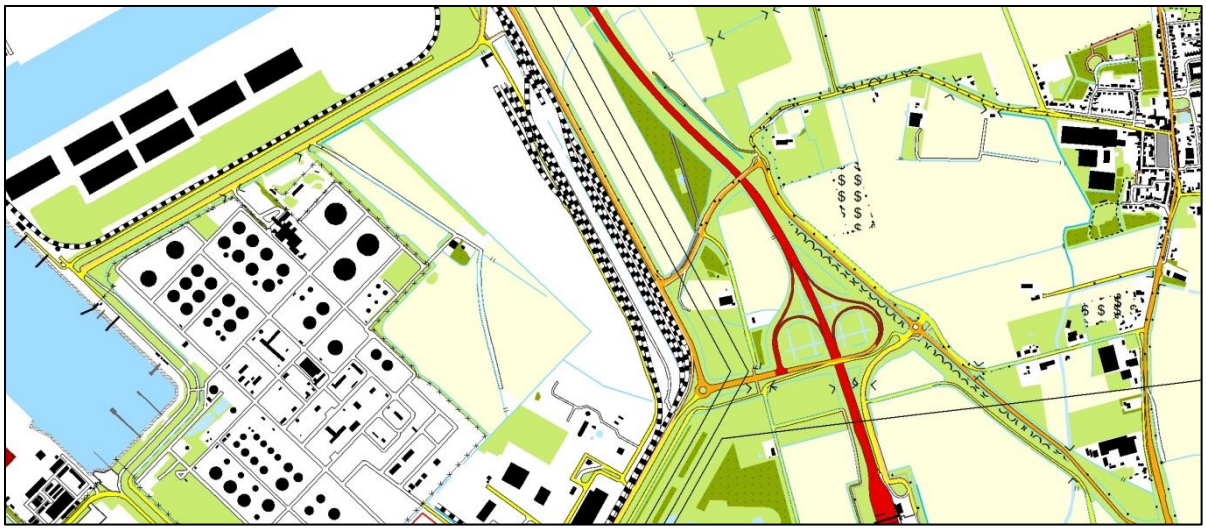


Fig. 15 Reference data (TOP10NL) for the accuracy assessment of feature extraction results.

For the assessment of buildings another reference dataset, the BAG is used. In Fig. 16 it is shown that the TOP10NL has a significantly higher degree of generalization and therefore the larger scale (1:1K) BAG reference dataset is used to assess the building extraction. From the BAG dataset objects representing small stone structures such as sheds are removed as well as objects that no longer existed according to its attributes, but are geometrically maintained in the dataset.

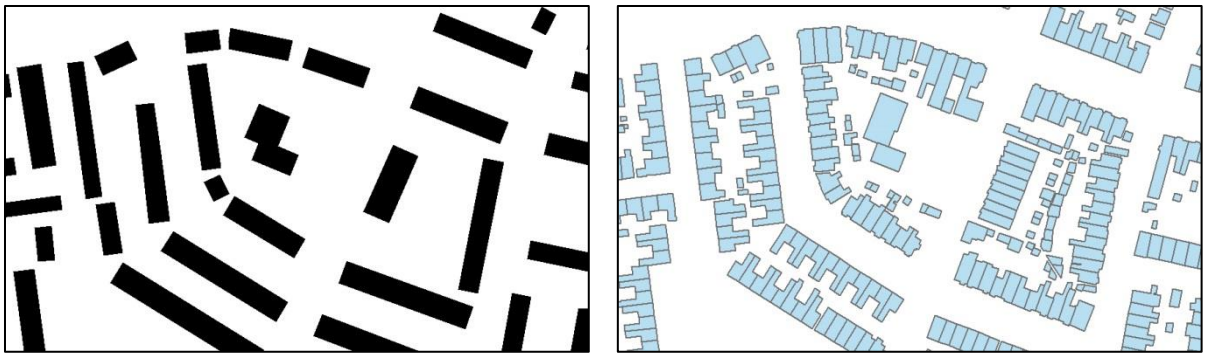


Fig. 16 Building extraction reference data (left: TOP10NL, right BAG).

In case the reference dataset has a linear horizontal deviation from the generated imagery and subsequently the extracted features as well, the reference dataset is repositioned in order to create better overlap with the extracted features. In the case of Vlissingen both reference datasets have been shifted horizontally around 7 meters for the purpose of feature extraction accuracy assessment.

The accuracy of the extraction is assessed for all the objects roads, forest, water and buildings based on 4 bands and 8 bands imagery. For the resulting 4 building layers corresponding layers that had a delta height value assigned to it were also included in the assessment to show its added value.

The best extraction strategy is determined through trial and error during the extraction process. A comparison between different extraction strategies using Kappa statistics is not done because the need to determine if DSM to DTM filtering can be improved with spectral information outweighs the need to determine the best strategy per object feature in this research.

5 DSM TO DTM FILTERING

In the last major production phase DSM to DTM filtering process is conducted. First of all this phase help answering the first research question, as to what suitable geometry based DSM to DTM filtering techniques are described in literature for photogrammetrically derived DSMs? Secondly, in this phase the combination of spectral- and geometric information in the DSM to DTM filtering process is put to the test in order to provide answers to the second research question, as to whether it is possible to combine the two different types of information in the DSM to DTM filtering process. Lastly a process of accuracy assessment is conducted in order to find the DSM to DTM filtering strategy that yields the best results, and thereby providing an answer to the last research question. DSM to DTM filtering is conducted using only geometry based information, spectral feature based information and a combination of both.

Digital Terrain Models (DTMs) are used in a wide variety of applications. Some examples are: volumetric difference computations to study the effects of erosion and deposition on a landscape over time, the creation of contour lines in large scale mapping productions, hydrographic analysis and urban building reconstruction. Various applications require different input DTMs e.g. for flood modeling, a DTM that does include houses but no vegetation could be desired. However, in this research the term DTM is used to describe terrain from which all above ground objects have been removed. In this research a DTM is required for simulation modeling. The representation of bare earth allows for manual or procedural modeling of above ground objects in later simulation modeling processes, whereas the use of a DSM in simulation modeling will result in unrealistic simulated environments (e.g. forests included as hills). Procedural modeling is the automated content creation (e.g. buildings) that uses a designed procedure, instead of manual production (Smelik et al., 2009). Procedural modeling often incorporates random variables e.g. a set of 3D building models is distributed randomly using 2D building footprints or point locations.

Jung et al. (2004) mentioned that there has been much attention to the automation of surface and terrain reconstruction since the 70's but because of limited success, especially in urban areas the attention faded away. More recently the rise of LIDAR and an increase in spectral and spatial resolution found in satellite imagery have given a new boost to the attention for the automation of terrains. However, the production of bare earth elevation models so far has proven to be a challenging process with no perfect solution. In this chapter various DSM to DTM filtering methods found in literature are described in more detail as well as the general problems associated with the DSM to DTM filtering process. Next the applied DSM to DTM filtering techniques in this research are described in detail and finally methodologies found in literature and applied methodologies for the DTM accuracy assessment are described. In the DSM to DTM filtering background section and the section on suitability of filtering techniques for photogrammetrically derived DSMs, an answer is provided for the first research question, whether there are suitable existing techniques for geometry based DSM to DTM filtering described in literature for photogrammetrically derived DSMs.

5.1 DSM to DTM filtering background

The amount of papers written over the last two decades on the subject of DSM to DTM filtering, and the amount of methodologies proposed clearly show that obtained results in DSM to DTM filtering leave room for improvement and no holistic approach to DSM to DTM filtering is found so far. The description of filtering techniques for urban areas are not uncommon, but filtering techniques applied on mixed landscapes covering relatively large areas and with locally large changes in elevation are less common. This makes the choice of the appropriate filter for military geo-specific terrain modeling a complicated one, given the need to perform well in various types of environments. Another notable difference is the larger focus on filtering techniques using LIDAR data instead of photogrammetrically derived surfaces. For the above reasons the focus has been on the identification of filtering methods that can be implemented using photogrammetrically derived DSMs, and secondly the suitability given the various landscapes targeted by military simulations.

At this point it is important to notice that it is not attempted to give a complete overview of all available methods for DSM to DTM filtering found in literature but merely to create a better understanding of existing filtering techniques. A larger overview of existing techniques and comparisons using the same study sites can be found in Sithole et al. (2004) and Pingel et al. (2013). The two existing types of DSM to DTM filtering found in literature are semiautomatic filtering and manual filtering (Jensen, 2007). Given the need for fast production cycles for military geo-specific terrain modeling, only semiautomatic methods are considered in this research.

The majority of literature found focusses on the bare earth extraction using surface models created by LIDAR techniques, and the DSM to DTM filtering through the use of geometric information only. LIDAR surface models are unaffected by weather conditions, more accurate, dense and better show discontinuities in the terrain in comparison to photogrammetrically derived DSMs (Baltsavias, 1999). The better representation of discontinuities in LIDAR surfaces is a direct result from the density in elevation points. Since lower elevation densities are obtained by photogrammetrically derived DSMs, more interpolation between points takes place that causes the smoothing of object edges. Another reason the edges of objects are better represented in LIDAR is that it is less susceptible to lighting conditions as is the case with photogrammetrically derived surfaces (Priestnall et al., 2000). The density of points in a surface derived through photogrammetric processes increases when more overlapping images are used in the image matching process. However, this results in a substantial increase in costs and computation time. The reason a high density of points is required is that the sudden changes in elevation are used in the classification of above ground objects (AGOs) when performing DSM to DTM filtering on geometry alone. According to Vosselman (2000), the decrease in point density results in a type II error increase, meaning less objects are classified as AGO and therefore erroneously remain in the surface model. It is interesting to note that although point clouds are typically generated by LIDAR and digital photogrammetry, the majority of researches use a gridded format. This is believed to be related to the large amount of algorithms that have been studied and that can be applied on these tessellated raster files. One such algorithm is mathematical morphology.

5.1.1 Mathematical morphology

Mathematical morphology is a DSM to DTM filtering technique recurrently found in literature over the last two decades that is otherwise known as mathematical transformation (Haala et al., 1998), (Sithole et al., 2004), (Joshi, 2010) and (Acqua et al., 2011). Mathematical morphology acts upon the relation between a surface raster and a horizontal structuring element (kernel or window). More specifically the mathematical morphology operator called opening is used in the process of DSM to DTM filtering. Mathematical opening is the process of erosion followed by dilation. Mathematical morphology is conceptualized more easily as a process using a binary raster containing only two values, namely bare earth and AGOs, although the whole process is equally suited for a grey scale surface raster. When the horizontal structuring element is completely contained within an AGO the output value for the central point of the structuring element is AGO in the output raster. The structuring element is never fully contained within an AGO around its edges, and therefore these edges are not output as AGO values, causing the erosion of AGOs. In the process of dilation the structuring element is used to determine whether the value is higher than zero when intersecting the structuring element with an AGO. This means that even when the structuring element is overlaid on the edge of an AGO the intersected value is higher than zero, because of the presence of AGO pixels. In this case the central point in the structuring element is used to output an AGO value, resulting in a new surface raster in which AGOs are expanded. When using mathematical opening with large structuring elements or in an iterative manner, it is possible to remove AGOs from the surface. From the above it becomes clear that the size and shape of the structuring element is essential in this process of mathematical morphological filtering.

Haala et al. (1998) used the process of mathematical opening to create a so-called initial DTM from LIDAR data. By subtracting the DTM from the DSM a normalized DSM (nDSM) is created. This contains the approximate above ground objects and their elevation above the terrain. The created nDSM is added as a spectral band to available multispectral imagery. This modified multispectral imagery is used in the extraction of both buildings and elevations and the results are used for DTM refinement. Acqua et al. (2011) used the process of mathematical opening and LIDAR data in a similar way except in this research a threshold was enforced on the nDSM in order to remove the buildings from the latter and to create a refined DTM. It is interesting to note that the operation of mathematical morphology can serve several different purposes, e.g. Joshi (2010) used the, through mathematical opening, derived nDSM for the improvement of the surface model instead of the terrain model, and was able to show the mutual benefit of high quality surfaces and feature extraction.

A problem found with the use of mathematical morphology is that although small – medium sized isolated buildings and trees might be relatively easily removed from the surface, this is not the case for very large buildings (e.g. greenhouses and industrial buildings) and large patches of forests. These would take very large structuring elements in order to eliminate them from the surface or would take too much iteration in order to remove these AGOs. As Ma (2005) mentions, large window sizes will have the negative side effect of small bare earth feature removal, while a small window will not succeed in the removal of large buildings, let alone large areas of forest. The above shows that the determination of parameters for filtering can be considered a bottleneck in the process of mathematical morphology and that the automatic determination of the optimal parameters such as window size and object filtering thresholds is desirable (Acqua et al., 2011). Other attempts are also

made in using various sizes of structuring elements instead of a single fixed structuring element (Kim et al., 2013).

A similar problem is found in the determination of a slope threshold parameter. The early morphological operations applied made use of the assumption that a terrain is relatively flat and consequently large local changes in elevation are considered as an AGO resulting in the increase of type I errors. The fact that small terrain details such as ridges or man-made hills might be mistaken for an above ground object makes mathematical morphology unsuitable if the modeling of these small features in the DTM is desired or the landscape is characterized with a large change in relief in relative close vicinity. Therefore modifications of the morphological operators were made in order to take the slope of the terrain into account. This is done by applying thresholds on the slope of a terrain. Vosselman (2000) developed and described the slope based filtering method in which both heights and distances were used in the discrimination between AGO's and terrain. Basically large difference in height over a short distance indicate the presence of an AGO and based on this assumption slope values can be subjected to a threshold. The AGO likelihood decreases as distance increases between two points. Although terrain features were better preserved when compared to regular mathematical morphology operations, the usage of this method is only optimal on landscapes with gentle slopes and therefore another modification is made by Sithole (2001) to make the filtering technique more versatile in regards to dealing with various levels of slope in a terrain. This method is referred to as the adaptive slope filter, and as the name implies the filter parameters are adapted to best fit the local slopes in the terrain. The adaptive character of the filter has been found to reduce the type I errors found in other methods, however an increase of type II errors was noticed (buildings were not completely removed) and the choice of parameters became increasingly complicated. In hindsight, dealing with slopes more effectively has not alleviated the difficulty in finding optimal parameters for the DSM to DTM filtering process. Terrains vary tremendously, creating difficulties in the determination of globally applicable threshold for slopes or structuring elements. Suggestions have been made to define more or less homogeneous regions (Vosselman, 2000) in order to adjust filters or the differentiation between forests and buildings for the same reason (Sithole, 2001). This could for instance be done by fusing spectral information and geometric information.

5.1.2 Planar surface segmentation

Parallel to the application of morphological operators over the last two decades is the development of DSM to DTM filtering techniques that implement segments or planar surfaces for the differentiation between terrain and above ground objects(Priestnall et al., 2000), (Van de Woestyne et al., 2004), (Ma, 2005), (Miliaresis et al., 2007) and (Kim et al., 2013). The term segment or planar surface is used interchangeably throughout this research. Planar surfaces are raster objects grouped together based on a certain characteristic. In contrast to mathematical morphology the use of planar surfaces does not suffer from the difficulty of determining optimal parameters, and is not influenced by the size or shape of an above ground object. Planar surfaces can be created using textures or height values. For instance if the height difference between a surface subjected to mean filtering and the original surface is larger than a certain threshold the output is AGO and otherwise bare earth. Another way is to segment the surface in segments with little variation in height. The AGOs can

subsequently be differentiated from the bare earth by determining the difference in height between the segment and its close surroundings. A large height indicates an AGO.

Baillard et al. (1998) segmented a photogrammetrically derived surface in altimetric regions and used an adjacency graph and Markovian labeling scheme to describe a segment as either bare earth or above ground. Spectral information is used to differentiate between buildings and vegetation. This relatively simple technique showed relative good results over urban areas. Priestnall et al. (2000) used both LIDAR and imagery in the DSM to DTM filtering process. First a mean filter is applied on the LIDAR generated surface in order to create a reference DTM. After subtraction the resulting nDSM and the imagery is used to segment the layer into bare earth and AGO using artificial neural networks. The AGOs are buffered and removed from the input DSM as well as all gradients in excess of 50^{o}. Van de Woestyne et al. (2004) argues that the height variable alone is not sufficient for the identification of bare earth and that in order to identify bare earth, connectivity plays a larger role than the difference in heights between AGO and bare earth. Segments build from a surface that are larger than the largest building and that are connected are a good indication of bare earth. Ma (2005) used the segments larger than the largest building to create an initial DTM. The initial DTM is refined by looking at the difference between the initial DTM and the original surface. If the difference between the initial DTM and the original surface is below a certain threshold the points from the original surface are reclassified as bare earth points and are added iteratively to the DTM until finally no more points are added to the DTM. Miliaresis et al. (2007) used the representative height values for buildings in his study as seed points for region growing and unsupervised classification for the distinction between AGO and bare earth. Again the determination of the right parameters was considered a tedious process that could not be automated and required user interaction.Kim et al. (2013) states that by the predominant focus on filtering urban areas the determination of bare earth on mixed urban and forested landscapes is neglected, and therefore developed a new method that combines planar surfaces derived from LIDAR, with spectral imagery. In this method a mean filter is applied on the surface and the resulting averaged surface is compared to the original surface. Whenever the difference was larger than a given threshold the output was considered AGO and otherwise bare earth. From this planar surface an initial DTM is constructed and a refinement step is performed similar to Ma (2005). In addition, the vegetation is identified from the spectral information and a specific strategy is implemented to insert the so-called last returns from the LIDAR beam. These last returns represent bare earth elevation values. It was shown that this method performed better in mixed landscapes than other methods.

One similarity that is found in both DSM to DTM filtering strategies is the interpolation of a DTM after the removal of AGOs. Although various interpolation techniques can be applied such as Kriging and inverse distance weighting (IDW), the most common method in this context is the interpolation through the application of triangulated irregular networks (TIN).

5.2 Photogrammetrically derived DSM to DTM filtering techniques

As mentioned earlier, the choice for the use of stereo satellite imagery is most of all based on arguments such as safety, restrictions and access to the targeted environments, and to a far lesser extent the potential quality of the sensors data. In order to represent the process of geo-specific terrain modeling as accurately as possible stereo satellite imagery is similarly used for the study sites.

LIDAR elevation data has different characteristics in comparison to photogrammetrically generated DSMs such as accuracy and density. The density of a surface is a critical element in the extraction of bare earth terrain from a surface. However, the point density is traditionally better in LIDAR derived surfaces in comparison to photogrammetrically derived surfaces. LIDAR data also has more clearly distinguishable building edges and higher accuracy. Jung et al. (2004) mentions that the classical surfaces derived from stereo imagery are unsuitable for DSM to DTM filtering, but states that the use of VHR imagery as well as the use of more overlap in the imagery from more perspective views will increase the reliability drastically.

In answer to the first research question it can be stated that if a photogrammetrically derived DSM has an elevation point density and accuracy that is similar to that of a LIDAR derived DSM, the geometry based DSM to DTM filtering techniques can be applied with equal success from a theoretic point of view. However, the actual success depends on the quality of the generated surface. Perhaps the disadvantage in geometric properties found in photogrammetrically derived surfaces can be compensated by the data fusion of the orthorectified imagery and the use of its spectral information in the filtering process. Some examples of the relative successful implementation of spectral imagery in the filtering process have been given for the mathematical morphology as well as the planar surface segmentation. However, examples where spectral imagery is used in conjunction with geometric elevation information often make use of LIDAR as a source of elevation and co-register the imagery horizontally to the surface (Kim et al., 2013).

A feature that is problematic in the photogrammetrically derived DSM to DTM filtering has always been the forested areas, especially on sloped terrain. No ground points are being recorded as is the case with laser altimetry where beams can (in the right conditions) penetrate the canopy, and return bare earth elevation values. Forest areas are usually too large to be filtered out completely when using structuring elements as is required in mathematical morphology. When using spectral information only to filter these areas out, all elevation information is lost as well. The problem is illustrated in Fig. 17. When these areas are extracted and need to be interpolated for later modeling a significant landmark (hill) can disappear from the terrain model and is considered an undesired result. Although forests are present in one of the study sites it is expected that the effect of the described problem is minor in the study site given the size of the forest and the relatively flat terrain. If large forest areas are present and located on hills, hopefully there are enough open spaces found in the forest to extract bare earth values so that the general shape of the hill remains in the final DTM. If this is not the case the choice would seem to either accept the cut off as shown in Fig. 17, or lower the elevation of the forests as presented in the surface with the average difference in height between the forest inner edge and a buffer outside the forests inner edge. This would cause correlation problems between the lowered elevation values and the neighboring bare earth elevation values that subsequently have to be overcome. Resolving this issue is clearly a trade-off between the representation of these

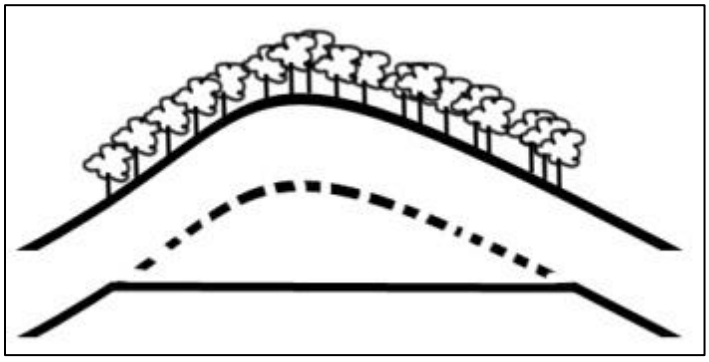


Fig. 17 Illustration of forest cover related DSM to DTM filtering.

prominent landscape features and the level of acceptance for assumptions being made in the process.

5.3 Applied DSM to DTM filtering methodology

For the process of DSM to DTM filtering a DSM is used that differs from the earlier mentioned DSMs. As mentioned before the output from the DSM extraction is an interpolated DSM and a point cloud. However, the interpolated DSM causes loss of information. Although this seems a paradox, the interpolation between ground earth points and building edges creates a more or less smoothed surface, making the discrimination between objects and bare earth significantly harder. Therefore, the generated point cloud is converted into a raster without the interpolation of the spaces between the measured elevation points. In Fig. 18 examples of the non-interpolated and the interpolated photogrammetrically derived DSM are shown for the Amersfoort study site. At the white spaces in the non-interpolated DSM example, no image matching took place during the photogrammetric processes. Note that the Amersfoort study site has a substantial larger density in matched points in comparison to the Vlissingen study site (see results).

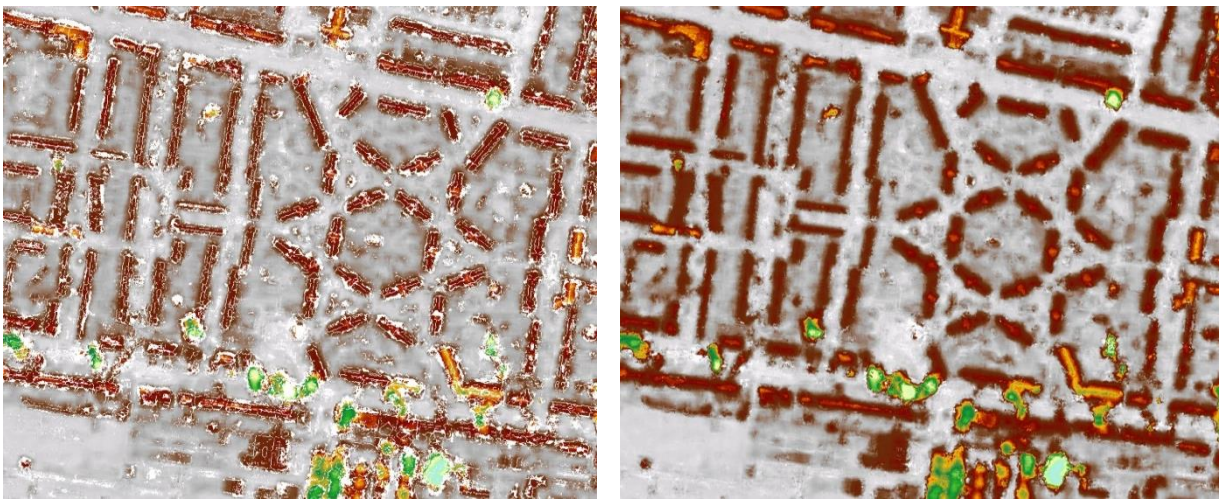


Fig. 18 Example of an interpolated and a non-interpolated photogrammetrically derived DSM (left: non-interpolated DSM Amersfoort, right: interpolated DSM Amersfoort).

In order to answer the last research question three different DSM to DTM filtering strategies are implemented and finally compared in terms of accuracy. These strategies are the DSM to DTM filtering solely based on spectral based feature input information, based on geometry based information and a combination of spectral based feature and geometry based information. The spectral based feature filtering, as well as the geometry based filtering strategy is shown schematically in Fig. 19.

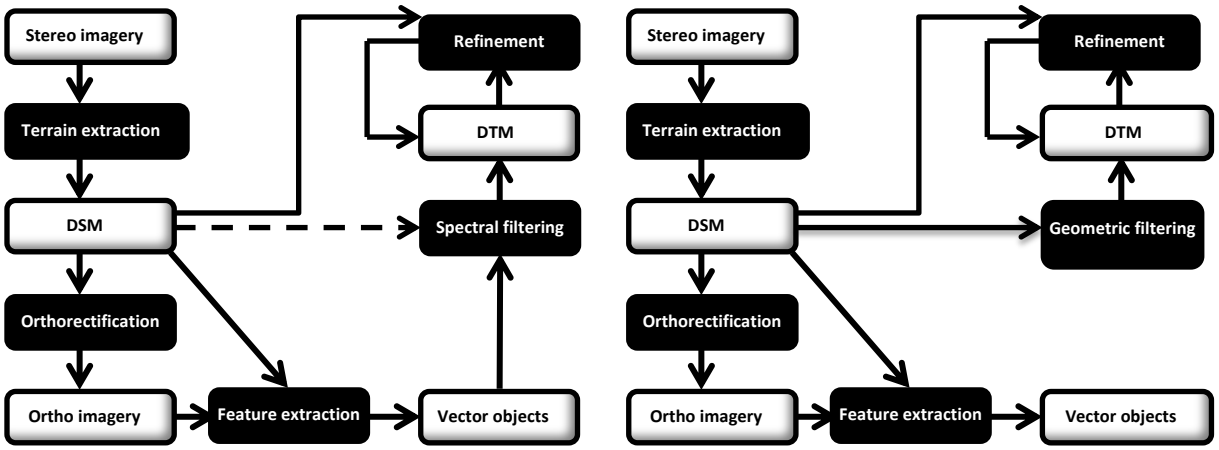


Fig. 19 DSM to DTM filtering strategies (left: spectral feature based DSM to DTM filtering, right: geometry based DSM to DTM filtering).

The left part of the schematic representations shows the photogrammetric processes and their output. The output DSM and orthorectified imagery are the same for all strategies since orthorectified imagery is a required layer for geo-specific terrain models. Features are also a requirement for geo-specific terrain models and therefore features are extracted using OBIA regardless of the input requirement for the specific DSM to DTM filtering strategy. The process of feature extraction is shown at the bottom of the schematic representations. It is during the process of DSM to DTM filtering that the three strategies start to differ from each other.

The spectral feature based DSM to DTM filtering strategy is the first implemented strategy. This strategy is not expected to yield the best results and the absence of examples of DSM to DTM filtering solely based on spectrally based features in literature helps confirm these assumptions. There are various reasons why spectral based feature DSM to DTM filtering is not expected to yield the best results. First of all, there are spikes in the photogrammetrically derived DSM which will not be removed when using spectral features only in the DSM to DTM filtering process. Secondly, there are substantial type I and type II errors found for certain extracted features such as roads (see results).This makes the use of these specific features with large errors in the DSM to DTM filtering process doubtful at least. Considering the above and the techniques used in literature it seems inevitable to implement geometric information in the DSM to DTM filtering process one way or the other. Therefore the removal of AGOs is done solely based on the extracted features, in order to create an initial DTM. The dashed line is used to indicate that the DSM is used in the filtering process as a source, although it has not been used in the identification of above ground objects. However, in subsequent processing geometric information is used to refine the initial DTM.

In the second strategy, the DSM to DTM filtering solely based on geometric information, the DSM itself is used to remove the above ground objects from the image, and therefore a solid line is used instead of a dashed line, indicating that besides a data source it is also the information used for AGO removal. Refinement takes place similarly to the first strategy.

The last strategy is the combined use of spectral based features and geometric information in the filtering process. This strategy is presented schematically in Fig. 20. Here the spectral based features and the geometric information are used in the production of an initial DTM. The methodology applied for each strategy is described in further detail in the subsections below.

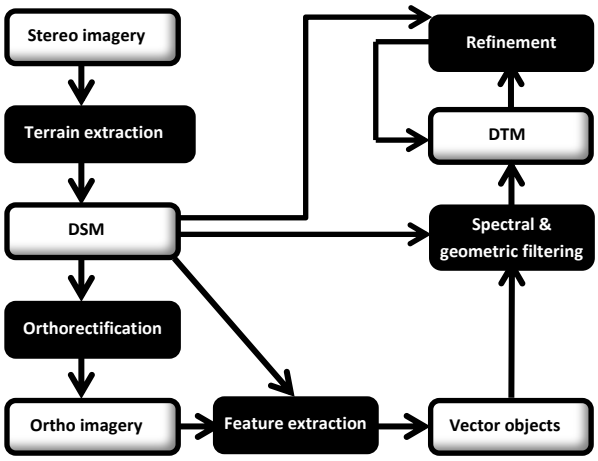


Fig. 20 Geometry and spectral feature based DSM to DTM filtering.

5.3.1 Applied spectral DSM to DTM filtering methodology

The goal of this strategy is to explore the potential of spectrally extracted features in the process of DSM to DTM filtering. As mentioned before, given the quality of the extracted features (see results), this strategy is not expected to yield the best results in DSM to DTM filtering. More specifically the type II errors found indicate that substantial numbers of AGOs will not be filtered out solely using spectrally based features. However, it is interesting to see if the errors present in the spectrally based features can be compensated by a refinement process. Another reason why it is interesting to research this strategy is that although the spectrally extracted features are prone to error, if the density of a photogrammetrically derived surface is very low, spectrally extracted features might be the best option you got to remove above ground objects. Decisions made in terms of applied methodology have been influenced by the outcome of preceding processing phases, and although some of these preceding process results are mentioned for the sake of clarification of the applied DSM to DTM filtering methodology, a complete overview is given in the results section.

For the spectral based feature DSM to DTM filtering the same vector layers are used that would otherwise be required for the geo-specific terrain modeling as well. These layers are forests, roads, buildings and water. Considering the high type I and type II errors found in the road layer, this layer is disregarded in the DSM to DTM filtering process. Water and Forest feature extraction was relatively successful considering the low type I and relatively small type II errors. Small holes are present in the forest feature layer that are most likely due to the fact that they are considered as shadow instead of forests in the OBIA process. In order to prevent that surface elevation values remain inside these holes, the latter is closed using the mathematical closing operation with a structuring element of 9 * 9 pixels. The building feature layer is buffered with 2 meters to make sure that building edges are removed from the surface more effectively. In Fig. 21, the process of spectral based feature DSM to DTM filtering is shown schematically.

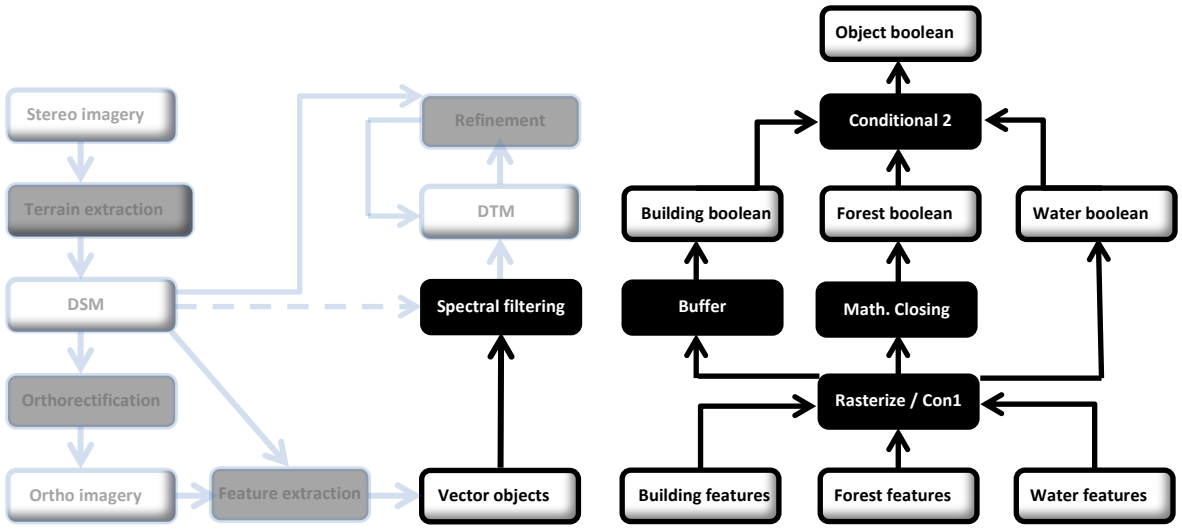


Fig. 21 Spectral based feature DSM to DTM filtering process (left: filtering process in generic workflow, right: specific DSM to DTM filtering workflow).

As shown in Fig. 21, the vector object layers are first rasterized, and transformed into a Boolean raster with the values AGO or bare earth. The building and forest layer are modified and finally all three object layers are combined into one Boolean raster with the values AGO or bare earth. Using a conditional function the bare earth values as defined by the Boolean raster are extracted from the DSM. In Fig. 22 the Boolean raster layers and the extracted elevation values are shown.

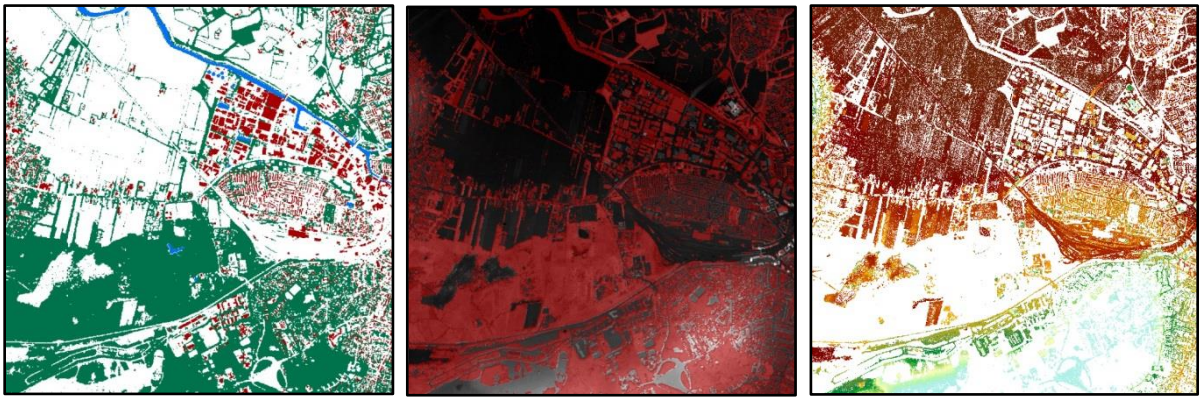


Fig. 22 Spectral based feature DSM to DTM filtering inputs. From left to right: used Boolean feature layers (forests are green, buildings red and water blue), combined Boolean showing AGO in red, extracted DTM elevation points from surface using a color ramp to visualize elevation (dark is relatively low and light blue is relatively high elevation).

As mentioned before, there are substantial levels of type II errors found in the extracted features. For both study sites the features extracted with 8 bands had the lowest amount of error, or the difference between 4 bands and the 8 bands were negligible. Therefore, the 8 bands feature layers were used for both study sites. The building layers produced using the delta height in the extraction process were used since they had substantially less type I errors and relatively smaller increases in type II errors when compared to the same layer where the delta heights were not implemented in the extraction process. The remaining type II errors indicate that a substantial amount of AGOs will remain in the initial DTM. Refinement of the initial DTM will be implemented as shown schematically in Fig. 23.

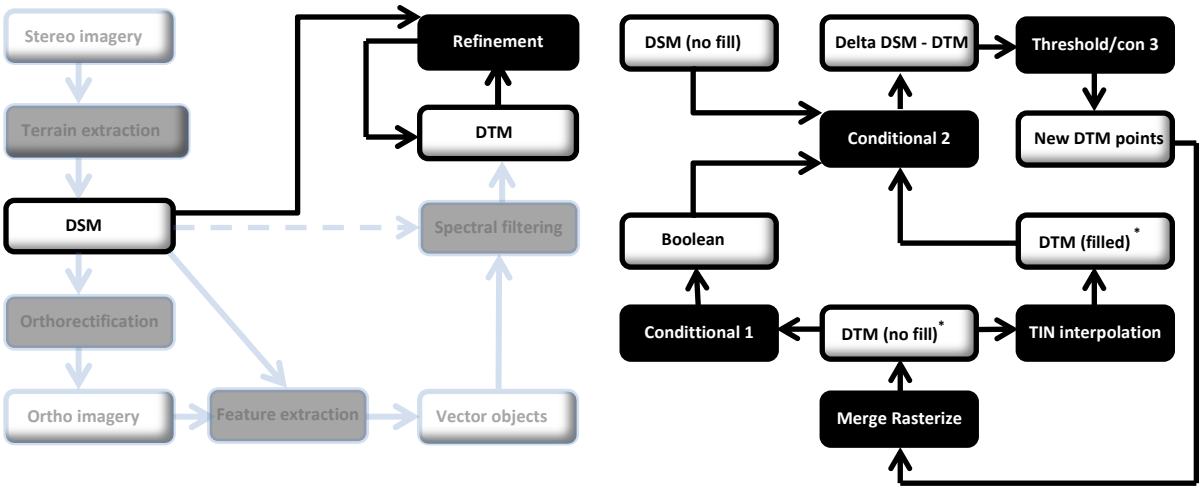


Fig. 23 Spectral based feature DSM to DTM filtering refinement process (the interpolated / unfilled DTM is the starting point and the interpolated / filled DTM is the final result).

The extracted elevation points as shown in Fig. 22 are referred to as the initial DTM. In the first refinement iteration the initial DTM is transformed into a Boolean raster layer using a conditional statement. The values are true in case of the presence of an elevation value and otherwise false. This Boolean raster layer ensures that only elevation values from the DSM are considered that are not already in the initial DTM. The initial DTM is also interpolated through the creation of a TIN structure. The Boolean raster, the filled DTM and the original unfilled DSM are used in a conditional function, in which the DTM values are subtracted from the DSM values if the Boolean raster value is false. If the resulting difference in the resulting value is smaller than 0.5 meters the corresponding value found in the DSM is added to the initial DTM. The merging of new elevation values to the DTM is done in the vector point domain and therefore rasterization takes place at the end of the process cycle. When the amount of new DTM points found becomes low, the iteration stops and the final DTM is interpolated. For all study sites 2 refinement iterations were applied.

5.3.2 Applied geometric DSM to DTM filtering methodology

The majority of literature found on the topic of DSM to DTM filtering applies to the geometry based DSM to DTM filtering. Although most of the research found focusses on the use of LIDAR derived surfaces, the filtering techniques should be equally applicable to photogrammetrically derived surfaces if the latter has a similar point density and accuracy. Since the goal is not to optimize existing geometry based DSM to DTM filtering techniques but study the potential benefit of using spectral information in the process, existing approaches are chosen in order to create a DTM based on geometric arguments only.

Initially the method as described by Kim et al. (2013) was chosen. This method applies planar surface segmentation. A large benefit associated with this type of DSM to DTM filtering is the fact that it doesn't suffer from the difficulties of parameter determination as is the case with mathematical morphological operators. In this method a mean filter is applied on the surface. Afterwards the mean surface is subtracted from the original surface. The resulting difference in elevation raster is converted into a Boolean raster in which a planar surface is output when the value is below a given threshold and otherwise the value is non-planar surface. The found planar surfaces can be

transformed into raster objects and therefore the delta height can be established from the average inner height and the average height found in a given vicinity of the object (buffer).

However, this methodology was aborted due to the geometric limitations caused by the elevation point density in the photogrammetrically derived surfaces. The photogrammetrically derived surfaces have a point density that is found too low for the creation of suitable planar surfaces. In Fig. 24 an example of the difference between the generated and reference surface is shown.

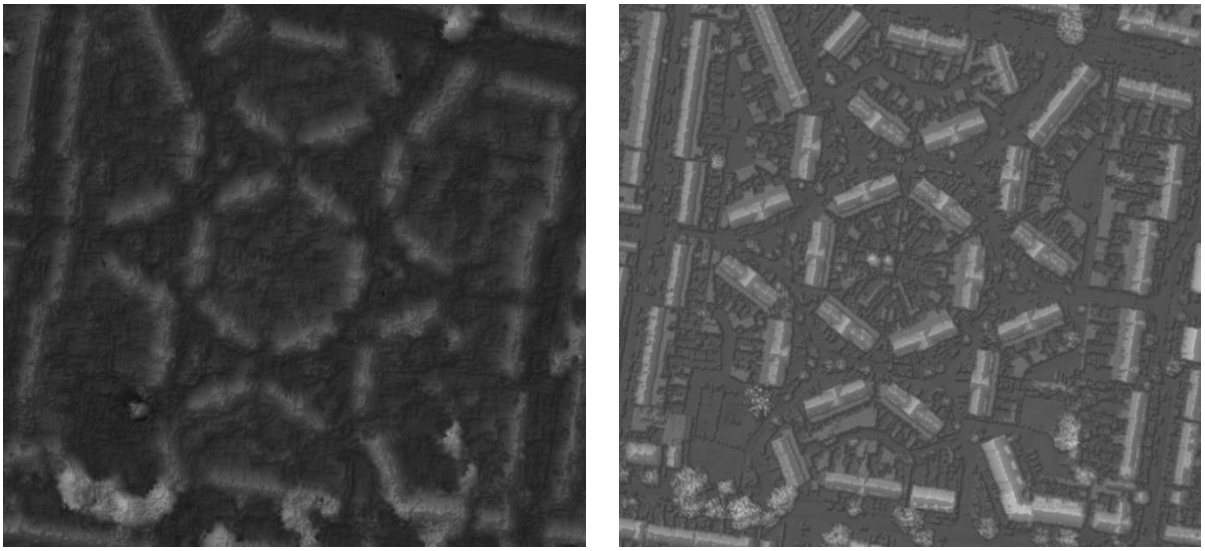


Fig. 24 Example of differences in geometric characteristics between photogrammetrically derived surface using triplet stereo images (left) and LIDAR derived reference surface (right).

Additionally, the vertical accuracy of the derived elevation points is lower in comparison to LIDAR derived surfaces resulting in more local elevation fluctuations. This in turn results in more elevation values output as non-planar although they should be classified as planar. In Fig. 25 an example is shown of the difference in planar surfaces between the photogrammetrically derived planar surface and the planar surface derived from the LIDAR derived AHN2 reference surface, and the transformation of the latter into raster objects. The same area is represented, and the same parameter settings have been applied on both surfaces, however there is no successful planar surface derivation from the photogrammetrically derived surface. Note that the point density obtained over the Amersfoort study site is substantially larger than the obtained point density at the Vlissingen study site (see results).

Fig. 25 Planar surfaces derived from photogrammetrically derived DSM (left), AHN2 LIDAR DSM (middle) and transformed AHN2 LIDAR DSM planar surface raster objects (right) in the Amersfoort study area.

Considering the limitations in the photogrammetrically derived surface in terms of planar surfaces, mathematical morphology was implemented instead in order to derive the DTMs from the DSMs.

Four iterations of mathematical morphological opening were applied with a 30 * 30 meter window in order to remove the above ground objects from the surface. The large window size is related to the presence of large industrial buildings in both study areas. After the second and fourth iteration a mean filter has been applied on the resulting initial DTM. The initial DTM is subtracted from the DSM in order to produce an nDSM. The nDSM has been subjected to a threshold in which every object present in the nDSM with a height larger than 2 meters is considered an above ground object and captured as such in a Boolean raster. The determined AGOs are buffered with 1 meter in order to reduce any edge effects. The resulting Boolean raster is used to extract the elevation values from the original unfilled DSM in case there is an indication of an AGO. The resulting elevation raster is converted into a TIN structure and interpolated in order to create the final DTM. The procedure is shown in Fig. 26.

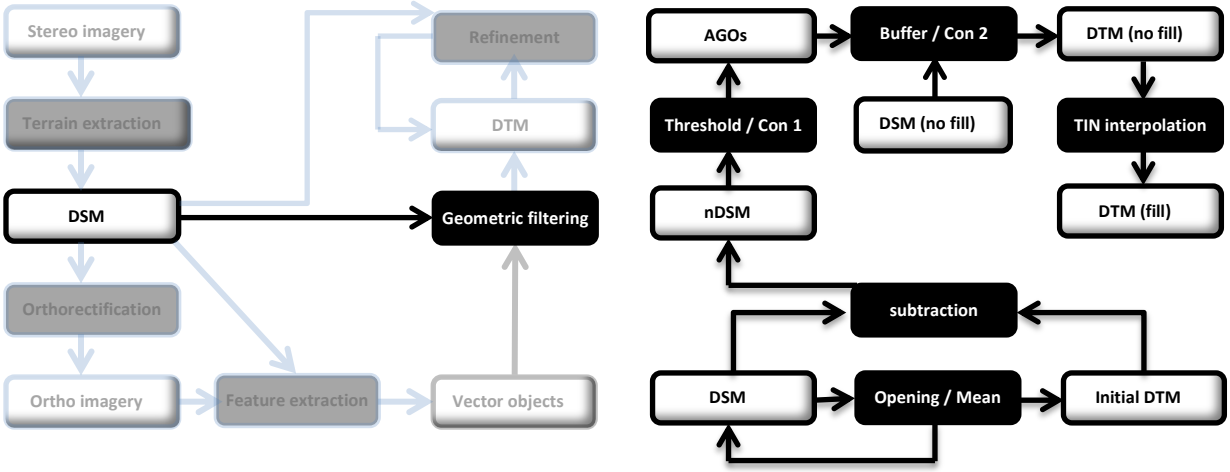


Fig. 26 Geometry based DSM to DTM filtering workflow.

The resulting DTM has been subjected to the same refinement implemented during the spectral based DSM to DTM filtering. 2 Iterations of refinement were applied.

5.3.3 Applied mixed spectral and geometric DSM to DTM filtering methodology

The combination of spectral feature based and geometric DSM to DTM filtering is not commonly found in literature and therefore the applied methodology is mostly based on suggestions made in literature with regard to dynamic and adaptive parameters determination and data source fusion. There are various ways of incorporating spectral information in the DSM to DTM filtering process. The extracted objects can be used to dynamically determine window sizes in mathematical morphology operators based on urban density, building sizes and the presence of forested areas and water bodies. If static window size are used, the window size parameter is determined based on the presence of the larger AGOs in the study site. The removal of these large objects would require large window sizes that would cause the loss of small terrain features. The extracted features representing a type of above ground object can also be removed from the DSM directly, assuming the remaining high elevated points are noise that can be removed by a final morphological opening operation or smoothing with substantially smaller windows sizes in comparison to the window size used during

the geometry based DSM to DTM filtering. The extracted features can also be used in the application of different DSM to DTM filtering strategies in one study area. For instance, the problematic forest land cover features in an area can be relatively easily identified and targeted specifically in a filtering strategy. When large patches of forests are covering the terrain and the removal of forest would cause the loss of a landmark such as a hill, the extracted features can be used in the approximation of the ground elevation located underneath the forest canopy.

Initially the methodology of dynamic parameter determination based on spectral features was chosen as the most suitable methodology. However, a substantial amount of type II errors present in the building feature extraction results indicated that this dynamic or adaptive parameter determination would result in undesirable results from the DSM to DTM filtering process over the areas where buildings were not identified during the feature extraction. For the same reason the direct removal of building objects from the surface would not allow for the use of smaller window sizes in subsequent mathematical morphology operations, since some large buildings would still be present in the surface, which would still require large window sizes in order to be removed. Because the results of both road and building feature extraction were found to be prone to errors, these features are not used in the spectral feature based and geometry based DSM to DTM filtering. In Fig. 27 an example is given of type II errors found after the building feature extraction.

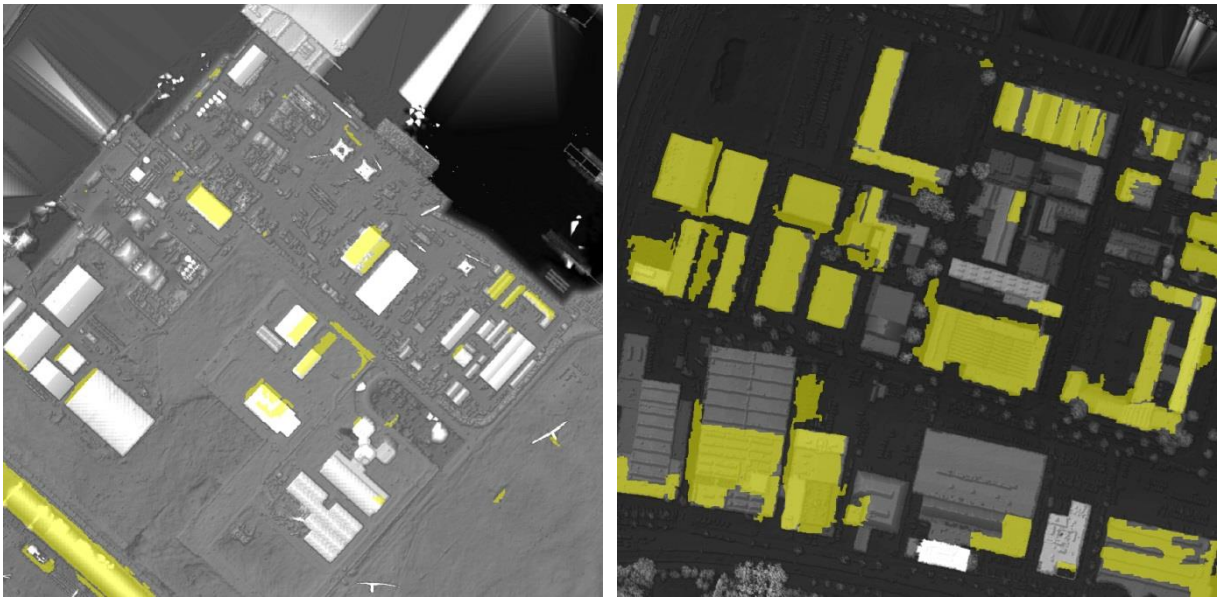


Fig. 27 Example of type II errors found after building feature extraction. The extracted buildings are symbolized in yellow and are superimposed over the hill-shaded reference surfaces (left: Vlissingen, right: Amersfoort).

In the applied method the forest and water features are used in order to perform the first iteration of surface filtering. The features were converted into raster Booleans and subjected to the same preprocessing as performed during the spectral feature based DSM to DTM filtering. After the removal of surface elevation points covered by the forest and water features the surface is interpolated. This interpolated surface is in turn used as input in the DSM to DTM filtering process as performed during the geometry based DSM to DTM filtering. The main advantage of this methodology is that the forests can be handled more effectively in this combined DSM to DTM filtering methodology in contrast to geometry based DSM to DTM filtering given the size of the forest features.

Because it was found that some bare earth features were removed in the geometric and combined DSM to DTM filtering process (see results) another approach to combined DSM to DTM filtering was investigated. This method involved the extraction of grass features which is not necessarily required for the purpose of simulation, but these features clearly cover bare earth parts of the surface. Successful extraction of grass features would lead to identification of ground elevation points that could otherwise mistakenly have been identified as an object in the mathematical morphological filtering because it is elevated above its surroundings, such as a dike. This method has also been abandoned because the type I errors in the grass feature extraction negatively affected the results. These type I errors were mostly deciduous tree canopies, which are spectrally similar to grass features. Another issue is that even though grass features were relatively well extracted over some areas where they could provide valuable elevation points, no elevation points were generated during the image matching process covering the same features. In Fig. 28 two examples are given of bare earth terrain features that are directly related with grass features.

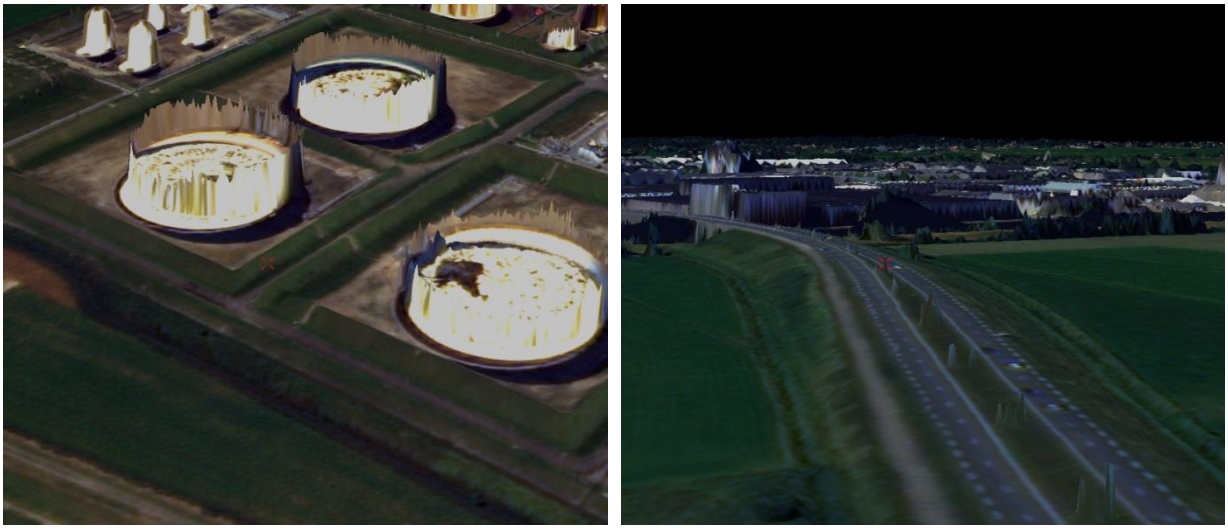


Fig. 28 Examples of typical bare earth terrain features that directly relate to grass land cover (left: Vlissingen, right: Amersfoort).

5.4 Applied DSM to DTM filtering accuracy assessment

In order to provide answers to the third research question, which DSM to DTM filtering strategy yields the best results in terms of accuracy, the DSM to DTM filtering assessment is performed similarly to the assessment of DSM and feature extraction accuracies. Two aspects of DSM to DTM filtering are investigated. Firstly, what is the performance in terms of AGO removal and secondly, to what extent are the bare earth elevation values of the points allocated to the DTM influenced by the applied DSM to DTM filtering strategy. The assessment of the effects of DSM to DTM filtering on the vertical accuracy complements the accuracy assessed in terms of type I and type II errors (AGO removal). While assessing the percentages of error types caused by the applied filtering strategies, its effect on the bare earth elevation values allocated to the resulting DTM are not considered. When assessing the vertical accuracy through the use of sample points the filtered AGOs as well as residual type II errors in the DTM are not taken into account.

The reference DTM for both study sites is subjected to preprocessing in order to adjust the noData value, projection and vertical datum. More information on the preprocessing of the reference elevation layers is found in appendix V. The reference DTM for the Vlissingen study site has been

subjected to an additional horizontal shift in the same way the reference feature layers were in the accuracy assessment of the feature extraction. This is because the generated data for Vlissingen are found to have been shifted horizontally in a linear way.

After the preprocessing of the reference DTMs, the discrepancy between the generated and reference DTMs is mapped. The discrepancy mapping differs from the one performed for the surfaces since negative differences between the two terrains are maintained. Negative values indicate a type I error in the DTM results, meaning a bare earth terrain feature (e.g. a dike or ramp) has mistakenly been considered an above ground object and as such is removed from the DTM. A large positive value indicates the presence of a type II error, meaning an object should have been considered an above ground object but was not and therefore remains in the DTM.

The quantitative amount of type I and type II errors is determined similarly to the applied accuracy assessment during the feature extraction. In this method an nDSM is generated from the reference DSM and DTM and is assumed to include all above ground objects. This reference nDSM is compared to the nDSMs that are generated from the DSM to DTM filtering strategy results and the photogrammetrically derived surfaces. This provides information on the percentages of type I and type II errors found in the results of the DSM to DTM filtering strategies. An advantage of this method is that when comparing the generated nDSM and reference nDSM the results are unaffected by the possible vertical inaccuracies found in the photogrammetrically derived DSMs. It should be noted at this point that the visual inspection of the discrepancy mapping is believed to provide a better understanding of the undertaken DSM to DTM filtering processes and their resulting quality. Small amounts of type I errors would seem optimal merely looking at the quantified accuracy. However, if the small amount of type I errors are related to the removal of typical bare earth landmark features, even a small amount of type I errors could be considered undesirable.

Beside looking at the success made in the removal of above ground objects, the effect of DSM to DTM filtering on the bare earth elevation values included in the resulting DTM are assessed. Ideally the process of DTM generation would cause minor changes in the elevation values found at bare earth locations as represented by the sample points. For the assessment of these vertical accuracies, the same point set is used that has been created for the purpose of vertical and horizontal accuracy assessment of the generated surfaces. All points are located on bare earth locations. Although statistically the amount of sample points is considered low, this method is considered a quick and pragmatic means to an end that provides some information on the levels of discrepancy between the photogrammetrically derived bare earth elevation values and the elevation values at corresponding locations in the generated DTM.

6 RESULTS

In this section the results of all three main production phases are described in detail. These results help provide the answers to the second, third and fourth research question. Whether spectral based features can be combined with geometric information in the DSM to DTM filtering process (second research question) is determined through literature study and the successful application of combined DSM to DTM filtering, regardless of the level of obtained accuracy. An answer to the third and fourth research question, which DSM to DTM filtering strategy yields the best results and whether there is a significant difference in automated feature extraction when using four or eight spectral bands, respectively, is directly based on the results as described in the section below.

6.1 Photogrammetry results

The results of the photogrammetry phase are described in terms of internal output quality as provided by the software, visual inspections including the production of slopes in order to identify systematic errors, discrepancy mapping, horizontal accuracy assessment and vertical accuracy assessment. The orthorectified reference imagery and AHN-2 LIDAR derived reference surface are used for the external accuracy assessments. The photogrammetric results are subjected to such an extensive accuracy assessment considering that the photogrammetrically derived information will be the source of input for further processing and will therefore provide the necessary insights that will be required for further production phases.

6.1.1 Internal DSM quality output

In Table 5 and Table 6, the software's internal quality output is summarized for the overall output surface quality and the image matching output quality. Due to computation reasons the Amersfoort study site was divided into four processing blocks and the Amersfoort values in Table 5 and Table 6 represent the average values from these four processing blocks, apart from the Amersfoort output- and matched points which represent the sum of the four processing blocks.

In Table 5 it is shown that there is a substantial difference in the RMSE and LE90 values between the two datasets that are related to the data input into the photogrammetric processes. The Vlissingen data did not have GCPs measured in the field, and therefore aerial orthorectified imagery was used to create horizontal control points to be used in the triangulation of the images. For the Amersfoort data GCPs containing Z-values were available and used in the triangulation process. Using these measured Z-values in the triangulation process forces other points to be stretched towards these values. As mentioned no measured Z-values were available for the Vlissingen area during the triangulation and therefore there was no big external influence on the Z-values during the image matching process. The absence of external influences and systematic errors resulted in low RMSE en LE90 values. It is believed that the larger errors found in the Amersfoort dataset are caused by the influence of the GCPs, since they cause the elevation points found during image matching to deviate more from its relative surface towards the measured elevation values. A more detailed inspection as described in more detail in section 6.1.3 showed that the measured GCP values are on average 0.5

meters higher than the corresponding elevations in the reference surfaces. This could be caused by the offset between the GPS device and the surface or because of the applied vertical datum transformation of the reference surface. Assuming the GCPs are indeed located 0.5 meters above the surface on average explains the higher RMSE and LE90 values and would cause the generated surface to have a general offset with regard to the reference surface. It should be noted that although this is undesirable since it will complicate the accuracy assessments, it will not be of influence on the applied DSM to DTM filtering techniques nor its accuracy assessment.

Table 5 Internal photogrammetry output surface quality provided by the used Leica Photogrammetry Suite (Erdas LPS).

Area (25 KM^2)	RMSE	LE90	Output points	Matched points	Output points (%)	Water (%)	Point / M^2
Vlissingen	1.338	2.201	11567881	122230021	9.464	8.08	0.46
Amersfoort	17.461	29.328	185380349	348753398	53.155	1.64	7.42

Another clear difference between the Vlissingen data and the Amersfoort data is the amount and percentage of output points. These are significantly lower for the Vlissingen area. The Vlissingen area includes more (6.4% more) water than the Amersfoort area, lowering the number of matched points. Vlissingen also only had two overlapping images that were used in the image matching process, whereas Amersfoort had three. However, the Amersfoort image matching process was divided into 4 processing blocks, all covering the entire study area, but not taking all images into account. One processing block matches all three images and the remaining three processing blocks, used as internal matching quality checks, only included 2 overlapping images. The lowest amount of output points from two images for Amersfoort is still at least twice as much as the Vlissingen dataset, taking the extra water into account.

In Table 6, the software's quality output in terms of matching output quality is given. Again there is a clear difference between the Vlissingen en Amersfoort data. It is shown that the larger amount of points from the Amersfoort dataset is also of relatively better matching output quality.

Table 6 Internal photogrammetry matching output quality provided by the used Erdas LPS.

Matching output quality	Excellent (%)	Good (%)	Fair (%)	Isolated (%)	Interpolated (%)
Vlissingen	35.280	48.862	15.858	0	0
Amersfoort	76.461	20.474	3.401	0	0

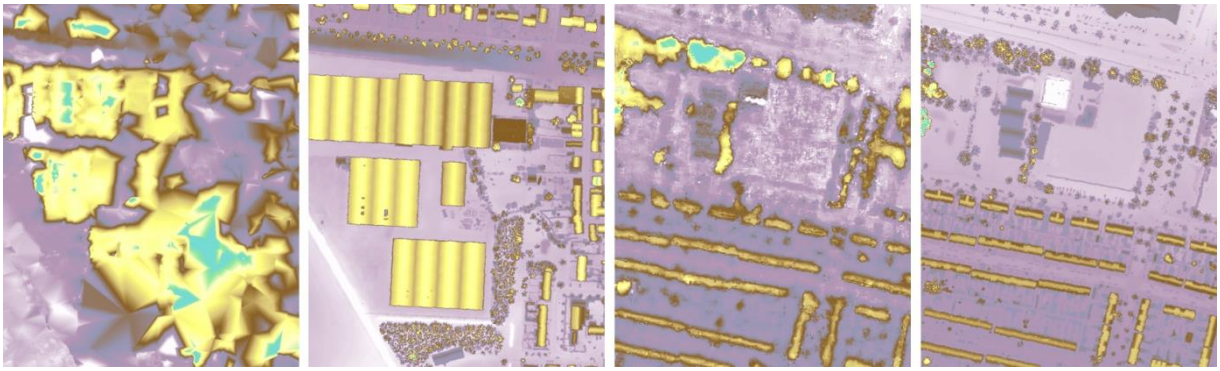


Fig. 29 Comparison of elevation point densities generated over the study sites and the LIDAR derived reference surfaces (left to right: Vlissingen generated surface, Vlissingen reference surface, Amersfoort generated surface, Amersfoort reference surface.

In Fig. 29 an example of the generated results is shown in which the clear difference between the two generated surfaces in terms of point density becomes apparent and for the sake of clarity these are also visually compared to the densities found in the LIDAR derived reference surfaces. It becomes clear that the density of elevation points found in the Vlissingen study site compromises the ability to capture building geometries.

The internal output quality indication of low RMSE and LE90 values of Vlissingen is believed to be the result of the absence of external influences in the triangulation process such as measured GCPs and systematic errors. Therefore, the absolute accuracy of the surface is still susceptible to the inaccuracies of the WorldView-2 sensor. This is an important reason why internal quality output alone is insufficient to determine the quality of output data and the availability of external reference information is so valuable. The high amount of RMSE and LE90 are believed to be caused by the GCPs and the subsequently applied accuracy assessment techniques will help determine the source and nature of the errors found in the Amersfoort internal output quality report. The internal output surface quality indicates that the point density obtained at the Amersfoort study site is superior with regards to the point densities obtained at the Vlissingen study site.

6.1.2 Visual DSM inspection

A visual inspection cannot provide answers in terms of absolute quality, but visual inspections are really helpful in identifying erroneous data at an early stage. As explained in the methodology section, visual checks are performed on the DSM data, by creating slopes and by looking at the point clouds (LAS files) in a triangulated state in 3D. Two examples of the generated data are shown in Fig. 30. None of the visual checks indicated that errors in the correlation between images have been made during the photogrammetric processes. Correlation errors are the result of the misalignment of one of the images during the photogrammetric processes and would create very “spiky” surfaces. Observing relative smooth surfaces such as shown in Fig. 30 indicate no errors were made.

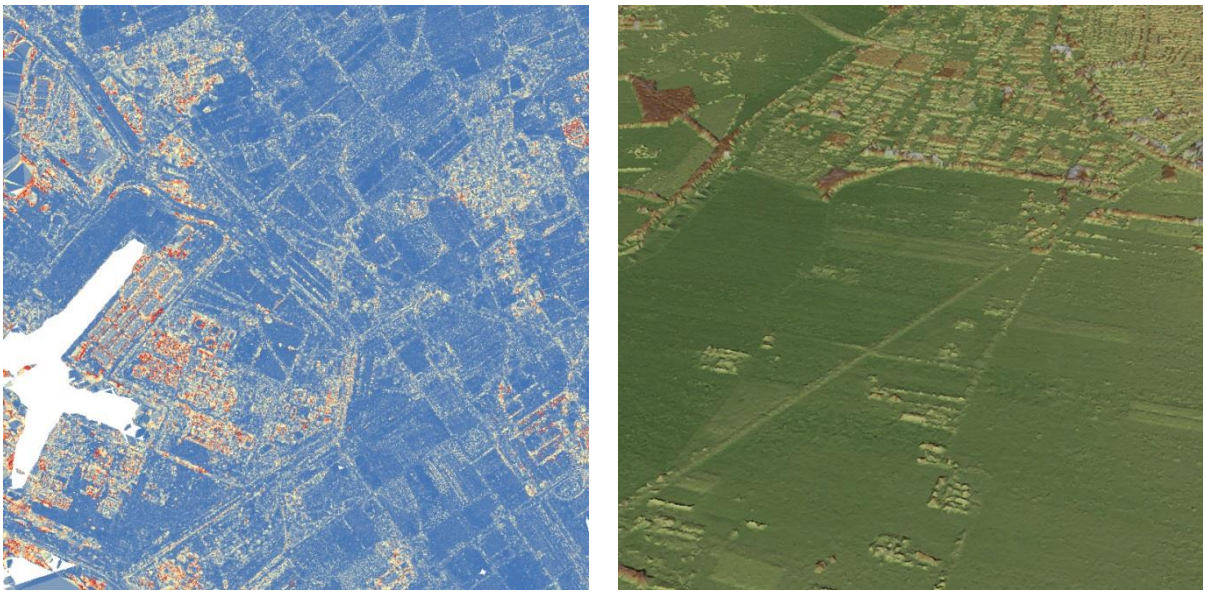


Fig. 30 Visual slope inspection over Vlissingen study site (left) and visual accuracy assessment over Amersfoort study site using TIN structures (right).

The visual inspection of the orthorectified imagery also did not show any clear anomalies resulting from the orthorectification process, which could be caused by errors in the surface model. A substantial difference in relief displacement between the Vlissingen study area and the Amersfoort study area was found. This has two reasons. First of all, the images that have been selected for orthorectification are those with the smallest off-nadir viewing angle. For Vlissingen the image used had an off-nadir viewing angle of 26.9 ^{o}, whereas Amersfoort had an off-nadir viewing angle of 4.1 ^{o}. The relatively larger off-nadir viewing angle is the first source of relief distortion. Secondly, only a small portion of the images has been used for DSM extraction and orthorectification. These subsets have been created looking at its representative nature for the whole area. For Vlissingen this resulted in a subset located further off the image center (Fig. 5). From the image centers there is an outward radial increase in relief displacement, also causing a larger relief displacement for Vlissingen in contrast to the more centrally located Amersfoort study area.

6.1.3 Surface discrepancy mapping

The results from the discrepancy mapping are visualized in Fig. 31. The largest discrepancies are located over forests areas and buildings. All major discrepancies have been further investigated and have all been related to the difference in acquisition time between the generated DSM and reference elevation data or the presence of forests. One exception is the discrepancy in the two surfaces as a result of the removal of military terrains and objects from the reference surface. This is indicated by the large dark red area in the southern part of the Amersfoort study site. Overall all large discrepancies had logical explanations and did not indicate production errors.

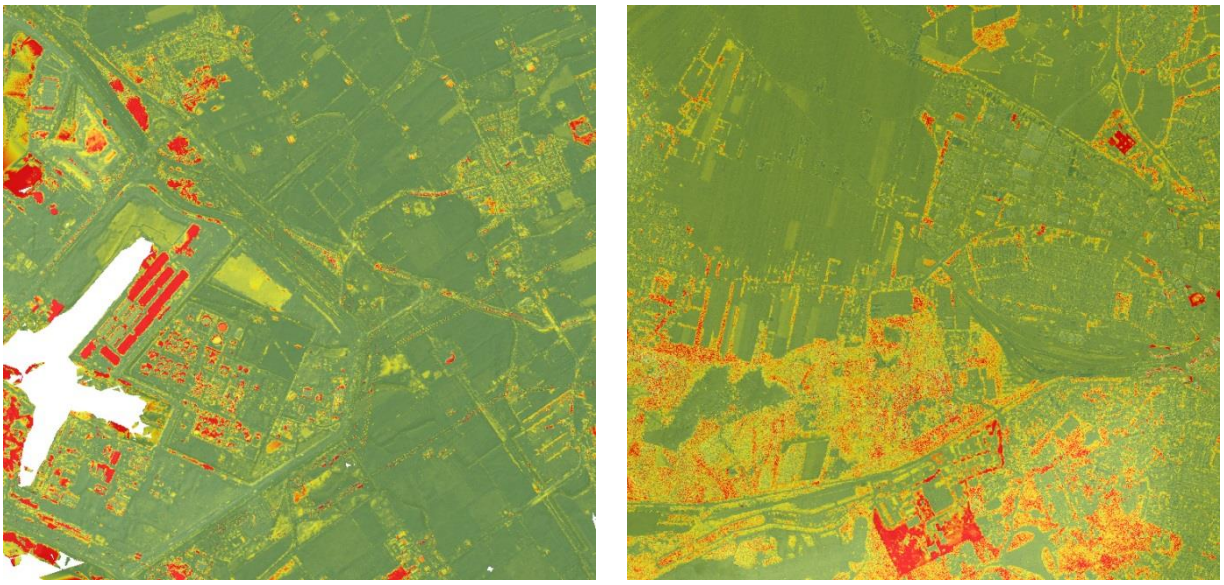


Fig. 31 Surface discrepancy mapping results indicating none to minor discrepancies in green and large discrepancies in red (left: Vlissingen, right: Amersfoort).

By looking at the elevation discrepancy distribution plotted as a histogram (Fig. 32), it becomes evident from the negative discrepancy distribution that on average the generated surfaces have higher elevation values in comparison to the reference surfaces. The larger difference found in the Amersfoort dataset seems to contradict with the use of GCPs. However, after investigation on the 19 GCPs it became clear that all GCPs have on average elevation values 0.5 meter above the reference surface. One GCP was present in the 5 * 5 km Amersfoort study site, which was located 1.7 meters

higher than the elevation found in the reference surface. This is believed to be the reason that there is a larger difference between the generated and reference surface in the Amersfoort study site. The reason for the relatively large difference in elevation values found between the GCPs and reference surface could be related to the offset between GPS device and surface that might not been taken into account or the loss of information during the vertical datum transformation. However, the true reason remains unknown. From visual inspection of the discrepancy mapping these differences found in the Amersfoort study site do not seem to cause any non-linear error.

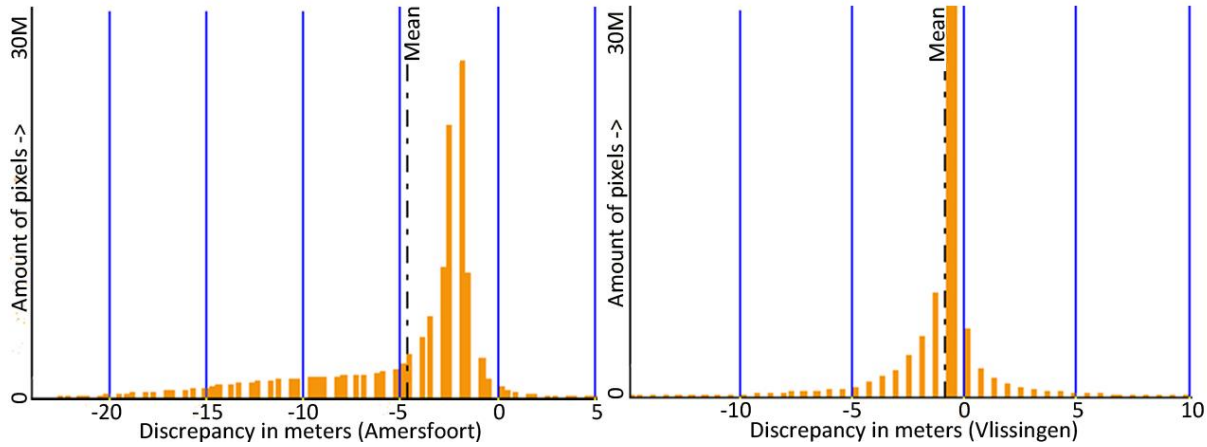


Fig. 32 Distribution of discrepancy values plotted as histograms (right: Amersfoort, left: Vlissingen). The negative distribution of discrepancies indicates that the majority of generated elevation points are located at higher elevations in comparison to the reference surface.

Furthermore Amersfoort has an average discrepancy of 4.5 meter, however when not taking discrepancies over forested areas into account the average discrepancy is lowered to 2.5 meters. The Vlissingen study site has an average discrepancy of 1.6 meters. The larger discrepancy over the Amersfoort study site is partly caused by the difference between the measured elevation values used as GCPs during the photogrammetric processes and the reference elevation values. The 1.6 meter average discrepancy at the Vlissingen study site corresponds well with the stated WorldView-2 sensor accuracy which has a maximum linear error with 90% confidence (LE90) of 4.5 meters (Bresnahan, 2011).

6.1.4 Horizontal DSM accuracy assessment

In order to determine the vertical output quality of the DSM it is first necessary to look at the horizontal output quality since it shows if the generated DSM correlates to the reference DSM. In order to determine the horizontal accuracy, the orthorectified satellite imagery was used. 10 points were created for each study area, distributed to cover the entire study area. After that similarly 10 points were created in the reference datasets, the orthorectified aerial imagery. These points were assigned to the same visual locations in the different images, and by looking at the difference in horizontal position we again see a clear difference in the two datasets. In Table 7, it is shown that Vlissingen has a larger horizontal error, which is practically linear in nature given the narrow range in angles of dislocation, whereas Amersfoort has a much smaller horizontal error which is oriented to the east side, but is not linear. An interesting detail is that the obtained horizontal accuracies are similar to the mentioned potential horizontal accuracies found in WorldView-2 with or without the use of GCPs. The horizontal errors found in the Vlissingen study site are an indication that the use of

derived horizontal GCPs from aerial orthorectified imagery did not provide an increase in horizontal accuracy. The horizontal errors found in the Amersfoort study site are an indication that the large internal errors found at the Amersfoort study site, as described in the internal DSM quality output section (6.1.1), are not related to the horizontal coordinates of the GCPs, but are solely caused by the Z coordinates of the GCPs.

Table 7 Horizontal DSM accuracy assessment statistics describing the angles (degrees) and distances (meters) of dislocation between generated and reference data.

Area	Mean distance (m)	Distance range (m)	Mean angle (°)	Angle range (°)	CL 95% (m)
Vlissingen	6,69	5.53 – 7.69	-87.13	-94.32 – -77.22	6.3 – 7.1
Amersfoort	1.26	0.65 – 2.01	-9.07	-43.07 – 91.82	1.1 – 1.5

The smaller horizontal error in the Amersfoort imagery is the result of the availability of more accurate GCPs in terms of X and Y coordinates. The horizontal error will be between 1.1 and 1.5 meter with a 95% confidence level. Since the generated imagery and surface of Amersfoort is relatively accurate and correlates well with the reference data the vertical DSM accuracy of the generated DSM can be determined without any preprocessing. The Vlissingen data, however, had a horizontal error that is between 6.3 and 7.1 meter at a 95% confidence level, and is therefore less accurate. Since that Vlissingen data has such a notable linear horizontal shift, all the reference data required for all further accuracy assessment processes (AHN2, TOP10NL & BAG) is shifted linearly in horizontal position to minimize errors caused by the correlation of generated data and reference data. Note that the horizontal shifting of the data is only done for a more reliable accuracy assessment of the Vlissingen dataset, and that obtaining a better horizontal accuracy for the Vlissingen dataset is not a goal in itself, since we accept this horizontal error for remote areas without GCPs.

6.1.5 Vertical DSM accuracy assessment

As mentioned earlier the linearly distributed horizontal error found in the generated surface and generated orthorectified satellite imagery requires the horizontal shift of the reference data in order to reduce correlation errors during the accuracy assessment. After relocating the reference surface, the differences in elevation between the generated and reference surface are investigated statistically using the same sets of sample points used in the assessment of the horizontal accuracy. The results are shown in Table 8. A negative value in the table indicates that at one of the sample points the reference surface is elevated higher than the generated surface.

Table 8 Vertical DSM accuracy assessment describing the differences between generated and reference surfaces.

Area	Mean difference (m)	Difference range (m)	CL95% (m)
Vlissingen	0.75	-0.12 – 2.79	0.01 – 1.52
Amersfoort	1.84	1.25 – 3.00	1.52 – 2.15

Two points have been removed in the Vlissingen point set. These points were mistakenly located in such close vicinity to objects that they produce distortion in the vertical accuracy assessment and were removed. After the removal of the two outliers it is shown that the deviation in elevation is between 0 and 1.5 meters with a 95% confidence level for the Vlissingen study site and between 1.52 and 2.15 meters with a 95% confidence level for the Amersfoort study site. This deviation is always

positive, meaning the generated surfaces have higher elevation values than the reference surface. Again it is clearly evident that the Amersfoort study site has a substantially larger deviation in elevation values with regard to the reference surface. However, taking into consideration the 0.5 meter that the measured GCP are elevated above the reference surface on average and the 1.7 meters that the specific GCP found in the Amersfoort study site is elevated above the reference surface, the larger difference in elevation found in the Amersfoort study area can be explained.

6.2 Automatic feature extraction accuracy assessment

In this section the quality of the extracted feature objects is determined using high quality reference data available for the study areas. A visual impression of the feature extraction results is shown as maps covering the study sites in appendix VI and appendix VII.

6.2.1 Type I and Type II errors

As mentioned in the methodology section, the quality of feature extraction can be described using type I, and type II errors. Type I errors represent falsely extracted features, and type II errors represent features that should have been extracted, but are not. The percentages of area that overlap between the extracted feature objects and their representative counterparts in the reference data is determined along with the two types of errors. In order to correlate the reference data to the extracted features the reference data has been subjected to a horizontal shift at the Vlissingen study site, and all reference datasets have undergone the same preprocessing steps as described in the methodology. The statistical results of the feature extraction accuracy assessment are shown in Table 9 and Table 10.

Table 9 Automatic feature extraction accuracy assessment results. Vlissingen is abbreviated as V. The abbreviations EF and Ref. stand for the extracted features and reference features, respectively. The symbol Δ is used to indicate the extracted building feature layers where the delta height extraction strategy has been implemented.

Extracted Features (EF)	EF area (M sq.)	Ref. dataset area (M sq.)	EF & ref. area overlap (M sq.)	EF & ref. area overlap (%)	Type I E. (area M sq.)	Type I E. (area %)	Type II E. (area M sq.)	Type II E. (area %)
V. forest 8b	2985552	1433795	1209194	**84**	1776358	**59**	229195	**16**
V. forest 4b	3037587	1433795	1143149	**80**	1894438	**62**	304794	**21**
V. water 8b	1481678	2020428	1462669	**72**	19008	**1**	557696	**28**
V. water 4b	1492249	2020428	1473477	**73**	18772	**1**	546952	**27**
V. roads 8b	1632932	1487169	587978	**40**	1045811	**64**	899190	**60**
V. roads 4b	1392160	1487169	487137	**33**	906072	**65**	1000032	**67**
V. building 4b	986872	1589976	653316	**41**	662756	**67**	936660	**59**
V. building 4bΔ	309059	1589976	507758	**32**	92422	**30**	1082218	**68**
V. building 8b	1285561	1589976	817767	**51**	828212	**64**	772209	**49**
V. building 8bΔ	429109	1589976	555930	**35**	154524	**36**	1034046	**65**

Table 10 Automatic feature extraction accuracy assessment results. Amersfoort is abbreviated as A. The abbreviations EF and Ref. stand for the extracted features and reference features, respectively. The symbol Δ is used to indicate the extracted building feature layers where the delta height extraction strategy has been implemented.

Extracted Features (EF)	EF area (M sq.)	Ref. dataset area (M sq.)	EF & ref. area overlap (M sq.)	EF & ref. area overlap (%)	Type I E. (area M sq.)	Type I E. (area %)	Type II E. (area M sq.)	Type II E. (area %)
A. forest 8b	8491843	5149358	4793334	**93**	0	**0**	356024	**7**
A. forest 4b	8256247	5149358	4655833	**90**	0	**0**	493525	**10**
A. water 8b	278754	414166	274777	**66**	3976	**1**	139389	**34**
A. water 4b	288095	414166	278749	**67**	9345	**3**	135417	**33**
A. roads 8b	2379497	2555069	760342	**30**	1620112	**68**	1794728	**70**
A. roads 4b	1840129	2555069	610046	**24**	1230655	**67**	1945024	**76**
A. building 4b	1119105	2736128	964685	**35**	356150	**32**	1771443	**65**
A. building 4bΔ	775485	2736128	844307	**31**	105818	**14**	1891820	**69**
A. building 8b	1609419	2736128	1261350	**46**	599406	**37**	1474778	**54**
A. building 8bΔ	1038253	2736128	1066153	**39**	168828	**16**	1669974	**61**

From these tables it becomes clear that in terms of the percentage in area overlapping and the percentages in type I errors and type II errors, forest and water features are extracted well using either 4 or 8 multispectral bands. In fact water seems to be extracted better using 4 bands imagery, although the difference is too small to be called significant. The extracted forest features of Vlissingen show a high percentage of Type I error, when compared to the reference data. However, when visually inspecting the results it becomes clear that the features are not falsely extracted, but the forest features are not classified as forests or trees in the reference data. This could be caused by semantic differences in data capturing, cartographic reasoning in the reference data, or seasonal differences at time of data acquisition. In Fig. 33, the first image shows the orthorectified satellite imagery for the Vlissingen area, the second image shows the extracted forest features, which seems realistic when looking at the original imagery. However, these forest areas are not present in the third image, which shows the forest features present in the reference dataset. This difference causes a high Type I error for forest features extracted in the Vlissingen area.

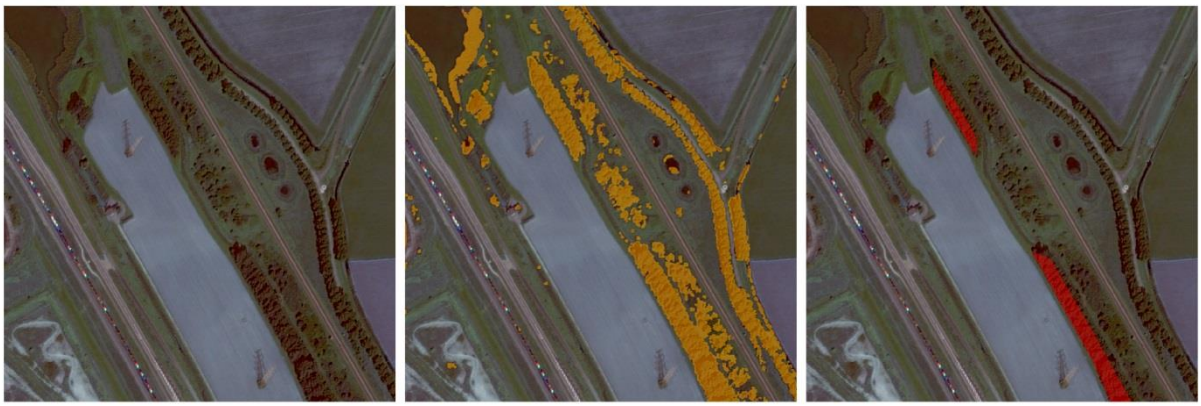


Fig. 33 Example of type I errors found in forest feature extraction over Vlissingen study site (from left to right: orthorectified satellite imagery, extracted forest features and reference forest features).

The relative success in the extraction of water is also related to the relatively large areas of water in both study areas. When looking at the results of the water extraction and the reference data in more detail, it becomes clear that large water bodies have a high percentage in overlap and low

percentages of type I- and type II errors, but very small (narrow) water features are not extracted. An example of this is shown in Fig. 34.

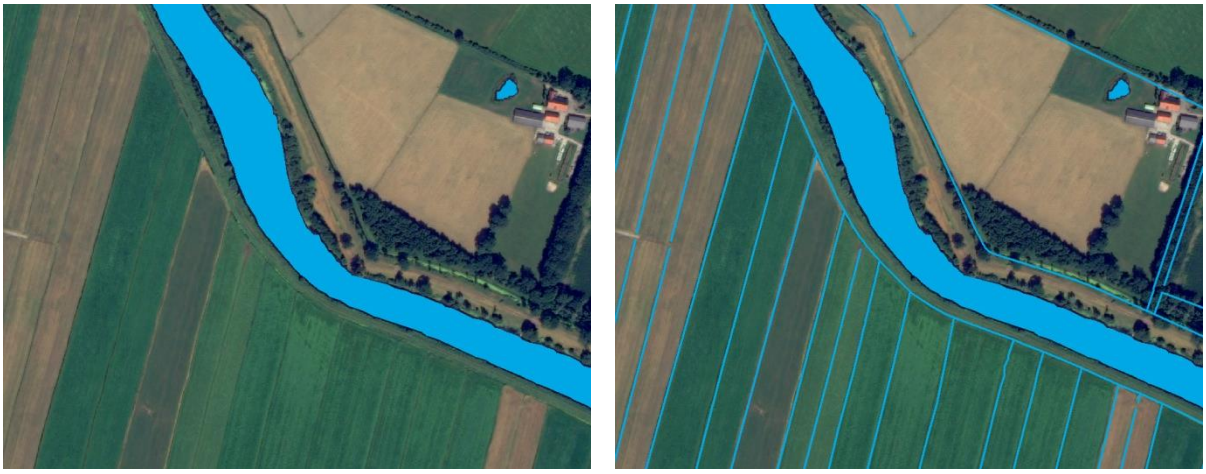


Fig. 34 Example of water extraction over Amersfoort study site showing narrow water features are not extracted automatically (left: extracted water features, right: reference water features).

In the left image the extracted water feature is shown. The right image shows the available reference features. The narrow water features, available in the reference data are line objects and are therefore not taken into account during the QA. Throughout both study areas the extraction of these narrow strips is minimum, which is understandable given its narrow nature. Note that the narrow feature can hardly be identified by the eye, and that knowledge of the studied area or large scale reference data is required to identify these errors.

There are no large differences found in the feature extraction output quality of buildings and roads. Both feature types show limited success. In both study sites the roads and building features overlap less than 50 percent with the reference data, while having high amounts of type I and type II errors (>50 percent). An example of the extracted buildings is shown in Fig. 35. These building extraction examples show the scale of inclusion and omission of building features. An interesting detail is that the extracted buildings in the Amersfoort study area have 50 percent less type I errors when compared to the Vlissingen study area, although the Amersfoort area is built up to a greater extent.

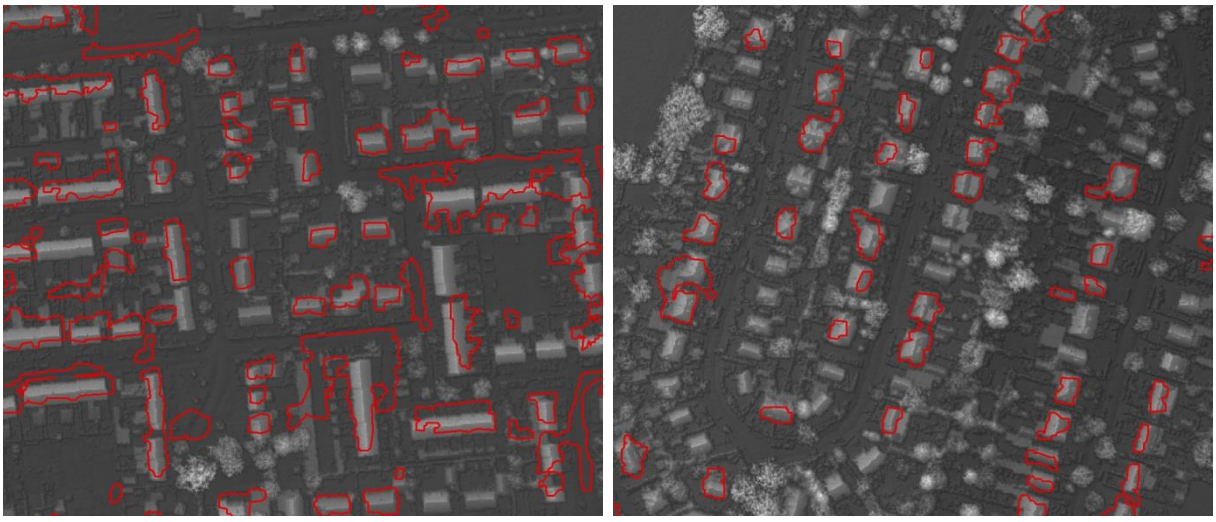


Fig. 35 Examples of building feature extraction results (superimposed over the reference hill-shaded surface) over suburban areas located in the study sites (left: Vlissingen, right: Amersfoort).

The obtained building and road feature extraction results have been found to be related to the complexity of the landscape as captured by the satellite images, in combination with the available feature extraction techniques. Spectral differentiation between buildings and roads is difficult since they can be constructed from similar materials such as concrete. A large variety of rooftops makes the capturing of building objects extremely difficult using spectral information alone, since they are constructed from different materials and therefore do not inhabit the same spectral niches. The buildings have been divided into three different classes in order to extract these specific buildings types more effectively as described in the methodology section. More information on the experiences on conflicting feature extraction strategies, errors in feature extraction and the complexity of the targeted feature objects is shown in appendices II & III.

6.2.2 Difference in feature extraction quality between 4- and 8 bands imagery

The differences in the output quality of feature extraction are determined for the 4 and 8 bands multispectral imagery that has been used in the feature extraction process. As mentioned in the methodology section, the feature extraction strategies for every object remained the same, while only changing the underlying imagery in order to determine if there are significant differences in output quality. In Table 11, the differences in output quality are described in terms of percentages overlap between the extracted features and the reference features and the percentages in type I and type II errors for the 4 and 8 bands imagery. In the table it is shown that there is 5 percent more overlap between the extracted features and the reference data when applying the extraction strategies on 8 bands multispectral imagery, 0.5 percent more type I errors and the expected reduction in type II errors of 5 percent, that relates to the 5 percent increase in overlap.

Table 11 Automatic feature extraction result differences between 4- and 8 bands imagery (referred to in the table as 4b and 8b, respectively). The abbreviations EF and Ref. stand for the extracted features and reference features, respectively. The symbol Δ is used to indicate the extracted building feature layers where the delta height extraction strategy has been implemented.

	EF & ref overlap (%)			Type I error (%)			Type II error (%)		
	4b	8b	overlap increase	4b	8b	Type 1 increase	4b	8b	Type II increase
Vlissingen buildings	41	51	**10**	67	64	**-3**	59	49	**-10**
Vlissingen Buildings Δ	32	35	**3**	30	36	**6**	68	65	**-3**
Vlissingen water	73	72	**-1**	1	1	**0**	27	28	**1**
Vlissingen forests	80	84	**4**	62	59	**-3**	21	16	**-5**
Vlissingen roads	33	40	**7**	65	64	**-1**	67	60	**-7**
Amersfoort buildings	35	46	**11**	32	37	**5**	65	54	**-11**
Amersfoort buildings Δ	31	39	**8**	14	16	**2**	69	61	**-8**
Amersfoort water	67	66	**-1**	3	1	**-2**	33	34	**1**
Amersfoort forests	90	93	**3**	0	0	**0**	10	7	**-3**
Amersfoort roads	24	30	**6**	67	68	**1**	76	70	**-6**
Averages	51	56	**5**	34	35	**0.5**	50	44	**-5.1**

An interesting result is that the biggest improvements in output quality when using 8 bands imagery are visible in the features that had limited overall extraction success, namely roads and buildings. The overlap between extracted features and reference features increase on average 6.5 percent for roads and 10.5 percent for buildings when using 8 bands imagery in the feature extraction process.

6.2.3 Delta height implementation

During the feature extraction process an operator is available that helps determine the delta height of an extracted object and its surroundings as predefined by a buffer distance around the object. This value can potentially resolve Type I errors in the automatic feature extraction of buildings, considering that a low delta height might indicate that the extracted feature is not a part of the targeted building features. A negative delta height for a feature might indicate that the extracted feature is not a building but part of a road for instance considering that the surrounding of the feature has a higher elevation value. It could also mean that a low building is surrounded by larger buildings and therefore the negative delta heights should be treated with care in the feature extraction process.

The delta height operator, however, did not work in the available feature extraction software package for larger areas at once, and therefore an alternative approach has been designed and implemented in order to create delta heights and to remove features that fall outside a given delta height threshold. This alternative approach is described in further detail in appendix IV.

In Fig. 36, the two study areas are shown. In green the extracted features are shown after all features with a negative delta height, and all features with a delta height lower than 1 meter are removed from the set of extracted features. The green features are superimposed over the black colored features that represent the same set of extracted features without the use of the delta height threshold. A lot of the black features visible in Fig. 36 are indeed no building features, but agricultural fields, water areas and roads. Therefore, a visual inspection of the datasets indicates that using the delta height, helps removing large erroneous features from the building features, causing the reduction in type I errors.

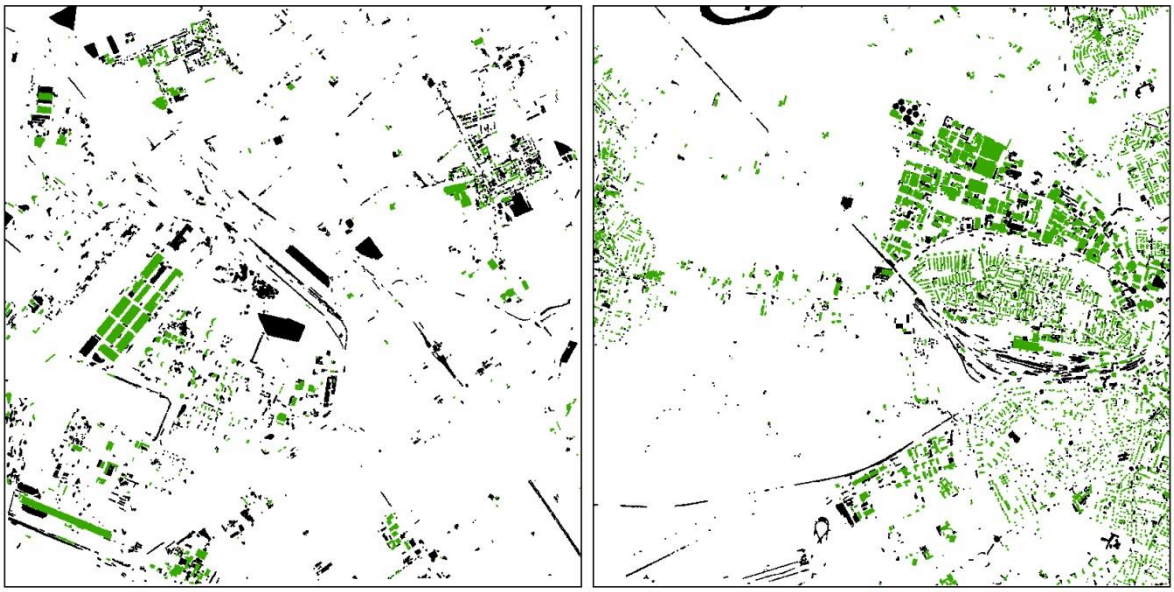


Fig. 36 Building extraction results, in which the delta height operator is used during the extraction process. The buildings extracted without the use of the delta height operator are shown in black, and the buildings extracted with the delta height operator in green (left: Vlissingen, right: Amersfoort).

In Table 9 and Table 10 it is shown that there is a significant effect on the type I errors when applying the delta height threshold in the extraction process. While the percentage in overlap between the

extracted features and the reference data is reduced roughly by 10 percent on average, and the type II errors increase with the related average of roughly 10 percent, the type I errors are reduced by 25 percent, which further confirms the assumptions made from the visual inspection.

6.3 DSM to DTM filtering accuracy assessment

The generated DTM accuracy assessment is performed visually and statistically. Type I and type II errors are identified visually from the discrepancy mapping results and determined statistically by comparing the reference nDSMs and the nDSMs generated from the different DSM to DTM filtering strategy results. Finally, the same points used for the vertical accuracy assessment of the generated surfaces are used to assess the effect of DSM to DTM filtering on the vertical accuracy. It must be noted that statistical analysis is not enough to judge the results in terms of quality. Visual inspection of the generated DTMs is crucial considering that the omission of typical terrain features can outweigh the overall statistical accuracy measures.

6.3.1 Terrain discrepancy mapping

Discrepancies are mapped for all three DSM to DTM filtering strategies and both study sites. In contrast to the discrepancy mapping conducted during the surface accuracy assessment the discrepancy map is not adjusted to include only positive values. Keeping the negative as well as the positive values gives clear indications of where type I and type II errors occurred. A color ramp is used to visualize the types of errors. Negative values are colored dark green indicating type I errors, gradually turning light green to yellow indicating minor discrepancies with the elevation values found in the reference DTM, and finally turning dark red where there are large positive differences indicating the presence of type II errors. In Fig. 37 the discrepancies between the spectral feature based filtered DTM and the reference DTM are shown.

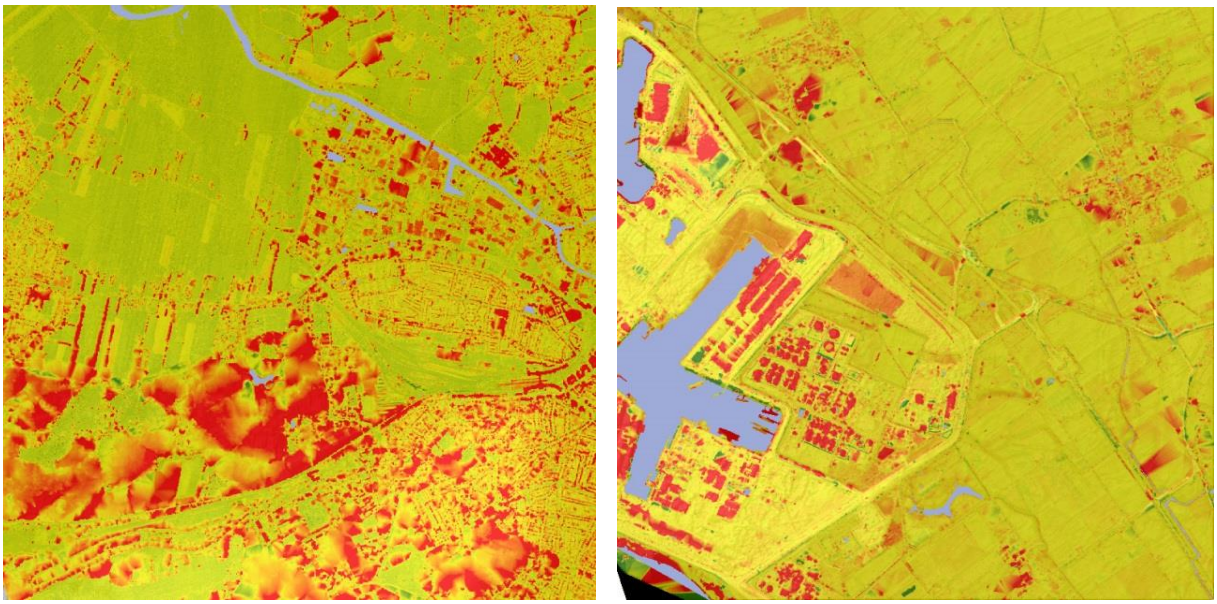


Fig. 37 Spectral feature based DSM to DTM filtering discrepancy mapping, indicating deviations from the reference DTM in dark green (type I error) and dark red (type II errors) at both study sites (left: Amersfoort, right: Vlissingen).

Visual inspection of the discrepancy mapping results clearly show that the amount of type II errors is the largest when using spectral feature based DSM to DTM filtering. These type II errors are related to the unsuccessful removal of forest and building features using only the automatically extracted features in the DSM to DTM filtering process. Only a small number of type I errors is present in the results.

The implemented geometry based DSM to DTM filtering discrepancy mapping results indicate that although there is a large reduction in type II errors, the amount of type I errors increases substantially. The geometry based DSM to DTM filtering discrepancy mapping results are shown in Fig. 38. Notice that in the Amersfoort area a large amount of type II errors still remain which is related to the presence of forests in the study area. The forest cover at Amersfoort is too large to be removed by the mathematical morphological opening operator. The implementation of the mathematical morphology operator is also the cause of the increase in type I errors, since it is unable to differentiate between an above ground object and an elevated object that is part of the terrain, such as dikes. Close up inspection of the type I errors show that the majority is associated with man-made infrastructure that is elevated above the rest of the terrain, such as a dike or a ramp. At the Amersfoort study site the majority of the remaining type I errors are found on locations where forest is removed in the filtering process. The LIDAR generated reference DTM includes bare earth elevation points measured underneath canopy cover, whereas the elevation values found underneath canopy cover in the photogrammetrically derived DTM are approximations of elevation created through the process of interpolation.

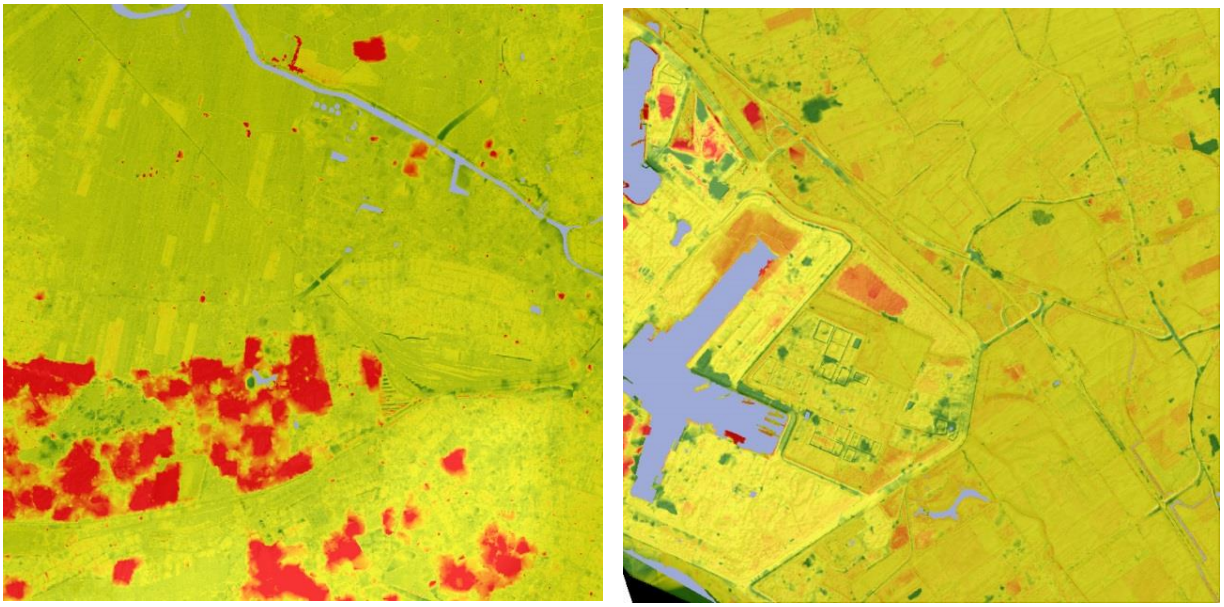


Fig. 38 Geometry based DSM to DTM filtering discrepancy mapping, indicating deviations from the reference DTM in dark green (type I error) and dark red (type II errors) at both study sites (left: Amersfoort, right: Vlissingen).

Similar amounts of type I error are found in the combined spectral feature based and geometry based DSM to DTM filtering results for both study sites. This is caused by the mathematical morphological operation that is implemented similarly (using the same window size and the same amount of iterations) to the geometry based DSM to DTM filtering. It is also clear that using the combined DSM to DTM filtering strategy provides a substantial decrease in type II errors when

compared to geometry based DSM to DTM filtering at the Amersfoort study site. A similar decrease is not apparent at the Vlissingen study site, which is logical given the low amount of type II errors present in the geometry based DSM to DTM filtering. The decrease in type II errors found in combined spectral feature based and geometry based DSM to DTM filtering is related to the presence of large forests that have been successfully targeted in the filtering process by the extracted forest features. The combined DSM to DTM filtering discrepancy mapping results is shown in Fig. 39. Recapitulating the above it can be stated that visual inspection of discrepancies indicate that combined DSM to DTM filtering produces the best results in study sites with forest cover, and that geometry based DSM to DTM filtering produces similar results compared to combined DSM to DTM filtering at areas where there is no forest cover. Both geometry and combined DSM to DTM filtering strategies are prone to type I errors that are caused by the inability of mathematical morphology to discriminate between elevated objects in the terrain. In appendix VIII and appendix IX, some detailed examples of the type I errors are shown for both study sites.

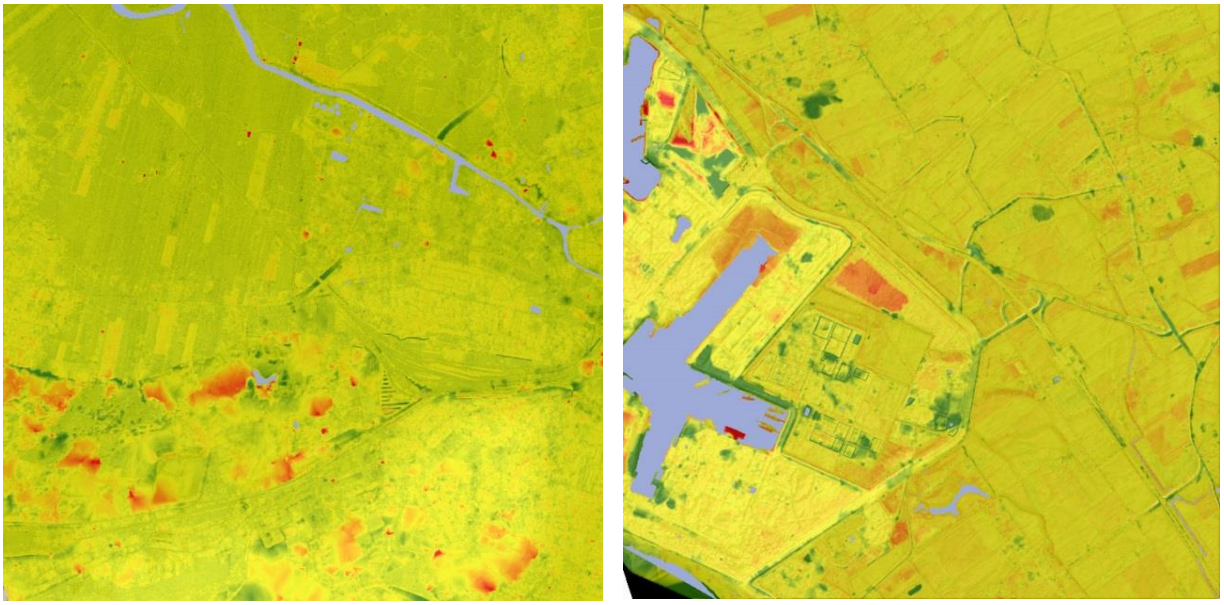


Fig. 39 Combined spectral feature based and geometry based DSM to DTM filtering discrepancy mapping, indicating deviations from the reference DTM in dark green (type I error) and dark red (type II errors) at both study sites (left: Amersfoort, right: Vlissingen).

The results from visual discrepancy mapping are supported by the generated discrepancy statistics. These are shown In Table 12. Clearly the difference between the generated and reference DTM become smallest when using combined DSM to DTM filtering at the Amersfoort study site, and geometry based filtering at the Vlissingen study site. The larger difference between the generated DTM and reference DTM is related to the elevation values found in the GCPs as described in the vertical surface accuracy results.

Table 12 Statistical discrepancies between the various generated DTMs and the reference DTM including mean discrepancy, range of discrepancy and standard deviations. Negative values indicate that generated elevation values are lower than the reference elevation values.

Area	Mean diff. (m)	Diff. range (m)	Standard deviation
Vlissingen spectral DSM to DTM filtering	0.95	-39.88 – 44.39	3.00
Vlissingen geometric DSM to DTM filtering	-0.18	-46.25 – 25.53	2.31
Vlissingen combined DSM to DTM filtering	-0.19	-33.62 – 27.36	1.91
Amersfoort spectral DSM to DTM filtering	4.16	-29.61 – 57.78	3.66
Amersfoort geometric DSM to DTM filtering	2.98	-14.77 – 37.88	3.01
Amersfoort combined DSM to DTM filtering	2.05	-14.63 – 51.98	1.17

6.3.2 Terrain type I and II errors

In order to quantify the amount of type I and type II errors, the reference nDSMs are compared to the nDSMs that are generated using the resulting DTMs from the various applied DSM to DTM filtering strategies and the photogrammetrically derived DSMs. This is similar to what has been done during the accuracy assessment of the feature extraction results. The quantification of errors is shown in Table 13. The determined type I and type II errors support the findings from the visual inspection of discrepancy mapping results. The lowest percentage in type I errors is found in the results from the spectral feature based DSM to DTM filtering strategy. This is related to the selective filtering of the surface based solely on the extracted features. In contrast to geometric DSM to DTM filtering, spectral feature based filtering only affect the surface where features were indeed extracted and as a result large parts of the study site are not taken into consideration during the filtering and therefore cannot result in type I errors. This selective filtering is highly influenced by the errors made during the feature extraction. The errors made during the extracted features resulted in the lowest amount of overlap between the reference and generated nDSMs which represent the filtered above ground objects.

The amount of type I errors increases substantially when incorporating geometry based information through mathematical morphological operators. The amount of type I error is substantially larger in the Vlissingen study site. However, this larger amount of type I errors in the Vlissingen study site is related to the presence of more elevated bare earth terrain features in that area, and not to the applied DSM to DTM filtering strategy. Type II errors show a strong decrease by using either geometry based or combined DSM to DTM filtering at the Vlissingen study site and combined DSM to DTM filtering at the Amersfoort study site.

Table 13 Quantified type I and type II errors (type I E. & type II E.) found after comparing the reference nDSMs (nDSM ref.) with nDSMs generated (nDSM gen.) from the different applied DSM to DTM filtering strategy results and the photogrammetrically derived surface.

DSM to DTM filtering strategy & area	nDSM ref. & gen. overlap (%)	Type I E. (%)	Type II E. (%)
Spectral feature based DSM to DTM filtering Amersfoort	68	23	32
Geometry based DSM to DTM filtering Amersfoort	90	33	10
Combined DSM to DTM filtering Amersfoort	91	31	9
Spectral feature based DSM to DTM filtering Vlissingen	22	68	78
Geometry based DSM to DTM filtering Vlissingen	59	76	41
Combined DSM to DTM filtering Vlissingen	61	76	39

In general it can be stated that the combined spectral feature based and geometry based DSM to DTM filtering strategy yielded the best results at both study sites. DSM to DTM filtering was most successful at the Amersfoort study site given the substantially larger nDSM overlap and smaller type I and type II errors.

6.3.3 Vertical DTM accuracy assessment

The vertical accuracy of the generated DTMs is assessed using the same sample points as used during the vertical accuracy assessment of the surfaces. Optimally a DSM to DTM filtering strategy would cause only minor changes in the original elevation values identified as bare earth and represented by the sampling points. Therefore, the difference between the surface elevation values and the generated terrain values provide valuable information on vertical accuracy. These statistical differences are shown in Table 14. As mentioned in the methodology section it should be taken into account that the sampling points are well distributed, but are relatively low in numbers for statistical analysis. However, it is believed that these statistics provide additive information on the effects of the applied DSM to DTM filtering strategies and is therefore included.

Table 14 Vertical DTM accuracy assessment statistics describing the statistical difference in terms of mean discrepancy, discrepancy ranges and the discrepancy range with 95% confidence level between all produced terrains and reference surface elevation at sampled bare earth locations.

Area & DSM to DTM filtering strategy	Mean diff. (m)	Diff. range (m)	CL95% (m)
Vlissingen spectral DSM to DTM filtering	-0.02	-0.20 – 0.05	-0.07 – 0.04
Vlissingen geometric DSM to DTM filtering	1.07	-0.01 – 2.78	0.30 – 1.84
Vlissingen combined DSM to DTM filtering	1.17	0.01 – 2.78	0.39 – 1.95
Amersfoort spectral DSM to DTM filtering	0.21	-0.67 – 0.81	-0.02 – 0.44
Amersfoort geometric DSM to DTM filtering	0.96	0.04 – 2.45	0.43 – 1.50
Amersfoort combined DSM to DTM filtering	0.86	0.04 – 2.54	0.38 – 1.35

Spectral feature based DSM to DTM filtering affects only the identified objects. Therefore, it only causes small differences as a result of the interpolation process. The inclusion of geometric information in the filtering process causes an average discrepancy of 1 meter between the generated surface and the generated terrains at the sampled bare earth locations. The ranges in which the bare earth elevation points differ is similar for both the Amersfoort and Vlissingen study site.

7 DISCUSSION

In this section the meaning of the obtained results are discussed along with their limitations and implications on the stated research questions. It is discussed whether the questions can be answered well enough to reach the overall research objective and provide the required insights for future productions of geo-specific terrain models. First, the specific production results are discussed along with the associated production limitations found in the process. Secondly, the general research limitations are discussed. Finally some suggestions for future work as identified by this research are described in more detail.

7.1 Photogrammetry results

Generating surfaces from stereo imagery in order to create the geo-specific terrain models used in military simulations is more a necessity than a choice. Stereo satellite imagery is the only guaranteed source of information when creating geo-specific terrain models in foreign territories since they are commercially available. Photogrammetrically derived surfaces are typically inferior to those derived from LIDAR, and in contrast to LIDAR data, satellite information is influenced by atmospheric conditions and is unable to penetrate forest canopies to detect bare earth. This research shows that using state of the art stereo satellite imagery produced by the WorldView-2 satellite, and enhanced automated terrain extraction techniques were not sufficient to generate the point density required to visually recognize buildings at the same resolution as the input imagery. Low densities of elevation points, such as obtained at the Vlissingen study site, could result in the complete omission of a terrain feature in the interpolated surface. However, using stereo overlap from three images, significant improvements can be made to the elevation point density, making recognition of building objects possible at high resolutions. Although recognition of buildings is possible from the surface derived from a triplet pair of stereo imagery, the geometry of above ground objects was not captured as precisely when compared to the reference LIDAR derived surface. Perhaps the statement that photogrammetric processes and the use of LIDAR data are no alternatives to one another but can only complement each other (Baltsavias, 1999) is still valid today for the use of VHR satellite imagery.

The low density elevation points found at the Vlissingen study site is believed to be partly related to the offset between the angles under which the stereo images were captured. The large offset between the angles cause relatively large differences is relief displacement between the images making the image matching increasingly difficult. Relief displacement is typical over urban environments. Another negative effect that is directly related to the capturing angles and surface quality is that large amount of relief displacement remains in the orthorectified imagery. This makes accuracy assessment using reference data difficult since the correct alignment of bare earth features not necessarily implies that building features align properly. Another difficulty is to determine to what degree the success of photogrammetric product derivation is determined by the input data or the user settings. A vast amount of parameters are available to tweak during the various photogrammetric processes, of which some are more influential than others. This jungle of parameters makes it hard to determine whether you have set all parameters correctly, or that improvements can still be made by making slight parameter adjustments. Determining the possible

photogrammetric output quality beforehand is not possible and therefore determining whether the best possible surface is generated from the available imagery is done through trial and error with various settings and is for a large part based on the available reference documentation included with the software and user experience.

7.2 Feature extraction results

The success of feature extraction depends for a large amount on external factors such as time of data capturing, seasonal differences, atmospheric conditions and the complexity of the targeted area. For example in the Amersfoort study area tree foliage occasionally covered roads completely, making spectral extraction impossible. The images covering the Amersfoort site were captured around noon. At the day of capturing the Royal Netherlands Meteorological Institute measured a maximum temperature of 33 degrees Celsius. As a result the sunscreens attached to buildings reflected large amounts of sunlight making the feature extraction of rooftops increasingly difficult.

The process of feature extraction is best started after determining what successful extraction implies and establishing a desired level of success. This might differ on a case to case basis. The goal can be to extract features with a minimum of type I errors, thereby allowing an increase in type II errors, or the opposite, the maximum extraction of objects resulting in lower type II errors but higher type I errors. Considering the low amount of error found in the results of the forest and water feature extraction, these objects were found suitable for further geo-specific terrain production processes. They were successfully extracted using 4 spectral bands as well as 8 bands of WorldView-2. It is found in this research that in order to use the extracted building and forest features in the DSM to DTM filtering process it is important that the type II errors present in the feature extraction results are minimum, otherwise large objects remain in the final DTM. The amount of type II errors in the extracted buildings was too large for successful implementation in the filtering process. Similar amount of errors were found in the extraction of road features and although the extraction results improved for the building and road objects using 8 spectral bands instead of four, thereby supporting the findings of Novack et al. (2011) on the matter, they were still found unsuitable for DSM to DTM filtering purposes considering the remaining levels of error.

The problematic extraction of roads and buildings is believed to be related to the object complexities found in the targeted urban environments. Roads are constructed from various materials, sometimes obscured by tree canopies or buildings and covered with various markings. Buildings have different purposes of use and therefore are built in a large variety in sizes, structures using a large variety of building materials, occasionally coinciding with material used for roads. The feature extraction software used in this research offers various operators that provide aid in the extraction; however, these are often found to conflict at other parts considering the variety of object characteristics. It is believed that geometries play a crucial role in the successful feature extraction of buildings. As shown in Fig. 40 buildings are better captured by its geometry than based on spectral characteristics.

Although geometric building extraction is considered out of scope for this research, this example is included since geometries are found to play a crucial role in building feature extraction. The example shows the generated nDSM after mathematical morphologic operators are applied in the DSM to DTM filtering process. The forest features were used to eliminate objects in the nDSM that intersect with the forest layer. What remains are the building features and the type I errors found in the DTM

results. This method is unaffected by the differences in shape or spectral characteristics of building materials.

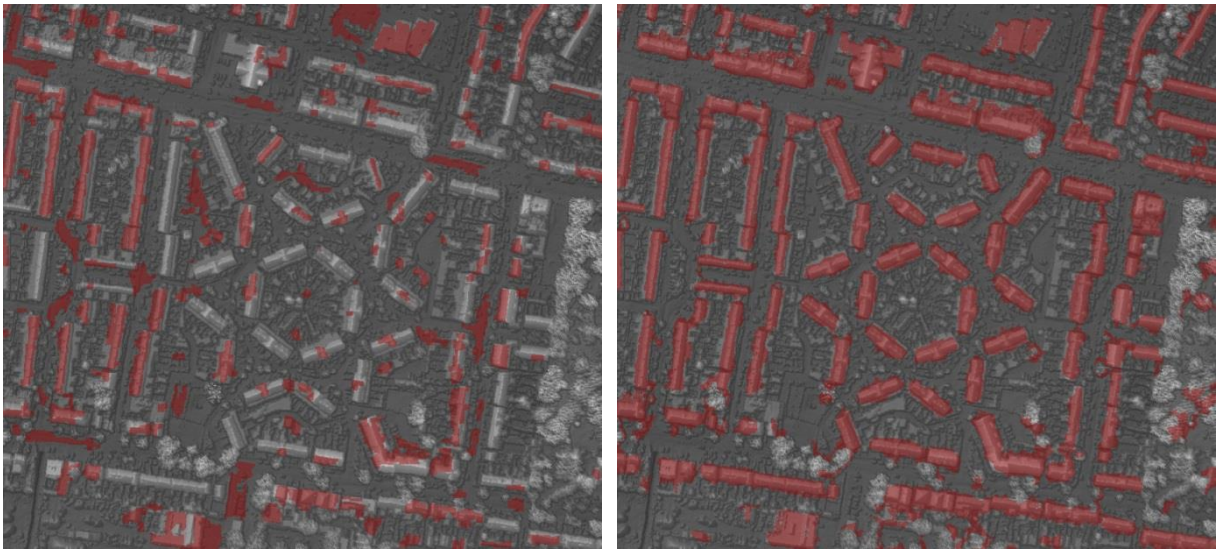


Fig. 40 Examples of spectral building feature extraction (left) and geometric building extraction (right). Building features are colored in red and superimposed over the hill-shaded reference surface.

7.3 DSM to DTM filtering results

The spectral DSM to DTM filtering results show that in order to create a DTM successfully all above ground objects such as buildings and forests should be captured during the automatic feature extraction process. The diverse spectral characteristics of the building rooftops present in the study areas are such that the geometric building footprints could not be captured successfully. The large amount of type II errors in buildings feature extraction and the small amount of type II errors present in the forest feature extraction results cause (parts of) above-ground objects to remain in the processed DTM. The mathematical morphologic opening operation was implemented during the geometry based DSM to DTM filtering process. Given the presence of large buildings in both study sites, large window sizes were required to remove these buildings from the surface using various iterations. Although the removal of large buildings succeeded, large forested areas could not be completely removed from the surface. Above that, the rigorous removal process caused the removal of small terrain features. The inability of the mathematic morphologic opening operation to discriminate between an AGO and a man-made terrain feature that should remain in the final DTM is the reason that there is a substantial increase in type I errors found in the generated DTM results. The successful implementation of mathematical morphologic operators in the process of geometry based DSM to DTM filtering therefore depends on the nature and characteristics of the targeted area. In the Vlissingen and Amersfoort study site the type I errors found in the DTM results were related to characteristic landmarks of which the removal is undesirable. These features could be manually inserted back into the DTM although there is an increased risk of correlation problems considering small vertical differences that can occur between the original surface and the generated DTM. Moreover, the identification of type I errors becomes increasingly difficult without available reference data. Various approaches for combined spectral feature based and geometry based DSM to DTM filtering have been investigated (see section 5.3.3), of which most have been abandoned due to the levels of error found in the extracted features. Eventually only the forest features were used in

the combined DSM to DTM filtering process. This methodology proved successful in the reduction of type II errors in the Vlissingen and Amersfoort study site although there are no apparent improvements in terms of type I errors. Therefore, geometry based DSM to DTM filtering is equally successful in areas with none to low amounts of forest cover in comparison to combined spectral feature based and geometry based DSM to DTM filtering.

DTM refinement was implemented during all DSM to DTM filtering processes. During this process elevation points are iteratively added to the generated DTM if the differences between the DTM and the original surface for that specific point are smaller than a given threshold. Two iterations are implemented in all cases with a threshold of 0.5 meters. As an effect the final DTM is built from a denser collection of elevation point. It must be noted that this method has a negative side effect in case man-made AGOs smaller than 0.5 meters in height are present in the study area since these are inserted back into the DTM.

The statement that photogrammetrically derived surfaces are insufficient for AGO removal made by Haala et al. (1998) is believed to be no longer valid when using state of the art satellite imagery. The applied geometry based and combined spectral feature based DSM to DTM filtering were successful in the removal of AGO. However, the methodologies and data used did not prevent the removal of man-made terrain features. It is believed that the presence of type I errors in the generated DTMs cannot be prevented using mathematic morphologic opening operations in the DSM to DTM filtering process given its inability to discriminate between an AGO and a man-made terrain feature, nor can it be resolved using spectral information given the feature extraction success rates in terms of type I and type II errors. The best approach for DSM to DTM filtering is believed to be one in which slopes are taken into account as described by Vosselman (2000) and Sithole (2001) and where surface connectivity, contingency or traversability are taken into account. This is in line with the statement made by Van de Woestyne et al. (2004) that looking at heights alone does not suffice and connectivity plays a larger role in determining bare earth.

7.4 Research limitations

Limitations of this research that should be taken into account are related to the study sites and the data used. The choice for the study sites is related to the availability of high quality reference data providing valuable insights in the obtained results. However, large differences in elevation were not present in the study sites. The implemented filtering techniques are not expected to be equally successful if very steep terrains or mountainous regions are present in the study area, considering the implemented rigorous filtering of elevated objects. Another problem is expected to occur if these elevated terrain features are covered by large forested areas, because the loss of data when using forest features to remove elevation points from the surface in the DSM to DTM filtering process is expected to be severe, resulting in a large increase in type I errors. The alternative would be to maintain the forests in the surface resulting in a high amount of type II errors, or to use the forest features to implement specific strategies in order to approximate bare earth elevation under forest canopies.

Even when problems resulting from forests or mountains are resolved, geo-specific terrain modeling using photogrammetrically derived data is not likely to be a success for every location on earth. Satellite imagery is influenced by atmospheric conditions. As a result, clouds can occur in the images

and clouds are mostly present at some locations around the equator. For clouds the same problem occurs as found with forest. When removing these areas from the surface all information with regard to the elevation is lost and should be approximated through interpolation techniques. This is a potential source of error that should always be taken in to consideration. No clouds were present in the images used in this research.

Perhaps to a lesser extent a research limitation, but more a consideration, is the vast amounts of data generated in the process of geo-specific terrain modeling. Considering that the study sites were relatively small, these amounts of data are believed to become a major bottleneck when processing large areas. This bottleneck is partly caused by limitations in hardware, but would also require the redesigning of the implemented strategies.

7.5 Future research

Given the identified limitations of mathematical morphology in this research, studies on the implementation of other existing LIDAR derived DSM to DTM filtering strategies applied to photogrammetrically derived surfaces would provide valuable information for the purpose of geo-specific terrain modeling. For instance, further investigation of the application of segmentation in the DSM to DTM filtering process could help identify strategies less prone to the typical type I errors found in this research.

Extracted forest features provide a practical advantage in resolving forest cover associated problems in photogrammetrically derived surfaces and successful building extraction is expected to be equally suitable for DSM to DTM filtering processes and as input data for the simulation. Successful building extraction from inferior surfaces, as generated using two or three overlapping stereo images, would be of great value for the purpose of geo-specific terrain modeling. For example, studying the extraction of buildings based on geometric characteristics aided by spectral information instead of a primary focus on spectral information aided by geometric information is therefore suggested as future research. If building extraction is successful it can be used in studies on adaptive DSM to DTM filtering or combined spectral feature based and geometry based DSM to DTM filtering.

Although photogrammetry is a well-established and well-documented in literature, new technological advancements as well as new algorithms custom made for VHR imagery are developed (e.g. HCS resolution merge, semi-global image matching) that offer new possibilities in terms of stereo imagery processing. Given the earlier stated problem that it is difficult to determine the potential success of photogrammetric product derivation from the input satellite images beforehand, more research on the recent do’s and don’ts of photogrammetry in this new and changed environments is believed to be a welcome addition. Examples could be the effect of using pansharpened imagery when applying various existing image matching algorithms or the effects of the offset between the angles at which images where captured. Are larger offsets or smaller offsets better for the image matching process?

8 CONCLUSIONS

The overall objective of the research is to improve future geo-specific terrain modeling results by investigating the potential benefit of using the spectral information that is available when using VHR resolution stereo satellite imagery as a single source of input for an integral approach in geo-specific terrain modeling. In doing so, extra emphasis is given to the potential benefit of using the extra non-typical spectral bands made available by the WorldView-2 satellite. In order to reach the objective, four research questions have been investigated and answered below.

Answering the first research question, as to what suitable methods and techniques exist in literature for the geometry based DSM to DTM filtering of photogrammetrically derived surfaces is not easily answered. Very little research focusses on the DSM to DTM filtering of photogrammetrically derived surfaces created with VHR satellite imagery. On the contrary a large amount of geometry based DSM to DTM filtering methodologies is found in literature over the last two decades describing its use on LIDAR derived surfaces. The large amount of different methodologies implies that there is no consensus in terms of optimal methodology amongst researchers. Most of these researches claim a relative success at a specific study area, but none provide a panacea for geometric DSM to DTM filtering. In that sense all methodologies are suitable and flawed into a certain extent. As such it can be stated that literature does not provide an answer as to what methodology is best fit for the purpose of DSM to DTM filtering of photogrammetrically derived surfaces and remains an open and challenging theme for future research.

An answer to the second research question, whether spectral feature based and geometry based DSM to DTM filtering can be combined in order to improve the photogrammetrically derived DSM to DTM filtering process is found after the successful implementation of extracted features in the combined DSM to DTM filtering process. This research shows that the successful implementation of extracted features is related to the success of the extraction process. Given the relatively high accuracies of the extracted water and forest features, these can be used to improve the DSM to DTM filtering substantially. The extracted features can be used in an initial DSM to DTM filtering iteration, removing for example large forest patches from the surface. As a result subsequent geometry based DSM to DTM filtering can be implemented more effectively. Especially these large sized above ground objects prove challenging to remove from surfaces when implementing mathematical morphological operators as a means of geometry based DSM to DTM filtering. Therefore, the implementation of extracted features such as forests in the DSM to DTM filtering process provides the advantage required to overcome the inability of mathematical morphological operators to remove large above ground objects without negatively affecting the rest of the surface. The implementation of buildings and roads do not show an improvement in the filtering process given their low feature extraction accuracies. However, indications are found in this research that buildings can be extracted more effectively by focusing primarily on its geometric aspects and secondarily on its spectral characteristics.

In answer to the third research question, as to which DTM generation strategy yields the best results in terms of accuracy, it can be stated that combined feature based and geometry based DSM to DTM filtering proved to yield the best results for both study sites. The methodology was most successful in the removal of large forested areas as found at the Amersfoort study site. It must be noted that it is

believed that results could be further improved by the implementation of other geometry based DSM to DTM filtering methodologies that do take slopes and connectivity into account. Furthermore, the accurate above ground object feature extraction, in any given area, would ultimately enable the extraction of most above ground objects from surfaces through combined spectral feature based and geometry based DSM to DTM filtering without the risk of introducing type I errors to the resulting DTM.

The difference in feature extraction output quality when using eight or four bands WorldView-2 imagery is determined in order to answer the fourth research question. It is found that the use of eight bands resulted in higher accuracies for all targeted features in general. The biggest improvements were made in the building and road feature extraction. The type II errors found in the road extraction results was reduced with 6.5 percent and a reduction of 10.5 percent was found in the building extraction results when using eight bands spectral imagery causing a substantially lower increase in type I errors. The features used during spectral feature based DSM to DTM filtering were all extracted using the eight spectral bands. However, during combined spectral feature based and geometry based DSM to DTM filtering only the water and forest features were used. These features showed similar extraction successes using four spectral bands. As mentioned before, indications are found that building extraction could be significantly improved when incorporating geometric information. This research shows that this is equally valid for the use of eight spectral bands. Therefore, the combination of this rich source of spectral information in combination with accurate geometric information may provide the means for successful building extraction over complex urban environments and as such is an interesting theme for future research.

REFERENCES

Acqua, F. D., Gamba, P., & Mainardi, A. (2011). Digital terrain models in dense urban areas. *International Archives of Photogrammetry and Remote Sensing, XXXIV-3/W4*, 195-202.

Baillard, C., Dissard, O., & MaÎtre, H. (1998). *Segmentation of urban scenes from aerial stereo imagery.* Paper presented at the Fourteenth International Conference On Pattern Recognition, Brisbane, Australia.

Baltsavias, E. P. (1999). A comparison between photogrammetry and laser scanning. *ISPRS Journal of Photogrammetry & Remote Sensing, 54*, 83-94.

Blaschke, T. (2010). Object based image analysis for remote sensing. *ISPRS Journal of Photogrammetry & Remote Sensing, 65*, 2-16.

Bresnahan, P. C. (2011). *Geolocation accuracy evaluations of WorldView-1 and WordView-2*. NGA Retrieved from http://calval.cr.usgs.gov/JACIE_files/JACIE11/Presentations/WedPM/115_Bresnahan_JACIE_11.005.pdf.

Cheng, P., & Chaapel, C. (2010). Pan-sharpening and geometric correction WorldView-2 satellite. *GEOInformatics June 2010,* 30-33.

Digitalglobe. (2009a). White paper: The benefits of the 8 spectral bands of WorldView-2 Retrieved 1 Oct, 2012, from http://www.digitalglobe.com/downloads/WorldView-2_8-Band_Applications_Whitepaper.pdf

Digitalglobe. (2009b). *Worldview-2 datasheet*. Digitalglobe. Retrieved from http://www.satimagingcorp.com/media/pdf/WorldView-2_datasheet.pdf

Drağuţ, L., Tiede, D., & Levick, R. (2010). ESP: a tool to estimate scale parameter for multiresolution image segmentation of remotely sensed data. *International Journal of Geographical Information Science, 24*(6), 859-871.

ERDAS. (2010a). IMAGINE Objective user's guide. 1-47.

ERDAS. (2010b). LPS Automatic terrain extraction user's guide. 1-151.

ERDAS. (2010c). LPS project manager user's guide. 1-421.

Farr, T. G., Rosen, A. P., Caro, E., Crippen, R., Duren, R., Hensley, S., et al. (2007). The Shuttle Radar Topography Mission, Rev. Geophys., 45, RG2004, doi:10.1029/2005RG000183. 1-43. Retrieved from http://www2.jpl.nasa.gov/srtm/srtmBibliography.html

Fisher, P. F., & Tate, N. J. (2006). Causes and consequences of error in digital elevation models. *Progress in Physical Geography, 30*(4), 467-489.

Fong, G. (2004). *Adapting COTS games for military simulation.* Paper presented at the 2004 ACM SIGGRAPH international conference on Virtual Reality continuum and its applications in industry Singapore.

Gamba, P., & Casella, V. (2000). *Model independent object extraction from digital surface models.* Paper presented at the XIXth ISPRS Congress 2000, Technical Commission III: Systems for Data Processing, Analysis and Representation, Amsterdam.

Haala, N., Stallmann, D., & Stätter, C. (1998). *On the use of multispectral and stereo data from airborne scanning systems for DTM generation and landuse classification.* Paper presented at the ISPRS commission IV Symposium on GIS: Between Visions and Applications, Stuttgart, Germany.

Habib, F. A., Kim, E., & Kim, C. (2007). New methodologies for true orthophoto generation. *Photogrammetic Engineering & Remote Sensing, 73*(1), 25-36.

Hay, G. J., & Castilla, G. (2006). *Object-based image analysis: Strengths, Weaknesses, Opportunities and Threats (SWOT).* Paper presented at the Bridging Remote Sensing and GIS, 1st International Conference on Object-based Image Analysis (OBIA 2006), Salzburg University, Austria.

Hay, G. J., & Castilla, G. (2008). Geographic Object-Based Image Analysis (GEOBIA): A new name for a new discipline. In T. Blaschke, S. Lang & G. J. Hay (Eds.), *Object-Based Image Analysis: Spatial concepts for knowledge-driven remote sensing applications* (pp. 75-89). Heidelberg, Berlin: Springer.

Hobi, M. L., & Ginzler, C. (2012). Accuracy assessment of digital surface models based on WorldView-2 and ADS80 Stereo Remote Sensing Data. [open access article]. *MDPI Sensors, 12*, 6347 - 6368.

Intergraph. (2013). Imagine Objective - The future of feature extraction, update & change mapping (pp. 8). www.geospatial.intergraph.com: Intergraph / Hexagon.

Jensen, J. R. (2007). *Remote sensing of the environment: an earth resource perspective* (2nd edition ed.). New Jersey: Pearson Prentice Hall.

Joshi, J. R. (2010). *Improving the Quality of digital surface model generated from very high resolution satellite imagery by using object oriented image analysis technique*. thesis. International Institute for Geo-Information Science and Earth Observation (ITC). Enschede.

Jung, F., Paparoditis, N., Bretar, F., & Maillet, G. (2004). *Surface and terrain reconstruction from very high resolution imagery: where are we now?* Paper presented at the 24th Symposium of the European Association of Remote Sensing Laboratories, Bubrovnik, Croatia.

Kasser, M., & Egels, Y. (2002). *Digital photogrammetry* (1st ed.). London: Taylor & Francis.

Kerle, N., Janssen, L. L. F., Huurneman, G. C., & (Editors). (2004). *Principles of remote sensing: an introductory textbook*. Enschede: The International Institute for Geo-Information Science and earth observation (ITC).

Kiema, J. B. K. (2002). Texture analysis and data fusion in the extraction of topographic objects from satellite imagery. *International Journal Remote Sensing, 23*(4), 767 - 776.

Kim, Y. M., Eo, Y. D., Chang, A. J., & Kim, Y. I. (2013). Generation of a DTM and building detection based on an MPF through integrating airborne lidar data and aerial images. *International Journal of Remote Sensing, 34*(8), 2947-2968.

Kuijper, F., van Son, R., van Meurs, F., Smelik, R., & de Kraker, K. J. (2010). *Techniques for automatic creation of terrain databases for training and mission preparation*. Paper presented at the IMAGE 2010 Conference, Scottsdale, Arizona.

Lane, M. A. (2009). Multi-class automated and semi-automated feature extraction using ERDAS IMAGINE Objective Retrieved 1 Nov, 2012, from http://rssportal.esa.int/img/wiki_up/image/Events/2009_ESA-EUSC-JRC/Papers/Pap_Lane.pdf

Li, Z., Zhu, Q., & Gold, C. (2005). *Digital terrain modeling; principles and metodology*. Boca Raton, Florida, America: CRC Press.

Lillesand, T. M., Kiefer, R. W., & Chipman, J. W. (2008). *Remote sensing and image interpretation* (6th ed.). New Jersey: John Wiley & Sons, Inc.

Ma, R. (2005). DEM generation and building detection from lidar data. *Photogrammetric Engineering and Remote Sensing, 71*(7), 847-854.

Mason, S., & Baltsavias, E. (1997). Image-Based Reconstruction of Informal Settlements. *Automatic extraction of man-made objects from aerial and space images, II*, 97–108.

Miliaresis, G., & Kokkas, N. (2007). Segmentation and object-based classification for the extraction of the building class from LIDAR DEMs. *Computers & Geosciences, 33*, 1076 - 1087.

Monmonier, M. (1996). *How to lie with maps* (2nd ed.). Chicago, USA: The University of Chicago Press.

Naseer, A., Eldabi, T., Jahangirian, M., & Stergioulas, L. (2008). *Potential applications of simulation modelling techniques in healthcare: lessons learned from aerospace & military.* Paper presented at the European and Mediterranean Conference on Information Systems 2008, Dubai.

Niemeyer, K. (2003). Modelling and simulation in defense. *Information & Security. An International Journal, 12*(1), 19-42.

Novack, T., Esch, T., Kux, H., & Stilla, U. (2011). Machine learning comparison between WorldView-2 and Quickbird-2 simulated imagery regarding object based urban land cover classification. *MDPI Remote Sensing, 3*, 2263 - 2282.

OGC. (2007). Candidate OpenGIS® CityGML Implementation Specification OGC 07-062: Open Geospatial Consortium Inc. (OGC).

Padwick, C., Deskevich, M., Pacifici, F., & Smallwood, S. (2010). *WorldView-2 pan-sharpening.* Paper presented at the ASPRS 2010 Annual Conference, San Diego, California.

Palà, V., & Arbiol, R. (2002). *True orthoimage generation in urban areas.* Paper presented at the 3rd International Symposium Remote Sensing of Urban Areas, Istanbul, Turkey.

Passini, R., & Jacobsen, D. (2002). *Filtering of digital elevation models.* Paper presented at the ASPRS annual convention 2002, Washington, USA.

Pingel, T. J., Clarke, K. C., & McBride, W. A. (2013). An improved simple morphological filter for the terrain classification of airborne LIDAR data. *ISPRS Journal of Photogrammetry and Remote Sensing, 77*, 21-30.

Priestnall, G., Jaafar, J., & Duncan, A. (2000). Extracting urban features from LIDAR digital surface models. *Computers, Environment and Urban Systems, 24*, 65-78.

Rodríguez, E., Morris, C. S., Belz, J. E., Chapin, E. C., Martin, J. M., Daffer, W., et al. (2005). An assessment of the SRTM topographic products, Technical Report JPL D-31639. 1 - 143. Retrieved from http://www2.jpl.nasa.gov/srtm/srtmBibliography.html

Sithole, G. (2001). *Filtering of LASER altimetry data using a slope adaptive filter.* Paper presented at the International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences XXXIV-3/W4, Annapolis, Maryland, USA.

Sithole, G., & Vosselman, G. (2004). Experimental comparison of filter algorithms for bare-earth extraction from airborne laser scanning point clouds. *ISPRS Journal of Photogrammetry & Remote Sensing, 59*, 85-101.

Smelik, R. M., De Kraker, K. J., & Groenewegen, S. A. (2009). *A survey of procedural methods for terrain modelling.* Paper presented at the Proceedings of the CASA Workshop on 3D Advanced Media in Gaming and Simulation (3AMIGAS), Amsterdam, the netherlands.

Smelik, R. M., Tutenel, T., de Kraker, K. J., & Bidarra, R. (2008). *A proposal for a procedural terrain modelling framework*. Paper presented at the EGVE 2008 - 14th Eurographics Symposium on Virtual Environments, Eindhoven, the Net5herlands. http://graphics.tudelft.nl/~ruben/RMSmelikEGVE08.pdf

Smelik, R. M., Tutenel, T., de Kraker, K. J., & Bidarra, R. (2010). Declarative terrain modeling for military training games. *International Journal of Computer Games Technology, 2010*, 1 - 11.

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

Sowmya, A., & Trinder, J. (2000). Modelling and representation issues in automated feature extraction from aerial and satellite images. *ISPRS Journal of Photogrammetry & Remote Sensing, 55*, 34-47.

Tian, Y., Guo, P., & Lyu, M. R. (2005). *Comparative studies on feature extraction methods for multispectral remote sensing image classification*. Paper presented at the 2005 IEEE International Conference on Systems, Man and Cybernetics, Waikoloa, Hawaii.

Van de Woestyne, I., Jordan, M., Moons, T., & Cord, M. (2004). *A software system for efficient DEM segmentation and DTM estimation in complex urban areas.* Paper presented at the ISPRS Congress 2004, Istanbul.

van der Zon, N. (2012). Kwaliteitsdocument AHN-2. Retrieved from AHN website: http://www.ahn.nl

Vosselman, G. (2000). *Slope based filtering of laser altimetry data.* Paper presented at the XIXth ISPRS Congress 2000, Technical Commission III: Systems for Data Processing, Analysis and Representation, Amsterdam.

APPENDIX I FEATURE EXTRACTION STRATEGIES

Table 15 Feature extraction strategies applied on Vlissingen study site (single feature probability abbreviated as SFP).

Vlissingen Objects	Raster Pixel Processor	Raster Objects Creator	Raster Objects Operator	Raster to Vector Conversion	Vector Objects Operators	Vector Objects Processors	Vector Cleanup Operators
Forests	SFP	Threshold and clump	Probability filter, Dilate, Size filter	Polygon trace	Island filter, Generalize, Split	-	-
Water	SFP	Segmentation	Focal, Probability filter, Size filter, ReClump	Polygon trace	Island filter, Smooth	-	Dissolve, Generalize, Outlier clipper
Roads	SFP	Threshold and clump	Size Filter, Dilate (2*), Erode2*), Probability filter	Polygon trace	Generalize, Island filter	Perimeter sq./Area	Probability filter
Buildings small red	SFP	Segmentation	Probability filter, ReClump Erode, Dilate, Size filter	Polygon trace	Generalize, Island filter	-	-
Buildings small grey	SFP	Segmentation	Probability filter, ReClump, Erode, Dilate, Size filter	Polygon trace	Generalize, Island filter	-	-
Buildings large	SFP	Segmentation	Probability filter, ReClump, Size filter	Polygon trace	Generalize, Island filter	Perimeter sq./Area Orthogonality	Probability filter, Dissolve, Generalize, Island filter

Table 16 Feature extraction strategies applied on Amersfoort study site

Amersfoort Objects	Raster Pixel Processor	Raster Object Creator	Raster Object Operator	Raster to Vector Conversion	Vector Object Operators	Vector Object Processors	Vector Cleanup Operators
Forests	SFP	Threshold and Clump	Probability filter, Dilate, Erode, Size filter	Polygon trace	Generalize	-	-
Water	SFP	Segmentation	Focal, Probability filter, Size filter, ReClump	Polygon trace	Island filter, Smooth	-	Dissolve, Generalize, Outlier clipper
Roads	SFP	Segmentation	Size Filter, Probability filter	Polygon trace	Segment merge, Generalize, Island Filter	Perimeter sq./Area	Dissolve Island filter
Buildings small red	SFP	Segmentation	Probability filter, ReClump, Erode, Dilate, Size filter	Polygon trace	Generalize, Island filter	-	-
Buildings small grey	SFP	Segmentation	Probability filter, ReClump Erode, Dilate, Size filter	Polygon trace	Generalize, Island filter	-	-
Buildings large	SFP	Segmentation	Probability filter, Size filter	Polygon trace	Generalize, Island filter	Perimeter sq./Area, Orthogonality	Probability filter, Dissolve, Generalize, Island filter

APPENDIX II CONFLICTING FEATURE EXTRACTION OPERATORS

In this appendix some examples are given of operators that can be used during the process of feature extraction, and its potential conflicting effect for different study sites. As stated earlier the extraction of features is increasingly complicated in areas with diverse landscapes and levels of urbanization. One reason that contributes to this effect is that the use of an operator in feature extraction can have conflicting results for the same target feature object in different areas. Four examples are given.

Perimeter/area operator

During the feature extraction process one useful cue is the perimeter divided by area (perimeter^2/area) metric. This metric can identify areas that are long and thin, because these would have a large perimeter but have relatively little area. In Fig. 41 and Fig. 42 it is shown that the use of this cue can have positive as well as negative effects on the feature extraction. In Fig. 41, image one, two and three show the relative successful application of this operator since it clearly removes the thin elements such as roads that are not part of the building features. The red features are spectrally recognized building features and the features resulting from the perimeter/area operator are shown in green. In order to eliminate features with this operator a threshold value has to be input into the operator. By only taking into account spectral information the distinction between certain buildings and road parts is hard to make.

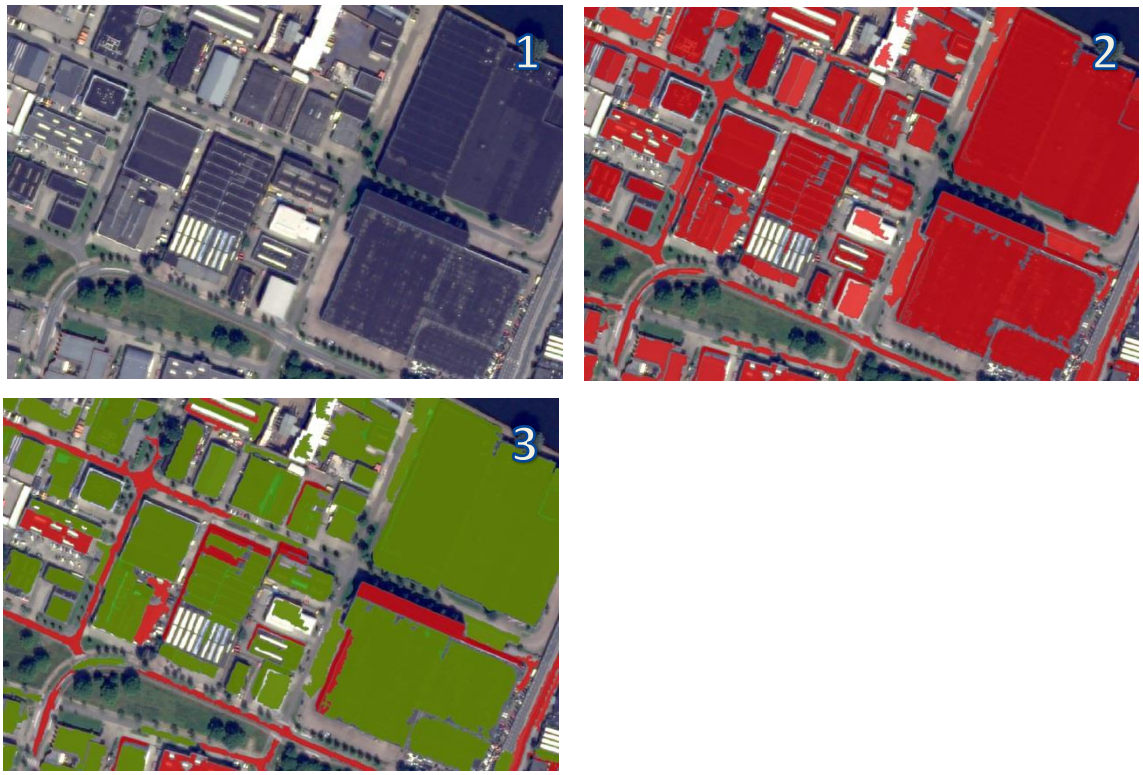


Fig. 41 Example of positive effects from perimeter^2/area metric operator implementation. 1: orthorectified imagery, 2: extracted buildings (red) before implementation, 3: extracted buildings (green) and removed features (red) after implementation.

In Fig. 42, image four, five and six, an example is given where the same operator is used, but its effects are relatively negative on the outcome of the extraction. It is shown that because the rooftops are made up of two distinct feature types, one has been spectrally identified whereas the other is not, creating a comb shaped building feature. Because of its comb shaped geometry there is a relatively high perimeter/area value, resulting in the elimination of the feature in the extraction process.

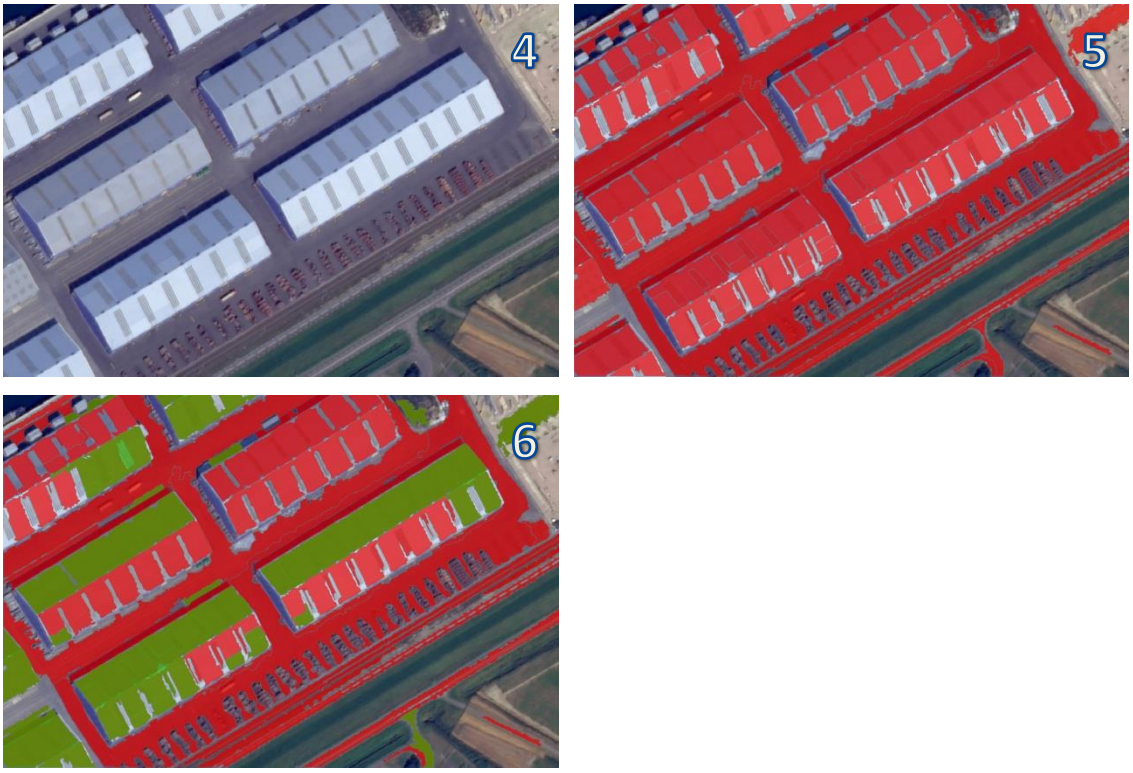


Fig. 42 Example of negative effects from perimeter^2/area metric operator implementation. 1: orthorectified imagery, 2: extracted buildings (red) before implementation, 3: extracted buildings (green) and removed features (red) after implementation.

Oriented bounding box

A vector object can be described by a primary and a secondary axis. The primary axis is the longest straight line that can be drawn between two nodes belonging to a vector object. The secondary axis is the second longest line that can be drawn. These metrics can be useful in various feature extraction operators. It can for instance be strategically implemented in the extraction of roads, since a very long primary axis and a small secondary axis would help identify a long and thin feature. These two axes are also used in the oriented bounding box vector operator. This operator creates an oriented bounding box in which the direction and length is defined by the longest axis, and the width is defined by the second axis.

In Fig. 43 two examples of the operator implementation are given. In image one, two and three, it is shown that the bounding box operator can be used with relative success for long strips of attached houses. Only the small vector objects have an oriented bounding box which differs from the building's natural orientation. In image four, five and six, an example is given in which the use of the oriented bounding box is relatively unsuccessful. L-shaped objects will always have an oriented bounding box that deviates from the buildings natural form, unless these vector objects are divided

into rectangular blocks. In image 5 the primary axis is drawn from the lower left corner of the object to the upper right corner.

Fig. 43 Example of positive effects (1-3) and negative effects (4-6) of oriented bounding box operator implementation. Building features before implementation colored in red and after implementation in green.

It should be noted that the extracted vector objects describing the individual houses are very often little rectangular / orthogonal in shape, because of tree canopies obscuring the rooftops, or the effect of mixed pixels around the edges of an object. Therefore, using the vector object's primary and secondary axis to create an oriented bounding box will lead to undesirable results for these individual buildings.

Island filtering

The island operator closes the gaps within the extracted vector objects, considering a predefined minimum and maximum size threshold for the gaps to be filled. When extracting road features this operator is very useful to close the small gaps that are not extracted as a road due to a difference in the spectral characteristics. This could for instance be a shadow, a car on the road or the markings on a road. In Fig. 44, image one, two and three show that the island filter is implemented successfully for the closing of the small gaps present within the extracted feature. At another part of the same study area a more undesirable effect of the same operator took place as shown in image four, five and six. Here the small gap in the middle of the roundabout, is indeed not a part of the road network, but gets filled by the operator because of the gap size.

Fig. 44 Examples of successful (1-3) and unsuccessful (4-6) island filter operator implementation. Road features before implementation are colored in red, and after implementation in green.

Orthogonality & size filter

The orthogonality object cue metric is a measure describing the perpendicularity of the line segments belonging to a vector object. The value ranges from 0 to 1, in which the value 1 indicates a

higher level of orthogonality. The level of orthogonality in combination with the measure of vector object size is very useful in the extraction of large industrial buildings. These objects tend to be orthogonal in shape and are relatively large. In Fig. 45, image one, two and three show the successful implementation of the orthogonality cue metric and the size filter cue metric. Both cues are used in combination with a minimum threshold value, causing only large orthogonal shapes to remain. As it becomes clear from image four, five and six, the same strategies cannot be applied to circular shapes which unfortunately are not uncommon at the Vlissingen study site. These circular shapes require different strategies.

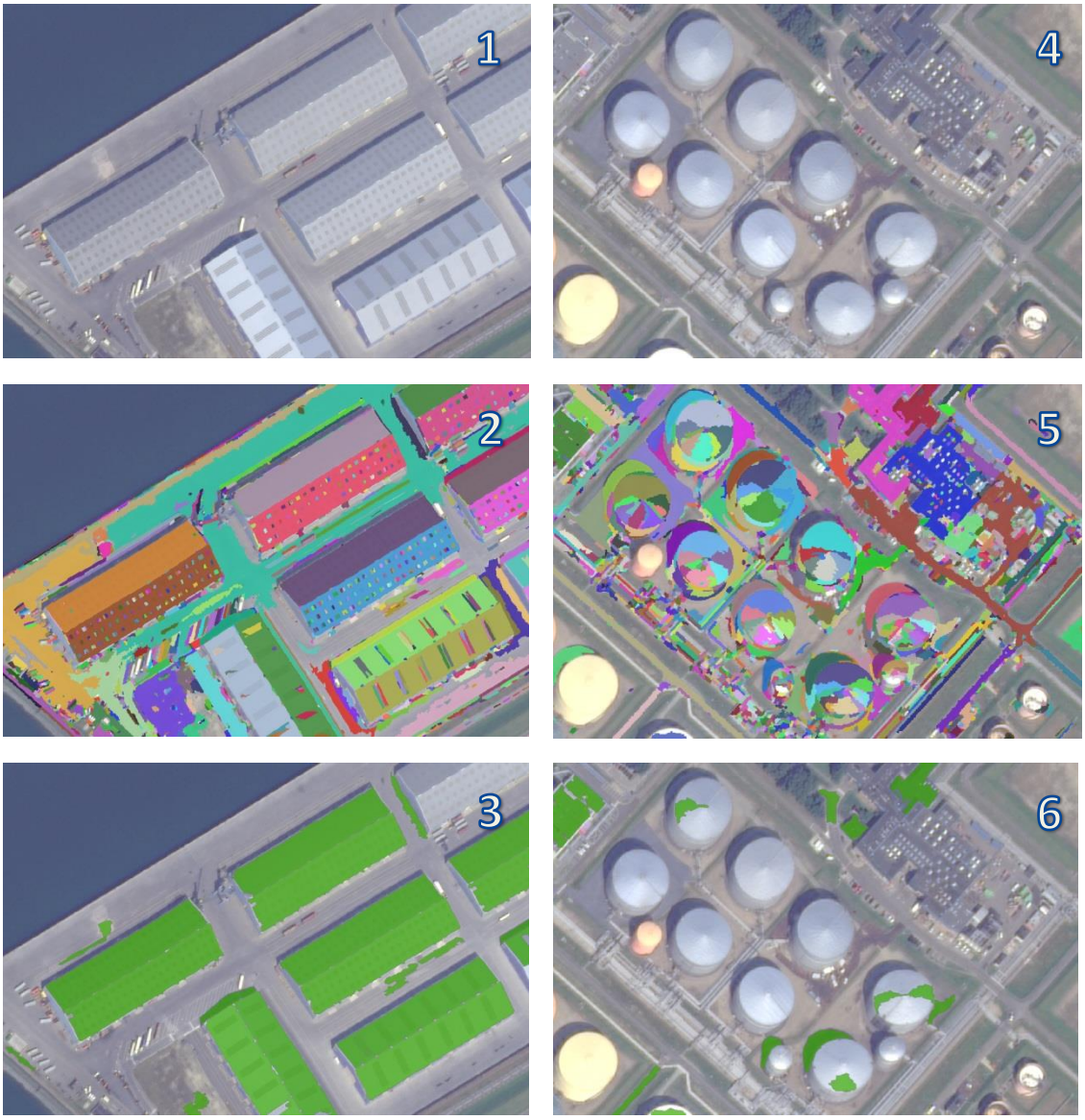


Fig. 45 Examples of successful (1-3) and unsuccessful (4-6) implementation of orthogonality & size filter operation. Image 2 and 5 represent the segments before implementation and results after implementation are shown in green in image 3 and 6.

APPENDIX III ERRORS OCCURRING IN FEATURE EXTRACTION

In this appendix some errors found during the feature extraction process are explained in more detail. These errors are considered as such because the reason for the undesired behavior could not be identified from both theory and software documentation. These errors are not necessarily software related errors, yet remain unresolved within this project.

Erroneous segment merge

During the process of feature extraction various operators can be applied in both the pixel and vector domain. One such operator is the segment merge operator that is useful in forming aggregate vector objects by merging adjacent vector objects into an object with a higher probability (ERDAS, 2010a). This operator was particularly useful in the roads feature extraction although an erroneous side effect has been identified. In Fig. 46 it is shown that the use of this operator can cause islands, although its use is the merging of adjacent vector objects into higher probable vector objects.

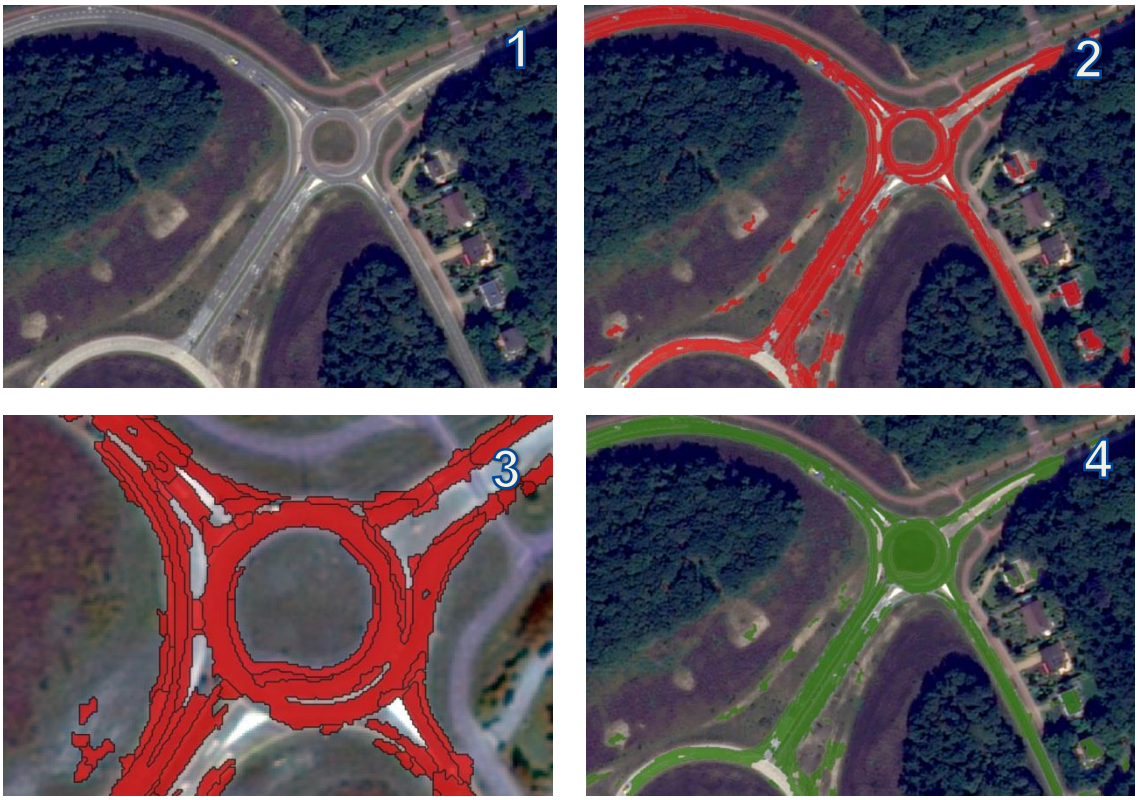


Fig. 46 Example of undesired side effects of segment merge during road feature extraction. 1: orthorectified imagery, 2: road raster object, 3: road vector object, 4: road after segmenet merge.

In Fig. 46 the first image shows the study area in which the error was found to occur. Image two and three show the road vector objects prior to the segment merge operator. The fourth image shows the vector objects after the segment merge operator, clearly showing that the island located in the middle of the roundabout is filled and now considered a road vector object. The reason for this error is unknown and is not resolved within this project.

Pixel candidates not converted to the vector domain

In the feature extraction workflow, training samples and quantified cues at the pixel level are used to find probable pixel candidates for a certain feature. As can be seen in Fig. 47 some clustered pixel candidates that were identified in the pixel domain are not converted to the vector domain for unknown reasons. In Fig. 47 the first image shows an example area in which this error occurred, the second image shows the clustered pixel candidates that have already been subjected to various filtering processes using cues at the pixel level and training samples, the third image shows the results of Imagine Objective's Raster Object to Vector Object Operator, which is used for the conversion of clustered candidate pixels to candidate vector objects. In the fourth image the red areas show the pixel candidates that are not converted to the vector domain. The reason for this error is unknown and is not resolved within this project.

Fig. 47 Examples of errors in the pixel to vector domain conversion during feature extraction process. 1: orthorectified imagery, 2: building raster objects, 3: building vector objects, 4: building vector objects superimposed on building raster objects.

APPENDIX IV ALTERNATIVE DELTA HEIGHT WORKFLOW

Delta height is the average difference in heights between an extracted feature and its surroundings given a certain buffer size around the extracted object. The use of the delta height object cue is very useful in finding falsely extracted buildings (type I errors), because it can be assumed that in most cases a low difference in average height between an extracted object and its surroundings indicates the feature is not a building. A delta height cue is available in the Imagine Objective suite. However, the use of this cue is limited by the available hardware derived physical memory in a computer, and in practice is limited to relatively small areas when using Imagine Objective.

The study areas are both 5 km * 5 km and are therefore too large to produce a delta height cue for the whole area, using Imagine Objective. An alternative approach is chosen in order to produce this cue and to reduce the type I errors in the extracted building features. Using the COTS software package ArcGIS, a 5 meter buffer was created around the extracted features, as can be seen in Fig. 48.

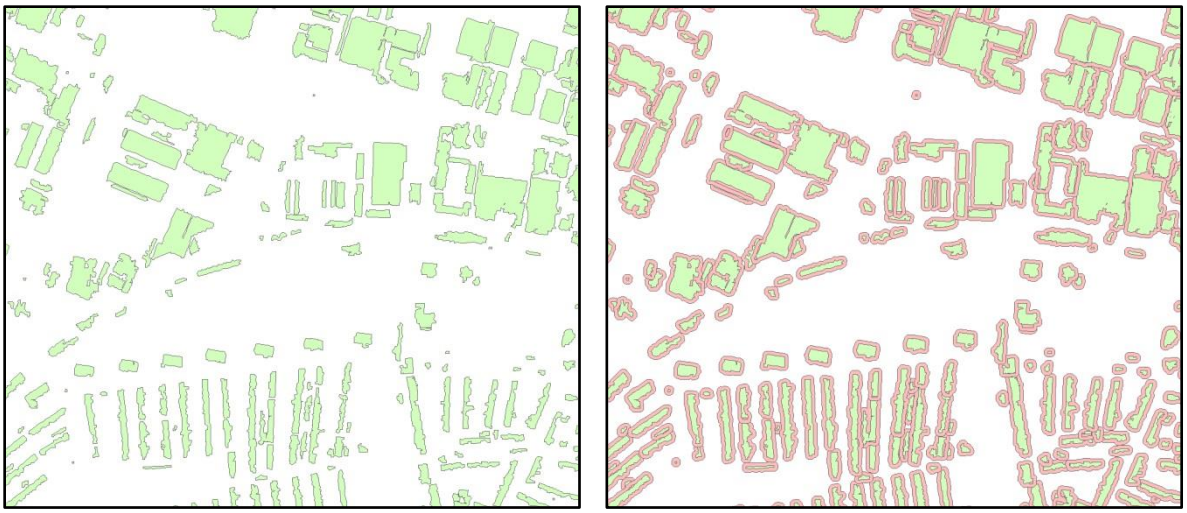


Fig. 48 Extracted (left) and buffered building features (right).

In the next step the point cloud that was derived in the process of DSM extraction was converted to raster, in such a way that no interpolation took place between individual points. Every point was converted to a 0.5m * 0.5m pixel. This is done so that no interpolated height values influence the delta height procedure. All heights used for delta height extraction are heights obtained through the pixel wise image matching in the photogrammetry process.

Next the DSM is subdivided in a DSM containing only pixels that overlap with the extracted features, and a DSM containing only pixels in the buffered zones as shown in Fig. 49. The resulting two DSM files are converted back into to point features in which the height values are stored as an attribute for calculation purposes as shown in Fig. 50. For every extracted building feature and its corresponding buffer, the mean height value of all points enclosed within the boundary of that feature are determined and assigned to the attributes of that feature. Through relational links between attributes of the extracted buildings and the buffers, both the average height of the

building and its surrounding buffer are linked to the buildings. From these two values the difference in average height is calculated and stored in the building feature attributes as a delta height cue.

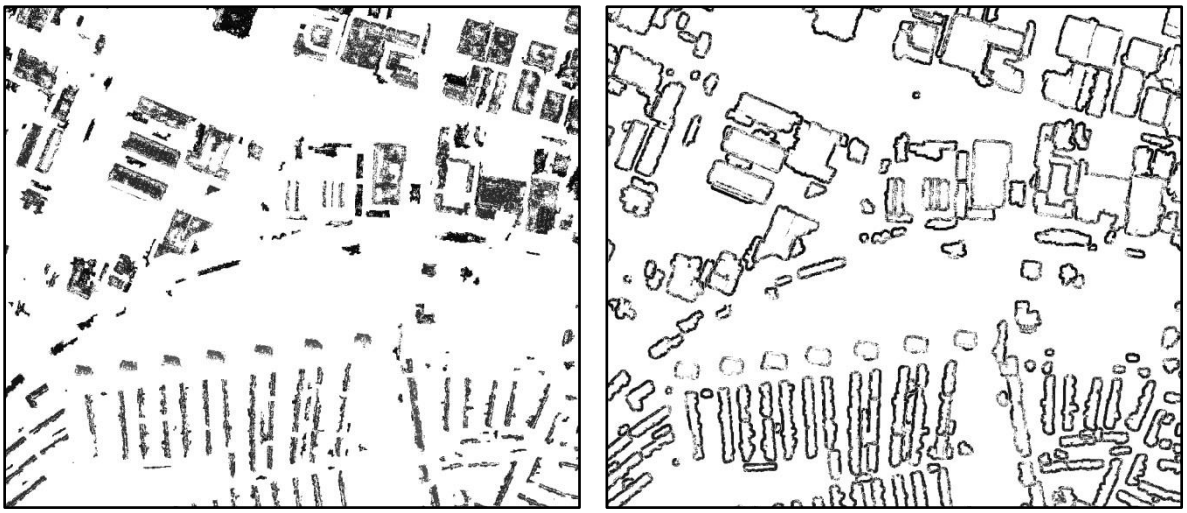


Fig. 49 Subdivided DSM (left: buildings and right: buffers) for extracted features and buffer zones.

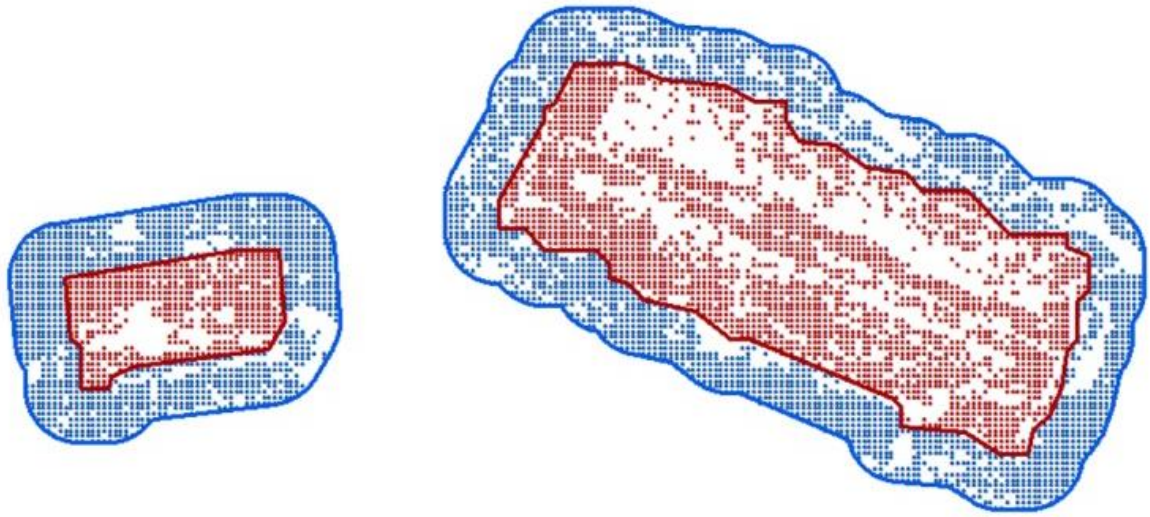


Fig. 50 Elevation points derived from DSM within extracted buildings and buffers.

Negative delta height values could indicate that a building is surrounded or next to a larger building, or it could be a faulty extracted building feature such as a road.

APPENDIX V PREPROCESSING OF AHN2 DATA FOR DEM AA

In order to use the available high quality elevation information such as the AHN-2, for accuracy assessment, preprocessing is required to make the datasets similar in certain aspects such as interpolation of (noData) values, projections and vertical datum's. The accuracy assessment of the DSM and DTM are similar and therefore both require the same preprocessing of the AHN-2 DSM and DTM.

First the unfiltered AHN-2 (DEM) is re-projected from the Dutch national projection into the same coordinate system as the photogrammetrically derived DEM which is projected in the coordinate system WGS84 UTM31N. The vertical datum of the AHN-2 is similarly converted from the Dutch national vertical datum to the global vertical datum WGS84. Maintaining different vertical datum's would result in differences up to about 75 meters for the same points in the different DSMs.

Although a local vertical datum better reflects the true elevation, conversion from the local datum to a global datum is done for practical reasons. For accuracy assessment purposes it is not the true elevation values that matter, but the relative difference in elevation values.

After the re-projection of the AHN-2 and the vertical datum transformation, the AHN-2 is converted back to point features in which the height values are stored in the attributes as can been seen in Fig. 51.

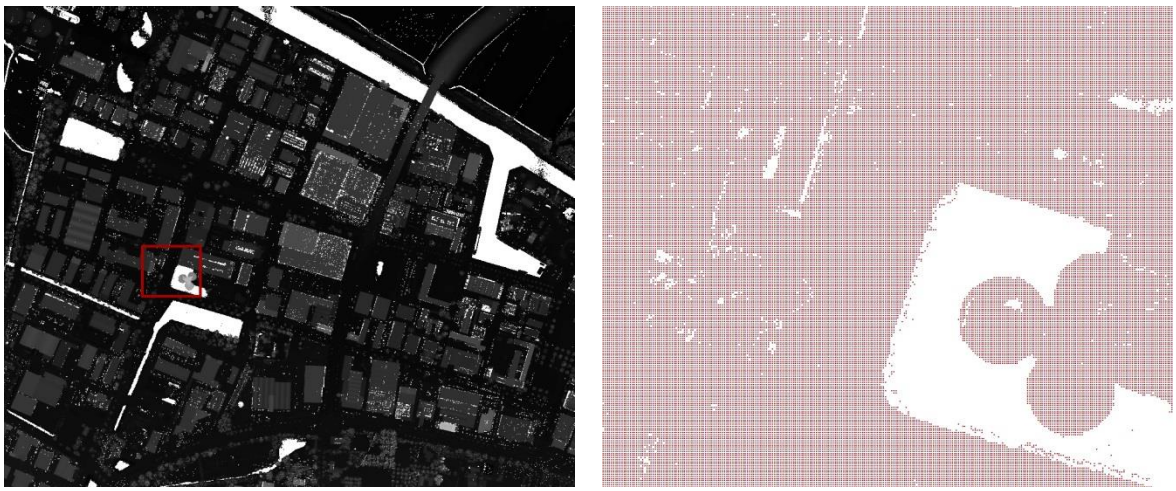


Fig. 51 AHN-2 DSM conversion from raster (left) to points (right).

The point features are used to create a Delaunay triangulation based triangulated irregular network (TIN) in which a surface or terrain is built from the triangular faces between the elevation points. Also larger areas with no elevation points will be covered by this surface. The TIN is in turn converted back to a raster allowing linear interpolation to take place between the original elevation points. This is shown in Fig. 52. At this point the AHN-2 DEM is ready for the process of accuracy assessment.

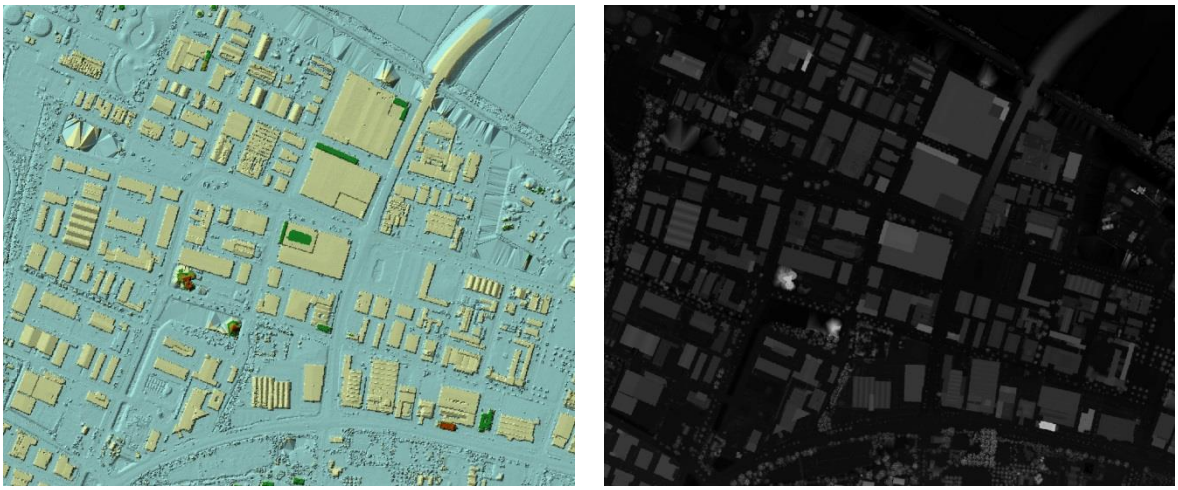
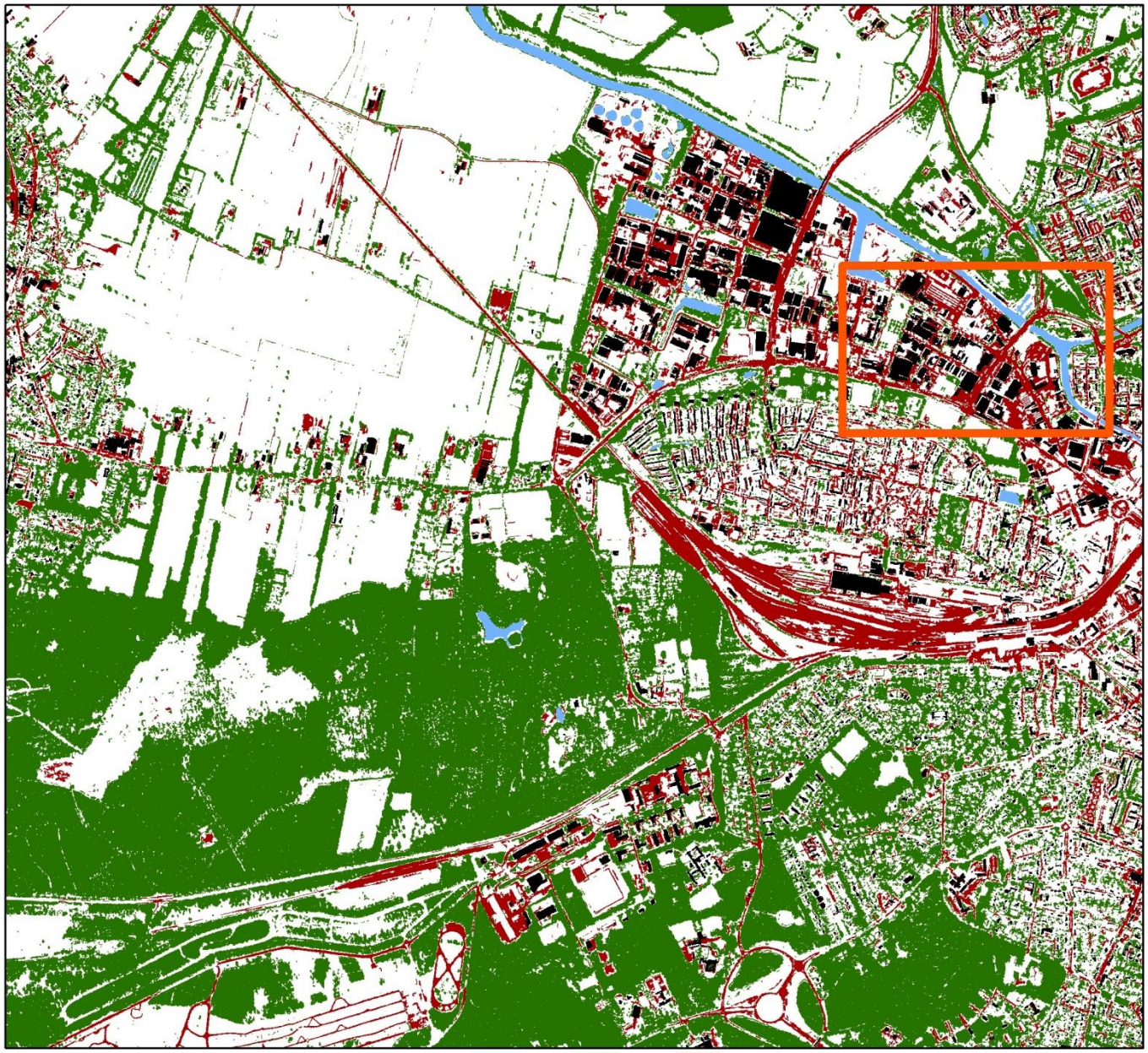


Fig. 52 Conversion from TIN surface (left) to raster DSM (right).

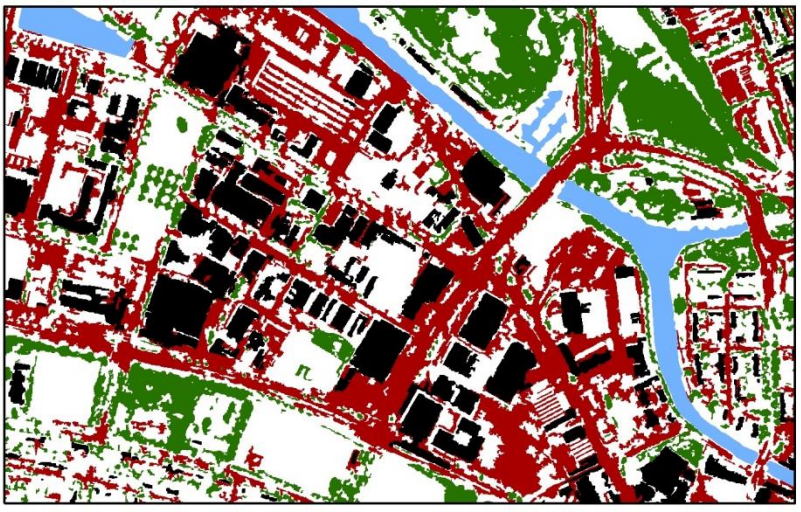
APPENDIX VI FEATURE EXTRACTION RESULTS AMERSFOORT



1:25.000

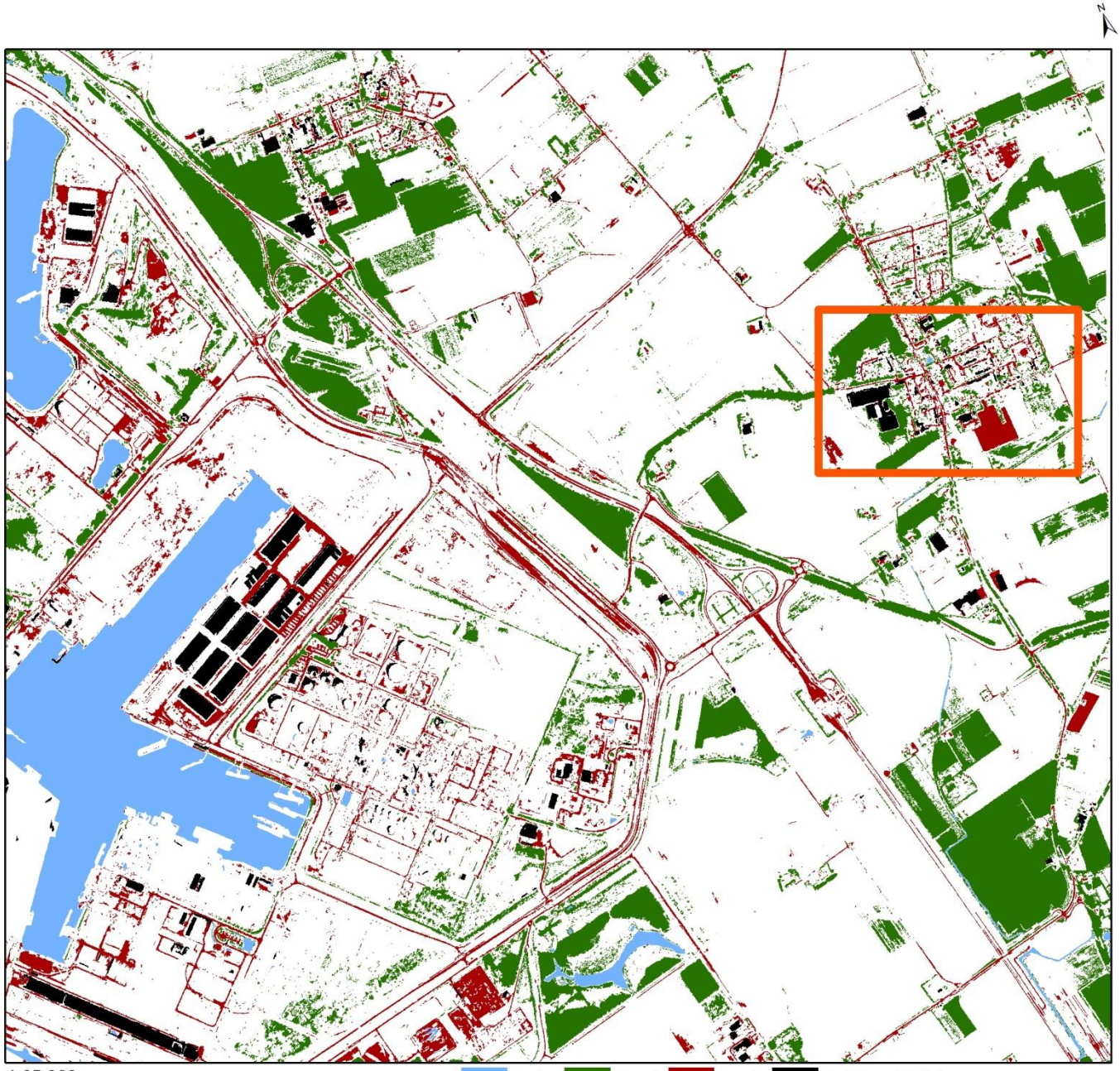
water forest roads buildings (delta)

1:65.000



1:10.000

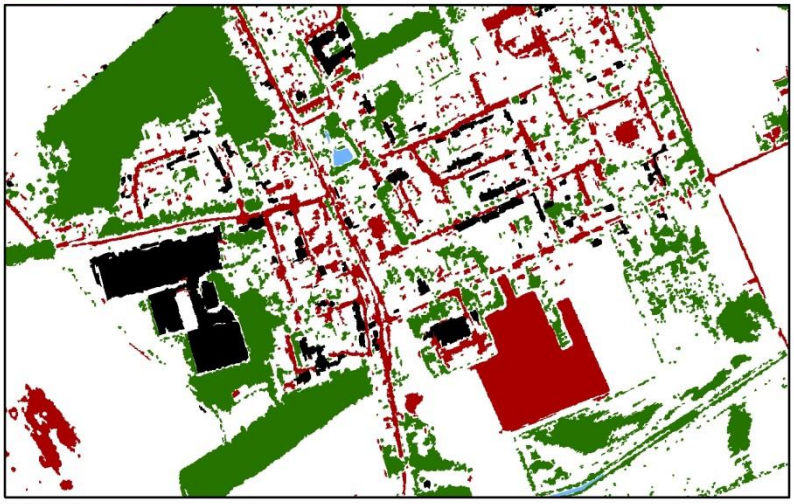
APPENDIX VII FEATURE EXTRACTION RESULTS VLISSINGEN



1:25.000

water forest roads buildings (delta)

1:65.000



1:10.000

APPENDIX VIII TYPE I ERROR EXAMPLES IN DTM DISCREPANCY RESULTS AMERSFOORT

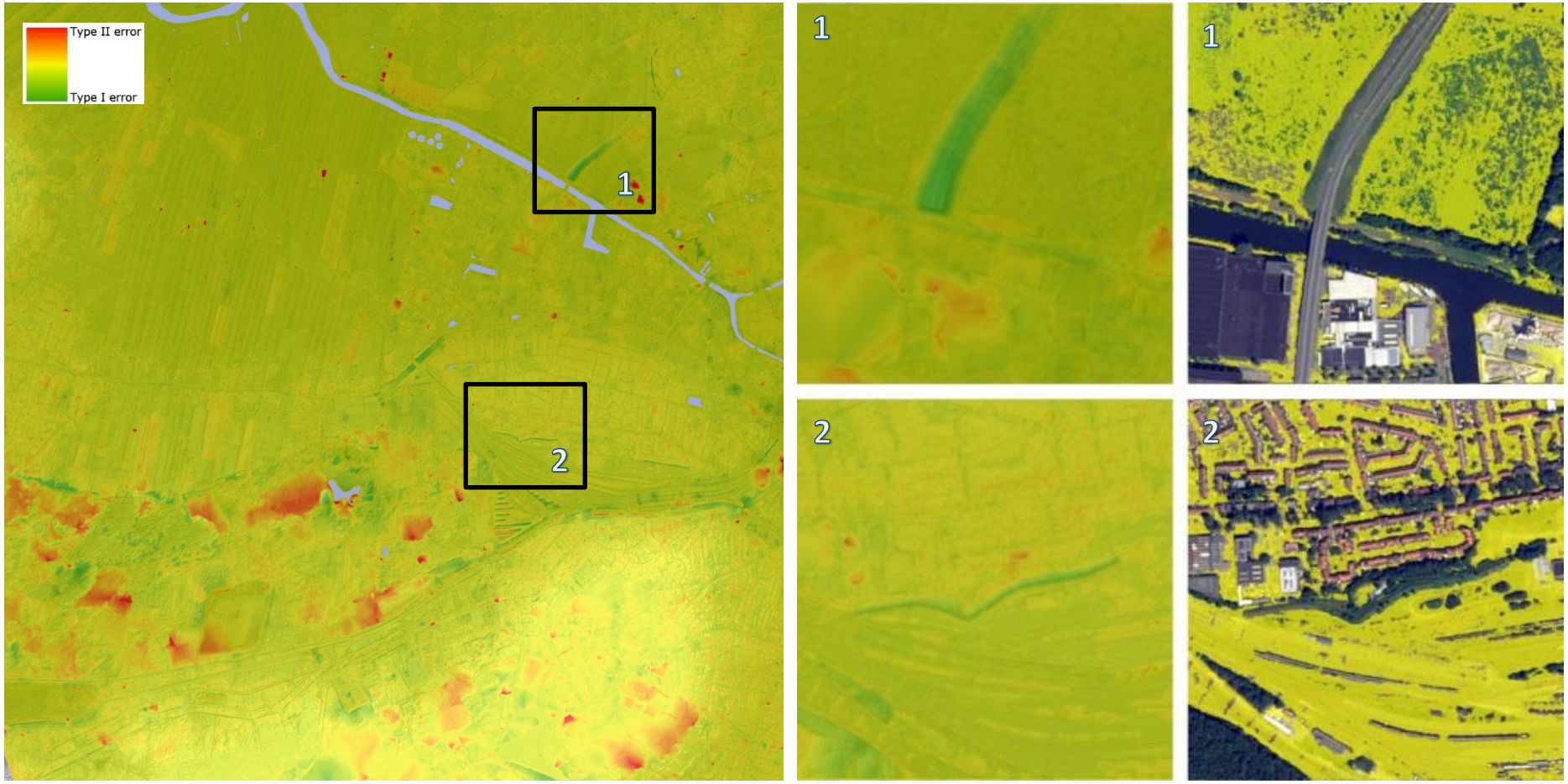


Fig. 53 Close-up examples of type I errors found in combined DSM to DTM filtering results at Amersfoort study site. Light green indicates the presence of type I errors. The non-interpolated bare earth points used for interpolation of the final DTM are colored yellow in the right column.

APPENDIX IX TYPE I ERROR EXAMPLES IN DTM DISCREPANCY RESULTS VLISSINGEN

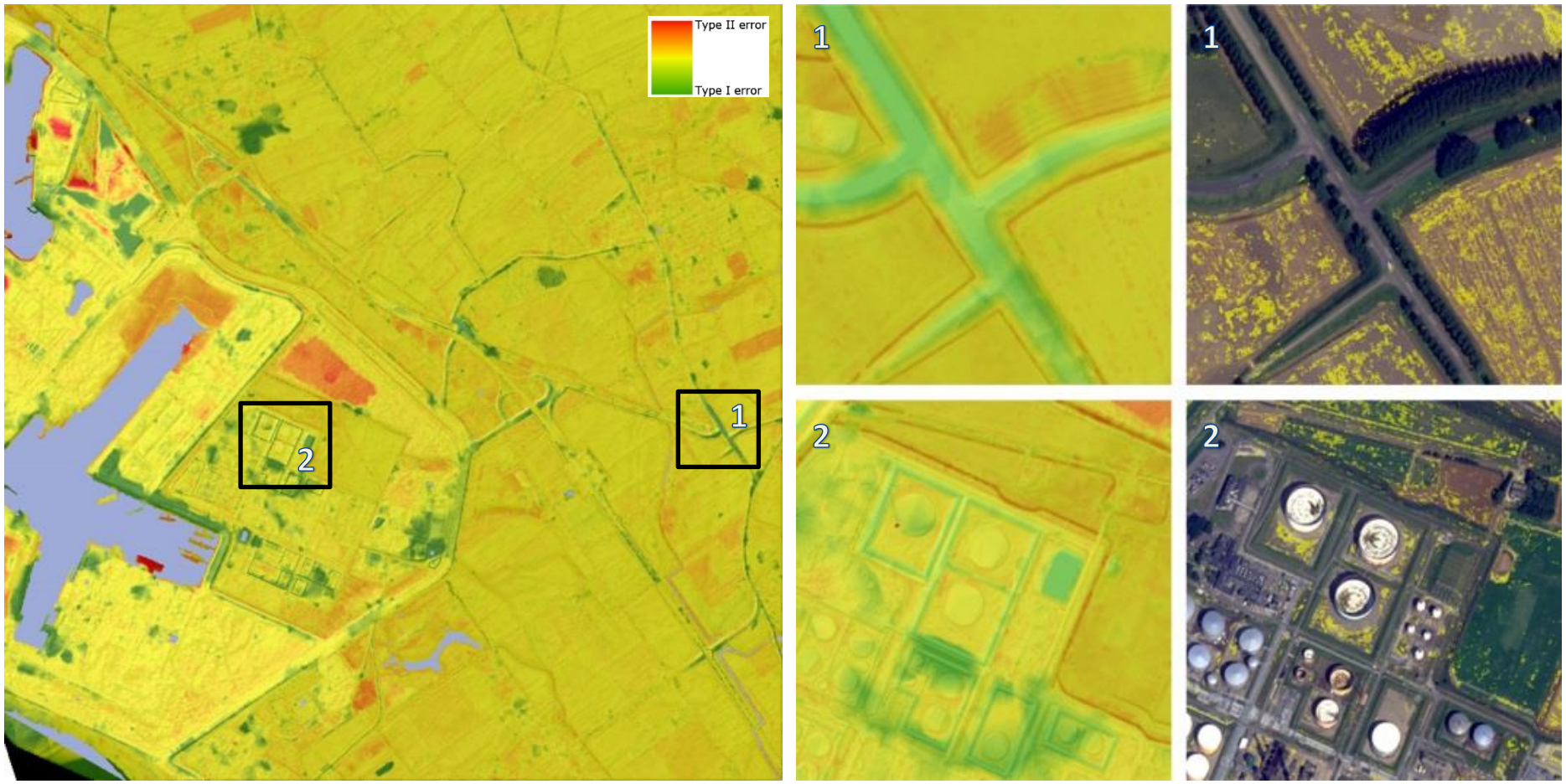


Fig. 54 Close-up examples of type I errors found in combined DSM to DTM filtering results at Vlissingen study site. Light green indicates the presence of type I errors. The non-interpolated bare earth points used for interpolation of the final DTM are colored yellow in the right column.