Technische Universität München

Ingenieurfakultät Bau Geo Umwelt

Signal Processing in Earth Observation

Prof. Dr.-Ing. habil. Xiaoxiang Zhu

3D Building Reconstruction from Spaceborne TomoSAR Point Cloud

Author: Yao Sun

Master Thesis

Earth Oriented Space Science and Technology - ESPACE

Supervisors:

Prof. Dr. -Ing. habil. Xiaoxiang Zhu, Dr.-Ing. Muhammad Shahzad

Prof. Dr. Qingyun Du (Wuhan Unerviserty)

15 February, 2016

Technische Universität München

Ingenieurfakultät Bau Geo Umwelt

Signal Processing in Earth Observation

Prof. Dr.-Ing. habil. Xiaoxiang Zhu

3D Building Reconstruction from Spaceborne TomoSAR Point Cloud

Author: Yao Sun

Master Thesis

Earth Oriented Space Science and Technology - ESPACE

Supervisors:

Prof. Dr. -Ing. habil. Xiaoxiang Zhu, Dr.-Ing. Muhammad Shahzad

Prof. Dr. Qingyun Du (Wuhan Unerviserty)

15 February, 2016

I confirm that this master thesis is my own work and I have documented all sources and material used.

Munich, 15 February, 2016   Yao Sun

Acknowledgments

First and foremost, I would like to thank my supervisor, Prof. Xiaoxiang Zhu, for the opportunity to work along with such an amazing team, for the supervision of the thesis, also for setting an excellent example for me to look up to. I want to thank Dr. Muhammad Shazhad, for all the discussions, for the suggestions on research, and for always being there for me. I also want to thank my supervisors for their high expectations, without which this thesis would not have reached the presented scope.

I want to express my thanks Prof. Qingyun Du, who supervised the thesis on part of Wuhan University (WHU), who is loved by all his students, for always supporting and trusting us and for being friend with us, and he inspires me to be a good teacher like him, if I will be one.

Besides my supervisors, I want to thank all the ESPACE lecturers, who inspired me to pursue academic life with their professional knowledge and their kindness. I want to especially thank Prof. Uwe Stilla, Prof. Urs Hugentobler and Prof. Michael Schmidt, who helped me a lot in studies and in making right decisions. I also would like to thank Dr. Marco Limberger, Dr. Janja Avbel and Peter Fischer, who supervised me in ESPACE seminars, helped me in relatively independent research.

Further thanks go to my dear friends, Abhi, Ben, Hunaiz, Jingwei, Muthu, Sasha, Tilo, for the good times we had together, and for more to come. Thank you for all the comforts and encouragements and belief in me, especially when I was confused. The past two and half years would be pale without you. I also want to thank Wei. You are an unexpected surprise. I could not finish this work without your support.

Last but not least, I want to thank my parents. For giving me life, for being my best

friends, for teaching me to be curious, open-minded, for providing me varies opportunities, for supporting my choices, for always listening to me and trusting me, for encouraging me to be who I am. I love you more than anything.

Abstract

Modern synthetic aperture radar satellites (e.g., TerraSAR-X/TanDEM-X and CosmoSky-Med) provides meter resolution data which when processed using advanced interferometric techniques, such as SAR tomography (or TomoSAR), enables generation of 3-D (or even 4-D) point clouds with point density of around 1 million points/km^2. Taking into consideration special characteristics associated to these point clouds e.g., low positioning accuracy (in the order of 1m), high number of outliers, gaps in the data and rich facade information (due to the side looking geometry), the thesis aims to explore for the first time the potential of explicitly modelling the individual roof surfaces to reconstruct 3-D prismatic building models from space. The developed approach is completely data-driven and except for vertical facades assumption, it does not impose any constraint on the shape of building footprint (or to its constituent roof segments) i.e., any arbitrarily shaped building could be reconstructed in 3-D with several roof layers. The workflow is modular and consists of following main modules:

Preprocessing and normalized DSM generation (Extraction of building regions): First a conventionally used ground filtering procedure is adopted to extract ground points from which a digital terrain model (DTM) is generated. Then among non-ground points, first the data gaps are filled using the contextual facade information and later digital surface model (DSM) is generated via nearest neighbor interpolation. Subtraction of the generated DSM with the DTM then gives us the normalized DSM (nDSM) containing the building regions/pixels which is further smoothed using BM3D (Block-matching and 3-D filtering) filtering method.

Segmentation of building roofs : In this module, first a gradient map is generated based on height jumps in the nDSM. Watershed segmentation is then adopted to oversegment

the nDSM into different regions. Subsequently, height constrained merging is employed to refine (i.e., to reduce) the retrieved number of roof segments by taking into account the height difference of two adjacent roof segments.

Reconstruction: Coarse outline of an individual roof segment is then reconstructed using alpha shapes algorithm. Due to varying and lower point density of TomoSAR points, alpha shapes however only define the coarse outline of an individual building which is usually rough and therefore needs to be refined/smoothed (or generalized). To this end, taking into account the average roof polygon complexity (APC), a regularization scheme based on either model fitting (i.e., minimum bounding ellipse/rectangle) or quadtree is adopted to simplify the roof polygons obtained around each segmented (or distinct) roof segment. The simplified roof polygons are then tested for zig-zag line removal using Visvalingam -Whyatt algorithm. Finally, height is associated to each regularized roof segment to obtain the 3-D prismatic model of individual buildings. The proposed approach is illustrated and validated over scenes containing two large buildings in the city of Las Vegas using TomoSAR point clouds generated from a stack of 25 images using Tomo-GENESIS software developed at German Aerospace Center (DLR).

Apart from the above mentioned processing scheme, a complimentary workflow that works directly over unstructured TomoSAR point clouds (i.e., without rasterization to DSM) has also been developed as part of this thesis. The workflow adopts a typical processing chain as employed using conventional airborne laser scanning point clouds and is comprised of RANSAC based recursive plane fitting and computation of adjacent planar intersections. In addition to this, preliminary ideas towards possible future improvements, e.g., joint exploitation of amplitude/intensity together with the 3-D spatial information of each point, aiming to increase the accuracy of reconstructed models from TomoSAR point clouds are also introduced and discussed in this thesis.

Contents

To my mom, 陈陇琴 ***Longqin Chen***

You are my dark matter

1 Introduction

1.1 Motivation

Reconstruction of 3D building models in urban areas has been a hot topic in remote sensing, photogrammetry, compute vision for more than two decades [Gruen et al., 2012] [Rottensteiner et al., 2012]. Numerous research papers have been published in this context that provide different reconstruction methods using variety of different data sources. [Kaartinen et al., 2012].

3D city models are used widely for urban planning [Verma et al., 2006], change detection [Rau and Lin, 2011], in commercial and public sector simulations for environmental research [Brenner, 2001] [Rau and Lin, 2011], telecommunication or solar potential analysis [Brenner, 2001], location based services [Wang, 2013][Brenner, 2005], 3D Geographic Information Systems (GIS) for navigation, driver assistance systems, virtual tourism, and many others [Zhou and Neumann, 2010].

Until a few years ago, large-scaled 3D city models were entirely measured manually [Brenner, 2005], since the reconstruction algorithms could not automatically produce enough detailed building models. The expectations for 3D building models are increasing along with the quality of input data[Haala and Kada, 2010], and the requirements for keeping city models up to date exist since the 3D spatial data is experiencing a high change rate as 2D maps [Brenner, 2001]. Continuing research is driven by the demand for accurate, automatically produced, and detailed 3D city models [Wang, 2013], also by the interest of using data with better quality, or from new sensors or advanced processing techniques, and by the urge to follow new ideas.

The data sources for reconstructing 3D building models are typically optical images (airborne or spaceborne), airborne LiDAR point clouds, terrestrial LiDAR point clouds and close range images. For airborne and terrestrial data, the relatively closer distance of collection guarantees certain details and accuracy. In contrast, satellite images tend to lose these details but allow reconstruction on large/global scale. Satellite images are suitable for reconstruction in large scaled area. Very High Resolution(VHR) satellite imagery nowadays also offers sub-meter resolution. However optical images are "passive", that the energy measured is naturally available, thus the measurements are influenced by illumination condition. While LiDAR system is "active", meaning that the sensor emits radiation toward targets, then detects and measures the reflected radiation from targets.

Synthetic aperture radar (SAR) is a side-looking radar instrument. It is also an "active" sensor. Comparing to optical images, SAR data is independent from day time due to the active emission of signals. Moreover, SAR data is almost independent of weather conditions because of the use of microwaves in radar signals, which is a major advantage comparing to sensors in the visible or infrared spectrum. Very high resolution (VHR) SAR sensors such as TerraSAR-X [Pitz and Miller, 2010] or COSMO-SkyMed [Lombardo, 2004] provide SAR data with spatial resolution up to 1 m, making it possible to extract and reconstruct man-made objects. Figure 1.1 shows an example of amplitude image of TerraSAR-X Data in spotlight mode, the spatial resolution is 1.1m in azimuth and 0.6m in range. As can be seen, buildings and urban structures are clearly visible.

Figure 1.2 shows the imaging geometry of a SAR sensor in a plane defined orthogonally to the azimuth direction. The axis perpendicular to azimuth and range is referred as cross-range direction or elevation. Conventional SAR imagery provides a projection of the 3D object reflection to the 2D azimuth-range plane. Due to the side-looking geometry, this projection introduces typical foreshortening, layover and shadowing problems which complicates the interpretation of SAR images(**??**).

SAR tomography (TomoSAR) is a method to solve this problem by exploiting stack of multiple SAR images acquired from slightly different looking angles. It reconstructs

Figure 1.1 TerraSAR-X amplitude image at las Vegas, USA. Resolution: 1.1m*0.6m (azimuth and range)
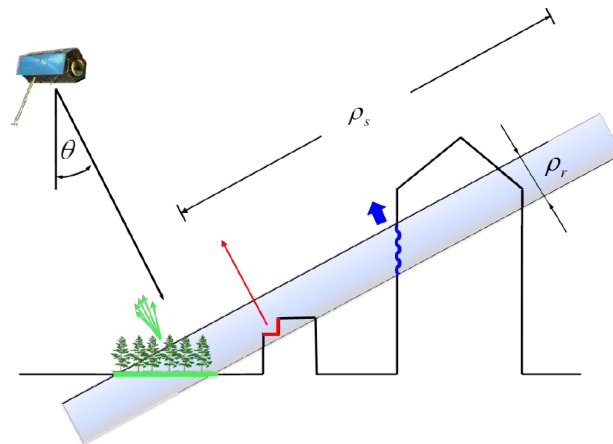

Figure 1.2 Different signal contributions in a VHR SAR image [Zhu and Bamler, 2010].

the reflectivity of the scattering objects along the elevation direction as well as the 3D position of the scatterers. The motions associated with each scatterer can also be retrieved by extending TomoSAR to differential TomoSAR (D-TomoSAR). Figure1.4 shows TomoSAR point clouds generated by DLR's Tomo-GENESIS system [Zhu et al., 2013].

(a) foreshortening (BC mapped to $B'C'$),
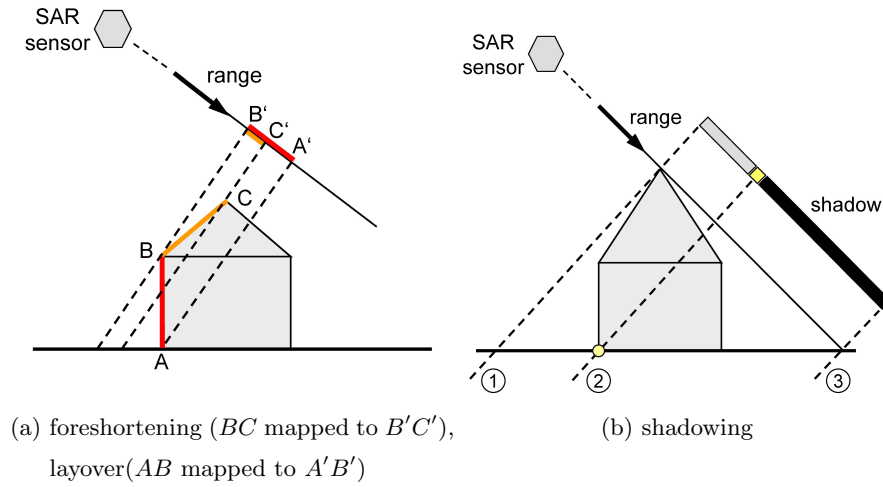layover(AB mapped to $A'B'$)

(b) shadowing

Figure 1.3 Three major geometrical effects in buildings [Auer et al., 2011b]

Modern synthetic aperture radar satellites (e.g., TerraSAR-X/TanDEM-X and CosmoSky-Med) provides meter resolution data which when processed using state-of-the-art To-moSAR methods enables generation of 3-D point clouds with point density of around 1 million points/km^2.

In the context of object reconstruction using TomoSAR point clouds, the difficulties are caused by the following characteristics:

1. Accuracy: 3D positioning accuracy of TomoSAR point clouds reconstructed from spaceborne data is in the order of 1m. The location error of TomoSAR points in elevation is higher than in range and azimuth, typically in one or two orders of magnitude. Ghost scatterers appear as outliers in TomoSAR point clouds [Auer et al., 2011a] that are far away from a realistic 3D position;

2. Points distribution: there are large variations of point density in TomoSAR point clouds. Data gaps exist in TomoSAR point clouds are of influence for objects reconstruction. The side-looking SAR geometry enables rich facade points in To-moSAR point clouds. Temporarily incoherent objects, e.g. vegetation, water body, are not contained in TomoSAR point clouds.

Despite of these considerations, object reconstruction from these TomoSAR point clouds
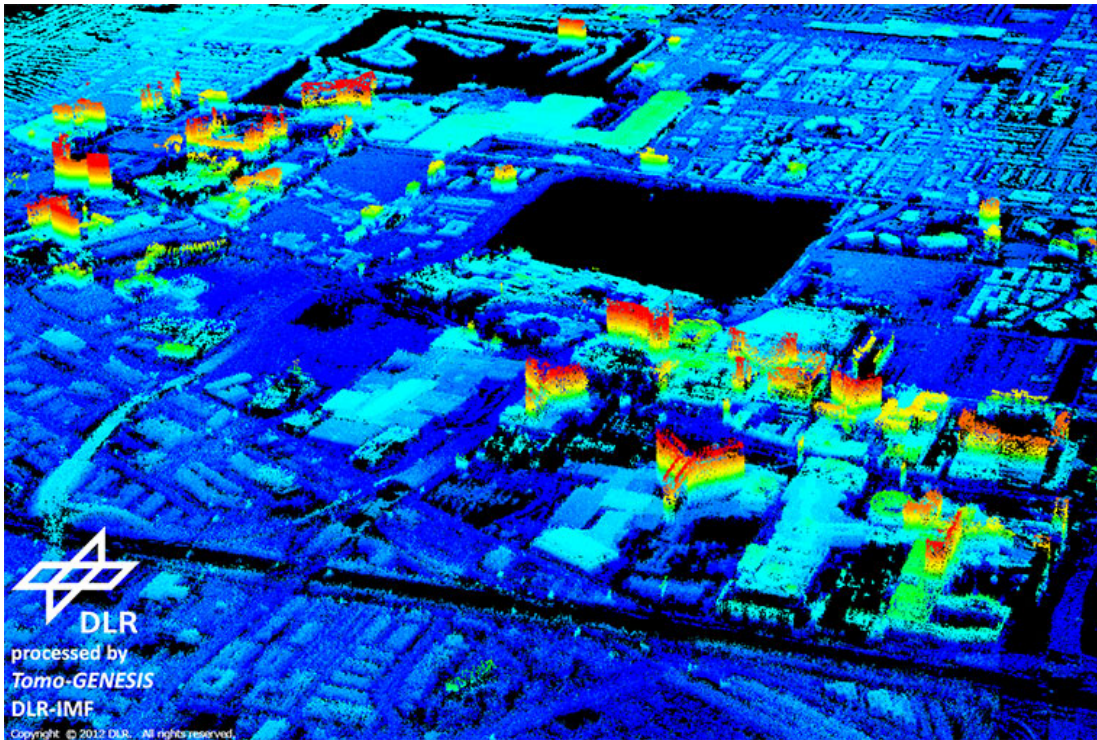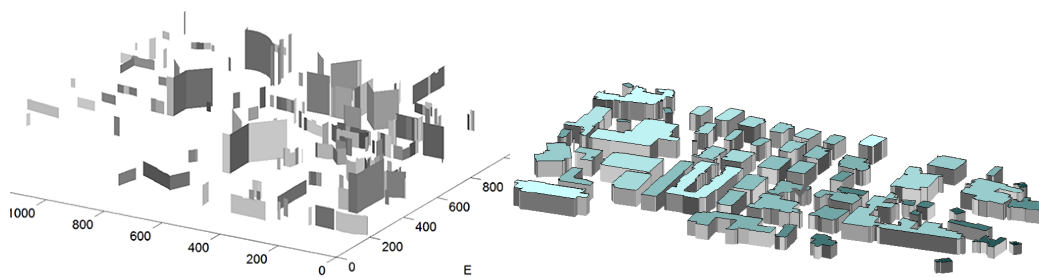
Figure 1.4 TomoSAR point clouds over las Vegas, USA, generated by DLR's Tomo-GENESIS system

(a) Facades from TomoSAR point clouds over las Vegas [Shahzad and Zhu, 2015b]

(b) Building footprints from TomoSAR point over las Vegas [Shahzad and Zhu, 2015a]

Figure 1.5 Previous building reconstruction results from TomoSAR point clouds

can greatly support the reconstruction of dynamic city models that could potentially be used to monitor and visualize the dynamics (i.e., long-term deformation in the mm- and cm-range, e.g. subsidence/uplift caused by earthquakes, bad construction, seasonal

changes etc.) of urban infrastructures in very high level of details.

3-D object modeling/reconstruction from TomoSAR data is still a new field and has not been explored much. Preliminary investigations towards object modeling/reconstruction using spaceborne TomoSAR point clouds have been demonstrated in [Shahzad and Zhu, 2015b] [Zhu and Shahzad, 2014], while TomoSAR point clouds generated over urban areas using airborne SAR data sites have been explored in [D'Hondt et al., 2012]. Although these approaches aim at 3-D building reconstruction but they are limited in handling building with several roof surfaces.

This work is motivated by chances and needs of using TomoSAR point clouds to reconstruct multi-roof buildings.

1.2 Scope and Objectives

Taking into consideration special characteristics associated to these point clouds e.g., low positioning accuracy (in the order of 1m), high number of outliers, gaps in the data and rich facade information (due to the side looking geometry), the thesis aims to explore for the first time the potential of explicitly modelling the individual roof surfaces to reconstruct 3-D prismatic building models from space. The developed approach is completely data-driven and except for vertical facades assumption, it does not impose any constraint on the shape of building footprint (or to its constituent roof segments) i.e., any arbitrarily shaped building could be reconstructed in 3-D with several roof layers.

The goals of this work are as follows:

Goals: explore the possibility of reconstructing multi-roof buildings using TomoSAR point clouds.

Procedure: a general building reconstruction procedure from TomoSAR point clouds shall be proposed and implemented. The building models should be LOD1 models, with right topological relationships between adjacent roof surfaces. The roof surfaces should be connected by step edges or connected to ground surface.

Test and evaluation: The proposed workflow shall be tested and evaluated use different data sites.

Discussion and future work: The proposed workflow and the results shall be discussed. Possible future improvements shall be rised.

1.3 Overview of the Contents

In chapter 2, brief introduction of the necessary background information on the TomoSAR technique, the definition of 3D building models, and the assumptions made about the building shape for the reconstruction procedure are presented.

Different approaches for building reconstruction from point clouds, especially from LiDAR data, are presented in chapter 3. The focus is on finding solutions for the sequential tasks of data-driven building reconstruction, from point clouds, or interpolated DSM.

In chapter 4, two data-driven workflows are proposed, one is based on segmentation of interpolated DSM, while the second directly works 3-D points is directly working on point clouds. After comparing the two workflows, decision is made to use DSM based workflow. Proposed solutions for presegmentation processing, segmentation, regularization are presented in detail.

The designed workflow is then tested and evaluated for two data sets in chapter 5.

In chapter 6, the proposed workflow is discussed, and potential improvements for the proposed segmentation and reconstruction methods are derived. Possible future improvements are also discussed, with preliminary experiments on extracting building masks and facades from SAR amplitude image, and match shapes in the building point clouds.

The proposed workflow, its performance and main characteristics are summarized in chapter 7, and the thesis is concluded by a brief statement on the research field of 3D building reconstruction from spaceborne TomoSAR point clouds.

2 Background Theory

2.1 SAR Tomography

The first concept for 3D imaging of volume scatterers using SAR tomography (To-moSAR) is presented in [Reigber and Moreira, 2000]. The basic principle is shown in Figure 2.1. A stack of SAR images is exploited in order to recover the reflectivity function in elevation direction. Red spots: orbits forming the synthetic aperture in elevation direction S. The enlarged red spot marks the master orbit. Yellow spots: scatterers to be detected. Azimuth direction: x. Range direction: r. SAR data captured in multi-baselines(red spots), the corresponding orbit positions of the SAR sensor form a synthetic aperture in elevation directions. The spatial baseline between orbit positions is required for providing angular information in the cross-range plane. Thereby, the elevation coordinates of two scatterers(yellow points), can be estimated with respect to a master (enlarged red spot). Thus, the layover problem can be resolved for each SAR image pixel.

For one image pixel, each SAR acquisition provides a spectrum sample of the reflectivity function in elevation.

$$g_n = \int_{\Delta S} \gamma(s) exp(-j2\pi\xi_n s)ds, n = 1, ..., N \qquad (2.1)$$

where g_n is the signal measured for an image pixel during pass n, ΔS is the extent of the imaged object in elevation, $\gamma(s)$ is the reflectivity function representing the distribution of backscattered intensities in elevation s, and ξ_n is the spatial frequency in elevation depending on the sensor position with respect to the master.
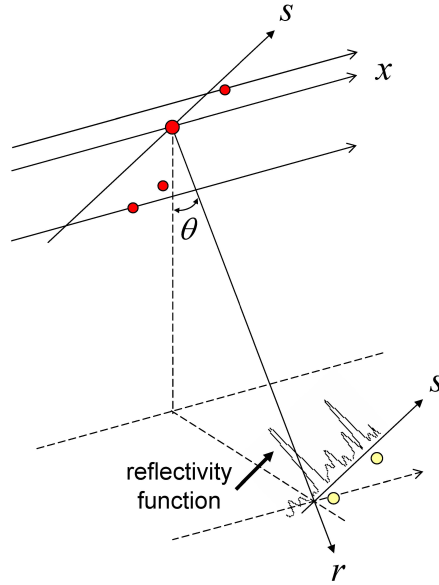
Figure 2.1 Synthetic aperture in elevation direction.[Auer et al., 2011b]

The basic aim of SAR tomography is to invert the equation to derive the intensity and position of signal responses in elevation. Different inversion methods are reported in the literature and may be grouped in parametric models and non-parametric models. In theory, parametric models, e.g. non-linear least square adjustment (NLS), provide the best solution. However, the definition of functional models requires a priori knowledge about the number of scatterer responses integrated into each resolution cell. By using penalized likelihood criteria, the maxima of $\gamma(s)$ was found in [Zhu et al., 2008].

In case of non-parametric models, limitations in tomographic processing occur due to the short length of the synthetic aperture in elevation as well as due to the low number and irregular distribution of samples. Compressive sensing has proven to be reasonable for overcoming these limits [Zhu and Bamler, 2010]. Moreover, tomographic methods have been extended to 4D-space including the velocity of scatterers [Zhu and Bamler, 2010]. Hence, the topography of urban areas can be provided in case of moving objects and object deformation can be monitored.

Figure 2.2 Level of Detail of 3D building model. [Biljecki et al., 2014]

2.2 3D Building Model: Definitions and Assumptions

3D building models are vector models, which are polygonal meshes representing the building with a significantly reduced number of data points than the original point clouds [Wang, 2013].

2.2.1 Level of detail

When speaking of building models, the first thing to clarify is how detailed the models should be. Reconstructed building models have different level of detail (LOD), differ from simple prismatic model (LOD1) with flat roof surfaces to detailed model with overhangs or balconies (LOD3), even with the interior of buildings (LOD4) (Figure 2.2). The LOD should fit user's requirements and acceptance criteria of 3D building models, as well as the quality of used data.

Though building models and reconstruction methods differ, for most authors who use high resolution images or LiDAR data, their definition agrees to the definition of LOD2 of the official OGC standard City Geography Markup Language (CityGML)[Gröger et al., 2008], an information model intending a standardized "representation, storage, and exchange of virtual 3D city and landscape models" [Krüger and Kolbe, 2012]. LOD2 includes detailed roof structures without roof overhangs, balconies, and consider buildings with footprint no smaller that 4*4 m^2. In this work, considering the accuracy of input data, we aim at reconstruct LOD1 models.

2.2.2 General Assumptions

Buildings are commonly assumed to be composed of planar shapes and the facades are vertical [Dorninger and Pfeifer, 2008] [Kada and Wichmann, 2012]. Roof segments are connected by intersection edges or vertical step edges to each other or to ground; overhangs and small building parts such as small dormers and chimneys are neglected. Some authors assume that all roof segments to be connected by step edges to each other [Poullis and You, 2009]. Rectilinear assumption is also accepeted by many authors that, the main directions of roof segments are constant with the main direction of the building. The building's outer boundary is often assumed to be a polygon consisting only of perpendicular and parallel edges [Haala et al., 1998] [Maas and Vosselman, 1999] [Matei et al., 2008] [Rau and Lin, 2011].

3 State of the Art

This chapter covers the relevant methods for reconstruction of 3D building models from TomoSAR data, mainly the methods dealing with LiDAR point clouds. First, the focus of 3D building reconstruction is clarified. Then the categories of the existing methods are briefly presented, i.e., 1) data driven or model driven methods, based on the start point of the workflow, and 2) point clouds based or DSM segmentation based methods, main data structure of the workflow. A literature survey is then given on building reconstruction approaches, in aforementioned categories.

3.1 Focus of 3D Building Reconstruction

Building models contain roofs (the uppermost part), facades (walls), and footprints. Building reconstruction workflows largely depend on nature of used data, particularly angle of view, i.e., on which side the data is collected. Methods emphasize roof reconstruction usually have data source collected from top view, which have much more information on roofs, e.g., ariel images, satellite images, airborne LiDAR data. While terrestrial data collection gives more information on facades, from pedestrian viewpoints, thus need more facade reconstruction. Oblique images provide information on both roofs and facades, while are typically used in providing elevation data and facade texturing. TomoSAR point clouds gives rich facade information due to side-looking geometry, while multi-view is needed to obtain the full structure of individual buildings.

Roofs are the focus in 3D building model reconstruction. "Differentiated roof structures and thematically differentiated surfaces" [Gröger et al., 2008] are the characters that distinguish LOD2 building model from LOD1 blocks model.

Footprints are usually assumed available or can be automatically extracted beforehand. As most research aiming at reconstruct LOD2 building models, which means no roof overhangs and balconies, footprints are considered as outline of the building from top view and often are not modeled separately but take outline of roofs or taken from 2D ground plans [Haala et al., 1998] [Durupt and Taillandier, 2006] [Schwalbe et al., 2005] [Kada and McKinley, 2009].

Vertical facade is a common assumption in 3D building reconstruction. So that authors dealing with LOD2 building model reconstruction usually focus on roof structures rather than reconstructing facades, and their 3D building shapes are generated by extruding the reconstructed roof shapes.

The methods presented mainly focuses on roof reconstruction. Facades reconstruction from terrestrial data collection for more detailed models and facade reconstruction from TomoSAR point clouds are presented.

3.2 Categories of existing methods

In general, based on the start point, building reconstruction approaches can be classified to two major categories: Data driven or non-parametric approach and Model driven or parametric approach. While based on the data structure to be segmented, the existing approaches can be classified into two types: point clouds based, and DSM(Digital Surface Model) based.

3.2.1 Data driven or Model driven

Data driven approach starts from data, thus are more flexible and suitable for different datasets in modeling complex roof shapes than model driven approache, thus data driven approaches are very popular in recent research [Rottensteiner and Briese, 2002] [Wang, 2013]. Building models from data driven methods fit most accurately the input data and are therefore suited for applications where the focus is put on accuracy and knowledge about small roof parts, such as for simulations or augmented reality. How-

ever, Point density varies in one dataset, thus details to be reconstructed vary, and how to bridge gap from features found in the laser data to final building models is the main problem. Thus data driven approaches require a high regularization effort.

Model driven approach selects best fitting parametric models from a prepared library of models and estimated the corresponding model parameters according to the best fit to the LiDAR data. They are usually too generalized to be able to reconstruct complex building shapes and cannot flexibly represent any roof shape. Since a great number of buildings in rural and suburban areas are rather simple and can be approximated by rectangular footprints and parameterized standard roof shapes [Haala and Kada, 2010], many approaches decompose the building point clouds into different parts and fit a model to each part [Tarsha-Kurdi et al., 2008]. Those models are then connected and corrected. The procedure is robust, effective and fast, and has small minimal regularization effort.

3.2.2 Point clouds based or DSM based segmentation

The first task in building reconstruction is to separate buildings from other objects in the whole data (presegmentation). Then the data is segmented to surfaces. Some authors directly work on point clouds, while some authors use interpolated grided Digital Surface Models (DSM) for building reconstruction.

Point clouds based

Most authors work directly on point clouds. The surfaces are usually extracted by plane fitting or rsgion growing technique. Then topological relations are build and the extracted surfaces are regularized. In model driven approach, the point clouds first are decomposed to small parts, to fit the basic shapes in library.

DSM based

Another widely followed approach is to reconstruct roofs by DSM simplification. The idea is, that the buildings are contained in highly detailed, meshed digital surface models (DSM) and that they only need to be simplified to the right abstraction level and

extracted if necessary [Haala and Kada, 2010], while above reconstruction approaches described, which construct building models from scratch to best fit the given elevation data.

3.3 Presegmentation

Presegmentation typically classifies the point cloud into building points and other points, mainly terrain and vegetation in LiDAR point clouds. If 2D building footprints are available beforehand, building point clouds can be directly extracted [Rau and Lin, 2011]. A popular way is ground filtering method, in which a Digital Terrain Model (DTM) is produced by morphological filter operations [Morgan and Tempfli, 2000] [Zhang et al., 2003] [Pingel et al., 2013], then a height threshold is set on the DTM. Another approach is to fit planes to points clouds, and clustering points. The largest cluster is assumed to be ground [Verma et al., 2006]. [Lafarge and Mallet, 2012] defineed expectation values for buildings, vegetation, ground and clutter by combining different covariance-based measures and height information by energy optimization. [Dorninger and Pfeifer, 2008] extracted all planar regions of the scene using region growing method in feature space and group the extracted points to buildings with a mean-shift algorithm.

3.4 Building roofs reconstruction

Building roofs are reconstruction from elevation data which can originate from various sources like LiDAR or image matching [Haala and Kada, 2010]. A typical process that reconstruction of building roofs is based on a segmentation process of the elevation data.

3.4.1 Data driven approach

3.4.1.1 Point cloud based approach

Region based methods and plane fitting methods are two main segmentation methods. Plane fitting methods fit surface planes to the point clouds. Region based methods group points according to their proximity in feature space. Region-based segmentation requires subsequent surface fitting processes.

1. **Region based methods**: Region-growing methods are used widely in segmentation. It start from a seed, based on some predefined criteria, each unclassified neighbor to the region's points is added to the region [Rottensteiner and Briese, 2002] [Elberink and Vosselman, 2009] [Verma et al., 2006] [Dorninger and Pfeifer, 2008].

   Region growing criteria decide whether a point should be added to a region or not from predefined thresholds on similarities. The most used criterion is the point's local plane parameters to the average region's plane parameters. [Verma et al., 2006].

   Region growing seeds can be determined by choosing randomly from the point clouds. [Alharthy and Bethel, 2002] choose a seed that located at centric within a roof segment. [Awrangjeb and Fraser, 2014] select the midpoints of the building boundary's edges as seed. Others estimate flatness using a covariance based local curvature measure, or according to the local neighborhood's RMSE from the local plane [Alharthy and Bethel, 2002].

2. **Plane fitting methods**: Planes can be fitted to a point clouds by maximizing the number of inliers, i.e. points whose orthogonal distances to the estimated surface are below a certain threshold. RANSAC and Hough Transform are popular method for plane fitting [Sohn et al., 2008] [Tarsha-Kurdi et al., 2008] [Ameri and Fritsch, 2000] [Brenner, 2000] [Vosselman et al., 2001] [Sohn et al., 2008] [Vosselman et al., 2004].

   RANSAC (RANdom SAmple Consensus) is an iterative model fitting procedure where in each iteration, a model is created from a randomly selecting a necessary number of samples from the data set. It maximizes the inliers and considers outliers

as noise. RANSAC tries many planes and then returns to the one that include maximum points as inliers. The quality of the random model is determined by counting the number inlier points using a distance threshold. In case the quality measure is better than in the previous iteration, the model is kept as the currently best estimate.

Hough Transform is a feature extraction technique in digital image processing. The classic Hough Transform transforms euclidean space coordinates of points to curves in hough space, and the transformed curves of points on same line in euclidean space intersect at one point in hough space . The more curves passing through the intersct point, the longer the original line segment is. 3D Hough Transform is an extension of classic Hough Transform, and has been used for the detection of planes [Vosselman et al., 2001] [Sohn et al., 2008] [Vosselman et al., 2004].

After segmentation, boundaries of roofs need to be traced and regularized. Connected boundary line segments can be determined from a triangulation of the segment points. Edges which belong to only one triangle are defined as boundary edges [Maas and Vosselman, 1999][Matei et al., 2008]. Zhou and Neumann [2008] define boundaries by tracing the closest LiDAR points to those edges. Rottensteiner [2003] define separation boundary lines between adjacent segments from the Delaunay triangulation: Differently segmented points connected by triangulation edges are boundary points, and the corresponding Voronoi edges form the boundary. Dorninger and Pfeifer [2008], Kada and Wichmann [2012] and Sampath and Shan [2007] use a modified convex hull approach called alpha shapes, in which each next boundary vertix is determined only from the local neighborhood of the previous vertex. If the local neighborhood is determined by a fixed radius, alpha shapes produce only satisfactory results if the point density is regular. Therefore, Sampath and Shan [2007] define the neighborhood with a rectangle whose extents and orientation depend on the along-track and across-track LiDAR sampling characterisitics. [Wang and Shan, 2009] identify unconnected boundary points by creating the convex hull of each point's local neighborhood. If the point is a vertex of this convex hull, it is chosen as a building boundary vertex. Lafarge

and Mallet [2012] determine each boundary point based on its distance to the line fitted through its neighborhood.

Boundary tracing delivers irregularly shaped polygons with a many polygon vertices. Regularization procedures simplify polygons by reducing vertices number. The Douglas-Peucker algorithm is often used for regularization. Line segments are regularized by connecting the farthest vertices in the irregular polygon, such that the distances of all skipped polygon points from the new line are below a threshold [Sampath and Shan, 2007][Sohn et al., 2012].

3.4.1.2 DSM based approach

Most of related researches focuses on the detection of building roof patches. Usually ground filtering methods are used to separate ground points and above ground points, then and the normalized DSM (nDSM) is obtained by subtracting the Digital Elevation Model (DEM) from the DSM. Based on the nDSM, buildings are detected according to the surface normals, elevation textures or shape structures.

[Rottensteiner and Briese, 2003] extracted roof races using seed regions and region growing in a regularized DSM. [Forlani et al., 2006] first use region growing on DSM to obtain a raw segmentation result, then apply a second segmentation based on gradient orientation analysis, which defines roof slopes. The two segmentation results are then overlaid to obtain the final roof segments. [Galvanin and Poz, 2012] construct DSM by rasterizing the LiDAR data. They first detect the non-ground DSM grids, recognize the object outlines, and then detect the building roof boundaries using the Markov random field approach. [Chen et al., 2014] used multiscale DSM grids instead of same sized DSM grids. In the large-scale grid, building seed regions are obtained, while in the small-scale grid, to detect the detailed features of building roofs with complicated top structures, a high-resolution depth image is generated by an iterative morphological interpolation using gradually increasing scales, and then segmented. Based on the building seed regions, detailed roof features are detected for each building and 3-D building roof models are then reconstructed according to the elevation of these features.

3.4.2 Model driven approach

Model driven approach has a library of models, usually too generalized to be able to reconstruct complex building shapes. If the library contains too many complex models, the flexibility and algorithm efficiency will be reduced. Thus most model driven reconstruction approaches decompose the building point clouds into different regions, and fit building parts with simple parametric models and then connect them to be a whole building model. Model driven approached usually contain decomposition of point clouds, building library of parametric models and model selection and parameter fitting.

3.4.2.1 Point cloud based approach

The model library need to contain the most common roof shapes in a parametric description, such as flat roofs, shed roofs, gabled roofs and hipped roofs [Kada and McKinley, 2009] [Haala and Brenner, 1999][Vosselman et al., 2001]. The library can also contain models for corners where basic roof shapes connect in a right angle or in a T-shape Kada and McKinley [Kada and McKinley, 2009]. [You et al., 2003] compose their models from standard computer graphic shapes such as planes, cubes, polyhedra, cylinders, spheres and ellipsoids.

The building point clouds can be segmented using roof segmentation methods. Footprint decomposition is achieved by analyzing the building boundary's line segments [Kada and McKinley, 2009, Haala et al., 1998]. You et al. [2003] require a user-input to determine the footprint regions and to select the appropriate model; only the model parameters are determined automatically.

Kada and McKinley [2009] select the model by computing the percentage of points whose local normal is similar to the predefined library model. Verma et al. [2006] determine model parameters by a RANSAC fitting procedure. For determining footprint extensions, building orientation and roof type, Maas and Vosselman [1999] compute height-weighted invariant moments from the point clouds, as well as from rasterized versions of the parametric models.

3.4.2.2 DSM based approach

The model driven approaches of DSM based segmentation first decompose the DSM to simple shaped cells, then for each cell a primitive is fitted. Then reconstruction results for all the cells are connected and combined to achieve final models.

[Brenner and Haala, 1998] segmented buildings into basic primitives based on the given ground plan and then fitted these primitives to the DSM data. Footprint decomposition is achieved by using previously detected step edges [Vosselman et al., 2001]. [Haala et al., 1998] estimate the model parameters by least squares minimization of the DEM pixel's vertical distances to the model.

Most of the DSM based approaches are studies in early years, when the quality of LiDAR data were good enough to apply region growing. In recent year, reconstrucion of building models from high resolution DSM is under studying [Sirmacek et al., 2012] [Arefi and Reinartz, 2013] .

[Arefi and Reinartz, 2013] proposed an approach based on the analysis of the 3D points of DSM from satellite images in a 2D projection plane. Parametric models are generated through single ridgeline reconstruction and subsequent merging of all ridgelines for a building. The edge information is extracted from the orthorectified image.

3.4.3 Other approaches

Some methods integrate data driven and model driven approaches. Oude Elberink and Vosselman proposed a target based graph matching method [Elberink and Vosselman, 2009], in which neighboring roof segments are represented as vertices connected by edges. These edges are labeled according to the relations of the segments' plane normals to each other. It has a library of target graphs (models), while constructed roof topology graphs after segmentation and labeling are to be matched. Geometric reconstruction then follows.

Zhou and Neumann proposed a 2.5D Dual Contouring method [Zhou and Neumann, 2010], which roots in computer vision background. It first sampling point clouds over a uniform

2D grid, then create geometry adaptively based on compute hyper-points by minimizing a 2.5D QEF in each quadtree cell. Then surface polygons and boundary polygons are generated.

3.5 Building facades reconstruction

In addition to the available footprint and roof shape, facades are the most important features that reflect the style and dimension of buildings. Airborne data collection mainly provides the outline and roof shape of buildings, while building facades information is limited, usually reconstructed models give planar facades and distinctive roof structures according to level of detail 2 within the OGC standard CityGML [Krüger and Kolbe, 2012]. Building facades can be modeled separately from terrestrial data collection or SAR images.

3.5.1 Facade models from TomoSAR

Zhu and Shahzad propose an approach for building facade detection and reconstruction from TomoSAR point clouds [Zhu and Shahzad, 2014]. First, the building facade regions are extractede by thresholding the Point density map. Then, the extracted points are segmented to different facades. After orientation analysis inside each segmented cluster, and 2-step k-means clustering is performed to segment and refine clusters. In reconstruction, the facade surfaces are decided to be falt or curved by analyzing derivatives of the local orientation angle. Polynomials are used to model the footprints in x-y plane. First-order and second-order(p =2) polynomials are used to reprsent footprints of flat facades and curved facades. Then the overall shape of the building footprint is described by identifying adjacent facade pairs and determining the intersection of their facade surfaces. Using TomoSAR point cloud, facades reconstruction over large area is presented[Shahzad and Zhu, 2015b]

3.5.2 Facade models from terrestrial data collection

Terrestrial data collection usually refers to terrestrial laser data and close range images, which provide complement dataset for reconstructing more detailed building models. They are used in applications that require for representing building facades with elements on them, such as windows and doors.

Many works make use of terrestrial laser points and images together to reconstruct facades, as laser points cannot texture facades, while image understanding for automatic facade reconstruction is not so easy[Pu and Vosselman, 2009] [Becker and Haala, 2009] [Frueh et al., 2005]. If larger areas need to be covered, ground-based mobile mapping systems with integrated terrestrial laser scanners are used to provide dense 3D point coverage at facades and the neighboring architecture.

4 Proposed Workflow

This chapter presents two data driven workflows that are designed for building recon-strcution from TomoSAR point clouds. After preprocessing, the first workflow works over interpolated DSM from point clouds and performs segmentation of distinct flat roof surfaces. While the second workflow works directly on point clouds and in contrast uti-lize plane fitting and clustering algorithms to perform segmentation of individual roofs which are subsequently used in generating 3-D building model.

The two proposed workflows are then compared. The DSM segmentation based workflow is considered to be more suitable for TomoSAR point clouds data.

4.1 Workflow based on DSM segmention

Inputs of the proposed workflow (Figure 4.1) are TomoSAR point clouds. Building points are extracted and modified in preprocessing step. Normalized DSM (nDSM) is then obtained and denoised. Segmentation of nDSM gives roof segments. In reconstruction step, the outlines of roof boundaries are reconstructed and regularized, then the height
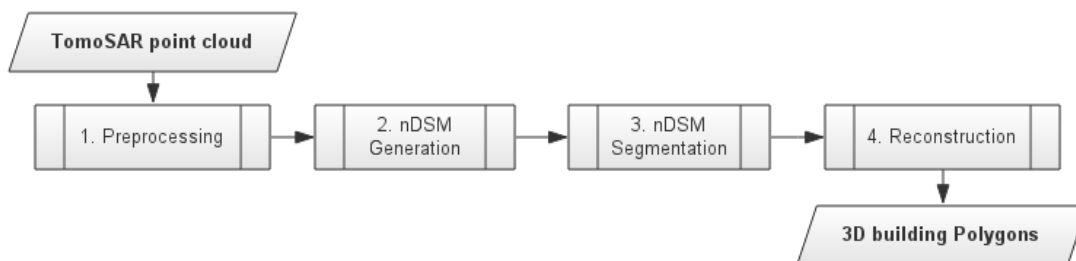
Figure 4.1 Overview of workflow (DSM based)
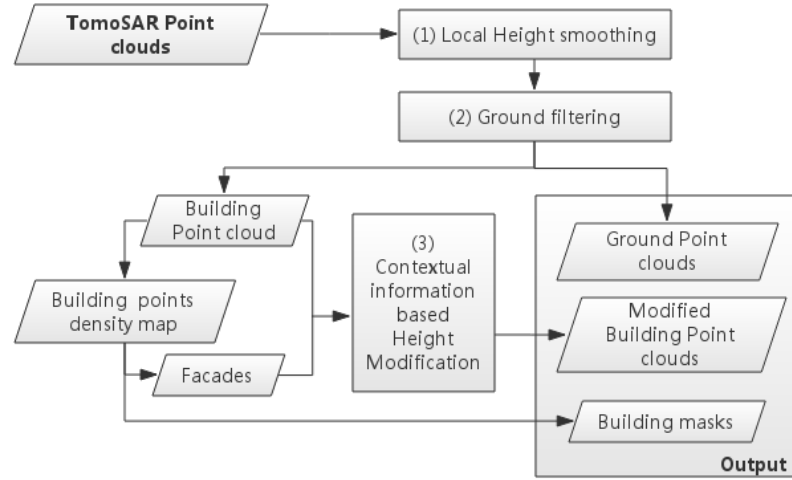
Figure 4.2 Overview of preprocessing workflow

is associated with each roof segment to build 3D building polygons.

4.1.1 Preprocessing

In order to obtain better interpolated DSM, before DSM generation, preprocessing is performed to extract building points, and to modify height of points. Firstly, L_G, the length of smallest grid unit in DSM is defined, and the study area is covered by grids G_s, so that all points fall in grids according to their X and Y coordinates. Preprocessing contains three steps: 1) Local height smoothing; 2) Ground filtering; and 3) Contextual information based height modification.

4.1.1.1 Local Height Smoothing

This step is designed to decrease the influence of facade points and non-surface point in interpolating DSM, without losing point density.

The step is performed "locally", in each grid in G_s. For each grid g, all the inside n points are $P(x_i, y_i, z_i)$, $i = 1...n$. Take first m highest of $P(x_i, y_i, z_i)$, and compute the

(a) before, top view   (b) side view   (c) after, top view   (d) side view
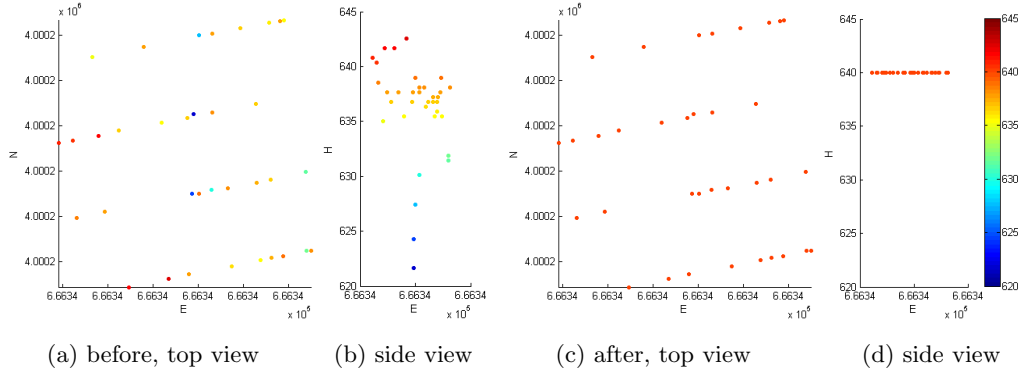
Figure 4.3 Local Height Reassigning example (in 1 grid). $L_G = 3m$, Height is color coded.

mean height of them, $mean(z_m)$. z_i is then assigned to $mean(z_m)$. An example is shown in Figure 4.3: (a)(b) show all original points in one grid; (c)(d) show points after height reassigning in the same grid.

4.1.1.2 Ground filtering

Ground filtering is used to separate ground points and above ground points (object points). The Simple Morphological Filter (SMRF) proposed by [Pingel et al., 2013] is used.

Mathematical morphology contains operations based on set theory to extract features from an image [Haralick et al., 1987]. Opening operation is an erosion of the image followed by a dilation. In the context of point clouds, for a point $p(x, y, z)$, the dilation of elevation z at (x, y) is defined as [Zhang et al., 2003]:

$$d_p = \max_{(x_p, y_p) \in w} (z_p).$$ (4.1)

where points (x_p, y_p, z_p) represent p's neighbors(coordinates) within a window w. The dilation output is the maximum elevation value in the neighborhood of p. While erosion of elevation z at (x, y) is defined as [Zhang et al., 2003]:

$$e_p = \min_{(x_p, y_p) \in w} (z_p).$$ (4.2)

(a) all points   (b) above ground points   (c) ground points
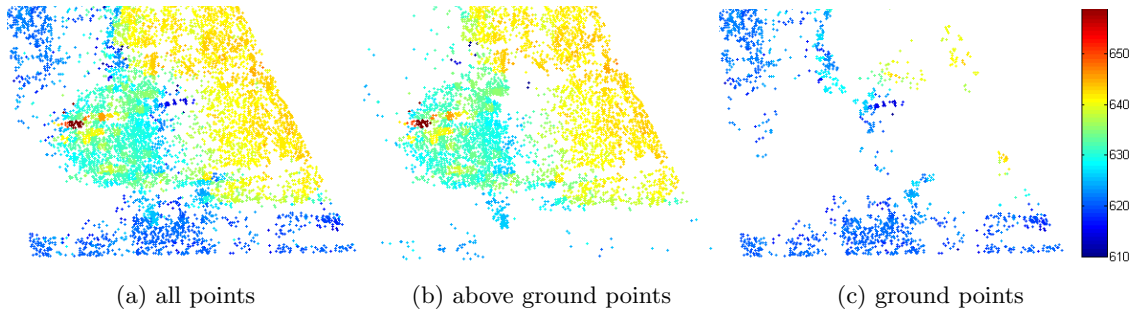
Figure 4.4 Ground filtering example (2D view) . Height is color coded.

In SMRF, after defining the grids, a Minimum Surface (ZI_{min}) is generated by interpolating the lowest elevation of all points within each grid and inpainting empty grids. Then, a progressive morphological filter is applied to ZI_{min}, so that ZI_{min} is iteratively smoothed by an image opening procedure to an opened DEM. At each iteration, all ZI_{min} pixels whose distance to the opened DEM is larger than an elevation threshold h are set to the values of the opened DEM. h increases at each iteration, defined as the value of the maximum slope tolerance parameter multiplied by the product of the window radius and the grid size. The opening filter's window size increased at each iteration, and the procedure stops when the filter is larger than the maximum expected building size. The DTM is created from an interpolation of all points whose difference between ZI_{min} and the smoothed surface is smaller than h. The original TomoSAR points then either belongs to ground or above ground object based on their relationship to the DTM. Figure 4.4 shows an example of ground filtering results.

4.1.1.3 Contextual information based height modification

One characteristic of TomoSAR point clouds is: rich points on the facades are available which are facing the sensor (i.e., not occluded). This information is utilized and point density is computed to extract facade regions which are then utilized as contextual information to further smooth the height estimates prior to DSM generation. This step is needed due to relatively low positioning accuracy of TomoSAR points, especially around facade points.
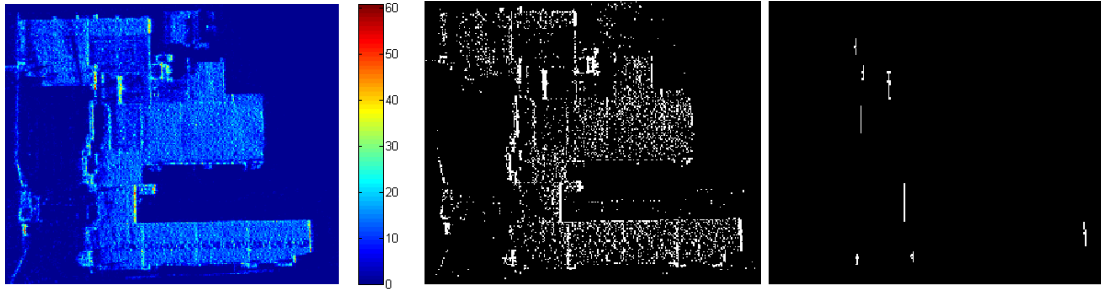
(a) PD map, point density is color codes

(b) Thresholding at PD=16

(c) Facades found after morphological operations

Figure 4.5 Thresholding point density map at example area

Find Facades using Point Density Map

Point Density means the number of points in a certain sized area, which is simply a 2D histogram of number of Points, when projected to the defined girds. An example is shown in Figure 4.5(a).

Due to the side-looking SAR geometry, the TomoSAR point clouds on vertical facades has higher Point density(PD) comparing to non-facade regions, because of the existence of strong corner reflectors, e.g., window frames on the building facades. Thus, thresholding the PD map will identify possible pixels representing facades. Morphological operations are then applied to refine the identified areas. One assumption is the "narrow" facades, meaning that the ratio of length and width of facade should be large. This assumption is applied to remove segments.

Thresholding converts the PD map into a binary image. Pixels whose value no smaller than the defined threshold T_{PD} are set to foreground, while the rest are set to background. Figure 4.5 shows an example of the extracted facade regions. Figure 4.5(a) is the PD map of above ground points in the area. Figure 4.5(b) is the thresholding result. In Figure 4.5(c), facades are found by applying morphological operations and the "narrow" facades assumption.

Contextual information of facades

In the desired building models, a basic assumption is that, one building facade connects two or more roof surfaces, or connect roof surface and footprint, meaning that the height difference $\Delta h > H_{min}$ exists between the two sides of the connecting facade. H_{min} is the minimal height difference to separate two roof surfaces. If the roofs are flat, the height of roof on the higher side equals to the height on top of the connecting facade (Figure 4.6(b)). If $\Delta h < H_{min}$, Δh is too small to be considered (Figure 4.6(c)), the facade should be neglected, since it is a vertical structure which does not belong to building, e.g., billboard.

Height modification

In point clouds, for a facade F and its two side zones Z_1 and Z_2 (Figure 4.6(a)), denote the higher one as Z_1 and the lower one as Z_2. Z_F is area of F. Points inside Z_1, Z_2, Z_F are: $P_1(x_i, y_i, z_i)$, $i = 1...n$, $P_2(x_j, y_j, z_j)$, $j = 1...m$, and $P_F(x_f, y_f, z_f)$. The mean height of all points inside Z_1, Z_2 are h_1 and h_2.

The height modification is then performed for the following two cases:

1. $\Delta h >= H_{min}$:

   1) Take first r highest points inside Z_F and compute the mean height h_r. For $P_1(x_i, y_i, z_i)$ inside Z_1 and $P_F(x_f, y_f, z_f)$ inside Z_F, reassigning z_i to h_r.

   2) For $P_2(x_j, y_j, z_j)$ inside Z_2, take first p lowest points and compute the mean height $mean(z_p)$, reassigning z_j to $mean(z_p)$.

2. $\Delta h < H_{min}$:

   Put $P_1(x_i, y_i, z_i)$ and $P_2(x_j, y_j, z_j)$ together, compute the mean height of all $P_1(x_i, y_i, z_i)$ and $P_2(x_j, y_j, z_j)$, h_{ij}. Reassigning z_i, z_j and z_f to h_{ij}.

By previous steps, most surrounding-facade points are modified, however the problem arises when there are limited number of roof points which make the comparison an incorrect estimate. If the building mask of the higher part of the building is available, the adjacency relationship of building mask and facade can be used as follows:

(1) Create a buffer cell around the facade. Select all points PT_{fb} inside the buffer cell.
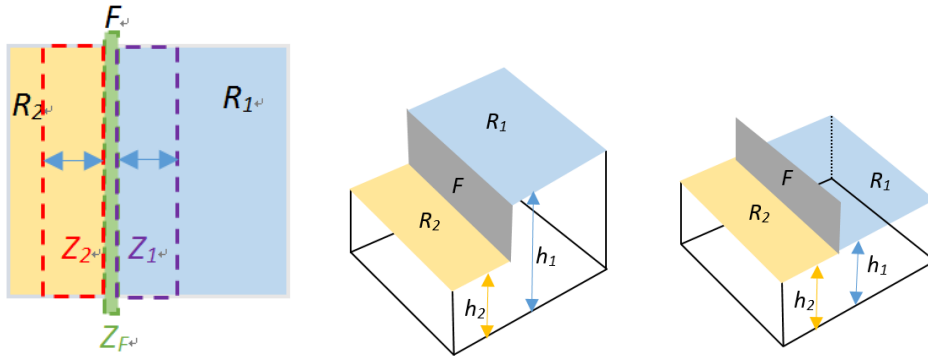
(a) top view of (a) and (b). Z_1 (b) F connects R_1 and R_2. (c) Δh is too small. F should and Z_2 are F's side zones, $\Delta h = h_1 - h_2$ be deleted. Z_F is area of F.

Figure 4.6 Contextual information of facades. R_1 and R_2: roofs. F: facade in between.

(2) In PT_{fb}, select points inside the out contour of building mask, set their heights to the local maximum value of surrounding points. Set the heights of remain points in PT_{fb} to the local minimum.

4.1.2 Generation normalized DSM of building area

Subsequent to incorporating contextual facade information to further smooth the height values, the next step is to interpolate the smoothed point cloud to generate DSM. The workflow is shown in Figure 4.7.

4.1.2.1 DSM interpolation

First, modified building point clouds and ground point clouds are put together to "Modified point clouds", and it is interpolated into DSM. The ground point clouds alone is used to interpolating DTM. The normalized DSM (nDSM) is obtained by subtracting DTM from DSM (Figure 4.8).

Building masks is then introduced to bound building areas in the nDSM. In this work, building masks is obtained by height thresholding of the nDSM.
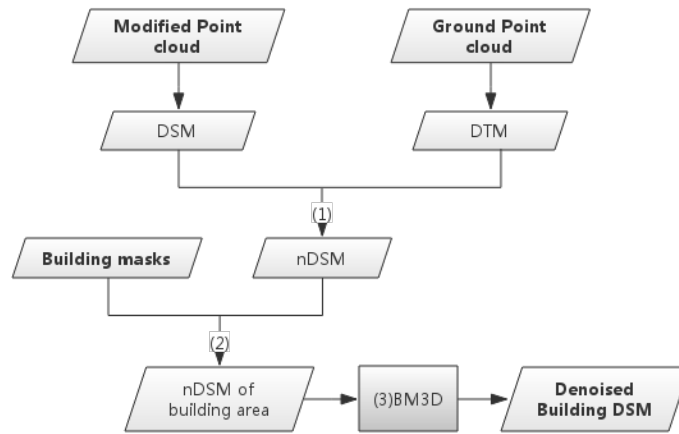
Figure 4.7 Workflow of DSM generation
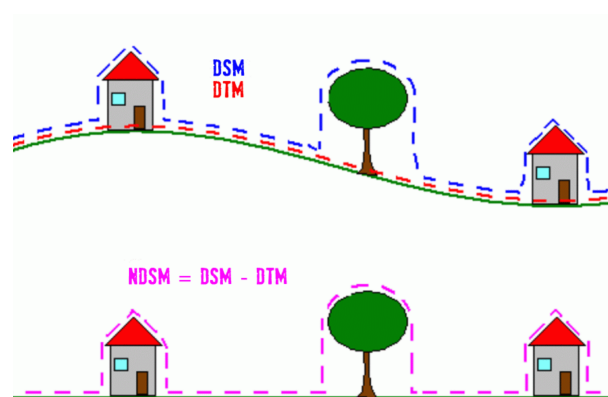
Figure 4.8 The principle for the generation of nDSM. [Atlas, 2014]

4.1.2.2 Building masks from point density map

Because of temporarily incoherent objects cannot be reconstructed from multi-pass spaceborne SAR image stacks, the TomoSAR point clouds on building regions generally have higher point density(PD) than ground regions, which usually have vegetation and water bodys inside.

The building masks are extracted from nDSM by thresholding the height, and removing small segments.

There are data gaps in point clouds. The can be located also by thresholding. Some data gaps have relatively regular shapes than others, and they are adjacent to facades. The reason behind is that occlusions are the causes of the facades-adjacent data gaps, while the coherent imaging nature cased the other type. This knowledge is employed to modify building masks and generate nDSM at building areas, in following 3 cases:

1. If the data gap is not adjacent to any facade, it should be caused by the coherent imaging nature. It is to be interpolated by surrounding points;

2. If the data gap is adjacent to "facade inside building masks", it is caused by occlusions of the facade. It should be included into building masks, and the height in nDSM at the data gap should be assigned to the heights of surrounding above-grounf points.

3. If the data gaps are adjacent to "facades at boundary of building masks", it is caused by occlusions of the facade. It should be outside of the building, so the height in nDSM at the data gap should be assigned to ground value.

The modified building masks are obtained. An example is shown in Figure 4.9.

4.1.2.3 DSM denoising

Prior to extraction and segmentation of different roof surfaces, it is necessary to further smooth/denoise the generated nDSM. To this end, BM3D is applied.
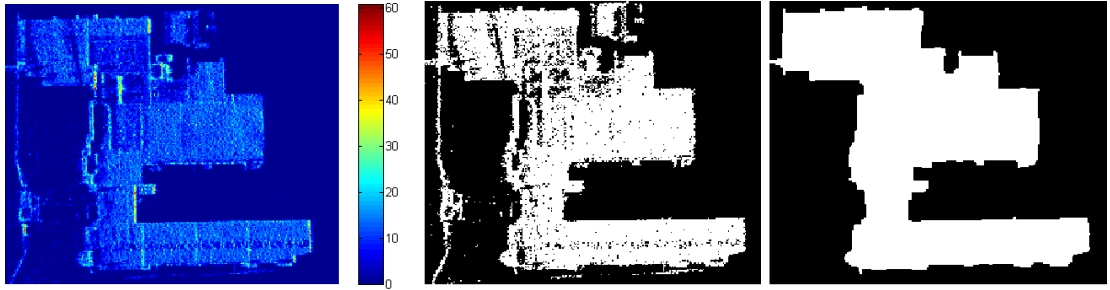
(a) PD map, point density is color codes

(b) Thresholding at PD=5

(c) Initial Building mask after morphological operations
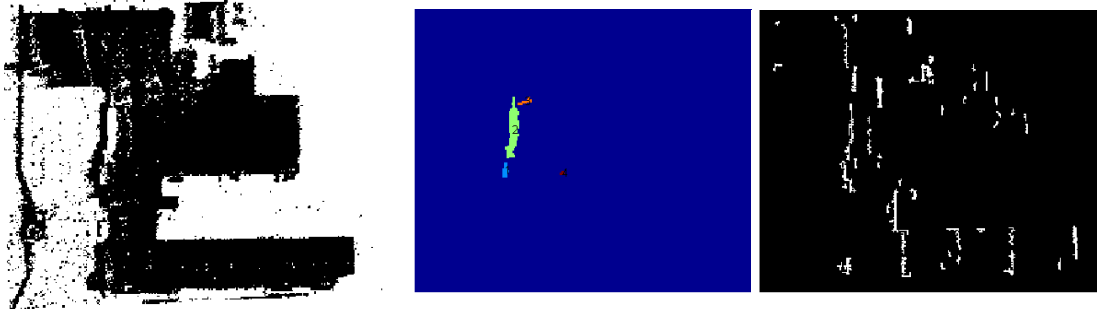
(d) Complement image of the thresholded image(PD=2)

(e) Data gaps inside Building mask

(f) Facade thresholding at PD=16, no "narrow" facade assumption applied)

(g) Data gaps and facades together

(h) Only data gaps adjacent to facades are selected

(i) Final building mask, modified using selected data gaps

Figure 4.9 Building mask extraction from Point density map at example area

BM3D (Block-matching and 3D filtering) is a powerful denoising method. It was developed by Dabov et al.[Dabov et al., 2007]. It is based on the fact that an image has a locally sparse representation in transform domain. It is realized in 2 steps, basic estimation and final estimation. Each of the two estimations contains 3 parts: Grouping, Collaborative filtering and Aggregation. The general concepts are: 1). Grouping: for a given image patch, finding the image patches similar to it and grouping them in a 3D block; 2) Collaborative filtering: Applying a 3-D transform to the formed group, attenuating the noise by shrinkage (Hard thresholding is use in step1 and Wiener filtering is used in step2) of the transform coefficients, inverting the 3-D transform to produce estimates of all grouped blocks, and then returning the estimates of the blocks to their original places. Because the grouped blocks are similar, BM3D can achieve a high level of sparse representation of the noise-free signal, thus, the noise can be set apart well from signal by shrinkage; 3). Aggregation: The output image is estimated by weighted averaging of all achieved block estimates that have overlap.

4.1.3 Segmentation

The resulting denoised nDSM is ready to be segmented. Our strategy is to use watershed transform to oversegment the nDSM, and applying constrained merging process to get the final segments.

4.1.3.1 Watershed segmentation

In grey scale mathematical morphology, the watershed transform was originally proposed by [Digabel and LANTUEJOUL, 1977]. As [Roerdink and Meijster, 2000] stated, the idea behind it comes from geography: a landscape is immersed in a lake, with holes in local minima. Catchment basins will fill up with water, starting at local minima. At points where water coming from different basins would meet, dams are built. When the water level has reached the highest peak in the landscape, the process is stopped. As a result, the landscape is partitioned into regions or basins are separated by dams, called watershed lines.

The gradient map is computed over the nDSM. nDSM contains height information, while gradient changes stand for height jumps. In our case, the gradient edges are the watershed lines, while the catchment basins are the homogeneous grey level regions of this image, i.e., the segments we want. Direct use of the gradient image usually produces oversegmentation results, due to noise or local irregularities. Thus smaller gradient was cut off before applying watershed segmentation.

A height difference threshold H_{min}, meaning, if local height difference H_L is smaller than H_{min}, it will be ignored. I_{min} is the intensity value in gradient image, corresponding to H_{min}. The local intensity value I_L smaller than I_{min} will be cut off. The left part in the gradient image denotes for regions where $H_L >= H_{min}$.

The exact choice of H_{min} can be tricky. Thus we set H_{min} to a value smaller than the demanding height difference H_{max}, then performing a constrained merging step on the oversegmentated result from watershed segmentation. H_{max} is the threshold of maximum merging height, meaning, if two segments has H_L bigger than H_{max}, the segments belong to different roofs and will be merged. Segments that lower than ground are removed before merging.

4.1.3.2 Constrained merging

The merging approach is based on the height difference of adjacent segments and the average polygon complexity (APC). The purpose is to get minimum amount of segments, and maximum segments' regularity.

The segment's regularity can be indicated by polygon complexity [Brinkhoff et al., 1995], is proposed in 1995. It is defined as following:

$$compl(pol) = 0.8 \cdot ampl(pol) \cdot freq(pol) + 0.2 \cdot conv(pol) \tag{4.3}$$

where
$freq(pol) = 16 \cdot (notches_{norm}(pol)) - 0.5)^4 - 8 \cdot (notches_{norm}(pol) - 0.5)^2 + 1,$
$notches_{norm}(pol) = notches(pol)/(verties(pol) - 3),$

$notches_{norm} = 0$

$notches_{norm} = \dfrac{14}{18-3} = 0.93$
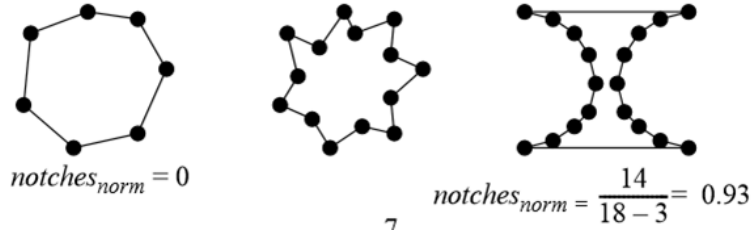
(a) notched and normalised notched

Föhr

Khalkidhiki

Isle of Lewis

$ampl = 0.02$

$ampl = 0.50$

$ampl = 0.60$

(b) ampl

Föhr

Khalkidhiki

Isle of Lewis

$conv = 0.04$

$conv = 0.47$

$conv = 0.28$
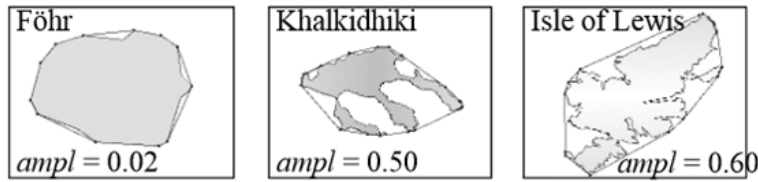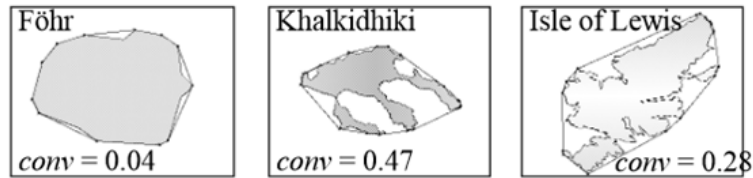
(c) conv

Figure 4.10 Definitions of polagon complexity [Brinkhoff et al., 1995]

$ampl(pol) = (boundary(pol) - boundary(convexhull(pol)))/boundary(pol),$

$conv(pol) = (area(convexhull(pol)) - area(pol))/(area(convexhull(pol)).$

The complexity $compl(pol)$ of polygonal objects, is measured by comparing a polygon with its convexhull. Notches are the polygon vertices which located inside of polygon's convexhull. $ampl(pol)$ measures the amplitude of notches vibration, and measures the frequency of it. $conv(pol)$ measures the area difference between a polygon and its convexhull.

For a group of segments, we define the average polygon complexity $compl(mean)$ as the weighted mean of polygon complexity, using area size as weight. $compl(mean)$ is in the interval$[0, 1]$. Smaller $compl(mean)$ indicates simpler polygon. As segment's area is the weight, and larger simple polygons are our goal, we propose following merging

approach:

(1) Define an area threshold T_a. Segments which are smaller than T_a are denoted as S_s, and bigger ones denoted as S_b;

(2) Merge S_s to S_m. Maximum merging height H_{mas} and step size h are defined. Hierarchically merging n times, $n = H_{max}/h$. In the ith merging, the maximum height difference is $H_i = H_{min} + (i - 1) \cdot h$. The adjacent matrix of all small segments S_s is computed, and the adjacent segments with $H_L < H_i$ are merged. After $i_t h$ merging, S_s is merged into S_m. Segments' height are computed from denoised DSM.

(3) Put S_m and S_b together, further merge segments. Repeat merging steps in 2), with the constrain that $comp(mean)$ is not allowed to increase, unless H_L of the adjacent segments is smaller than H_{min}.

(4) For segments whose area size smaller than minimal roof size T_a, merge them to their nearest largest segment, under constrain of minimum increasing of $comp(mean)$.

4.1.4 Reconstruction

Once the building pixels are segmented into individual roof segments, the next step is to reconstruct the outline of the distinct segment which are utilized to reconstruct the overall 3-D prismatic building model.

4.1.4.1 Regularization

Minimal bounding shape type detection

Minimal bounding shape type detection of each segment is performed to select a better bounding shape from ellipse and rectangle. The area difference between the segment and its bounding shapes is computed, and the shape with smaller area difference with the segment is chosen. Figure 4.11(a) shows the minimal ellipse bound(green) and minimal rectangle bound(red) for a segment(yellow). Area difference of the ellipse with segment is

(a) Minimal ellipse bound(green) and minimal rectangle bound(red) for one segment.

(b) Minimal bounding rectangles(red) and minimal bounding boxed(black) for 2 segments.
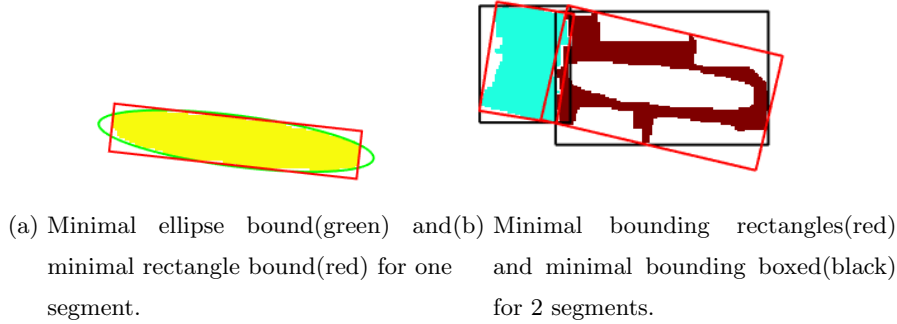
Figure 4.11 Minimal bound detection

smaller than it of the rectangle with the segment, thus ellipse is selected. For segments which are ellipse bounded, the minimal bounding ellipse is considered as its refined outline, thus no more regularization is needed.

Orientation computation

For segments which are rectangle bounded, the orientation for each segment is computed to find rotation angle for the segment before quadtree regularization. Orientation angle is defined as the smaller angle between image's x-axis and the major axis or minor axis of the segment's minimal bound ellipse. The segment is rotated, so that it's main orientations are aligned with the images $x - y$ directions. After quadtree regularization, they are rotated back.

To keep a building's the main orientation consistency, the segment's complexity is computed for every segment using equation 4.3. Segments are separated into two classes: (a) keep its orientation if the segment's complexity is smaller than a threshold T_c, with minimal bounding rectangle; (b) otherwise use image's x-axis direction as segment's orientation, with minimal bounding box, all the boxes have same orientation as the image. In Figure 4.11(b), minimal bounding rectangles(red) and minimal bounding boxed(black) for 2 segments. T_c=0.05, the blue segment: $compl(pol)$ =0.0482, fit rectangle to it; for the dark red segment: $compl(pol)$=0.0774, fit box to it.

Quadtree regularization

Quadtree decomposition is a powerful technique which divides an image into 2D homo-

(a) before refularization (b) qurdtree decomposition (c) after refularization
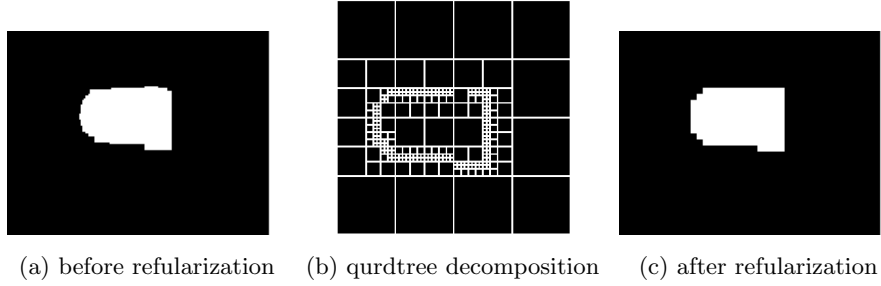
Figure 4.12 Quartree regularization

geneous regions, as shown in many works as [Samet, 1984]). The decomposition of a image builds a tree, with each node stands for a $2^m \times 2^m$ homogeneous block in the image. In our case, each segment is put in a $2^k \times 2^k$ image with empty background, then a quadtree is built by decomposing the image. Now, the root in the quadtree stands for the whole image, and the leave nodes are the smallest blocks in the image, which are all located along boundary of the segment in the image, and the locations are exactly where the to-be-refined small structures are. We take all the small structures in, as parts of the segment, results in a "dilated" segment, which need to be eroded later(Figure 4.12).

Quadtree regularization results in simplified polygons that have only corner vertices, with 2 perpendicular boundary directions. Thus the main directions of segment need to be aligned to image axis directions. Check segments' orientation computed in previous step. For segments whose orientation is not the image's x-axis, rotate them with their orientation angles, and rotate back after quadtree regularization. Segments with similar orientation angles are grouped together and rotated with one angle, to make the scene more regular.

Overlap cropping

The next step is to overlay all the segments together. Previous two adjacent segments $S1$ and $S2$ now are $S1_n$ and $S2_n$, with an overlap O. Compute average polygon complexity for two cases: (a) $S1_n$ and $S2_{n_O}$, $S2_{n_O}=S2_n-O$, and (b) $S1_{n_O}$ and $S2_n$, $S1_{n_O}=S1_n-O$. Overlap O is assigned to the case which gives smaller average polygon complexity. In Fig-

ure 4.13, (a) 2 segments and their overlap(in yellow); (b) first assignment, APC=0.0519; (c) second assignment, APC=0.0404. The second case is accepted because of the smaller APC.

| (a) | (b) | (c) |
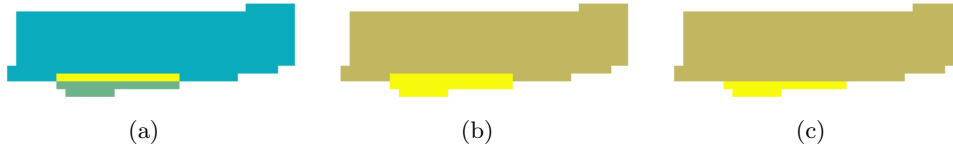
Figure 4.13 Assign overlap to one of the two neighbor segments based on Average Polygon Complexity

Zigzag line removal

Quadtree regularization takes out smaller grids along roof segment's boundary, leaving all boundary line segments perpendicular to their neighbors. Sequential short line segments form a "zigzag" line, which need to be removed. Our approach to detect zigzag lines is based on the Visvalingam -Whyatt algorithm[Visvalingam and Whyatt, 1993].

First, find "zigzag" which meets two conditions: 1) its effective area A is smaller than a predefined maximum removable area A_{max}, and 2) its percentage of effective area in whole polygon area P is smaller than a predefined maximum removable percentage P_{max}. The effective area of point is the area change of the polygon when adding and removing this point. Second, if at least 3 of such points are consecutive, a zigzag line is found.

Keep the two end points on the zigzag line and remove all other points. The zigzag lines are removed.

4.1.4.2 Modeling

For all roof polygons, corresponding height is introduced from DSM to construct 3D building polygons.

4.2 Workflow based on point cloud segmentation

4.2.1 Overview of workflow

Inputs of the point clouds (PC) based workflow (Figure 4.14) are TomoSAR point clouds. For roof segmentation, RANSAC algorithm is employed to obtain plane fitting results. Density based clustering (DB scan) and height local minimum based clustering are used to segment point clouds. Boundary polygon segments are created by convex hull collapsing. Boundary polygons are then regularized to building boundaries.

4.2.2 Preprocessing

Same as DSM based workflow, the building points need to be extracted out from TomoSAR point clouds first. Since plane fitting approach (subsection 4.2.3) is used in segmentation, instead of using ground filtering approach, the largest fitting plane in the scene is considered as ground plan and points in it are taken as ground points.

TomoSAR point clouds are noisy compare to LiDAR point clouds, thus, a smoothing procedure before segmentation is preferable. The procedure is similar to the local height smoothing (subsubsection 4.1.1.1), with a different neighbourhood definition: a radius R is predefined, for each point P, and a cylinder centered at P with radius R that extended vertically is the neibourhood of P.

4.2.3 Plane fitting based Segmentation

Input of segmentation step is building point clouds. The steps are as follows:

1. RANSAC (explained in methodology) plane fitting algorithm is iteratively used to extract planes which best fit most points in the point clouds;

2. A density based clustering algorithm is used on each plane extracted in 1st step. Density-based spatial clustering of applications with noise (DBSCAN) is a data clustering algorithm, given a set of points in some space, it groups together points
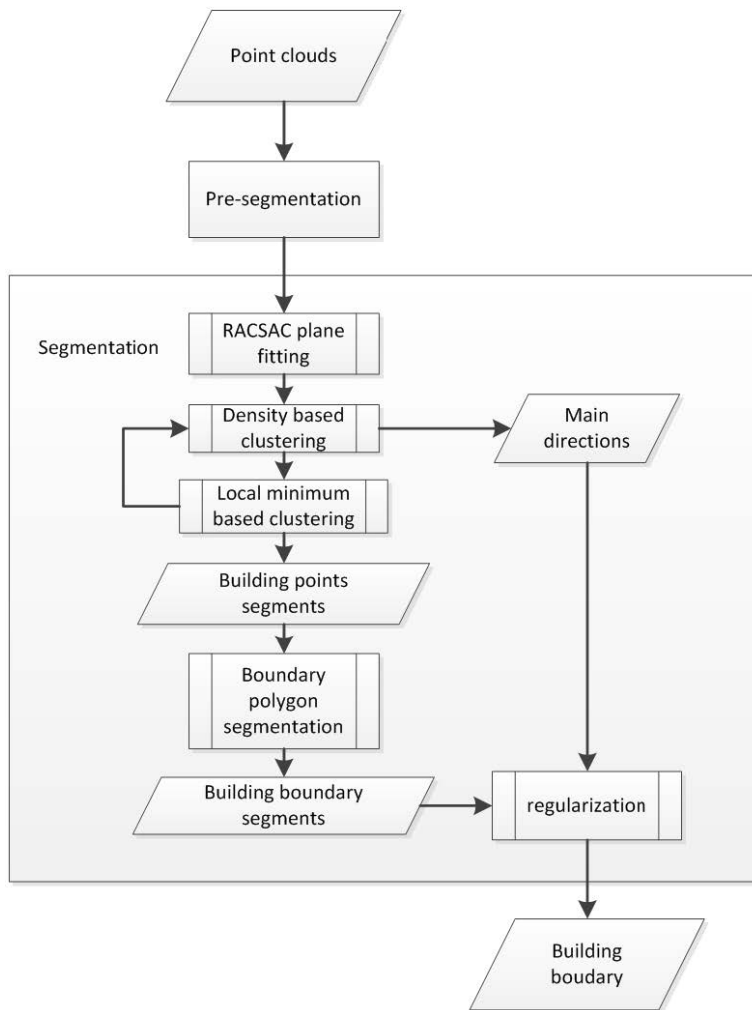
Figure 4.14 Overview of workflow (PC based)
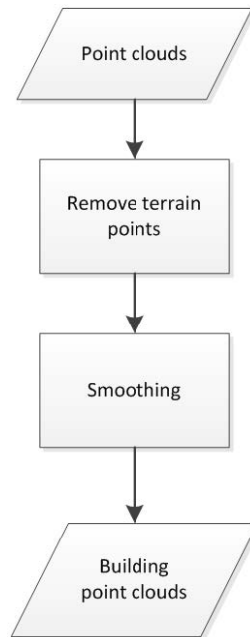
Figure 4.15 Preprocessing workflow (PC based)

that are closely packed together (points with many nearby neighbors), marking as outliers points that lie alone in low-density regions (whose nearest neighbors are too far away) . This step is used to separate different roof surfaces on the same RANSAC plane;

3. Height histogram of each cluster is then plotted (occurrence to height). In the height histogram, the shape of the bins should follow Gaussian distribution, if there exists only 1 roof plane. Additional peaks in height histogram indicate that there might be more roof planes exist that contain more points that none roof planes. To simplify the algorithm, local minimum is computed instead of local maximum. This has advantage in recognizing sub-clusters in cases like what is shown in Figure 4.17(b). In this case, the first peak is not a local maximum, thus might lead problem if we find local maximum. A minimum points threshold is defined to filter out clusters that have too less points. This step follows DBscan clustering;

4. After local minimum based clustering, the sub-clusters are examined to see if they

need DBscan again. This is realized by comparing the area ratio of the convex hull and the tightest out polygon areas. The bigger the ratio is, the more complex the boundary polygon is, the more likely the points in the cluster are distributed unevenly;

5. Now we look into height histogram again. Since each cluster now only has 1 peak in its height histogram, if the bins on the sides contain very less points, they are deleted as noise;

6. After clustering procedure, all 3 kinds of clusters are considered as roof segments;

7. Main directions of the scene are determined from histogram (Fig.4.19) analysis of the angle between boundary points' tangential direction and a certain direction, e.g. y-direction.

Figure 4.17 shows height histogram of two example clusters. In left case, 3 local minimums are shown, indicating the cluster can be further separated by local minimum into 4 sub-clusters. While in right case, 1 local minimum is shown, suggesting the cluster can be separated into 2.

In Figure 4.20, for two example segments, red polygon is the Convex hull, while blue polygon is the tightest out boundary polygon. In left case, points in cluster are more likely distributed unevenly than in the right case.

4.2.4 Reconstruction

4.2.4.1 Boundary extraction and boundary polygon segmentation

Building boundary polygons are created by iteratively collapsing the convexhull of the segment points.

Then, each boundary polygon is segmented by comparing the direction change and a predefined threshold at each vertex in the boundary polygon. If the change of direction is bigger, the polyline is split at this vertex.
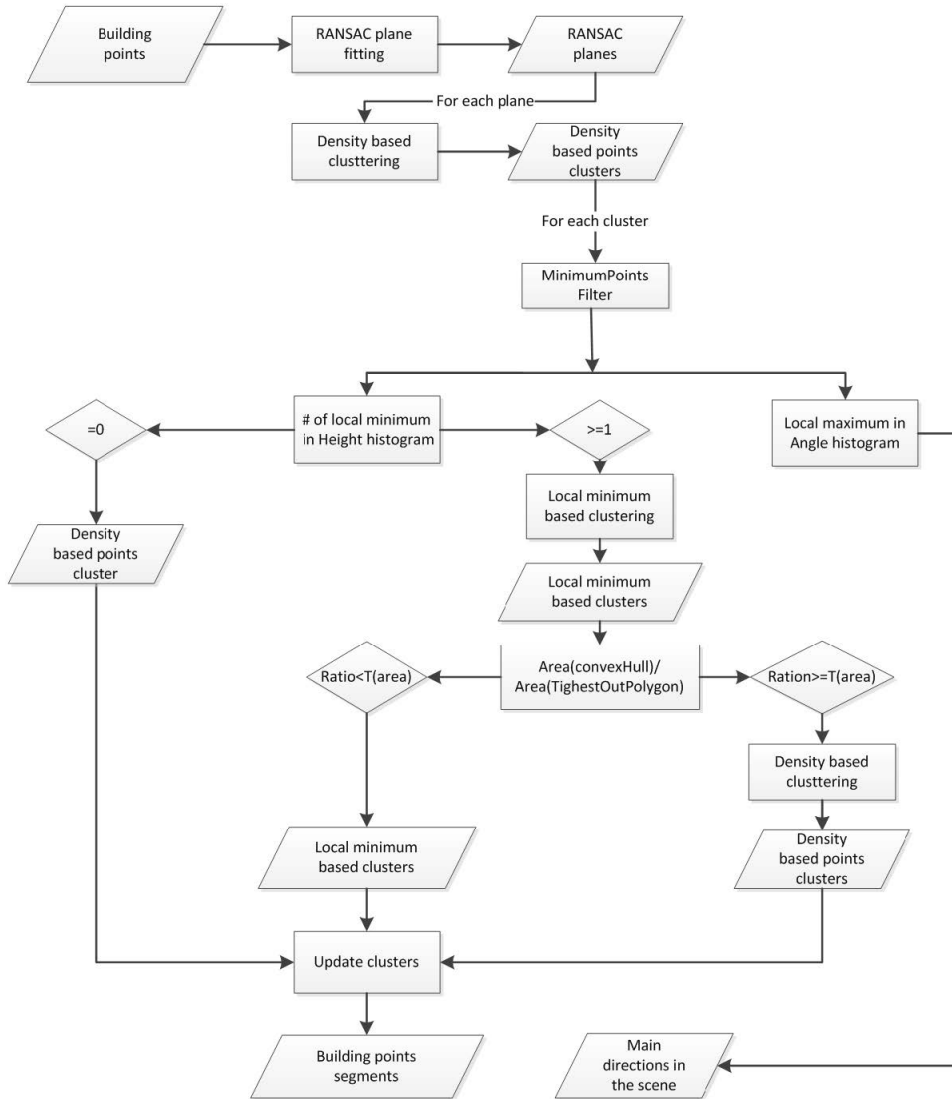
Figure 4.16 Segmentation workflow (PC based)

(a) (b)
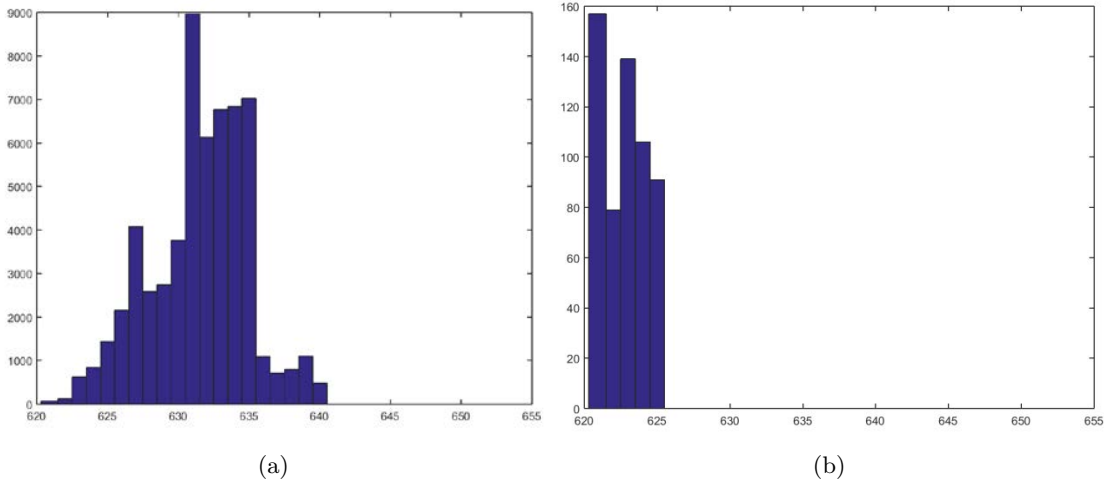
Figure 4.17 Height histogram of example clusters

4.2.4.2 Boundary refinement and regularization

For now, boundary polygon segments are available, as well as the main directions in the scene. The next step is to refine and regularize boundaries.

For each main direction, a set of line equation can be built to fit boundary polygon segments. For each boundary polygon segment, a straight line follows one main direction can be found to be used as best approximation of it. After all fitting lines for boundary polygon segments are found, intersubsection points between neighbor fitting lines are computed, further connected, to form building roof boundary.

In the main direction boundary, parallel lines need to be adjusted if the distance between them are small so that they can be presented as one main direction line. This is down by following: searching for all parallel boundary segments from the longest Boundary segment, if distance between them $< Td$, recomputed fitting line, until all parallel lines are adjusted.

In Figure 4.20, (a) extracted boundary for 1 cluster; (b) segmented boundary, different segments in different color; (c) main direction fitting lines to the boundary segments; (d) refined boundary.

(a) (b)
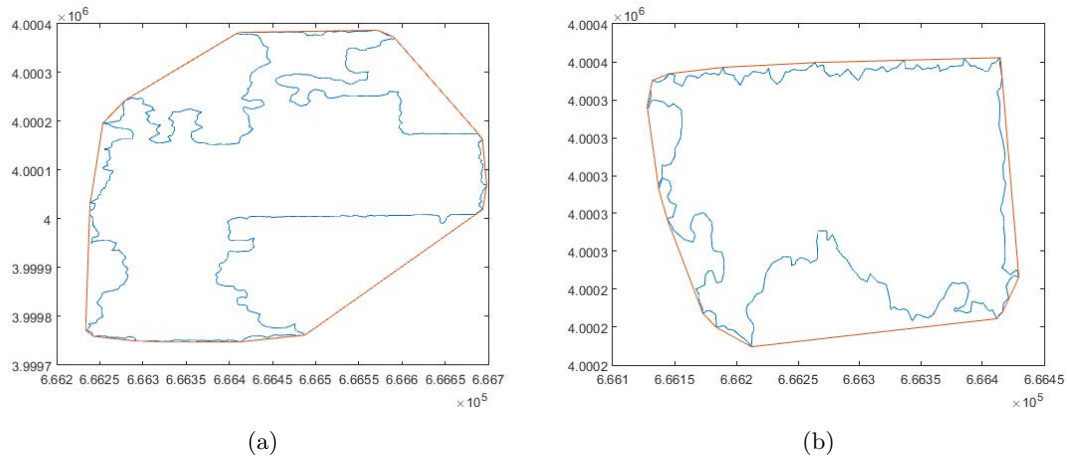
Figure 4.18 Convex hull and tightest out boundary polygon

4.2.4.3 Topology correction

1. Merging. If two of the extracted boundary polygons are intersected horizontally, or one contains the other one, and the vertical height difference between them is smaller than a threshold Tvd, they should be merged into one polygon. (Figure 4.24 (b))

2. Cropping. For polygon pairs $poly_i$ and $poly_j$, who horizontally intersect each other with vertical height difference bigger than the threshold T_{vd}, one of them need to be cropped. This is decided as following: take out all points inside their intersect part from TomoSAR point clouds, compute the vertical distance between each point to the two polygons: Dpi and Dpj, then compute total distance of all the points Di and Dj, by adding up all Dpi, adding up all Dpj. If $Di > Dj$, crop polygon i. Otherwise crop polygon j. (Figure 4.24 (c))

4.2.4.4 3D model reconstruction

Again the 2D model is extruded to 3D model by adding height information. An example is shown in Figure 4.25.
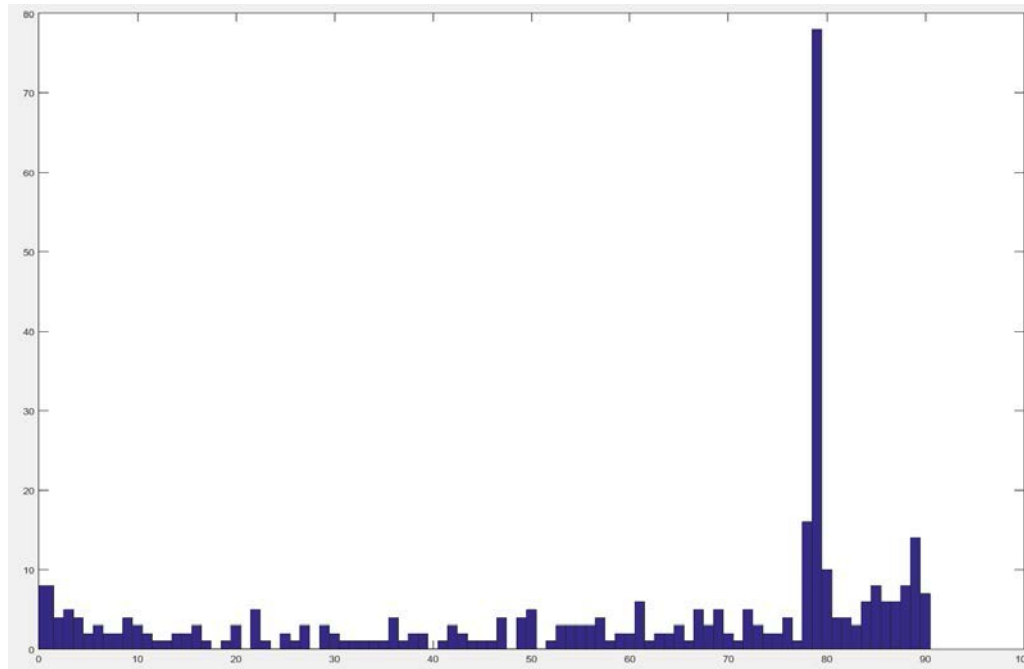
Figure 4.19 Angle histogram of boundary points)

4.3 Comparison and Decision of the workflow

Both the presented workflow are data driven. Decision to chose which workflow to choose depends upon several factors e.g., the accuracy of the reconstruction and computational efficiency.

In presegmentaion stage, more efforts are made in DSM based workflow: 1) ground filtering, 2) building masks, facades and data gaps extraction from point density map, 3) "non-surface" point height modification. The reasons for not including these steps in the point clouds based workflow are as follows:

1) for ground filtering: the biggest cluster found after plane fitting is assumed to be groud;

2) for information from Point density map: without the "non-surface" point height modification step, facades detection is not necessary. Since building regions are defined by convex hulls, there is no need to detect building masks and data gaps on building.

(a)   (b)

(c)   (d)
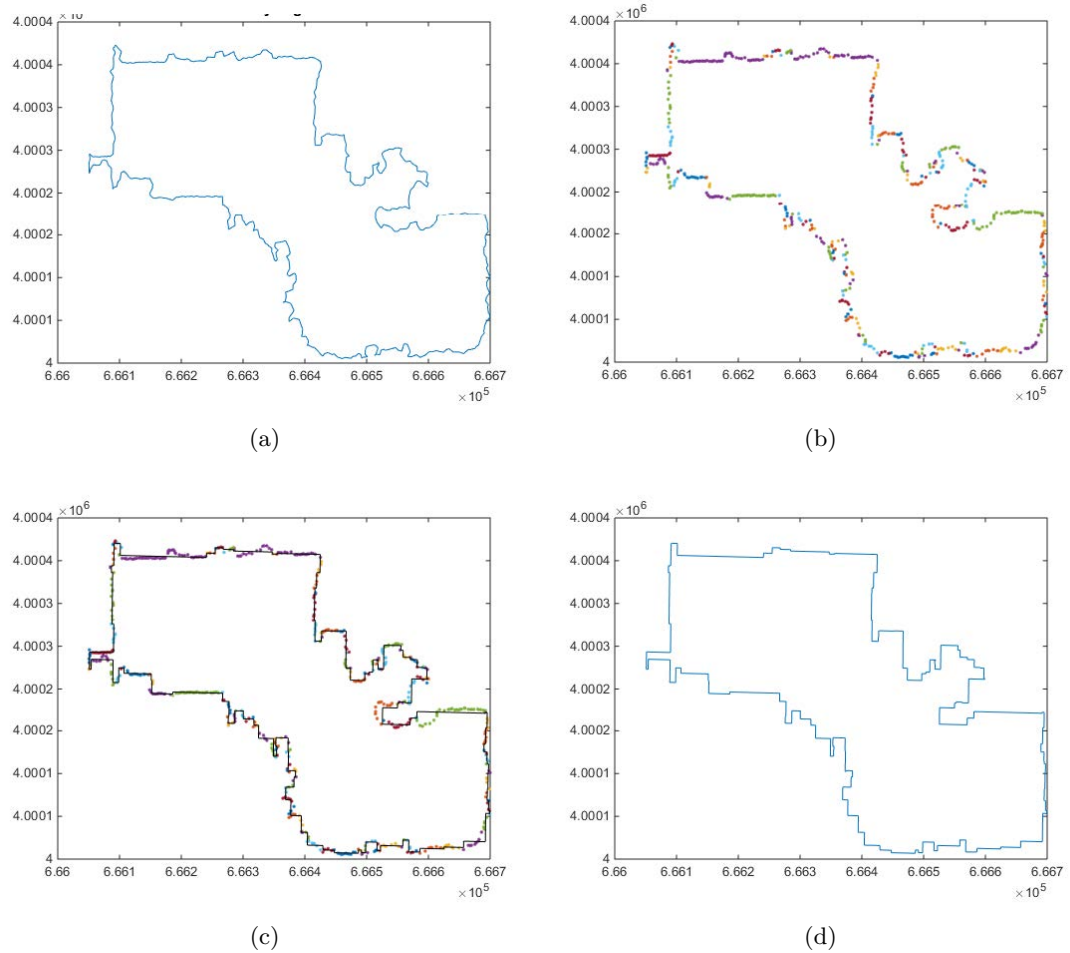
Figure 4.20 Refined polygon boundary

3) for points' height modification: the RANSAC plan fitting procedure can dealing with outliers.

However, the preprocessing in point based workflow may not be adequate. If the terrain of tested area is more complex than can be fitted with a plane, ground filtering is still needed. In some forms of buildings, the convex hull of points can not represent the real shapes of the buildings, without building masks, facades and data gaps detection. RANSAC algorithm fits planes to points, at places with rich facades points and very few roof points, the roof plane will not be fitted, thus the building shape is not closed.
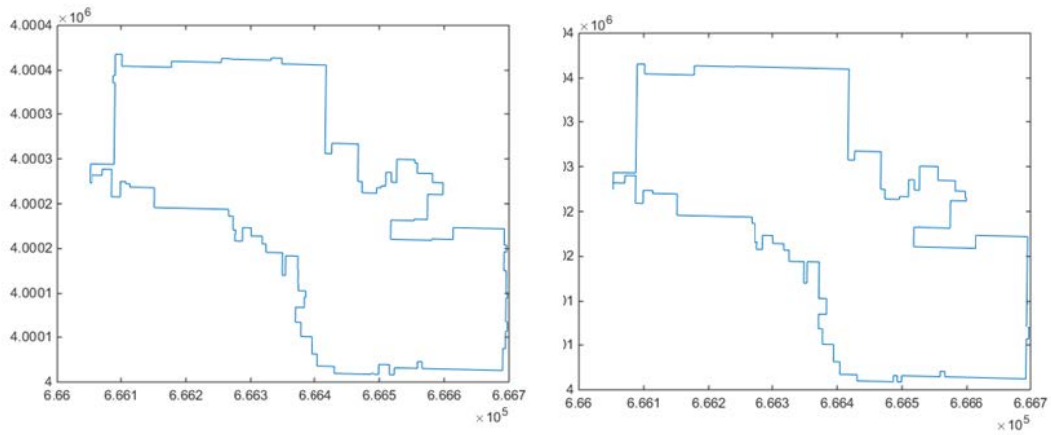
Figure 4.21 Adjusted boundary polygon: (left) Td=2m, (right) Td=5m
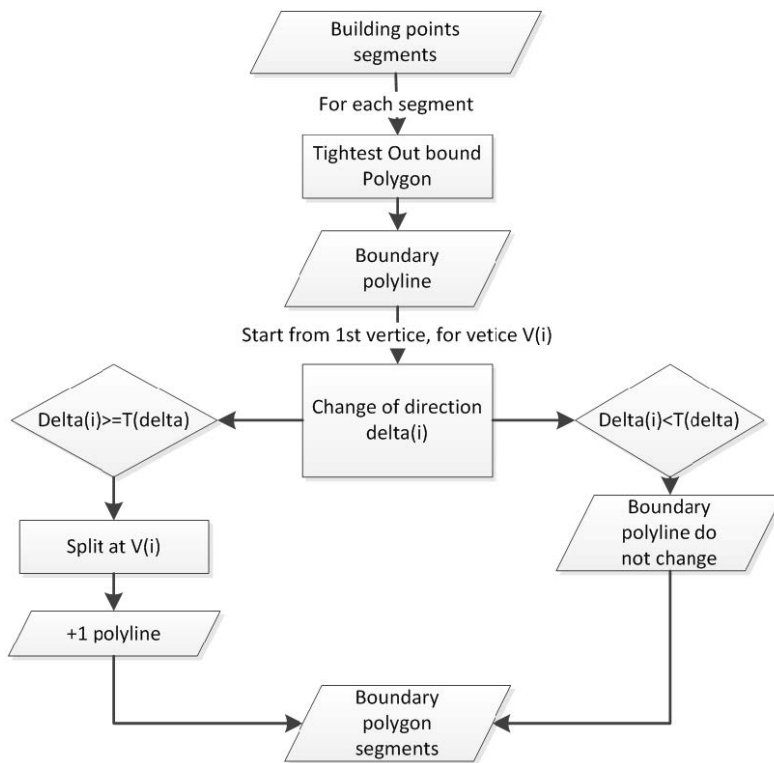
Figure 4.22 Boundary polygon segmentation workflow
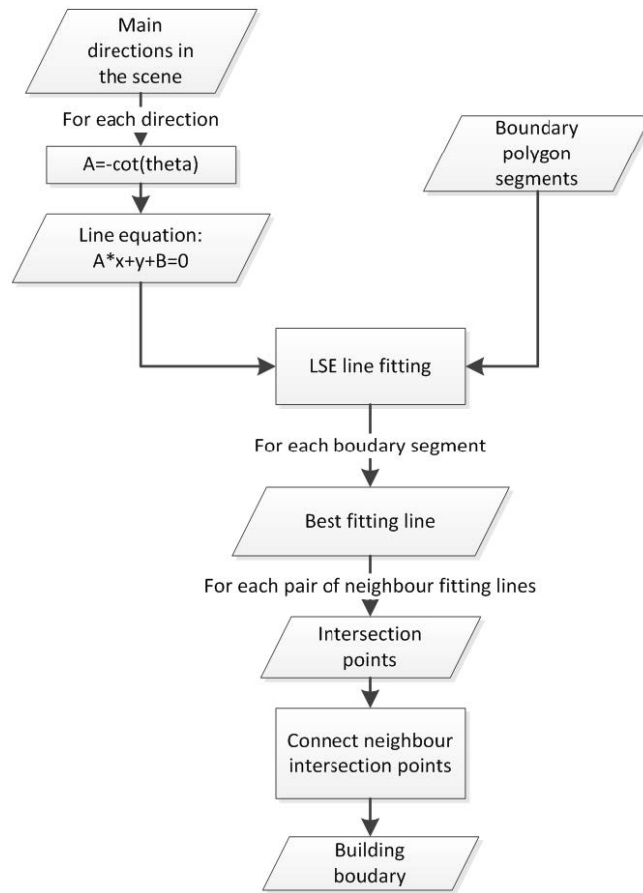
Figure 4.23 Boundary regularization

(a) Polygons before topology correction

(b) Polygons after merging

(c) Polygons after cropping. Polygons do not belong to target building are removerd.
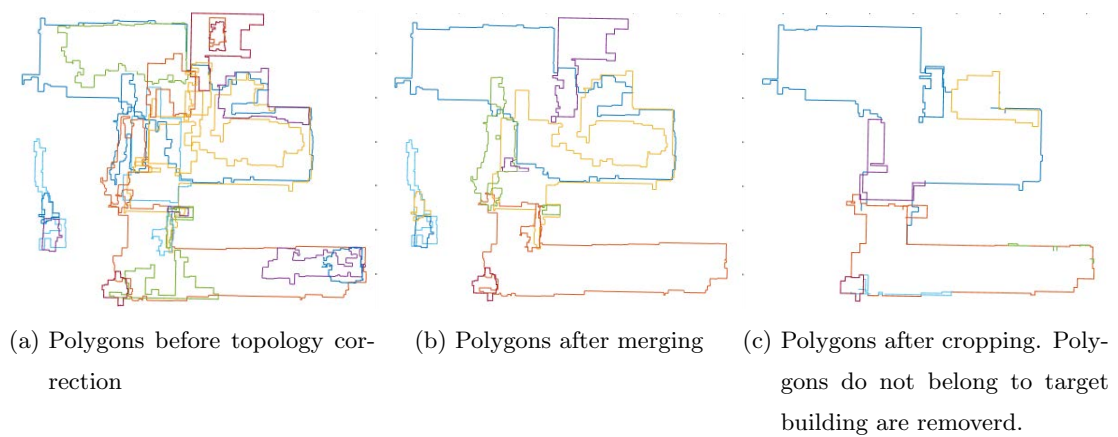
Figure 4.24 Example of topology correction
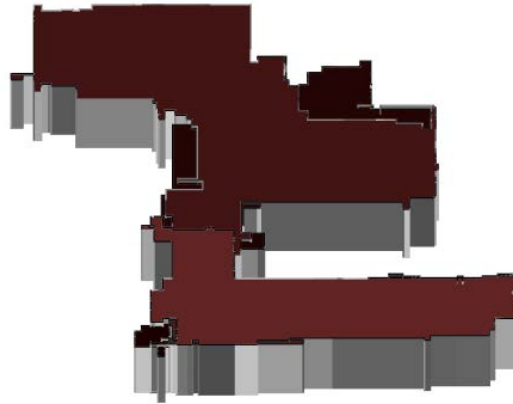
Figure 4.25 3D polygon model of the example building

In segmentation stage, image processing techniques are adoppted in DSM based workflow, while RANSAC plane fitting and clustering techniques are used in point based workflow. Comparing to DSM based workflow, the points segments are more "independent" to each other, as in clustering step, spatial relationships is not considered. While in DSM based workflow, the spatial adjacency is a hidden constrain.

In reconstruction, as each polygon is regularized separately, lacking of the spatial adjacency the topological relationships might be changed if they were right before, which can not be assured. When three or more planes intersect, how to decide the intersect nodes becomes more tricky. When the roof polygons are flat, they do not directly intersect to each other since they are connected by step edges. If the polygons have no overlaps in $x - y$ plane(top view), it is not possible to locate their intersecting edge, thus right topological relationships can not be reconstructed.

Computational efficiency is another important fact to consider. Though computational efficiency is of secondary importance, it still need to be stated that, the data amount of unstructured TomoSAR point clouds is considerable. When large amount of data are to be processed, not only computational efficiency, but also the computer performance should be taken into account.

Based on these considerations, the DSM based workflow is chosen.

5 Tests and Evaluation

5.1 Tests of DSM based reconstruction workflow

5.1.1 Input data

TomoSAR point clouds of two large building in las Vegas city are used to test. The test point clouds are generated from a stack of 25 images using Tomo-GENESIS software developed at DLR [Zhu et al., 2013]. Figure 5.1 shows the two point clouds in 3D.

5.1.2 Parameters Setting

Through experiments and visual interpretation, the optimal input parameters for segmentation and reconstruction were determined in following table.

Table 5.1 Parameters setting of tested data sites

Parameters	data site 1	data site 2
Ground filtering $[c, s, w, h]$	$[3, 0.1, 500, 8]$	$[3, 0.015, 400, 10]$ $[3, 0.015, 400, 20]$
Thresholding SD map $[T_{BM}, T_F]$	$[5 , 16]$	$[5, 16]$
Modifying "non-surface" points height $[N_{max}, N_{max}, H, L]$	$[20, 200, 1.8, 9]$	$[20, 200, 1.5, 8]$
Denosing Std. of the noise	3	3
Segmentation $[H_{min}, H_{max}]$	$[0.45, 3]$	$[0.45, 3]$
Minimal quadtree grid length	2	2
Removing zigzag line $[A_{max}, P_{max}]$	$[15, 0.1]$	$[15, 0.1]$
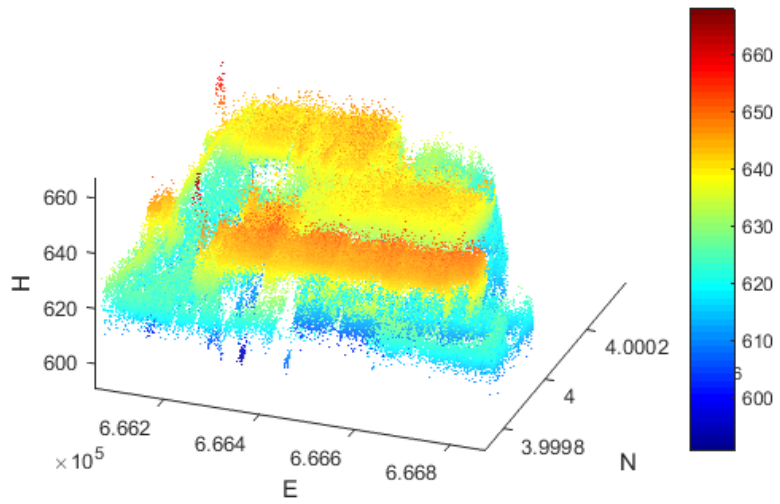
(a) Test site 1
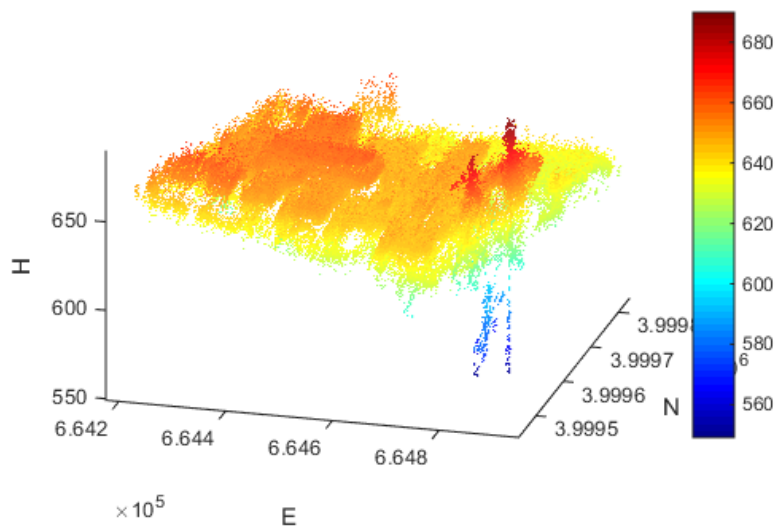
(b) Test site 2

Figure 5.1 TomoSAR point clouds on two data sites, height is corlorcoded

5.1.3 Results

The results of local height reassigning and ground filtering are shown in Figure 5.2 and Figure 5.3. Based on the hight histogram of above ground points(Figure **??**), in data site 2, the above ground points are further separated using ground filtering. The results are shown in Figure 5.3(e) and (f).

The Segmentation results before and after constrained merging are show in Figure 5.6. The regularization results are shown in Figure 5.7. Figure 5.8 show final 3D models reconstructed.

5.2 Evaluation

The building points are plotted with the reconstructed models in Figure 5.9. The root mean square (RMS) error of all points from respective planes are computed. The results are shown in Table 5.3 and Table 5.3.

The overall RMS for datasite 1 is 3.19 m, while for datasite 2 is 3.76 m. For datasite 1, the large RMSs in segment 2, 7, 10 come from very high local points. While for datasite 2, the large RMSs in segment 11 and 12 come from merging several small segments in that area. In point cloud, the range of height of points are large. In segmentation step, those small segments are merged together, and given a mean height.

For object extraction from amplitude image, the result of potential building masks is better in HMRF algorithm, while the result of potential facades is better in adaptive thresholding algorithm. However, the interpretation in both algorithms is difficult.

Roof index	Roof area(m^2)	Number of pts in roof region	Mean(m) of distance between pts and roof	RMS(m) of distance between pts and roof
All	214944,75	241624	3,77	3,19
1	1003,5	567	5,62	6,05
2	139,5	76	6,85	7,90
3	5019,75	3823	5,14	6,30
4	571,5	238	2,88	3,53
5	47371,5	47992	2,00	2,99
6	859,5	586	4,44	5,14
7	283,5	233	6,93	12,71
8	1867,5	1855	3,60	5,83
9	24187,5	27004	2,83	4,01
10	571,5	246	8,77	9,50
11	427,5	276	2,45	3,36
12	1147,5	1029	4,31	5,35
13	859,5	898	6,22	6,95
14	859,5	1016	2,53	3,26
15	43627,5	55928	2,02	3,06
16	20254,5	24261	1,69	2,41
17	283,5	382	2,89	3,34
18	139,5	25	3,51	4,46
19	139,5	242	3,63	4,54
20	3019,5	225	5,15	5,77
21	23107,5	31085	1,71	2,36
22	10030,5	12185	1,53	2,23
23	1579,5	1707	2,90	3,74
24	4603,5	4951	1,99	2,73
25	1003,5	768	3,44	4,28
26	139,5	130	4,62	5,79
27	21847,5	23896	2,02	2,84

Table 5.2 Evaluation of data site 1

Roof index	Roof area(m^2)	Number of pts in roof region	Mean(m) of distance between pts and roof	RMS(m) of distance between pts and roof
All	132635	129569	3,45	3,76
1	8612	3310	2,76	3,95
2	8633	8673	2,11	2,92
3	4233	753	2,80	3,93
4	5505	6321	2,44	3,25
5	18698	23224	2,99	3,82
6	6504	7542	1,78	2,68
7	46573	47098	1,68	2,36
8	7151	6539	1,97	2,54
9	2136	1564	1,95	2,67
10	4142	3281	3,26	5,08
11	4117	1607	8,99	13,19
12	1283	1905	13,07	15,44
13	1177	1296	3,15	3,87
14	7376	8214	1,77	2,44
15	6083	8047	1,34	1,85
16	412	195	3,17	4,16

Table 5.3 Evaluation of data site 2
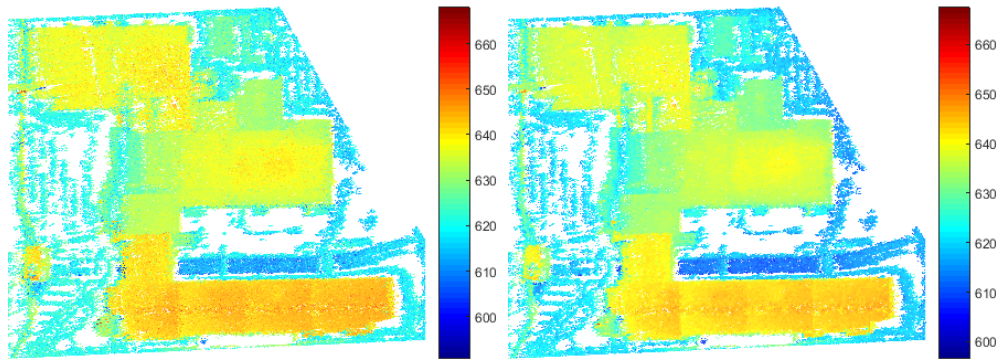
(a) Original point clouds

(b) Point clouds after local height reassigning
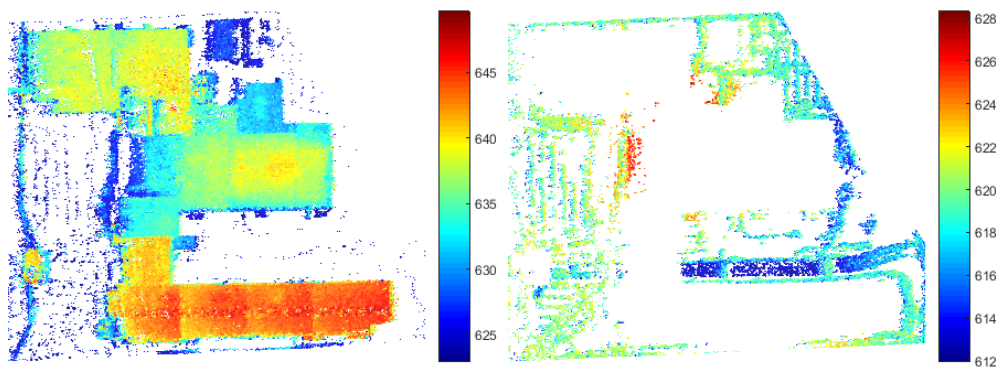
(c) Above ground point clouds

(d) Ground point clouds

Figure 5.2 From original point clouds to ground filtering results, test site 1
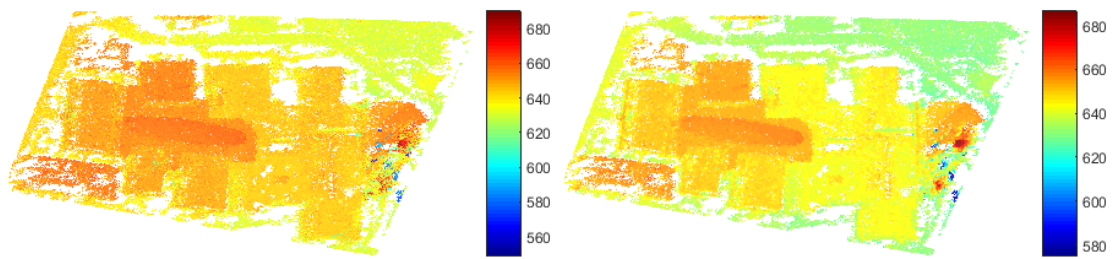
(a) Original point clouds

(b) Point clouds after local height reassigning
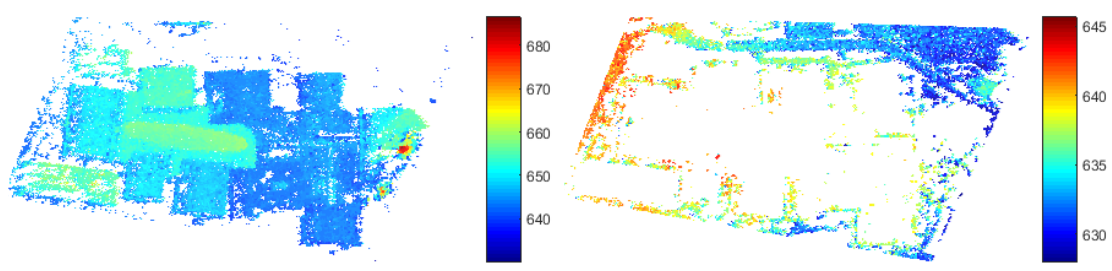
(c) Above ground point clouds

(d) Ground point clouds

Figure 5.3 From original point clouds to ground filtering results in separate layers, test site 2
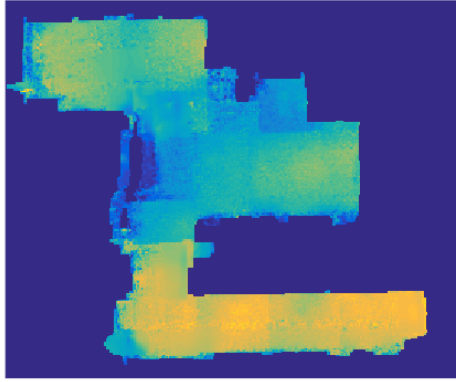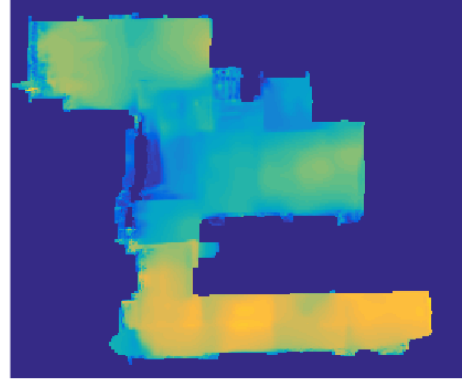
(a) Interpolated DSM   (b) Denoised DSM

Figure 5.4 DSM and denoised DSM over building area, test site 1
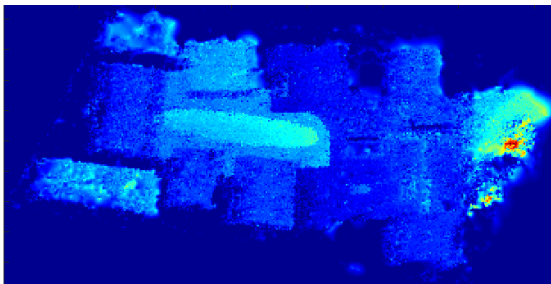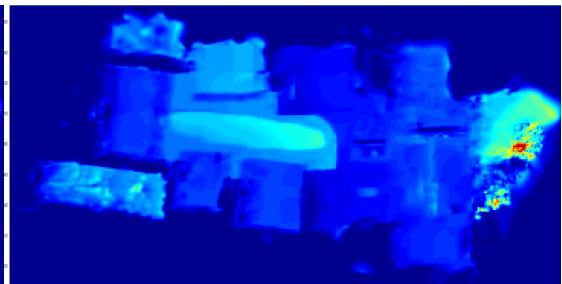
(a) Interpolated DSM   (b) Denoised DSM

Figure 5.5 DSM and denoised DSM over building area, test site 2
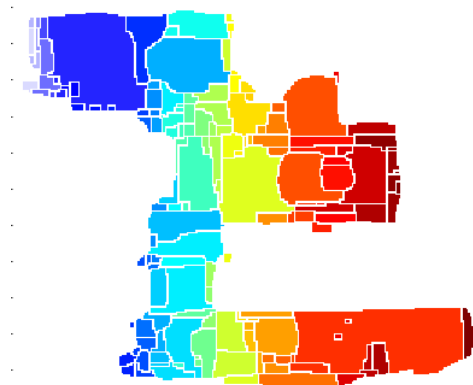
(a) datasite1, before

(b) datasite1, after
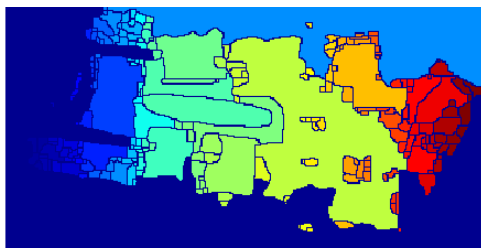
(c) datasite2,before
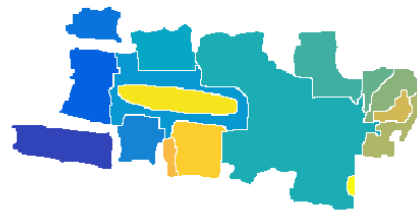
(d) datasite2, after

Figure 5.6 Segmentation results, before and after constrained merging
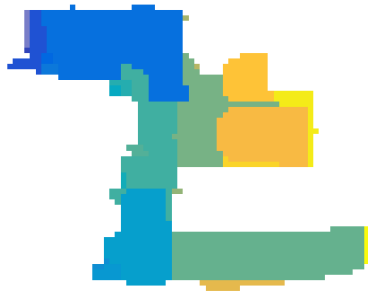
(a) datasite1
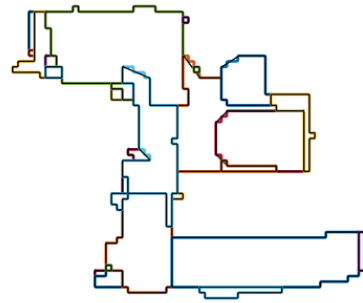
(b) datasite1

(c) datasite2
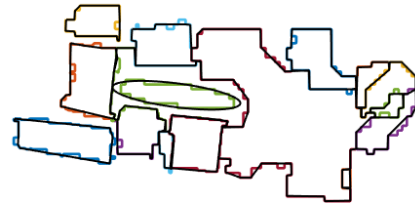
(d) datasite2

Figure 5.7 Quardtree regularization results and "zig-zag line" removal results
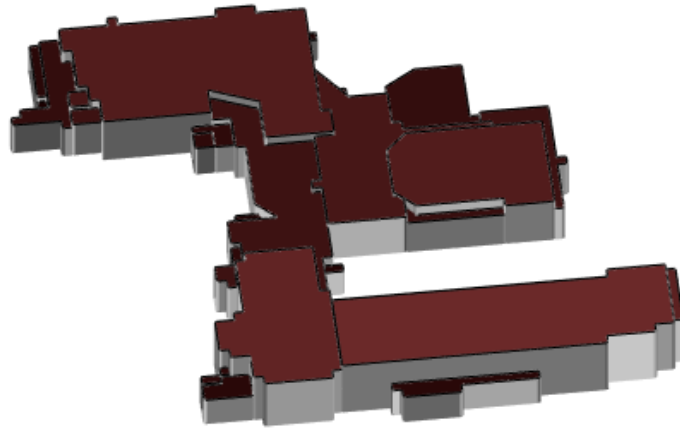
(a) data site 1
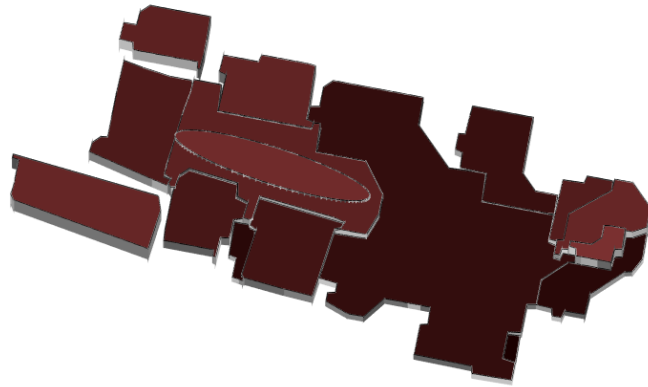
(b) data site 2

Figure 5.8 Reconstructed building models
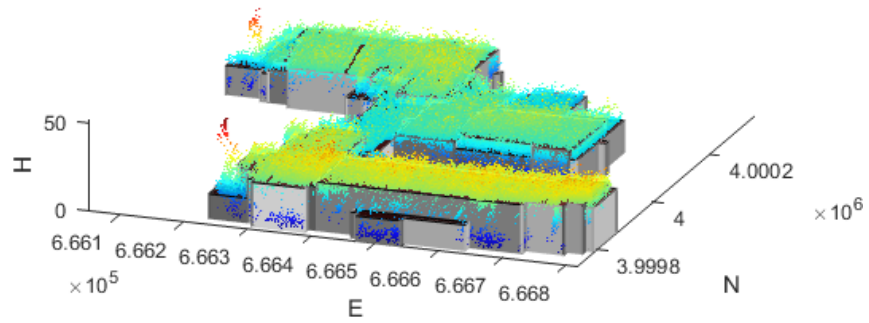
(a) data site 1
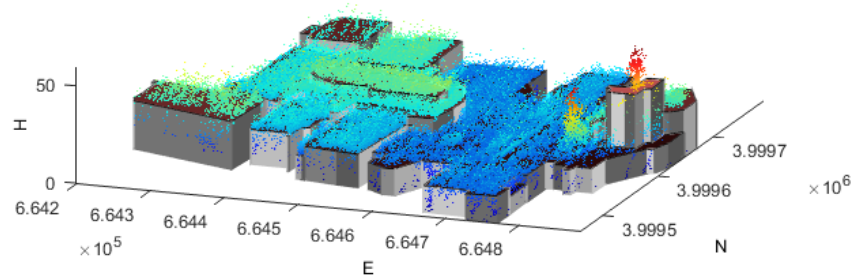
(b) data site 2

Figure 5.9 TomoSAR points in building range are plotted with the reconstructed models.
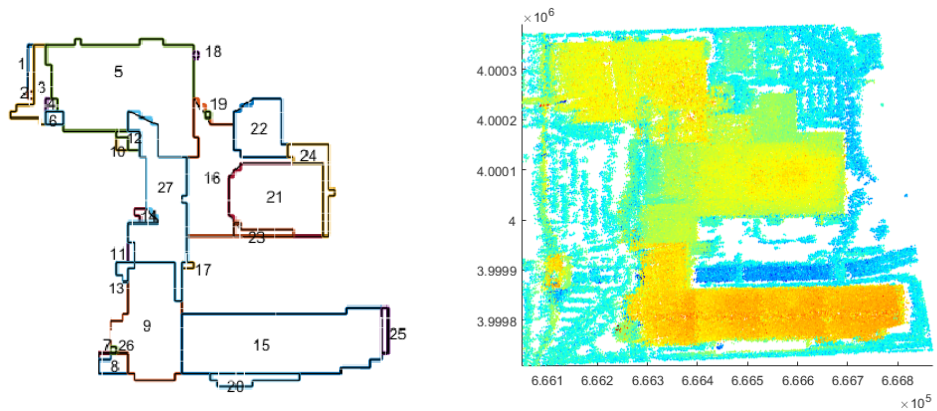
(a) Building polygons with index

(b) Point cloud in top view

Figure 5.10 Building roof polygons and original point clouds, data site 1.
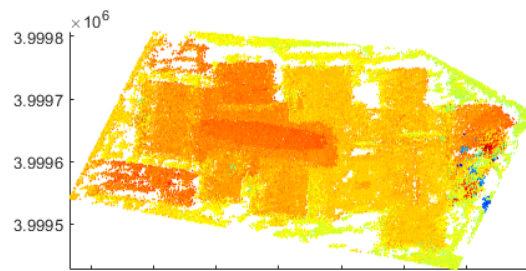
(a) Building polygons with index   (b) Point cloud in top view

Figure 5.11 Building roof polygons and original point clouds, data site 2.

6 Discussion

6.1 Discussion of the proposed workflow

6.1.1 Discussion of preprocessing

The purpose of the preprocessing is to provide better input data for DSM interpolation.

Building masks are very important since they control the area of buildings in DSM. Building masks are extracted from Point Density map in the proposed workflow, based on the assumption that there are more points on building parts than in ground. The contextual information between building masks, facades and data gaps are analysed and used to refine extracted building masks. However, the point density assumption is not always right. When there are few points on roofs, in the thresholding and morphological operations, the extracted building masks may not be complete.

The parameters setting in height modification is assigned manually. In future, an estimation may need to decide best parameters.

In the tests of the workflow, some of the parameters are not constantly chosen, meaning more analysis of parameters setting is needed to increase the level of automation of the proposed workflow.

The proposed workflow needs supervision, especially in morphological operations when processing Point Density maps. Instead of simply operations such as opening and closing, more reliable operations "Openging by reconstruction" and "Closing by reconstruction" are used. However, the quality of morphological operations depends on the shapes of

objects. Since usually multiple morphological operators are needed to give better results, the number of operators, the sequence of the morphological operations, the repeating time of each operator, and the parameter used in each operator are all not fixed. The effiency of the workflow would be largely increased if this part can be unsupervised with results no worse than current results.

6.1.2 Discussion of Segmentation

The building DSM is firstly oversegmented using watershed transform. Then constrained merging is applied to achieve final roof segments. One limit of watershed transform is that, the gradient contours are not closed, so the first guess of threshold of watershed segmentation is not easy to choose: too small lead to sever over-segmentation that is difficult merge, and too large lead to under-segmentation and lose of roof segments. In the test data sites, the threshold is set to several values and the decision is made manually. Better solutions are needed in future.

Watershed transform exploits gradient jumps, which correspond to step edges in roof structure. Besides step edges, another character that distinguishes different connected roof surfaces is the roof surfaces' gradient orientation. In the proposed workflow, the information of gradient orientation is not used yet. The building roofs in test sites are mainly flat, or tilted with very small slopes, so that the proposed workflow is sufficient. Future study of the gradient orientation is needed. The original point clouds are noisy, while after preprocessing and DSM denoising, if the inclination difference of connected roof surfaces is small, they might not be separable. The minimal separable inclination angle is of interest.

For now, only several large buildings are separately reconstructed. Small buildings with multi-roof structures are hard to reconstruct. The minimal size or the scale of multi-roof building that can be reconstructed is not yet explored.

6.1.3 Discussion of Reconstruction

In reconstruction step, minimal bounding ellipse selects roof segments that fit the shape, thus simplified regularization for these segments. The "ellipse" here is actually a shape model, and the advantage of "minimal regularization effort" in model driven approach is clearly shown. More simple shape models can be added, such as semi-circle. Moreover, combinations of simple shapes woule also be interesting.

The proposed quartree regularization modified with average polygon complexity, thus avoid the change of topological relationships between adjacent segments. In this approach, since all boundary pixels smaller than minimal size are assigned to foreground, small structures which should be deleted are actually amplified. To improve the results, up-sampling the image might be helpful. In this case, one pixel is up-sampled into several pixels, so that area controlled by each pixel is decreased, and results from quadtree regularization will be better.

6.2 Discussion of more general questions

6.2.1 Categories of existing methods: which is more suitable?

DSM or point clouds

Efforts have been made to reconstruct building models from both DSM segmentation and poing clouds segmentation. In LiDAR community, DSM segmentation based approaches are mostly researched in the early years of LiDAR building reconstruction, when LiDAR points were not available in sufficient point density to apply region-growing. From this aspect, DSM based segmentation is more suitable for TomoSAR point clouds.

Data driven or Model driven

In recent years, more authors use data driven approaches mainly because the point density and accuracy of input data are increased, so that and topological relationships between roof segments can be solved. While From this aspect, for TomoSAR point

clouds, data driven approach is not enough to reconstruct building models in larger area, where point density variation is large. Pure model driven approach requires prior knowledge of study area to define model library thus is not adaptive to different data sites. Combining data driven and model driven is a good way, i.e., data driven as main approach to process areas with large point density, and model driven as constrains to process the rest of data. In fact, the reconstruction results of dense area can be used as feedback to build the model library, so that areas covered with less amount of points may be reconstructed, and the effort of building model library is decreased since a "leaning" process is included.

In the proposed workflow, ellipse fitting is performed in reconstruction step, as ellipse shape is assumed to be in the "shape library" of building roofs. While the "shape library" need to be expanded in order to better model the reality as well as reconstruct roofs with lower point density. One possible way to dynamically train the "shape library" is showed following, with an example.

Notice that in Figure 6.1, two similar shapes exist in upper part of the point clouds. One of them is nearly closed in 2D view, with rich points each facades. Since the two shapes are near, they might have very similar structure.

In Figure6.2, two nearby structures are similar. Morphlogical operations closed the structure, and the more regular one is used as a feature to match the other. Matched results together with thresholded SD map are taken as final results for shape matching. Cross-correlation is used.

The assumptions behind are: 1) In the same region, buildings usually are constructed in similar style; 2)Very tall buildings are usually small in 2D view, thus have relatively simple roof structures.

To built a shape library suitable for studying area, upper layer of the points should be separated out and learned. Each founded shape will be added to the library. Constrains should be made, such as similarity to accept a shape, and longest distance to considered as "near". If such library can be build, it will be very helpful in large area reconstruction.

(a) All points (b) up layer
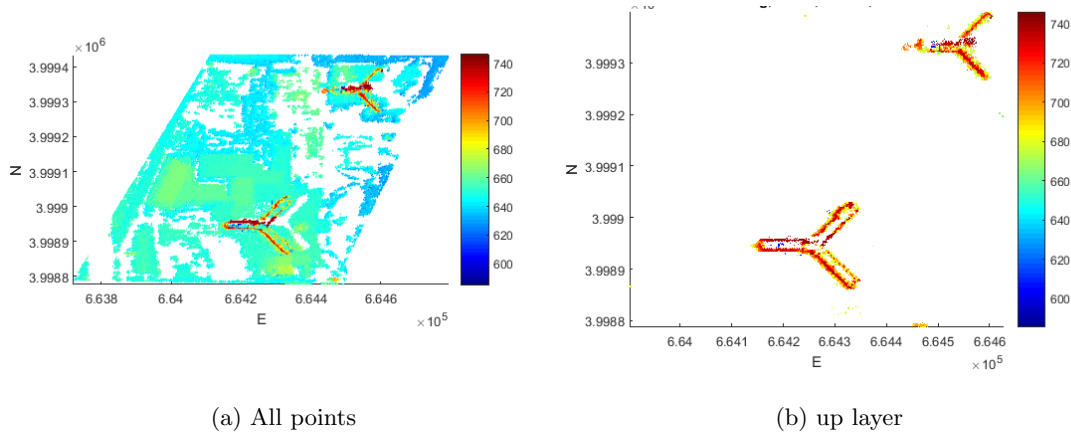
Figure 6.1 TomoSAR point clouds on at one site

6.2.2 Amplitude image

Figure 6.3 shows SAR amplitude image over the convention center in las Vegas. As can be seen, some bright lines and patches are clearly visible that are potentially edges of roofs or part of building masks.

The intention of using amplitude image is to extract lines/edges and jointly use this information together with the DSM- or point cloud-based workflows to improve the reconstruction. Preliminary experiments have been done in this regard, and are shown in following sections.

6.2.2.1 General steps in processing amplitude image

The amplitude image is first denoised by non-local mean algorithm. Then object of interest are extracted, and two independent methods are tested: 1)local adaptive thresholding, or 2)Hidden Markov Random Field Model and its Expectation-Maximization. Knowledge of building masks and facades are included as constrains in following morphological operations. Finally, the characters of detected objects are extracted using Hough Transform.
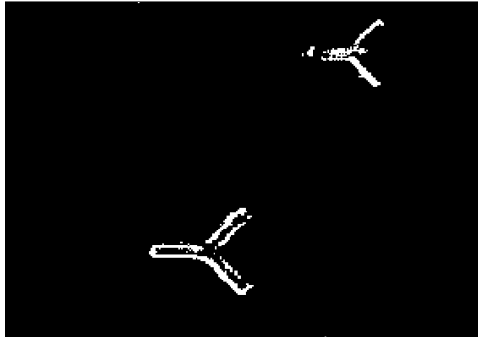
(a) thresholded SD of upper points
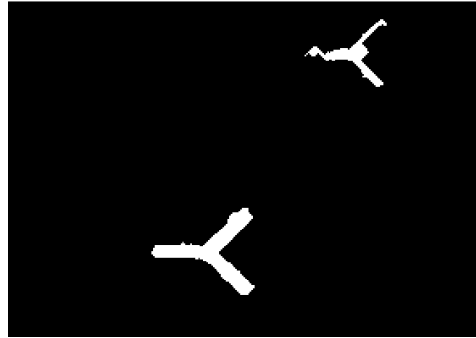
(b) closed shapes found
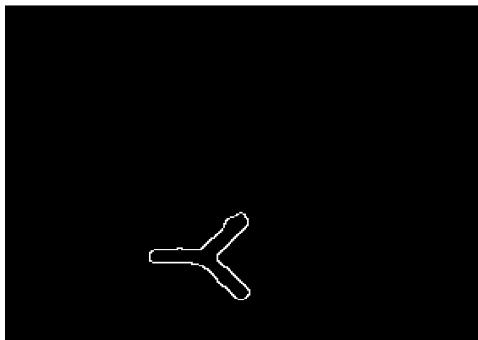
(c) contour of 1 found shape
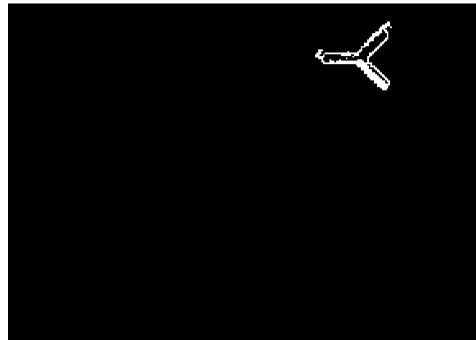
(d) feature matching result for another shape
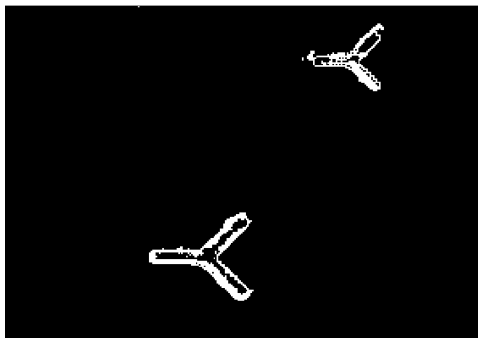
(e) Detected shapes with thresholded SD map
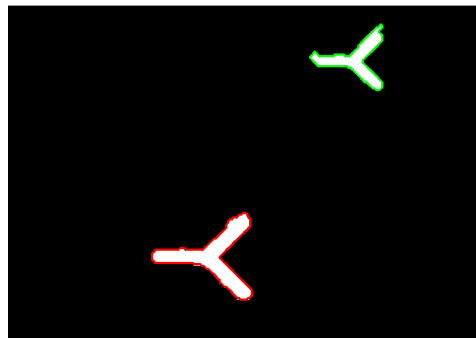
(f) Final shapes

Figure 6.2 Shape matching example

6.2.2.2 Object extraction

Two methods are test. They are both independent, and aiming at extract objects of interest.

Adaptive Thresholding

Thresholding is a commonly used technique in image processing which segment an image by setting pixels whose intensity values above a threshold to a foreground value and the rest to a background value. Conventional thresholding operators use a global threshold for all pixels, the threshold used in adaptive thresholding is changed over the image, depending on the local intensity change. The assumption is that smaller image regions are more likely to have approximately uniform illumination.

Local adaptive thresholding is used in proposed workflow. To finding the local threshold, the intensity values of the local neighborhood of each pixel are statistically examined. The statistic which is most appropriate depends largely on the input image. The mean of the local intensity distribution is used.

In local adaptive thresholding, the size of the neighborhood need to be large enough to cover sufficient foreground and background pixels, otherwise a poor threshold is chosen. On the other hand, choosing regions which are too large can violate the assumption of approximately uniform illumination. In this work, two neighborhood sizes A_1 and A_2 are chosen to extract potential building masks and facades of interest.

Hidden Markov Random Field Model and its Expectation-Maximization

Markov random fields (MRFs) have been widely used for computer vision problems. The HMRF-EM(hidden Markov random field and its expectation-maximization) framework was first proposed for segmentation of brain MR images[Zhang et al., 2001]. The used program is implemented by [Wang, 2012]. For detailed explain [Wang, 2012] is referred.

In the HMRF-EM framework, first an initial segmentation is generated using k-means clustering on the amplitude. Since we want bright edges, potential building masks, and the back ground, "k" is set to 3. The initial segmentation provides the initial labels, and

future refined in the HMRF.

6.2.2.3 Knowledge based Morphological Operations

Potential Building masks

First step after object extraction is to remove foreground objects that are smaller than typical building size in I_{AT1}. Opening by reconstruction is used to get I_{AT1O}. To remove unwanted objects in background, first image is converted to its compliment image I_{AT1C}, then two constrains are defined, which referred as additional "Knowledge".

The first constrain is the size of segment. Segments whose size is smaller than T_{aMin} should be removed, while those with size $>= T_{aMax}$ should be kept.

The second constrain is the segment complexity. Recall the process of data gaps classification in preprocessing of DSM processing, the data gaps inside building masks tend to be more complex than those who are outside of building masks, and usually adjacent to one facade. The segment complexity of all segments whose size fulfill first constrain is computed, and segments whose complexity $>= T_c$ should be removed.I_{AT1K} is the image after applying "knowledge".

The image is converted to its compliment image I_{AT1BM} to get the potential building masks.

Potential Facades

For the results from adaptive thresholding, since there are plenty of details in thresholded image I_{AT2}, the orientation of building is introduced as the first additional "Knowledge". While it is not needed for the results from HMRF.

The second "Knowledge" is the ratio of the length and the width of the segments. The assumption here is that the potential facades are narrow objects.

The orientation of building means the main directions of the buildings. To estimate it, first a global threshold is applied to the image to show only high intensity pixels. The threshold is set to a large value, so that only some of the facades will remain, because

of the edge effect. Then for each connected component S_i, its tangent direction $\vec{t_i}$ and normal direction $\vec{n_i}$ is computed. All the tangent directions and normal directions are the main directions. To simplify the following processes, the directions with small difference are merged to number N. $i = 1...N$.

Morphological structuring line elements are created using the detected main directions, described by the length of the line element L_i and the orientation θ_i of it. The number of the morphological structuring elements N is the number of main directions N. Then, each element is applied to open the image I_{AT2}, followed by an area open operator to remove small objects whose size $< T_f$, resulting in I_{AT2i}. All I_{AT2i} are put together to get I_{AT2N}, the detected potential facades.

Now the second "Knowledge" is applied, to remove objects whose ratio of the length and the width does not fulfill the constrain. The final image is denoted as I_{AT2F}.

6.2.2.4 Hough Transform

To get vectors of detected objects, so that the extracted objects can be transformed to the UTM (Universal Transverse Mercator) coordinate which contains point clouds, Hough transform is applied to both I_{AT1BM} and I_{AT2F} (explained in subsection 3.4.1).

The advantage of Hough transform is that it only extract straight lines, which is much more accurate than simplifying traced boundaries in raster images.

6.2.2.5 Preliminary experiment results

Input data

The input data is SAR ampiltude image over test site 1 at the convention center of las Vegas. The input data is shown in Figure 6.3.

Results

The results of extracting object using adaptive thresholding are shown in Figure 6.3 to Figure 6.8. The detection results of Potential Building Masks and Facades are shown in
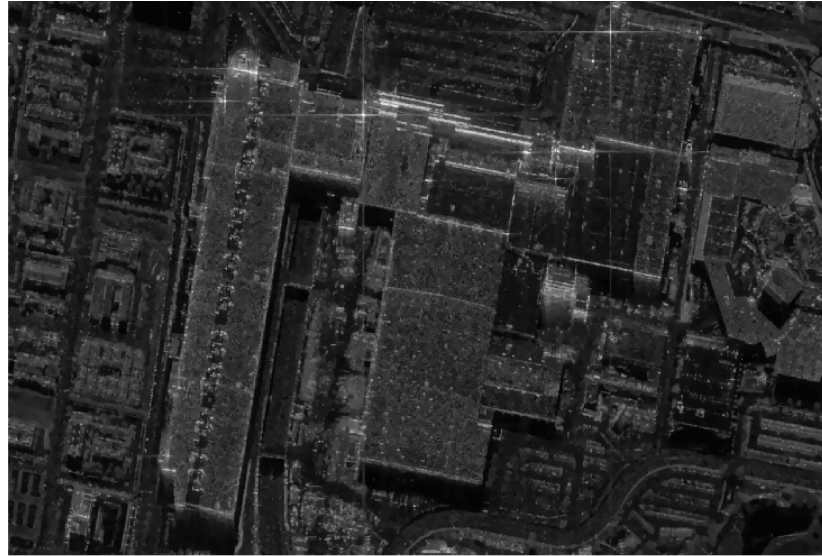
Figure 6.3 Amplitude Image of Convention Center, las Vegas

Figure 6.4, Figure 6.5.

The results of extracting object using HMRF are shown in Figure 6.6 to Figure 6.7.

Figure 6.8 shows the extracted potential facades and building masks together.

The information extracted from amplitude image is not yet combined with reconstruction results from TomoSAR point clouds. SAR coordinates can be transformed to UTM coordinate, if height is introduced. When putting the extracted straight lines together with reconstructed building models from TomoSAR point clouds, the straight lines can be used to increase accuracy of the building models. If the quality of extracted building masks and line segments is high, "Cell Decomposition" is also worth trying. However, because of the layovers and shadows in SAR images, building masks and line segments is not very easy to identify. For now, only the results of building regions without strong layovers are looking good.
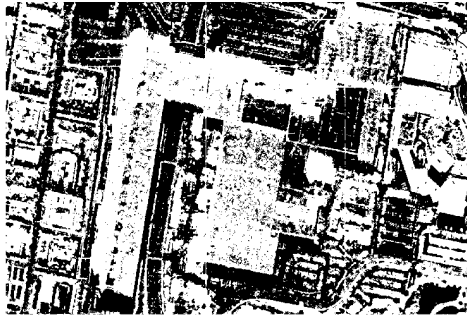
(a) Adaptive thresholding result

(b) Opening by reconstrucion

(c) Compliment image of (b)
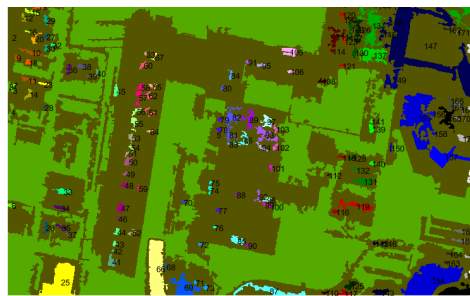
(d) All Segments bigger than T_a, labeled
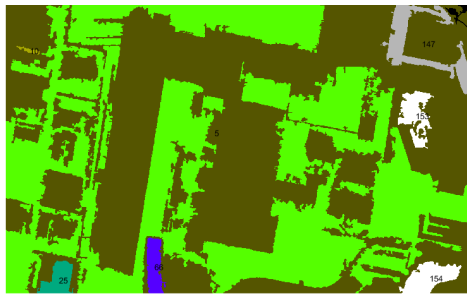
(e) Segments fulfill constrains

(f) Results for potential Building Masks extracted from Amplitude image
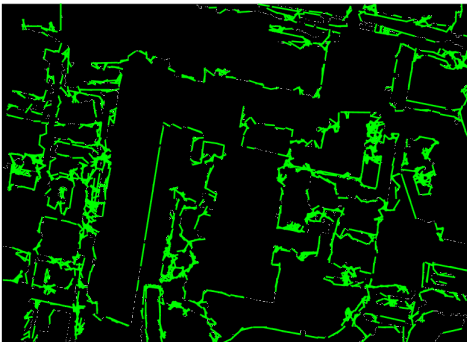
(g) Hough lines(green), setting 2
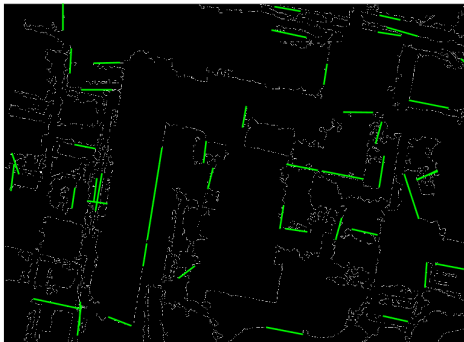
(h) Hough lines(green), setting 1

Figure 6.4 Results for potential Building Masks extracted from Amplitude image, Using Adaptive thresholding.

(a) Amplitude image
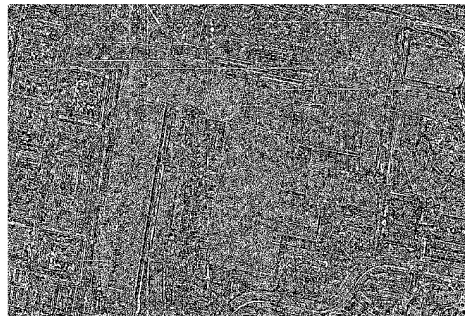
(b) Adaptive thresholding result
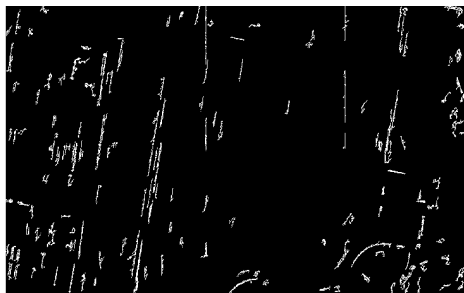
(c) Detected objects with orientation 1

(d) Detected objects with orientation 2
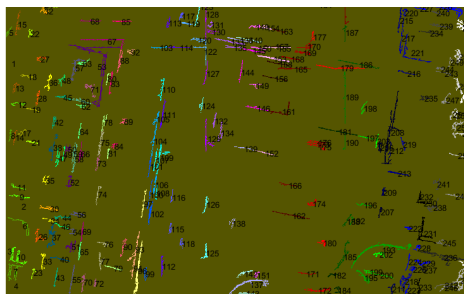
(e) All detected objects, labeled
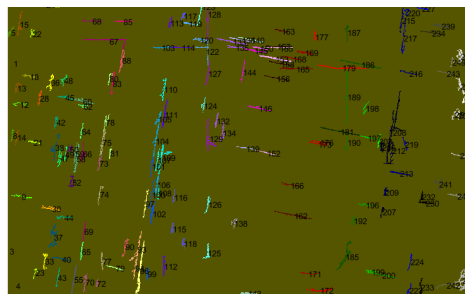
(f) Objects fulfill constrains
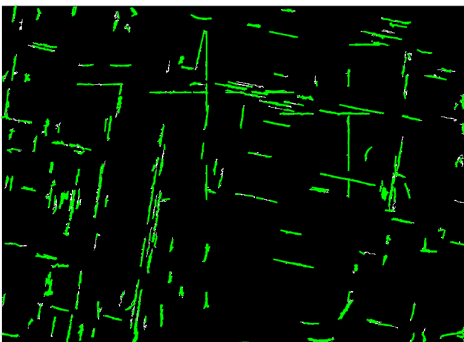
(g) Hough lines(green), setting 1
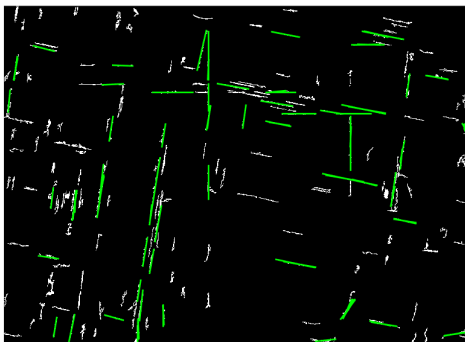
(h) Hough lines(green), setting 2

Figure 6.5 Results for potential Facades extracted from Amplitude image, using adaptive thresholding

(a) K-means segment result, k=3
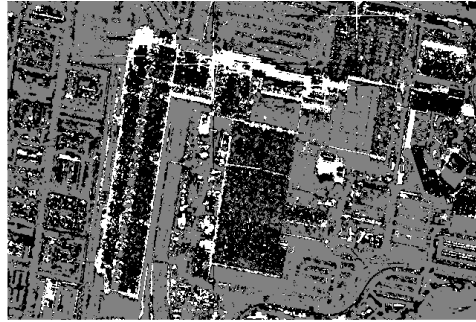
(b) Refined segments by HMRF.

(c) Two classes of segments are put together as foreground

(d) Complement image of (c), small segments are removed.
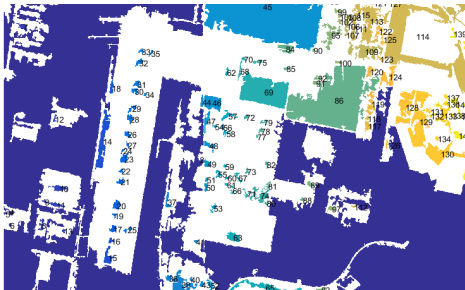
(e) Complement image of (c), small segments are removed.

(f) Segments fulfill constrains

(g) Results for potential Building Masks
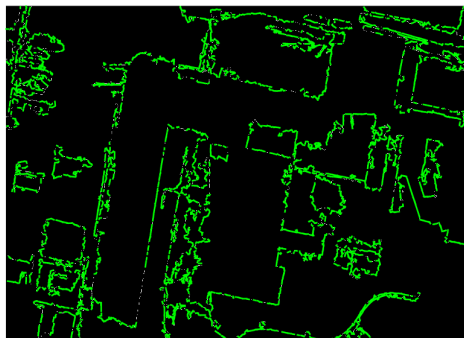
(h) Hough lines

Figure 6.6 Results for potential Building Masks extracted from Amplitude image, Using HMRF.
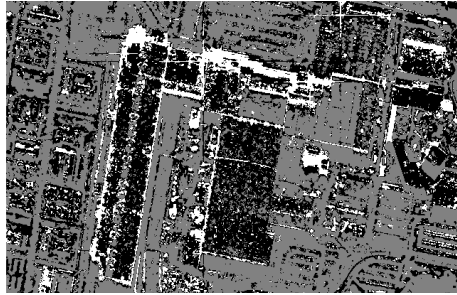
(a) Refined segments by HMRF.
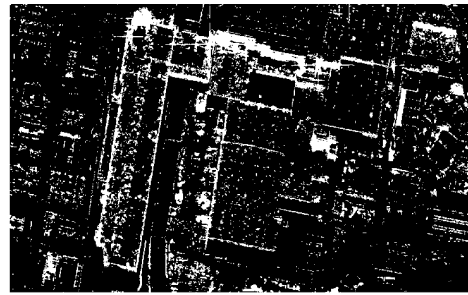
(b) One class of segments is considered as foreground
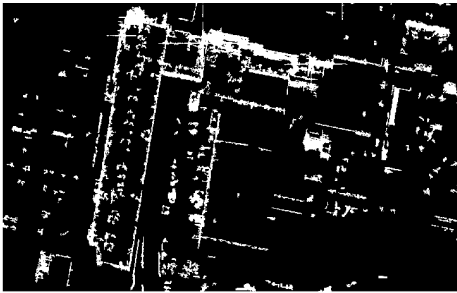
(c) Remove small segments from (b)
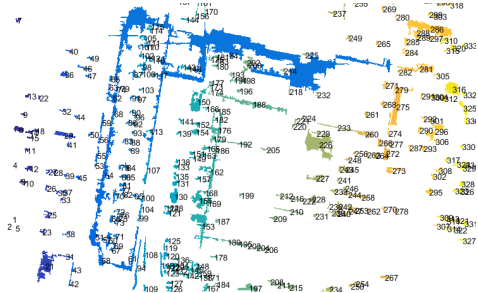
(d) Label connected components in (c)
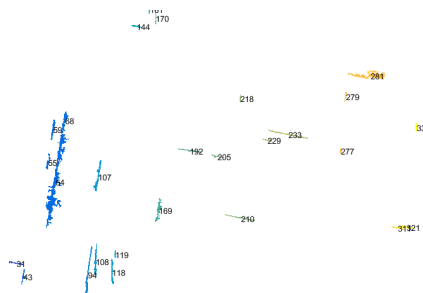
(e) Remove segments based on knowledge of facades
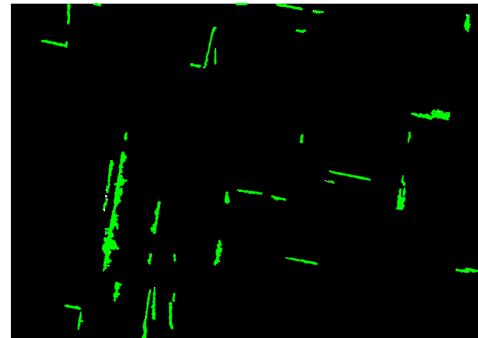
(f) Segments fulfill constrains

Figure 6.7 Results for potential Facades extracted from Amplitude image, Using HMRF.
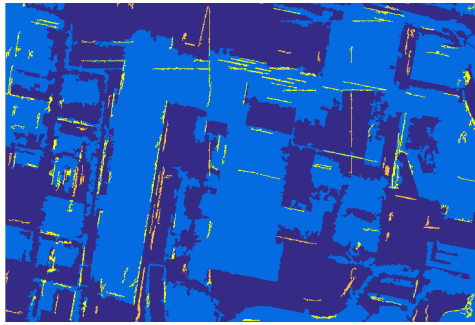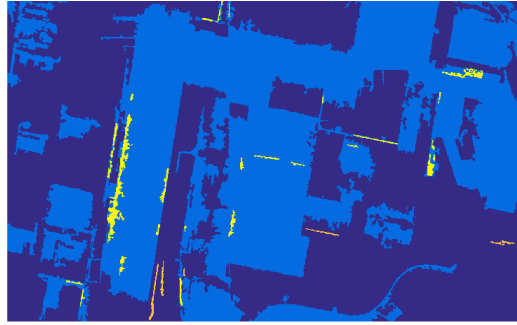
(a) From adaptive thresholding   (b) From HMRF

Figure 6.8 Results for potential facades and building Masks

7 Conclusion and Outlook

The research explored in this thesis aimed at exploring the potential of building roof reconstruction using unstructured spaceborne point clouds generated by tomographic processing of meter resolution TerraSAR-X spotlight image stacks. Two complimentary data-driven workflows, one working over rasterized DSM and other utilizing direct 3-D points, have been proposed. Both workflows are modular and reconstructs building models by segmenting individual roof surfaces. Taking into the account the computational efficiency, DSM based workflow is however preferred and has been detailed in this thesis.

Briefly, the developed DSM based workflow first extracts the building regions/pixels via ground filtering approach from interpolated and denoised DSM - generated by exploiting height and contextual facade information. Then, a novel watershed based segmentation via constrained merging scheme is employed is developed to segment individual roof surfaces. Quadtree regularization under constrains of polygon complexity is later utilized to simplify the outlines of these segments. Finally height is associated to each reconstructed polygonal segment to generate prismatic 3-D building model. Results are demonstrated over two large building complexes in the city of Las Vegas using TomoSAR point clouds generated from data stacks acquired from ascending orbit only.

In addition to exploitation of spatial geometric information, preliminary experiments using SAR amplitude (backscattering) information have also been done where local adapative thresholding and HMRF with its expectation-maximization algorithm are utilizedsed to extract possible building masks and facades by defining the sizes and orientations of different objects. The results have shown that, the bright edges and patches in amplitude images can be extracted, but the interpretation is quite challenging, especially in areas

with severe layovers. The edges can be potential facades where edge-effecct is obvious, while the near-edge regions may correspond to either parts of roofs or ground. These regions are strongly affected by geometrical effects caused by side-looking SAR consequently rendering the extraction and 3-D reconstruction of buildings using amplitude only very difficult.

In future, the emphasis would be to jointly fuse the spatial geometric information together with the corresponding SAR amplitudes to develop a more robust roof reconstruction methodology. The developed approach would then be tested over large areas containing buildings with different sizes/scales. In addition to this, the potential of taking a more general prismatic model to more specific polyhedral roof modeling could also be explored. In this direction, possible steps would to adopt a model-based approach where a defined grammar with a library of primitives representing different roof shapes could be employed.

List of Figures

List of Tables

Bibliography

[Alharthy and Bethel, 2002] Alharthy, A. and Bethel, J. (2002). Heuristic filtering and 3d feature extraction from lidar data. *International Archives of Photogrammetry Remote Sensing and Spatial Information Sciences*, 34(3/A):29–34.

[Ameri and Fritsch, 2000] Ameri, B. and Fritsch, D. (2000). Automatic 3d building reconstruction using plane-roof structures. *ASPRS, Washington DC*.

[Arefi and Reinartz, 2013] Arefi, H. and Reinartz, P. (2013). Building reconstruction using dsm and orthorectified images. *Remote Sensing*, 5(4):1681–1703.

[Atlas, 2014] Atlas, B. E. (2014). Building and vegetation heights (edition 2014).

[Auer et al., 2011a] Auer, S., Gernhardt, S., and Bamler, R. (2011a). Ghost persistent scatterers related to multiple signal reflections. *Geoscience and Remote Sensing Letters, IEEE*, 8(5):919–923.

[Auer et al., 2011b] Auer, S. J. et al. (2011b). *3D synthetic aperture radar simulation for interpreting complex urban reflection scenarios*. PhD thesis, Technische Universität München.

[Awrangjeb and Fraser, 2014] Awrangjeb, M. and Fraser, C. S. (2014). Automatic segmentation of raw lidar data for extraction of building roofs. *Remote Sensing*, 6(5):3716–3751.

[Becker and Haala, 2009] Becker, S. and Haala, N. (2009). Grammar supported facade reconstruction from mobile lidar mapping. In *ISPRS Workshop, CMRT09-City Models, Roads and Traffic*, volume 38, page 13.

[Biljecki et al., 2014] Biljecki, F., Ledoux, H., and Stoter, J. (2014). Error propagation

in the computation of volumes in 3d city models with the monte carlo method. In *ISPRS/IGU Joint International Conference on Geospatial Theory, Processing, Modelling and Applications, Toronto, Canada*, pages 31–39.

[Brenner, 2000] Brenner, C. (2000). Towards fully automatic generation of city models. *International Archives of Photogrammetry and Remote Sensing*, 33(B3/1; PART 3):84–92.

[Brenner, 2001] Brenner, C. (2001). City models–automation in research and practice.

[Brenner, 2005] Brenner, C. (2005). Building reconstruction from images and laser scanning. *International Journal of Applied Earth Observation and Geoinformation*, 6(3):187–198.

[Brenner and Haala, 1998] Brenner, C. and Haala, N. (1998). Fast production of virtual reality city models. *International Archives of Photogrammetry and Remote Sensing*, 32(part 4):77–84.

[Brinkhoff et al., 1995] Brinkhoff, T., Kriegel, H.-P., Schneider, R., and Braun, A. (1995). Measuring the complexity of polygonal objects. In *ACM-GIS*, page 109. Citeseer.

[Chen et al., 2014] Chen, Y., Cheng, L., Li, M., Wang, J., Tong, L., and Yang, K. (2014). Multiscale grid method for detection and reconstruction of building roofs from airborne lidar data. *Selected Topics in Applied Earth Observations and Remote Sensing, IEEE Journal of*, 7(10):4081–4094.

[Dabov et al., 2007] Dabov, K., Foi, A., Katkovnik, V., and Egiazarian, K. (2007). Image denoising by sparse 3-d transform-domain collaborative filtering. *Image Processing, IEEE Transactions on*, 16(8):2080–2095.

[D'Hondt et al., 2012] D'Hondt, O., Guillaso, S., and Hellwich, O. (2012). Automatic extraction of geometric structures for 3d reconstruction from tomographic sar data. In *2012 IEEE International Geoscience and Remote Sensing Symposium*.

[Digabel and LANTUEJOUL, 1977] Digabel, H. and LANTUEJOUL, C. (1977). Iterative algorithms, special issues of practical metallography, vol. 8.

[Dorninger and Pfeifer, 2008] Dorninger, P. and Pfeifer, N. (2008). A comprehensive automated 3d approach for building extraction, reconstruction, and regularization from airborne laser scanning point clouds. *Sensors*, 8(11):7323–7343.

[Durupt and Taillandier, 2006] Durupt, M. and Taillandier, F. (2006). Automatic building reconstruction from a digital elevation model and cadastral data: an operational approach. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 36(Part 3):6.

[Elberink and Vosselman, 2009] Elberink, S. O. and Vosselman, G. (2009). Building reconstruction by target based graph matching on incomplete laser data: Analysis and limitations. *Sensors*, 9(8):6101–6118.

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

[Frueh et al., 2005] Frueh, C., Jain, S., and Zakhor, A. (2005). Data processing algorithms for generating textured 3d building facade meshes from laser scans and camera images. *International Journal of Computer Vision*, 61(2):159–184.

[Galvanin and Poz, 2012] Galvanin, E. A. d. S. and Poz, A. P. D. (2012). Extraction of building roof contours from lidar data using a markov-random-field-based approach. *Geoscience and Remote Sensing, IEEE Transactions on*, 50(3):981–987.

[Gröger et al., 2008] Gröger, G., Kolbe, T. H., Czerwinski, A., Nagel, C., et al. (2008). Opengis city geography markup language (citygml) encoding standard, version 1.0. 0.

[Gruen et al., 2012] Gruen, A., Baltsavias, E., and Henricsson, O. (2012). *Automatic extraction of man-made objects from aerial and space images (II)*. Birkhäuser.

[Haala and Brenner, 1999] Haala, N. and Brenner, C. (1999). Extraction of buildings and trees in urban environments. *ISPRS Journal of Photogrammetry and Remote Sensing*, 54(2):130–137.

[Haala et al., 1998] Haala, N., Brenner, C., and Anders, K.-H. (1998). 3d urban gis

from laser altimeter and 2d map data. *International Archives of Photogrammetry and Remote Sensing*, 32:339–346.

[Haala and Kada, 2010] Haala, N. and Kada, M. (2010). An update on automatic 3d building reconstruction. *ISPRS Journal of Photogrammetry and Remote Sensing*, 65(6):570–580.

[Haralick et al., 1987] Haralick, R. M., Sternberg, S. R., and Zhuang, X. (1987). Image analysis using mathematical morphology. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, (4):532–550.

[Kaartinen et al., 2012] Kaartinen, H., Hyyppä, J., Kukko, A., Jaakkola, A., and Hyyppä, H. (2012). Benchmarking the performance of mobile laser scanning systems using a permanent test field. *Sensors*, 12(9):12814–12835.

[Kada and McKinley, 2009] Kada, M. and McKinley, L. (2009). 3d building reconstruction from lidar based on a cell decomposition approach. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 38(Part 3):W4.

[Kada and Wichmann, 2012] Kada, M. and Wichmann, A. (2012). Sub-surface growing and boundary generalization for 3d building reconstruction. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences I-3*, pages 233–238.

[Krüger and Kolbe, 2012] Krüger, A. and Kolbe, T. (2012). Building analysis for urban energy planning using key indicators on virtual 3d city models—the energy atlas of berlin. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 39(B2):145–150.

[Lafarge and Mallet, 2012] Lafarge, F. and Mallet, C. (2012). Creating large-scale city models from 3d-point clouds: a robust approach with hybrid representation. *International journal of computer vision*, 99(1):69–85.

[Lombardo, 2004] Lombardo, P. (2004). A multichannel spaceborne radar for the cosmoskymed satellite constellation. In *Aerospace Conference, 2004. Proceedings. 2004 IEEE*, volume 1. IEEE.

[Maas and Vosselman, 1999] Maas, H.-G. and Vosselman, G. (1999). Two algorithms for extracting building models from raw laser altimetry data. *ISPRS Journal of photogrammetry and remote sensing*, 54(2):153–163.

[Matei et al., 2008] Matei, B. C., Sawhney, H. S., Samarasekera, S., Kim, J., and Kumar, R. (2008). Building segmentation for densely built urban regions using aerial lidar data. In *Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on*, pages 1–8. IEEE.

[Morgan and Tempfli, 2000] Morgan, M. and Tempfli, K. (2000). Automatic building extraction from airborne laser scanning data. *International Archives of Photogrammetry and Remote Sensing*, 33(B3/2; PART 3):616–623.

[Pingel et al., 2013] Pingel, T. J., Clarke, K. C., and 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.

[Pitz and Miller, 2010] Pitz, W. and Miller, D. (2010). The terrasar-x satellite. *Geoscience and Remote Sensing, IEEE Transactions on*, 48(2):615–622.

[Poullis and You, 2009] Poullis, C. and You, S. (2009). Automatic reconstruction of cities from remote sensor data. In *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on*, pages 2775–2782. IEEE.

[Pu and Vosselman, 2009] Pu, S. and Vosselman, G. (2009). Knowledge based reconstruction of building models from terrestrial laser scanning data. *ISPRS Journal of Photogrammetry and Remote Sensing*, 64(6):575–584.

[Rau and Lin, 2011] Rau, J.-Y. and Lin, B.-C. (2011). Automatic roof model reconstruction from als data and 2d ground plans based on side projection and the tmr algorithm. *ISPRS Journal of Photogrammetry and Remote Sensing*, 66(6):S13–S27.

[Reigber and Moreira, 2000] Reigber, A. and Moreira, A. (2000). First demonstration of airborne sar tomography using multibaseline l-band data. *Geoscience and Remote Sensing, IEEE Transactions on*, 38(5):2142–2152.

[Roerdink and Meijster, 2000] Roerdink, J. B. and Meijster, A. (2000). The watershed

transform: Definitions, algorithms and parallelization strategies. *Fundamenta informaticae*, 41(1, 2):187–228.

[Rottensteiner and Briese, 2002] Rottensteiner, F. and Briese, C. (2002). A new method for building extraction in urban areas from high-resolution lidar data. *International Archives of Photogrammetry Remote Sensing and Spatial Information Sciences*, 34(3/A):295–301.

[Rottensteiner and Briese, 2003] Rottensteiner, F. and Briese, C. (2003). *Automatic generation of building models from LIDAR data and the integration of aerial images*. Citeseer.

[Rottensteiner et al., 2012] Rottensteiner, F., Sohn, G., Jung, J., Gerke, M., Baillard, C., Benitez, S., and Breitkopf, U. (2012). The isprs benchmark on urban object classification and 3d building reconstruction. *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci*, 1:3.

[Samet, 1984] Samet, H. (1984). The quadtree and related hierarchical data structures. *ACM Computing Surveys (CSUR)*, 16(2):187–260.

[Schwalbe et al., 2005] Schwalbe, E., Maas, H.-G., and Seidel, F. (2005). 3d building model generation from airborne laser scanner data using 2d gis data and orthogonal point cloud projections. *Proceedings of ISPRS WG III/3, III/4*, 3:12–14.

[Shahzad and Zhu, 2015a] Shahzad, M. and Zhu, X. (2015a). Reconstruction of building footprints using spaceborne tomosar point clouds. *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 1:385–392.

[Shahzad and Zhu, 2015b] Shahzad, M. and Zhu, X. X. (2015b). Robust reconstruction of building facades for large areas using spaceborne tomosar point clouds. *Geoscience and Remote Sensing, IEEE Transactions on*, 53(2):752–769.

[Sirmacek et al., 2012] Sirmacek, B., Taubenböck, H., Reinartz, P., and Ehlers, M. (2012). Performance evaluation for 3-d city model generation of six different dsms from air-and spaceborne sensors. *Selected Topics in Applied Earth Observations and Remote Sensing, IEEE Journal of*, 5(1):59–70.

[Sohn et al., 2008] Sohn, G., Huang, X., and Tao, V. (2008). Using a binary space partitioning tree for reconstructing polyhedral building models from airborne lidar data. *Photogrammetric Engineering & Remote Sensing*, 74(11):1425–1438.

[Tarsha-Kurdi et al., 2008] Tarsha-Kurdi, F., Landes, T., Grussenmeyer, P., et al. (2008). Extended ransac algorithm for automatic detection of building roof planes from lidar data. *The photogrammetric journal of Finland*, 21(1):97–109.

[Verma et al., 2006] Verma, V., Kumar, R., and Hsu, S. (2006). 3d building detection and modeling from aerial lidar data. In *Computer Vision and Pattern Recognition, 2006 IEEE Computer Society Conference on*, volume 2, pages 2213–2220. IEEE.

[Visvalingam and Whyatt, 1993] Visvalingam, M. and Whyatt, J. (1993). Line generalisation by repeated elimination of points. *The Cartographic Journal*, 30(1):46–51.

[Vosselman et al., 2001] Vosselman, G., Dijkman, S., et al. (2001). 3d building model reconstruction from point clouds and ground plans. *International archives of photogrammetry remote sensing and spatial information sciences*, 34(3/W4):37–44.

[Vosselman et al., 2004] Vosselman, G., Gorte, B. G., Sithole, G., and Rabbani, T. (2004). Recognising structure in laser scanner point clouds. *International archives of photogrammetry, remote sensing and spatial information sciences*, 46(8):33–38.

[Wang, 2012] Wang, Q. (2012). Hmrf-em-image: implementation of the hidden markov random field model and its expectation-maximization algorithm. *arXiv preprint arXiv:1207.3510*.

[Wang, 2013] Wang, R. (2013). 3d building modeling using images and lidar: A review. *International Journal of Image and Data Fusion*, 4(4):273–292.

[You et al., 2003] You, S., Hu, J., Neumann, U., and Fox, P. (2003). Urban site modeling from lidar. *Computational Science and Its Applications—ICCSA 2003*, pages 987–987.

[Zhang et al., 2003] Zhang, K., Chen, S.-C., Whitman, D., Shyu, M.-L., Yan, J., and Zhang, C. (2003). A progressive morphological filter for removing nonground measurements from airborne lidar data. *Geoscience and Remote Sensing, IEEE Transactions on*, 41(4):872–882.

[Zhang et al., 2001] Zhang, Y., Brady, M., and Smith, S. (2001). Segmentation of brain mr images through a hidden markov random field model and the expectation-maximization algorithm. *Medical Imaging, IEEE Transactions on*, 20(1):45–57.

[Zhou and Neumann, 2010] Zhou, Q.-Y. and Neumann, U. (2010). 2.5 d dual contouring: A robust approach to creating building models from aerial lidar point clouds. In *Computer Vision–ECCV 2010*, pages 115–128. Springer.

[Zhu et al., 2008] Zhu, X., Adam, N., and Bamler, R. (2008). First demonstration of space-borne high resolution sar tomography in urban environment using terrasar-x data. *Proceedings of CEOS SAR Workshop on Calibration and Validation*.

[Zhu and Bamler, 2010] Zhu, X. X. and Bamler, R. (2010). Very high resolution space-borne sar tomography in urban environment. *Geoscience and Remote Sensing, IEEE Transactions on*, 48(12):4296–4308.

[Zhu and Shahzad, 2014] Zhu, X. X. and Shahzad, M. (2014). Facade reconstruction using multiview spaceborne tomosar point clouds. *Geoscience and Remote Sensing, IEEE Transactions on*, 52(6):3541–3552.

[Zhu et al., 2013] Zhu, X. X., Wang, Y., Gernhardt, S., and Bamler, R. (2013). Tomo-genesis: Dlr's tomographic sar processing system. In *Urban Remote Sensing Event (JURSE), 2013 Joint*, pages 159–162. IEEE.