FIRST PRISMATIC BUILDING MODEL RECONSTRUCTION FROM TOMOSAR POINT CLOUDS

Y. Sun[a,c], M. Shahzad[a], X. X. Zhu[a,b,*]

[a] Signal Processing in Earth Observation (SiPEO), Technical University of Munich, 80333 Munich, Germany
[b] Remote Sensing Technology Institute (IMF), German Aerospace Center (DLR), Oberpfaffenhofen, 82234 Weßling, Germany
[c] School of Resource and Environmental Science, Wuhan University, 430079, Wuhan, China
yao.sun@tum.de, muhammad.shahzad@tum.de, xiao.zhu@dlr.de

Commission III, WG III/2

ABSTRACT:

This paper demonstrates for the first time the potential of explicitly modelling the individual roof surfaces to reconstruct 3-D prismatic building models using spaceborne tomographic synthetic aperture radar (TomoSAR) point clouds. The proposed approach is modular and works as follows: it first extracts the buildings via DSM generation and cutting-off the ground terrain. The DSM is smoothed using BM3D denoising method proposed in (Dabov et al., 2007) and a gradient map of the smoothed DSM is generated based on height jumps. Watershed segmentation is then adopted to oversegment the DSM into different regions. Subsequently, height and polygon complexity constrained merging is employed to refine (i.e., to reduce) the retrieved number of roof segments. Coarse outline of each roof segment is then reconstructed and later refined using quadtree based regularization plus zig-zag line simplification scheme. Finally, height is associated to each refined roof segment to obtain the 3-D prismatic model of the building. The proposed approach is illustrated and validated over a large building (convention center) in the city of Las Vegas using TomoSAR point clouds generated from a stack of 25 images using Tomo-GENESIS software developed at DLR.

1. INTRODUCTION

Three dimensional (3-D) building models reconstruction in urban areas has been a hot topic in remote sensing, photogrammetry, compute vision for more than two decades (Gruen et al., 1997, Rottensteiner et al., 2012). 3-D city models are used widely for urban planning, change detection, commercial and public sector simulations for environmental research, location based services and many others (Verma et al., 2006, Brenner, 2005, Rau and Lin, 2011). Numerous research papers have been published in this context that provides different reconstruction methods using variety of data sources.

Typical data sources for reconstructing 3-D building models include optical images (airborne or spaceborne), airborne LiDAR point clouds, terrestrial LiDAR point clouds and close range images. In addition to them, recent advances in very high resolution synthetic aperture radar imaging together with its key attributes such as self-illumination and all-weather capability, have also attracted attention of many remote sensing analysts in characterizing urban objects such as buildings. However, SAR projects a 3-D scene onto two native coordinates i.e., "range" and "azimuth". In order to fully localize a point in 3-D, advanced interferometric SAR (InSAR) techniques are required that process stack(s) of complex-valued SAR images to retrieve the lost third dimension (i.e., the "elevation" coordinate). Among other InSAR methods, SAR tomography (TomoSAR) is the ultimate way of 3-D SAR imaging. By exploiting stack(s) of SAR images taken from slightly different positions, it builds up a synthetic aperture in elevation that enables the retrieval of precise 3-D position of dominant scatterers within one azimuth-range SAR image pixel. TomoSAR processing of very high resolution data of urban areas provided by modern satellites (e.g., TerraSAR-X, TanDEM-X

and CosmoSkyMED) allows us to reconstruct points (scatterers) with a density of around 600,000-1,000,000 points/km^2 (Zhu and Bamler, 2012). Geocoding these scatterer positions from SAR geometry to world (UTM) coordinates provide 3-D or even 4-D (space-time) point clouds of the illuminated area. Object reconstruction from these TomoSAR point clouds can greatly support the reconstruction of dynamic city models that could potentially be used to monitor and visualize the dynamics of urban infrastructures in very high level of details.

Although few approaches, e.g., (Guillaso et al., 2015, Guillaso et al., 2013, D'Hondt et al., 2012), aiming towards information extraction exist, 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 (Zhu and Shahzad, 2014, Shahzad and Zhu, 2015, Shahzad and Zhu, 2016) while TomoSAR point clouds generated over urban and vegetation areas using airborne SAR datasets have been explored in (D'Hondt et al., 2012, Schmitt et al., 2015) respectively.

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 façade information (due to the side looking geometry), this paper demonstrates for the first time the potential of explicitly modelling the individual roof surfaces to reconstruct 3-D prismatic building models from TomoSAR point clouds.

The proposed methods work on Digital Surface Models (DSM) that are interpolated from point clouds, after extracting building points from the whole data. Comparing to methods that directly work on point clouds, which normally contains plane fitting (Sohn et al., 2008, Tarsha-Kurdi et al., n.d., Ameri and Fritsch, 2000, Brenner, 2000) or region growing (Elberink and Vossel-

---

[*]Corresponding author

man, 2009, Dorninger and Pfeifer, 2008, Vosselman et al., 2004), DSM based methods have higher computation efficiency due to simpler data structure, and can potentially benefit from many digital image processing techniques. Though loses of original data precision exist due to the interpolation process, many author use DSM for building reconstruction (Rottensteiner and Briese, 2003, Forlani et al., 2006, Galvanin and Poz, 2012, Chen et al., 2014).

The paper is structured as follows. Section 2 presents the proposed approach. The evaluation and validation of the approach is provided in Section 3. Section 4 outlines our conclusion and possible future improvements.

2. PROPOSED METHODS

The proposed methods consists of four main steps, including pre-processing of point clouds, nDSM generation from interpolating point clouds, segmentation of nDSM over building area, and 3-D model reconstruction. Detailed explanations are given in the following sections.

2.1 Preprocessing

In order to obtain better interpolated DSM, preprocessing is necessary. Two steps are involved: local height smoothing and Ground filtering. Local height smoothing is designed to reduce the influence of facade points and non-surface point in interpolating DSM. For this purpose, the whole dataset area is divided into grid cells of size $L_G \times L_G$. Then, for each grid, heights of all points inside are assigned to the mean height of highest m points. An example is shown in Figure 1.

Ground filtering is then used to separate ground points and above-ground points. In this work, the Simple Morphological Filter (SMRF) (Pingel et al., 2013) is employed. Above-ground points are considered as building points, since temporarily incoherent objects, e.g., vegetation, water body, are not contained in TomoSAR point clouds, and other small above-ground objects are ignored.

2.2 nDSM generation

The point clouds are now ready to be interpolated to generate normalized DSM (nDSM). nDSM is obtained by subtracting Digital Terrain Model (DTM) from DSM. DTM is interpolated using ground points extracted in ground filtering step, while the DSM is obtained by interpolating the whole point clouds after local height smoothing. Building masks are needed to find the building areas in nDSM. In this work, building areas are obtained by height thresholding of the nDSM.

Prior to extraction and segmentation of different roof surfaces, it is necessary to further smooth/denoise the generated nDSM. To this end, Block-matching and 3-D filtering (BM3D) (Dabov et al., 2007) is applied.

2.3 Segmentation

The resulting denoised nDSM containing only building regions are now segmented into different roof surfaces. Our strategy is to use watershed transform to oversegment the nDSM, and later apply a constrained merging process to get the final roof segments.

2.3.1 Watershed segmentation In grey scale mathematical morphology, the watershed transform was originally proposed in (Digabel and Lantuéjoul, 1978) and improved in (Beucher and Lantuéjoul, 1979). A gray scale image can be considered as a topographic surface, with intensity values as different elevation. If the surface is flooded from local minima, at points where water coming from different sources meet, watershed lines are built. As a result, the image is partitioned into catchment basins. For a nDSM image, its gradient map expresses height changes on it. In the gradient map, watershed lines are the gradient edges, and catchment basins are the homogeneous grey level regions of this image, i.e., the segments we want. An example is shown in Figure 2.

Direct use of gradient map usually produces severe over-segmented results, due to noise or local irregularities, thus too small gradients are "cut off", i.e., set to zero. h_{min} is defined as the "cut off" height value. I_{min} is the corresponding "cut off" intensity value in the gradient map. Intensity values that are smaller than I_{min} are set to zero. H_{min} denotes the minimal height difference between two adjacent surfaces, i.e., the two surfaces will be merged together if the height difference is smaller than H_{min}. The value of h_{min} is defined in the range of $[0.1 \cdot H_{min}, 0.2 \cdot H_{min}]$. Then a watershed segmentation is performed.

2.3.2 Constrained merging The previous step produces over-segmented results that to be merged. The purpose is to get minimum amount of segments, and maximum segments' regularity. Two merging constraints are employed:

Height difference (HD) constraint: Adjacent segments should be merged if the height difference between them is smaller than H_{min}.

Average polygon complexity (APC) constraint: APC of adjacent segments should not increase during merging to keep maximum regularity.

The segment's regularity can be indicated by polygon complexity (Brinkhoff et al., 1995). The polygonal complexity $compl(pol)$ is measured by comparing a polygon with its convexhull, using three characters of $notches$ (vertices that are located inside of the polygon's convexhull): $ampl(pol)$, amplitude of notches vibration; $freq(pol)$, frequency of notches vibration; and, $conv(pol)$, area difference between a polygon and its convexhull.

Quantitative definition of $compl(pol)$ is (Brinkhoff et al., 1995):

$$compl(pol) = 0.8 \cdot ampl(pol) \cdot freq(pol) + 0.2 \cdot conv(pol) \quad (1)$$

For multi-segments, we extend $compl(pol)$ to $compl(mean)$, the average polygon complexity, which is weighted mean of all segments' $compl(pol)$, with area size as weight. $compl(mean)$ is in the interval $[0, 1]$. Smaller $compl(mean)$ indicates more regular polygons. As larger simpler polygons are our aim, the constrained merging is carried out in two steps (segment'height is computed from nDSM):

(1) Merge small segments S_s to S_m, only considering HD constrain. S_s are segments that are smaller than area threshold T_a;

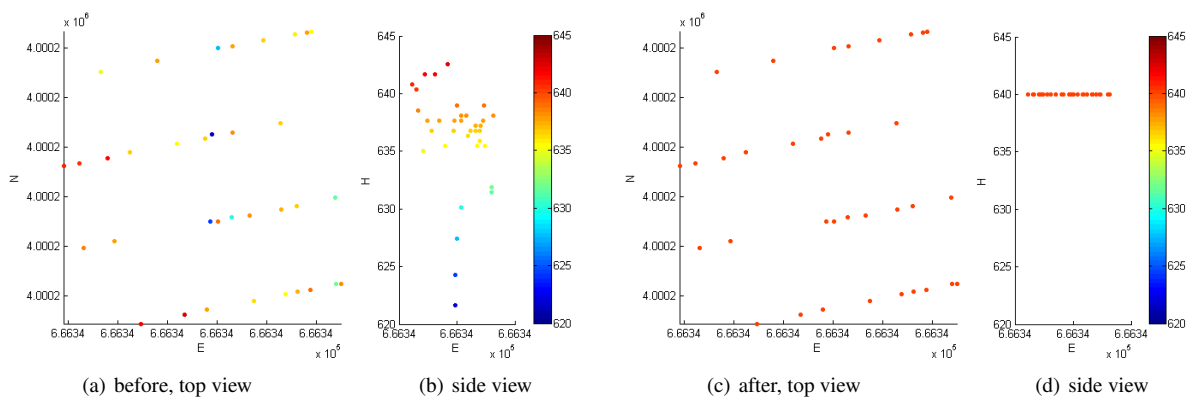(2) Merge S_m and other segments, considering both constraints.

(a) before, top view (b) side view (c) after, top view (d) side view

Figure 1: Local height smoothing example in 1 grid. (a)(b): before smoothing; (c)(d): after smoothing. $L_G = 3m$. Height is color coded.
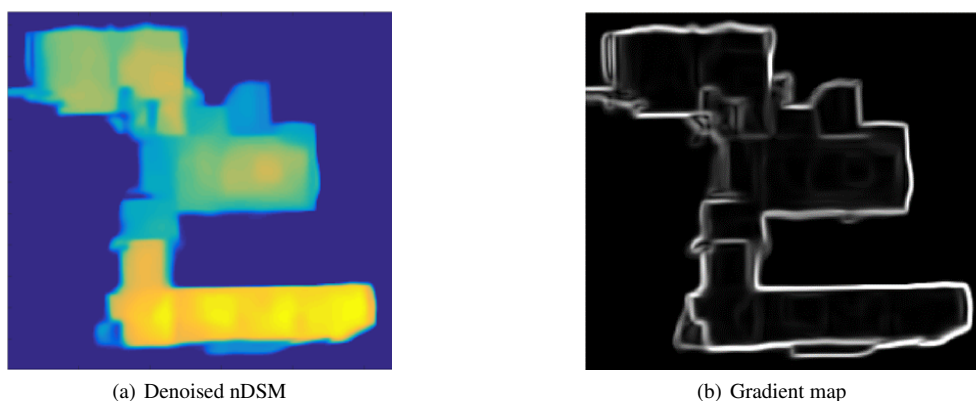

(a) Denoised nDSM

(b) Gradient map

Figure 2: Example of denoised nDSM and its gradient map.

2.4 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. A quadtree based regularization approach is proposed, and stated as follows.

2.4.1 Regularization Quadtree decomposition is a technique which divides an image into homogeneous regions (Samet, 1984). We put each segment in a $2^k \times 2^k$ sized image with empty background, by decomposing the image, a quadtree is built: the root stands for the whole image, while the leave nodes at lowest level are the single pixels in the image, which compose the outline since they are all located along the segment' boundary. We take all the single pixels in, as parts of the segment, results in a refined, enlarged segment.

Before applying quadtree regularization to one segment, the main directions of it are aligned to image axis directions, by rotating it with its orientation angles. After quadtree regularization, it is rotated back. Segments with similar orientation angles are grouped together and rotated with one angle to make the scene more regular.

The next step is to overlay all the segments together and correct their topological relationships. Previous two adjacent segments $S1$ and $S2$ now are $S1_n$ and $S2_n$, with an overlap O, due to the enlargement in quadtree regularization. APC is computed for the following two cases: (a) $S1_n$ and $S2_{n_O}$, with $S2_{n_O}=S2_n - O$; (b) $S1_{n_O}$ and $S2_n$, with $S1_{n_O}=S1_n - O$. Then O is assigned to the case which gives smaller average polygon complexity.

The outline of each segment now consists of boundary line segments that are perpendicular to their neighbors. Sequential short line segments form a "zigzag line", which needs to be removed to refine the segment's shape. Our approach to detect zigzag lines by computing the effective area A of each node, based on the Visvalingam-Whyatt algorithm (Visvalingam and Whyatt, 1993), and delete those with A smaller than the predefined maximum removable area A_{max}.

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

3. EXPERIMENTAL RESULTS

3.1 Tests

The proposed approach is illustrated and validated over a large building (convention center) in the city of Las Vegas using TomoSAR point clouds generated from a stack of 25 images using Tomo-GENESIS software developed at DLR (Zhu, 2011, Zhu et al., 2013). Figure 3 shows the TomoSAR point clouds of the area of interest.

3.2 Results and evaluation

Figure 4 shows results over the test building. The building model in Figure 4(f) consists 27 3-D polygons. For quantitative evaluation of the results, the root mean square (RMS) error of all points from respective planes are computed. The overall RMS computed for the building of interest is 3.19 m.
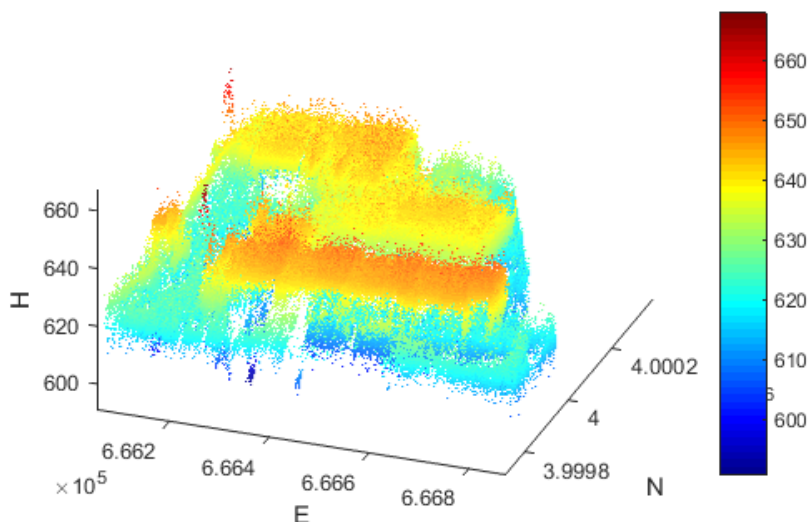
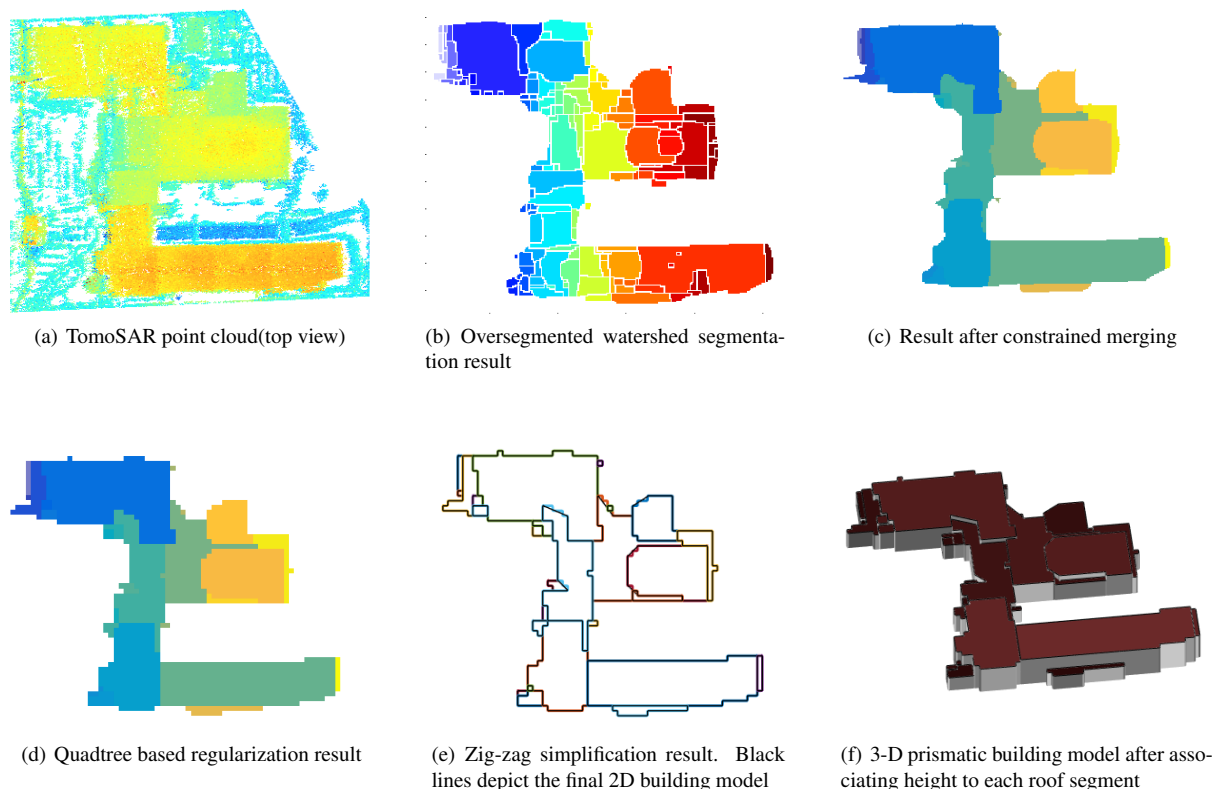Figure 3: TomoSAR point clouds of the test building, height is color coded.

(a) TomoSAR point cloud(top view)

(b) Oversegmented watershed segmentation result

(c) Result after constrained merging

(d) Quadtree based regularization result

(e) Zig-zag simplification result. Black lines depict the final 2D building model

(f) 3-D prismatic building model after associating height to each roof segment

Figure 4: From original point cloud to 3-D prismatic model over test building.

4. CONCLUSION AND OUTLOOK

This paper demonstrates for the first time the potential of using TomoSAR point clouds to reconstruct prismatic building models from space. The developed approach is modular and aims to systematically segment and reconstruct different flat roof surfaces to generate 3-D prismatic building models. The approach is validated over one large building. Although the approach seems to work well for the test site, there are few considerations that are worth to mention:

- In segmentation step, the building DSM is firstly over-segmented using watershed transform, then constrained merging is applied to achieve final roof segments. One limitation 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: a too small threshold would lead to sever over-segmentation that gets difficult merge while a too large threshold lead to under-segmentation which consequently result in detection loss of roof segments.

- In reconstruction step, the approach is totally data drive. While, with some simple shape models such as ellipse, semi-circle, regularization can be simplified for segments that fit corre-

sponding shape models, by using the shapes as refined segments' outlines instead of going through the quadtree regularization. Combination of data driven and model driven approach will improve the reconstruction workflow.

- Lastly, the tuning parameters in the approach are manually selected. In the future, a study on sensitivity analysis of these parameters and their selection criterion would be necessary.

ACKNOWLEDGEMENTS

REFERENCES

Ameri, B. and Fritsch, D., 2000. Automatic 3D Building Reconstruction Using Plane-Roof Structures. In: *Proceedings of the American Society for Photogrammetry and Remote Sensing Conference*, Washington, D.C, pp. 22–26.

Beucher, S. and Lantuéjoul, C., 1979. Use of watersheds in contour detection. In: *International Conference on Image Processing, Real-Time Edge and Motion Detection/Estimation*, Rennes, France, pp. 12–21.

Brenner, C., 2000. Towards Fully Automatic Generation of City Models. In: *International Archives of Photogrammetry and Remote Sensing*, Vol. 33, Part B3, Amsterdam, pp. 85–92.

Brenner, C., 2005. Building reconstruction from images and laser scanning. *International Journal of Applied Earth Observation and Geoinformation* 6(3-4), pp. 187–198.

Brinkhoff, T., Kriegel, H. P., Schneider, R. and Braun, A., 1995. Measuring the Complexity of Polygonal Objects. In: *Proceedings of the Third ACM International Workshop on Advances in Geographical Information Systems*, Baltimore, MD, USA, pp. 109–117.

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. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 7(10), pp. 4081–4094.

Dabov, K., Foi, A., Katkovnik, V. and Egiazarian, K., 2007. Image Denoising by Sparse 3-D Transform-Domain Collaborative Filtering. *IEEE Transactions on Image Processing* 16(8), pp. 2080–2095.

D'Hondt, O., Guillaso, S. and Hellwich, O., 2012. Automatic extraction of geometric structures for 3D reconstruction from tomographic SAR data. In: *IEEE International Geoscience and Remote Sensing Symposium*, Munich, pp. 3728–3731.

Digabel, H. and Lantuéjoul, C., 1978. Iterative Algorithms. In: *Proceedings of the 2nd European Symposium Quantitative Analysis of Microstructures in Material Science, Biology and Medicine*, Caen, France, 1977, Riederer Verlag, Stuttgart, Germany, pp. 85–99.

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), pp. 7323.

Elberink, S. O. and Vosselman, G., 2009. Building Reconstruction by Target Based Graph Matching on Incomplete Laser Data: Analysis and Limitations. *Sensors* 9(8), pp. 6101.

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), pp. 357–374.

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. *IEEE Transactions on Geoscience and Remote Sensing* 50(3), pp. 981–987.

Gruen, A., Baltsavias, E. P. and Henricsson, O., 1997. *Automatic extraction of man-made objects from aerial and space images (II)*. Birkhäuser Verlag, Basel.

Guillaso, S., D'Hondt, O. and Hellwich, O., 2013. Building characterization using polarimetric tomographic SAR data. In: *Urban Remote Sensing Event (JURSE)*, Sao Paulo, Brazil, pp. 234–237.

Guillaso, S., D'Hondt, O. and Hellwich, O., 2015. SAR tomography with reduced number of tracks: Urban object reconstruction. In: *IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, Milan, pp. 2919–2922.

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, pp. 21–30.

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), pp. S13–S27.

Rottensteiner, F. and Briese, C., 2003. Automatic generation of building models from LiDAR data and the integration of aerial images. In: *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, Vol. 34, Dresden, Germany, pp. 174–180.

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. In: *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences*, Vol. I(3), pp. 293–298.

Samet, H., 1984. The Quadtree and Related Hierarchical Data Structures. *ACM Computing Surveys* 16(2), pp. 187–260.

Schmitt, M., Shahzad, M. and Zhu, X. X., 2015. Reconstruction of individual trees from multi-aspect TomoSAR data . *Remote Sensing of Environment* 165, pp. 175 – 185.

Shahzad, M. and Zhu, X. X., 2015. Robust Reconstruction of Building Facades for Large Areas Using Spaceborne TomoSAR Point Clouds. *IEEE Transactions on Geoscience and Remote Sensing* 53(2), pp. 752–769.

Shahzad, M. and Zhu, X. X., 2016. Automatic Detection and Reconstruction of 2-D/3-D Building Shapes From Spaceborne TomoSAR Point Clouds. *IEEE Transactions on Geoscience and Remote Sensing* 54(3), pp. 1292–1310.

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), pp. 1425–1438.

Tarsha-Kurdi, F., Landes, T., Grussenmeyer, P. et al., n.d. Extended RANSAC algorithm for automatic detection of building roof planes from LiDAR data.

Verma, V., Kumar, R. and Hsu, S., 2006. 3D Building Detection and Modeling from Aerial LIDAR Data. In: *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Vol. 2, pp. 2213–2220.

Visvalingam, M. and Whyatt, J. D., 1993. Line generalisation by repeated elimination of points. *Cartographic journal* 30(1), pp. 46–51.

Vosselman, G., Gorte, B. G., Sithole, G. and Rabbani, T., 2004. Recognising structure in laser scanner point clouds. In: *International archives of photogrammetry, remote sensing and spatial information sciences*, Vol. 36,Part 8/W2, pp. 33–38.

Zhu, X. X., 2011. *Very High Resolution Tomographic SAR Inversion for Urban Infrastructure Monitoring – A Sparse and Nonlinear Tour*. Vol. 666, Deutsche Geodätische Kommission.

Zhu, X. X. and Bamler, R., 2012. Super-Resolution Power and Robustness of Compressive Sensing for Spectral Estimation With Application to Spaceborne Tomographic SAR. *IEEE Transactions on Geoscience and Remote Sensing* 50(1), pp. 247–258.

Zhu, X. X. and Shahzad, M., 2014. Facade Reconstruction Using Multiview Spaceborne TomoSAR Point Clouds. *IEEE Transactions on Geoscience and Remote Sensing* 52(6), pp. 3541–3552.

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)*, Sao Paulo, Brazil, pp. 159–162.