LOD3 BUILDING RECONSTRUCTION FROM MULTI-SOURCE IMAGES

H. Huang,*M. Michelini, M. Schmitz, L. Roth, H. Mayer

Institute for Applied Computer Science, Universität der Bundeswehr München, Germany - hai.huang@unibw.de

Commission II, WG II/4

KEY WORDS: 3D Reconstruction, Multi-View Stereo, Structure from Motion, Building Models, Generative Models

ABSTRACT:

We propose a pipeline for the detection as well as modeling of individual buildings based on multi-source images. It allows to consistently reconstruct whole buildings at Level of Detail 3 (LoD3): the roof from airborne images and the facades including elements such as windows and doors mainly from terrestrial images. We employ a parametrized top-down model – the "shell model" – with the roof as well as the facades semantically and geometrically integrated. This generative model fosters stability for building detection by enabling the use of multi-source data and offers flexibility in modeling by means of a fully CAD-compatible integration of building components. Experiments performed on imagery from different terrestrial and airborne (Unmanned Aerial Vehicle – UAV) cameras demonstrate the potential of the approach.

1. INTRODUCTION

3D building reconstruction is of great interest for many applications such as city planning, navigation, crisis management/emergency response, and tourism. In the last decades, it has been intensively studied and a large number of different approaches has been reported. Overviews of approaches before 2010 are given in (Brenner, 2005, Schnabel et al., 2008, Vosselman, 2009).

The approaches of the last decade include (Sampath, Shan, 2010), which segments and reconstructs complicated buildings from airborne LiDAR (Light Detection And Ranging) point clouds using polyhedral models. (Lafarge et al., 2010) presents building reconstruction from a Digital Surface Model (DSM) combining generic and parametric methods. For more sophisticated buildings, basic geometric primitives, e.g., planes, cylinders, and cones, are combined with mesh-patches to present irregular shapes (Lafarge, Mallet, 2012). (Huang et al., 2013) proposes a statistical approach for Level of Detail (LoD) 2 building model reconstruction from LiDAR data via generative models.

The topological consistency of rooftops is the focus of (Chen et al., 2017). (Zeng et al., 2018) proposes residential building reconstruction using deep neural networks applying shape grammar rules. In (Li et al., 2019), the point cloud is segmented using a Triangulated Irregular Network (TIN) and the boundaries of roofs are refined on a 2D grid.

(Partovi et al., 2015) presents an extension of a hybrid framework for data from stereo satellite imagery with ridge-linebased building mask decomposition. (Tutzauer, Haala, 2015) is concerned with 3D facade reconstruction based on point clouds derived from mobile mapping systems and oblique airborne cameras. Radiometric segmentation is used to overcome the limited accuracy of the point clouds.

A plane-based building model reconstruction and regularization approach is introduced in (Holzmann et al., 2018), employing

an improved stable plane detection approach based on 3D lines instead of points. The detection and reconstruction of buildings from point clouds derived from space-borne Synthetic Aperture Radar (SAR) data is reported in (Shahzad, Zhu, 2016).

A building reconstruction pipeline based on point clouds derived from UAV (Unmanned Aerial Vehicle) images is presented in (Li et al., 2016). Roof structures are determined using Markov Random Field optimization and fitted to the estimated building footprints. Another approach using point clouds from UAV images is (Nguatem, Mayer, 2017), which employs a nonparametric Bayesian framework and polygon sweeping. Starting from planar roof segments, (Zhou, Neumann, 2012) tries to organize them using "global regularities" in the form of orientation and placement constraints.

In recent years, the quality and availability of 3D point clouds from LiDAR and image-based reconstruction have been significantly improved. The approaches for building reconstruction reach a high level of automation and cover larger urban and suburban areas. In (Huang, Mayer, 2017, Huang et al., 2019), an approach for scene and building decomposition is proposed which improves the efficiency of complex building reconstruction in dense scenes. A reconstruction pipeline for sophisticated building models based on multi-source data including building footprints, mesh data, and terrestrial images is introduced in (Kelly et al., 2017) with facade elements detected from the imagery using deep neural networks. (Zhu et al., 2018) presents a large-scale urban modeling framework based on surface meshes derived by a multi-view-stereo system.

Besides the challenges by data flaws and complex building structures, previous approaches suffer from incomplete measurements for buildings. Terrestrial LiDAR data or images reveal details of buildings, particularly of the facades, but have poor coverage of the roofs and the ground due to the obtuse observation angle. Because of the viewing angle and flying height, airborne LiDAR data or imagery provide suitable measurements for the roof but not for the facades, even for oblique views. Many approaches are, therefore, limited to either incomplete building modeling without roofs or LoD2 building modeling

^{*}Corresponding author

with detailed roof shape but only an extrusion approximating the walls.

Because of the improvement of data acquisition technologies, especially the advent of UAVs, it is now possible to acquire data covering the whole building from all sides. High-resolution terrestrial images of the facades are complemented by UAV imagery taken from slightly larger distances using nadir and oblique views, with a clear view on the ground and the roofs. While this opens up a possibility for direct LoD3 building modeling, we also face new challenges concerning effective fusion and utilization of the data, i.e., the semantic and geometric integration of the roof and the facades as well as the facade elements including windows and doors.

To this end, we employ the "shell model", a generative statistical model for LoD3 building reconstruction based on the fusion of terrestrial and UAV imagery. It is a hybrid model combining concepts and elements of CSG (Constructive Solid Geometry) and BRep (Boundary Representation) models. The shell model consists of an outer and an inner layer defining a solid body model with a certain thickness in between. Under observation of measurement data including point clouds from LiDAR and image matching, we conclude that the data actually always only reveal the surface instead of the solid body of the objects. Due to the measurement uncertainty, the surface is also far from perfect. Shell models allow for a more suitable and practical geometrical modeling in comparison with conventional surface models, e.g., meshes or assembled planes, as well as solid body models. In this work, roof(s) and facades are modeled together by a hollow shell with a hypothetical thickness based on the multi-source images of the whole building.

All the components are integrated into a model with CSG operations. Experiments are performed on multi-source image data to demonstrate the potential of the proposed approach.

The paper is organized as follows. Section 2 describes the derivation of dense 3D point clouds from multi-source images. In Section 3 the concept of shell models is introduced. The use of shell models for both building detection and modeling is demonstrated based on a running example. The paper ends with a conclusion and recommendation for future work in Section 4.

2. MULTI-SOURCE IMAGERY AND 3D RECONSTRUCTION

The input multi-source images, as shown in Figure 1, stem from both UAV and terrestrial cameras. They are acquired with a hand-held DSLR (Digital Single-Lens Reflex) camera (Nikon D800, 36M pixels) and an UAV-mounted light-weight mirrorless camera (Sony ILCE-7R, 36M pixels). The images cover the whole building from the facades to the roof. These widebaseline images are fused by a precise and reliable orientation estimation approach and dense colored 3D point clouds are reconstructed.

2.1 Sparse 3D Reconstruction

A sparse 3D reconstruction of this image set is conducted using the Structure from Motion approach described in (Michelini, Mayer, 2016). This approach requires (approximate) camera calibration, which is obtained from the meta-data of the images. Images are matched using wide-baseline image matching (Mayer et al., 2012) which is required for strongly geometrically/radiometrically distorted images. To improve robustness



Figure 1. Input images from an aerial (UAV – top and center) and a terrestrial (bottom) cameras.

and accuracy, image triplets instead of pairs are employed. Finally, triplets are merged to even larger image subsets transforming the orientations into a common reference frame.

Figure 2 shows the estimated orientations (top). Images with detected overlap are linked with colored lines (bottom), which are reduced in number for a better visualization. Color indicates different cameras.

2.2 Dense 3D Reconstruction

Based on the accurately estimated orientations, dense depth maps are generated considering epipolar constraints. The foundation of our Multi-View Stereo (MVS) approach is Semi-Global Matching (SGM) (Hirschmüller, 2008) because of its potential for an efficient as well as effective processing, especially for high-resolution images. A pixel-wise uncertainty measure for the disparities (Kuhn et al., 2017) is the basis for a high quality probabilistic integration of the stereo models in voxel space. Figure 3 shows the target building (top) and the dense point cloud (bottom) reconstructed from the multi-source images shown in Figure 2.

3. BUILDING MODELS

We employ a generative model – "shell model" for both detection (cf. Section 3.1) as well as modeling (cf. Section 3.2) of buildings from the input point cloud. The concept of the shell model is inspired by two observations: (1) no matter which acquisition technique (e.g., LiDAR or 3D reconstruction based on image matching) has been employed, the "3D" measurement data represent only **the surface** and **not the solid body** of the objects and (2) because of measurement errors, noise, and lack of physical construction precision the underlying surfaces are represented by a layer with a certain thickness.

We, therefore, consider a "shell" a more reasonable and practically useful geometrical model for the parsing of measurement data than a simple surface or a solid body and propose a shell model for building modeling in LoD3. It consists of parallel inner and outer layers and the solid body defined by them. I.e., it is a hybrid model combining elements of BRep and CSG models. Figure 4 presents the concept of the shell model. Instead of a standard BRep or CSG model, the building is modeled as a "shell" with a certain thickness.

Please note that shell models have different definitions of the inner and outer layers depending on the application, namely The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XLIII-B2-2020, 2020 XXIV ISPRS Congress (2020 edition)



Figure 2. Orientation estimation for images from different sources: The pyramids show the orientations with the colors representing different cameras (calibration). Cameras are linked with colored line segments (bottom) if the corresponding images overlap.

building detection (cf. Section 3.1) and modeling (cf. Section 3.2).

3.1 Building Detection

The shell model for building detection is designed as shown in Figure 5 (top left). The outer layer (blue) and the inner layer (red) represent the tolerances around the model matching the data (green). The definition of layers is different for building modeling (Figure 5, top right), for which more details are given in Section 3.2.

3.1.1 Primitives Our statistical building detection employs generative primitives. Inspired by (Huang et al., 2013), we provide an extended library of primitives, as shown in Figure 6, by adding, e.g., the half-hipped roof for typical European houses (cf. the running example in Figure 7) and arched and butterfly roof. This empirically defined library is supposed to cover a majority of typical European residential and industrial buildings.

The primitives are parametrized as:

$$\theta \in \Theta; \Theta = \{\mathcal{P}, \mathcal{C}, \mathcal{S}\} , \tag{1}$$

where the parameter space Θ consists of position parameters $\mathcal{P} = \{x, y, azimuth\}$, contour parameters $\mathcal{C} = \{length, width\}$ (rectangular footprint), and shape parameters \mathcal{S} : ridge/eave height(s) and parameters of hips.





Figure 3. Top: A photo of the target building; Bottom: The reconstructed colored 3D point cloud.

3.1.2 Parameter Optimization The Maximum A Posteriori (MAP) estimate of Θ is employed to find the optimal model fitting the data:

$$\widehat{\Theta}_{MAP} = \underset{\Theta}{\operatorname{argmax}} \left\{ \frac{L(\mathcal{D}|\Theta)p(\Theta)}{P(\mathcal{D})} \right\} = \underset{\Theta}{\operatorname{argmax}} \left\{ L(\mathcal{D}|\Theta)p(\Theta) \right\}$$
(2)

 $L(\mathcal{D}|\Theta)$ is the likelihood function representing the goodness of fit of the model to the data \mathcal{D} and $p(\Theta)$ the prior for Θ . The prior is derived from empirical knowledge and can be incrementally improved based on the accepted models (pieces of evidence). I.e., parameter values of already found primitives (single buildings or building components) are used to update the priors. $P(\mathcal{D})$ is the marginal probability. It can be omitted from the goal function as it does not depend on Θ and is, therefore, constant in the optimization.

Reversible Jump Markov Chain Monte Carlo (RJMCMC) (Green, 1995) is used for the statistical search of the parameters, resulting in an efficient exploration of the high-dimensional (determined by the number of parameters) search space. The reversible jumps allow to switch between search spaces, i.e., different types of primitives with different numbers of parameters.

Figure 7 presents the detection of a building with a half-hipped roof. The shell model with inner (red) and outer (blue) layers (bottom) is fitted to the input point cloud (top). The model of best fit is then assumed to be the layer (green) between them.

In comparison to previous approaches (mostly BRep-based) including (Huang et al., 2013), the proposed shell models have the following advantages for building detection:

The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XLIII-B2-2020, 2020 XXIV ISPRS Congress (2020 edition)



Figure 4. Comparison of the concepts of BRep (left), CSG (right), and shell models (bottom).



Figure 5. Shell models for building detection and modeling.

- Roof and facades are defined by the same model and are detected jointly. There is no need to adjust and match roof and facade planes. The result is guaranteed to be a complete and watertight model.
- 2. The parameters of the model are determined through the consensus of all planes of the building and are, therefore, more precise and stable.
- 3. The detection is more efficient. A time-consuming part of the MCMC search is the calculation of the likelihood *L*, i.e., the evaluation of the goodness of fit of the proposed model. This calculation has to be performed a large number of times during statistical optimization. In previous work, the likelihood is calculated based on the distances of all individual points to the corresponding planes. For the shell model, the goodness of fit can be represented as the number of points inside the shell, which can be efficiently calculated as the difference of the number of points in the solids constructed by the outer and the inner layer.

3.2 Building Modeling

As shown in Figure 5, the shell model for building modeling is defined differently than for detection. After the optimally



Figure 6. Predefined primitive library.

fitting model (green) has been detected (cf. Section 3.1), it is employed in building modeling as the outer layer representing the facade, while the inner layer models the inside walls. In this case, the thickness of the shell simulates the thickness of the walls. Yet, we note that the thickness of the walls is, in most cases, only hypothetical, as there is usually no data available for the interior of the building.

In conventional building modeling, there is data available either solely for the roof or the facades. We can either choose "roofbased" modeling, i.e., facades modeled as an extrusion from the eave lines to the ground, or "facade-based" modeling, meaning that the exact roof geometry including the overhang is ignored. As we fuse the images from both aerial and terrestrial cameras (cf. Section 2), the reconstructed point cloud is available for both roofs and facades. The shell model shows its advantage by inherently integrating roofs and facades jointly taking all data from multiple sources into account:

- 1. Roof and facades are predefined jointly in the top-down model. The resulting building model is guaranteed to be watertight without any extra effort to assemble individual planes.
- 2. Both roof and facades are appropriately modeled with the possibility to detect and model roof overhangs.
- 3. The shell model allows to model windows and doors quite naturally as openings.

3.3 Integration of Building Parts and Facade Elements

The shell model has a tree structure compatible with standard CSG operations such as "Union" and "Difference". This means that the shell model can be extended by merging multiple building parts and refined by modeling facade elements. The integration does not only mean geometric combination but also semantic organization.

3.3.1 Roof Overhang Roof overhang refers to the part of a roof which extends over the facade planes. In almost all previous approaches for building reconstruction, the roof overhang of buildings has been ignored. On one hand, it is considered as a trivial part in comparison with the main body of the building. On the other hand, it is hard to detect because of (1) the



Figure 7. Building detection from point cloud (top) with shell model (bottom).

weak support by a limited number of points and (2) the inherent constraints in roof-based and facade-based modeling. The roof overhang, however, plays an important role in detailed building modeling, in which the contour of the roof (eaves) and the boundary of the walls, i.e., the building footprint, should be distinguished and modeled precisely.

Since we have data available for the roof as well as the facades, we can detect and model this relatively slim structure. We deal with the roof as an additional building part, which is a solid body (or an open shell) consisting of the planes of the roof with a certain thickness. The detection and modeling of the roof overhang can, thus, be performed separately without constraints by the facades. We particularly employ a rule-based "edge sweeping" method (Huang, Brenner, 2011):

Hypotheses are generated by extending the roof to fit the points that are supposed to represent the overhang. I.e., these points belong to the planes of the roof but are outside the boundary of the walls. Since the points of the overhang are few and noisy, we link the "sweeping" of both sides of the roof by joint constraints to ensure parallelism and symmetry. The roof is merged to the main body of the building with the CSG "union" operation (cf. also Figure 5, bottom).

3.3.2 Facade Elements While the aerial images cover mainly the roof, the terrestrial images with relatively higher resolution reveal the details of the facades. While we have worked with Implicit Shape Models (Reznik, Mayer, 2008) and Structured Random Forests (Rahmani, Mayer, 2018) in the past, here we detect the elements i.e., particularly windows and doors, in

rectified terrestrial images employing a Convolutional Network (ConvNet). A shell model has semantically defined planes for roof and facades. From the primitive model the facades can, thus, be directly derived in the form of 3D polygons. The original images are projected via a planar homography onto the facade planes representing the corresponding regions of the facades.

We employ FC-DenseNet56 (Jégou et al., 2017), trained for the semantic segmentation of rectified facade images. I.e., we generate pixel-wise proposals for the classes wall, door, window and occlusion. The results for projections from different images are merged and facade elements are fitted to the overall result (Schmitz, Mayer, 2016, Schmitz et al., 2019). The detected facade elements are projected back into 3D space for integration in the LoD3 model. Figure 8 presents this process.



Figure 8. Detection of windows and doors with a ConvNet on a planar facade. Image parts of the facade are projected onto the facade plane using corresponding 3D coordinates. The final results are projected back into the 3D model.

Since the walls have a certain thickness, the windows and doors can be modeled either as alcoves (left) or cavities (right) as shown in Figure 9.



Figure 9. Openings can be represented as alcoves (left) or cavities (right).

Figure 10 presents reconstruction results based on the point clouds (left) derived from multi-source images. The CSG models (right) contain the roof with overhang, the walls with a given thickness and the facade elements.

A combined (L-shaped) building, as demonstrated in Figure 11, is modeled by merging two individual primitives with the "union" operation of CSG, while the windows and doors are integrated with the "difference" operation as alcoves. The BRep The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XLIII-B2-2020, 2020 XXIV ISPRS Congress (2020 edition)



Figure 10. Input point clouds (left) and reconstruction results (right).

elements of shell models – the outer layers – can be used for texturing (cf. Figure 11, bottom). Based on the determined facade as well as roof planes, images are projected onto the corresponding facets (3D polygons with known plane parameters and vertices) as textures.

4. CONCLUSION AND FUTURE WORK

In this paper, we have presented a pipeline for the reconstruction of individual buildings from multi-source imagery. We have demonstrated the advantages of shell models in building detection as well as modeling using the fusion of airborne and terrestrial images. The main contributions of this paper can be summarized as follows:

- Introduction of a complete pipeline to utilize multi-source images from different platforms with different resolutions;
- Efficient and stable building detection based on the consensus of roof as well as facade planes;
- Detailed modeling of both roof and facade geometry considering roof overhangs;
- Water-tight and CAD-compatible vector models with optional alcoves/cavities for openings

We are aware that many challenges remain. Public and commercial buildings may have special shapes that cannot be represented by the introduced rectangular primitives. The superstructures on the roof and facades, e.g., dormers, chimneys and balconies, as well as annexes of the buildings, for instance, storage sheds and outer stairs, have not been tackled yet, even though they could, to a certain extent, be approximated with the existing primitives, i.e., flat-, shed- and gable-roofs.

Concerning future work, we first consider to upgrade the library of primitives with flexible geometric shapes (instead of



Figure 11. Reconstruction of an L-shaped building.

only rectangles) and to extend it with specific types for superstructures and annexes. Besides conventional classification of 2D images, ConvNets could also be used for direct parsing of the 3D geometry (Qi et al., 2016, Qi et al., 2017, Wang et al., 2018), i.e., the segmentation of point clouds into building parts and the detection of facade elements using both color and depth information.

REFERENCES

Brenner, C., 2005. Building Reconstruction from Images and Laser Scanning. *International Journal of Applied Earth Observation and Geoinformation*, 6(3-4), 187–198.

Chen, D., Wang, R., Peethambaran, J., 2017. Topologically Aware Building Rooftop Reconstruction From Airborne Laser Scanning Point Clouds. *IEEE Transactions on Geoscience and Remote Sensing*, 55(12), 7032-7052.

Green, P. J., 1995. Reversible Jump Markov Chain Monte Carlo Computation and Bayesian Model Determination. *Biometrika*, 82, 711–732.

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

Holzmann, T., Maurer, M., Fraundorfer, F., Bischof, H., 2018. Semantically aware urban 3D reconstruction with plane-based regularization. *European Conference on Computer Vision*, 487–503.

Huang, H., Brenner, C., 2011. Rule-based roof plane detection and segmentation from laser point clouds. *Joint Urban Remote Sensing Event (JURSE) 2011, 11-13 April*, IEEE, Munich, Germany, 293–296.

Huang, H., Brenner, C., Sester, M., 2013. A generative statistical approach to automatic 3D building roof reconstruction from laser scanning data. *ISPRS Journal of Photogrammetry and Remote Sensing*, 79(0), 29 – 43.

Huang, H., Kuhn, A., Michelini, M., Schmitz, M., Mayer, H., 2019. 3D urban scene reconstruction and interpretation from multisensor imagery. *Multimodal Scene Understanding: Algorithms, Applications and Deep Learning*, Elsevier, 307 – 340.

Huang, H., Mayer, H., 2017. Towards automatic large-scale 3D building reconstruction: Primitive decomposition and assembly. *AGILE Conference on Geographic Information Science*, 205–221.

Jégou, S., Drozdzal, M., Vazquez, D., Romero, A., Bengio, Y., 2017. The one hundred layers tiramisu: Fully convolutional densenets for semantic segmentation. 2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CV-PRW), 1175–1183.

Kelly, T., Femiani, J., Wonka, P., Mitra, N. J., 2017. BigSUR: Large-scale Structured Urban Reconstruction. *ACM Transactions on Graphics*, 36(6).

Kuhn, A., Hirschmüller, H., Scharstein, D., Mayer, H., 2017. A TV Prior for High-Quality Scalable Multi-View Stereo Reconstruction. *International Journal of Computer Vision*, 124(1), 2–17.

Lafarge, F., Descombes, X., Zerubia, J., Pierrot-Deseilligny, M., 2010. Structural Approach for Building Reconstruction from a Single DSM. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32(1), 135–147.

Lafarge, F., 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.

Li, M., Nan, L., Smith, N., Wonka, P., 2016. Reconstructing building mass models from UAV images. *Computers & Graphics*, 54, 84 - 93. Special Issue on CAD/Graphics 2015.

Li, M., Rottensteiner, F., Heipke, C., 2019. Modelling of buildings from aerial LiDAR point clouds using TINs and label maps. *ISPRS Journal of Photogrammetry and Remote Sensing*, 154, 127 - 138.

Mayer, H., Bartelsen, J., Hirschmüller, H., Kuhn, A., 2012. Dense 3D reconstruction from wide baseline image sets. *Outdoor and Large-Scale Real-World Scene Analysis*, Springer, Berlin, Heidelberg, 285–304.

Michelini, M., Mayer, H., 2016. Efficient Wide Baseline Structure from Motion. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, (III-3), 99–106.

Nguatem, W., Mayer, H., 2017. Modeling urban scenes from pointclouds. *International Conference on Computer Vision*, 3857–3866.

Partovi, T., Huang, H., Krauß, T., Mayer, H., Reinartz, P., 2015. Statistical Building Roof Reconstruction from WORLDVIEW-2 Stereo Imagery. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XL(3/W2), 161-167.

Qi, C. R., Su, H., Mo, K., Guibas, L. J., 2016. PointNet: Deep Learning on Point Sets for 3D Classification and Segmentation. *Computer Research Repository*, abs/1612.00593.

Qi, X., Liao, R., Jia, J., Fidler, S., Urtasun, R., 2017. 3D graph neural networks for RGBD semantic segmentation. *Computer Vision and Pattern Recognition*, 5199–5208.

Rahmani, K., Mayer, H., 2018. High Quality Facade Segmentation Based on Structured Random Forest, Region Proposal Network and Rectangular Fitting. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, IV(2), 223–230.

Reznik, S., Mayer, H., 2008. Implicit Shape Models, Self-Diagnosis, and Model Selection for 3D Facade Interpretation. *Photogrammetrie–Fernerkundung–Geoinformation*, 3, 187–196.

Sampath, A., Shan, J., 2010. Segmentation and Reconstruction of Polyhedral Building Roofs From Aerial Lidar Point Clouds. *IEEE Transactions on Geoscience and Remote Sensing*, 48(3), 1554–1567.

Schmitz, M., Huang, H., Mayer, H., 2019. Comparison of Training Strategies for ConvNets on Multiple Similar Datasets for Facade Segmentation. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLII-2/W13, 111–117.

Schmitz, M., Mayer, H., 2016. A convolutional network for semantic facade segmentation and interpretation. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLI-B3, 709–715.

Schnabel, R., Wessel, R., Wahl, R., Klein, R., 2008. Shape recognition in 3D point-clouds. *The 16th International Conference in Central Europe on Computer Graphics, Visualization and Computer Vision.*

Shahzad, M., 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), 1292-1310.

Tutzauer, P., Haala, N., 2015. Facade reconstruction using geometric and radiometric point cloud information. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XL-3/W2, 247-252.

Vosselman, G., 2009. Advanced point cloud processing. *Photo-grammetric Week '09*, 137–146.

Wang, Y., Sun, Y., Liu, Z., Sarma, S. E., Bronstein, M. M., Solomon, J. M., 2018. Dynamic Graph CNN for Learning on Point Clouds. *Computer Research Repository*. http://arxiv.org/abs/1801.07829.

Zeng, H., Wu, J., Furukawa, Y., 2018. Neural procedural reconstruction for residential buildings. *European Conference on Computer Vision*, 759–775. Zhou, Q.-Y., Neumann, U., 2012. 2.5D building modeling by discovering global regularities. *Computer Vision and Pattern Recognition*, 326–333.

Zhu, L., Shen, S., Gao, X., Hu, Z., 2018. Large scale urban scene modeling from MVS meshes. *European Conference on Computer Vision*, 640–655.