

COMPARING THE PERFORMANCE OF ROOF SEGMENTATION METHODS IN AN URBAN ENVIRONMENT USING DIGITAL ELEVATION DATA

B. Farajelahi^{1*}, M. Najaf¹, H. Arefi¹

¹ Department of Geomatics, University College of Engineering, University of Tehran, Tehran, Iran - (behaeenfarajelahi, mostafanajaf98, hossein.arefi)@ut.ac.ir

Commission IV, WG IV/3

KEY WORDS: Complex Roof, Segmentation, Region Growing, Clustering, MSAC, K-means, Digital Surface Model

ABSTRACT:

In this paper, we present a novel 3D segmentation approach using digital elevation data. Building detection has been emerging as an important area of research. It has attracted many applications, such as geomatics, architectonics, computer vision, photogrammetry, remote sensing, industry, disaster management, and city planning. Building detection techniques can basically be divided into two categories: the classical approach and the deep learning approach. The main goal of this study is to compare some commonly used detection techniques in photogrammetry, like segmentation-based and classification-based methods using digital elevation data as input. The 4 different methods of roof detection with their detailed analysis and their final results are presented in this paper. This study encourages researchers to further advance research in building detection techniques. Results show that the 2D region growing can successfully segment the building components like the main facades of the complex roof and provide accurate qualitative and quantitative results compared to the other methodologies used in this study.

1. INTRODUCTION

Buildings are the most important component of an urban environment. Nowadays, detection and reconstruction of buildings has attracted intensive attention for different applications. There are various methods to acquire the data, such as light detection and ranging (LiDAR) which produces a dense 3D point cloud that is used to reconstruct 3D models of an urban area (Adam, Chatzilari et al. 2018), (Lafarge and Mallet 2011), (Samadzadegan, Mahmoudi et al. 2010). Totally, aerial images, ground views, or laser scans are the first inputs for 3D modelling. Recently, a lot of methods have been proposed and it is challenging to compare these methods properly. These methods have been developed in various contexts (kinds of data, types of buildings, level of user interactivity, etc.) and use some evaluation criteria (Adam, Chatzilari et al. 2018), (Lafarge and Mallet 2011), (Lafarge, Descombes et al. 2008). Different methods are suggested to generate the 3D models automatically or interactively by using different data like aerial images or LiDAR point clouds (Zhang, Wang et al. 2014). Building modelling has four primary steps. The first step is separating the point cloud into 2 classes named ground points and non-ground-points. Arefi and Hahn proposed a novel approach to this separation by means of a morphological method (Arefi and Hahn 2005). The second step is removing the outliers or those unwanted points which are related to the trees, cars, and the walls of buildings. Numerous algorithms have been applied to remove these points. For example, by means of the analysis of multiple returns of the pulses or the use of different geometric features which are derived by the covariance matrix of the point neighborhood (Alharthy and Bethel 2002), (Sampath and Shan 2008). The third step is to specify the planar roof segments which have the same properties. Many methods have been studied, such as region growing (Rottensteiner and Briese 2003). Hough transform (Vosselman and Dijkman 2001) and RANSAC (Random Sample Consensus) (Tarsha-Kurdi, Landes et al. 2008).

The last step is to reconstruct appropriate building models, for which there are three techniques: Model-Driven, Data-Driven, Hybrid Methods.

The remainder of this paper is as follows: Section 2 describes detection methods. Section 3 expresses a brief introduction to 4 different segmentation techniques used in this study. Section 4 includes the comparison between the segmentation methods and their qualitative and quantitative results. Section 5 presents the final conclusion of these distinctive approaches.

2. DETECTION METHODS

Building detection is a classification issue which separates different urban objects like ground, roads, and trees. The urban objects will be categorized into two groups, like buildings and non-buildings.

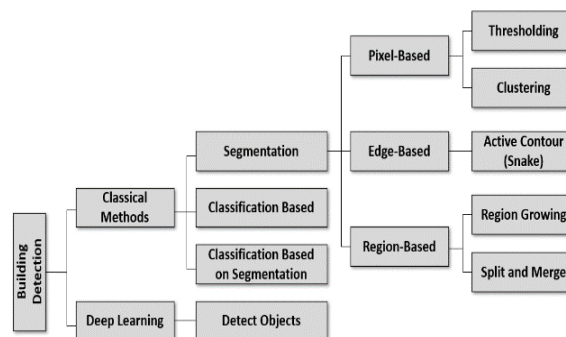


Figure 1. Detection Techniques.

One of the main applications of building detection is updating the maps and detecting changes. Building detection is a crucial process before extracting the building borders. One of the main challenges in detecting a building is the correct separation of tall trees or a group of trees from buildings and also the correct

separation of buildings from adjacent trees. Since most building detection methods are based on accurate height data with high spatial resolution, any error or violation in this data will cause errors in building detection (Vosselman 2010). Numerous building detection algorithms based on local features have been proposed in the last two decades. The building detection methods are totally grouped into two main approaches: the classical approach and the deep learning approach.

2.1 Classical Approach

In the classical approach, different building detection methods such as segmentation-based methods and classification-based methods, and so do segmentation-based classification methods, have been proposed and applied in different areas.

2.1.1 Segmentation Methods

Segmentation methods are mainly categorized into three categories: pixel-based, edge-based, and region-based. The pixel-based and edge-based separate the image based on the rapid changes in intensity. The region-based method is basically based on similarity and homogeneity by using various methods such as merging, region splitting, and region growing (Dubey, Gupta et al. 2016).

2.1.1.1 Pixel Based

Thresholding is a segmentation technique that splits the objects from the background (Sampath and Shan 2008), (Rottensteiner, Trinder et al. 2004). In this method, the pixel intensity is used by specifying a threshold. An image is grouped into two parts: the foreground and the background, so the detection and classification will be simpler because it includes the main information related to the position and the component shape (Sampath and Shan 2008). One of the main disadvantages of this technique is that threshold-based techniques are inefficient in blurring images, so region-growing algorithms are recommended (Sheng, Gao et al. 2016) and also the threshold should be good enough value to have a robust result. In order to choose a proper threshold for this approach, the histogram of the pixel intensities will be used and the mean of these values will be introduced as the threshold (Khaloo and Lattanzi 2017).

2.1.1.2 Edge Based

The edge detection technique is mainly based on the sudden discontinuities in intensity changes. In general, objects on boundaries tend to produce intensity changes. These operators are applied to make an edge image. Edge detection methods reduce the amount of data to be processed. The edge detection removes useless information and saves important structural properties (Vo, Truong-Hong et al. 2015). Gradient-based 1st order derivatives and Laplacian-based 2nd order derivatives are the two commonly used methods for edge detection (Fischler and Bolles 1981). Edge detection has a broad variety of applications like image compression, security, enhancement, computer vision, etc. It has a weak performance in the presence of noise (Fischler and Bolles 1981). In order to segment the point cloud by energy optimization, an energy function is used by means of building features, in which the function has a minimum value for building features and a maximum value for other urban objects.

2.1.1.3 Region Based

The region-based method is a similarity-based segmentation that, based on some characteristics such as color, intensity, texture, etc., partitions an image into sub-regions. Those pixels that have the same intensity characteristics and are close to each other will be grouped together and indicate the same object. Region growing, split and merge algorithms are the most commonly used methods of region-based technique. Region growing is a great way to detect objects in noisy images and it starts with an initial seed point. The region is formed by calculating some specific properties and comparing them with adjacent points based on certain topologic measures (Sampath and Shan 2008).

2.1.2 Classification Techniques

One of the most widely used methods for detecting and extracting buildings is the classification method. These methods are generally divided into two methods: supervised classification and unsupervised classification. In supervised classification, trained data is generated with user supervision, and in non-supervised classification, the intended classes are created using clustering methods (Rottensteiner, Trinder et al. 2004).

3. METHODOLOGY

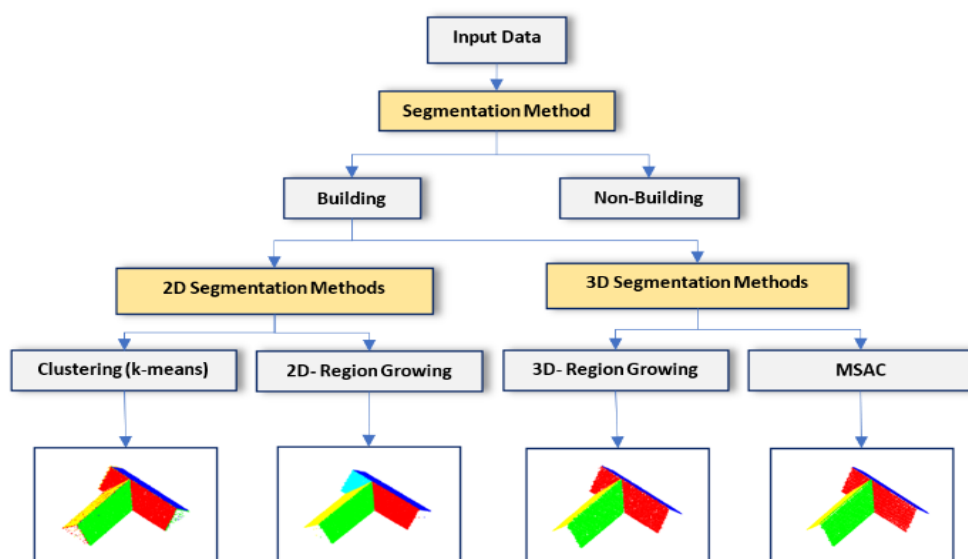


Figure 1. The flowchart of this study.

3.1 2D Region Growing

The 2D region growing approach starts from an initial single point given by the user. In this method, the region is iteratively grown by comparing all unassigned neighbouring pixels and the measure of similarity is the difference between the pixel intensity value and the region's mean. Therefore, the pixel with the lowest difference measured is assigned to the proper region. The process ends when the intensity difference between the mean intensity for each region and the new pixel is larger than a specified threshold (Louizi and Gammoudi 2011). The result of the 2D region growing approach is shown in Figure 3. (e) and (f). The result of the 2D region growing approach is shown in Figure 3. (e) and (f).

3.2 3D Region Growing

In this method, first, the normal vector and curvature of each point are estimated. Then, this segmentation algorithm starts by finding a seed point and grows by adding new points. When the first segment is completed, a new seed point is used for the growth of the next segment. The point P_i with the minimum curvature value, σ_p , was considered as the initial seed point to start the region growing procedure from a part of the scenery where the surface is smoother and the surface variation is lower (Khaloo and Lattanzi 2017). The result of the 3D region growing approach is shown in Figure 3. (i) and (j).

3.3 K-means

The K-means clustering technique is a pixel-based segmentation algorithm. This algorithm clusters the point nearest to its centroid. The average of all points is used as the centroid and has coordinates as the arithmetic mean for all points in the cluster. This algorithm is used to minimize the distance of the points to the center of an assigned cluster. The disadvantage of the k-means algorithm is that it is necessary to set the number of clusters from the beginning (Vo, Truong-Hong et al. 2015). The result of the k-means approach is shown in Figure 3. (k) and (l).

3.4 MSAC

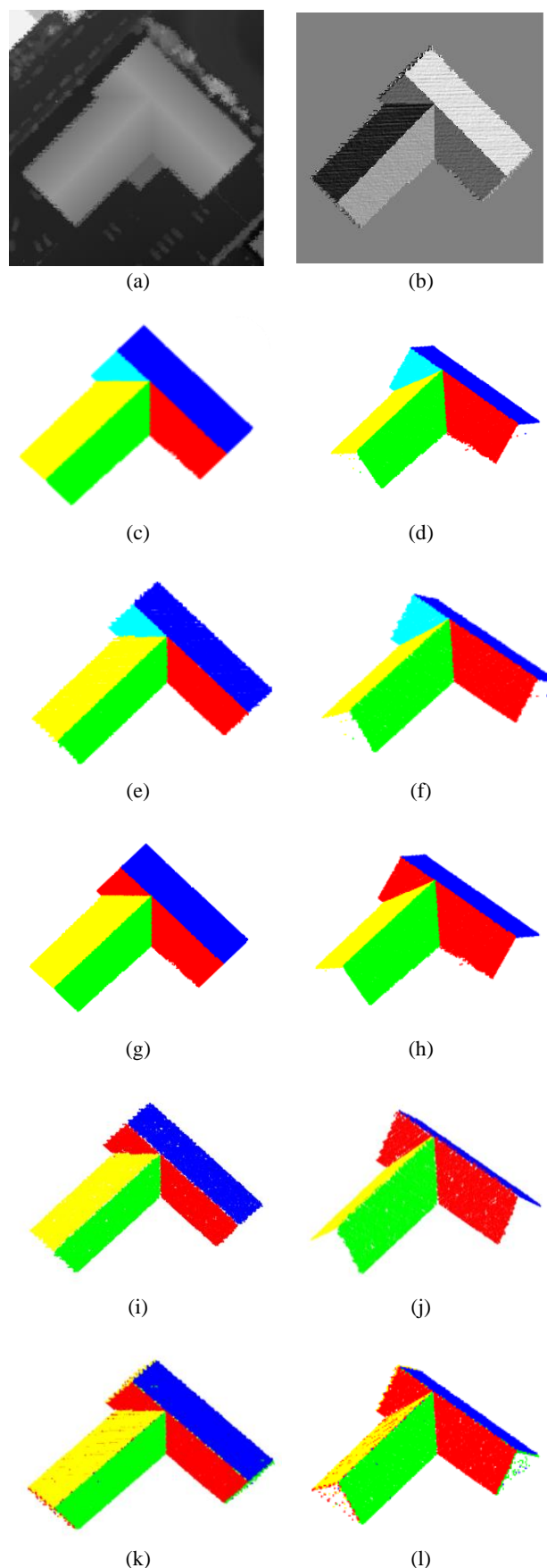
Random sample consensus (RANSAC) is a method to calculate the parameters of a mathematical model from a set of data with outliers that we use to calculate the parameters of the roof surface mathematical model. In this method, outliers do not affect the estimation of mathematical model parameters. Therefore, this method can also be used as a method to identify outliers. This algorithm is a non-deterministic method in that it outputs only possible results with a certain probability. The basic model of this method was published by Fischler and Bolles. MLESAC or MSAC is generalized of RANSAC. This method utilizes the same RANSAC sampling strategy to generate possible solutions. It chooses the solution that maximizes the likelihood, not just the number of inliers (Fischler and Bolles 1981). The result of the MSAC approach is shown in Figure 3. (m) and (n).

4. EXPERIMENTAL PROTOCOL

4.1 Overview of the Approach

In this article, we have used the Stuttgart dataset on four different methodologies. First, the point cloud is cropped then the DSM is generated with the spatial resolution of 1 meter in Cloud Compare v2.12 alpha. As shown in Figure 3. (a) our main focus is on a complex roof that contains different inclined facades. For each methodology, the result is shown in both 2D and 3D views

to clarify the difference between these four approaches visually. Each facet is shown with a different colour.



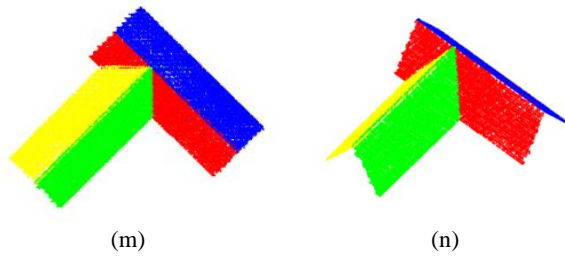


Figure 3. (a) The input dataset, (b) The Gradient image, (c) Ground Truth in 2D view (5 segments), (d) Ground Truth 3D in 3D view (5 segments), (e) 2D region growing in 2D view (5 segments), (f) 2D region growing in 3D view (5 segments), (g) Ground Truth in 2D view (4 segments), (h) Ground Truth 3D in 3D view (4 segments), (i) 3D region growing in 2D view (4 segments), (j) 3D region growing in 3D view (4 segments), (k) k-means in 2D view (4 segments), (l) k-means in 3D view (4 segments), (m): MSAC in 2D view (4 segments), (n) MSAC in 3D view (4 segments).

4.2 Evaluation metrics

The result of the segmentation methods is quantitatively evaluated and compared with the reference data, which is known as ground truth. The ground truth result is obtained from manual segmentation. In this study, we used precision, recall, quality, and F1-score measures. Precision refers to the number of correctly extracted elements. (Vo, Truong-Hong et al. 2015). To evaluate the results, it is important to make a declaration of precision, which is described as follows: The true positives are the points in a segment that have a corresponding point in the ground truth, and the false positives are the points in a segment that do not have a corresponding point in the ground truth segment.

$$precision = \frac{|TP|}{|TP| + |FP|} \quad (1)$$

where TP = True Positives
 FP = False Positives

The recall is the number of the ground truth elements that were properly extracted and is sensitive to the existence of ground truth segments that were not detected. False Negatives are the points in the ground truth segment without a correspondence in the segments of the results. Thus, recall is expressed as:

$$recall = \frac{|TP|}{|TP| + |FN|} \quad (2)$$

where FN = False Negatives

The following metric measures the absolute quality of the segmentation model. The quality is computed from equation (3):

$$quality = \frac{|TP|}{|TP| + |FN| + |FP|} \quad (3)$$

The F1 score combines two criteria: precision and recall. The F1 score is used as a unique index to determine the capacity of the proposed method. The F1-measure is computed from Equation (4):

$$F1 = \frac{2 \cdot precision \cdot recall}{precision + recall} \quad (4)$$

4.3 Ground truth segmentation

The ground truth data was segmented manually. In this study, a convention of clearly identifying the different structural components of a building's roof parts has been used to generate the ground truth data. Because we have four or five segments as a result of the approaches, we decided to create two of these ground-truth data for evaluation. In Figure 3. (g) and (h), the four manually segmented structural components are depicted. In Figure 3. (c) and (d), the five manually segmented structural components are depicted. The ground truth was used to compare the performance of MSAC, k-means, 2D region growing, and 3D region growing with the help of precision, recall, quality, and F1-score, as defined in Section 4.2.

4.4 Experimental Results

In this section, the quantitative and qualitative results comparing MSAC, k-means, 2D region growing and 3D region growing algorithms are presented. The segmentation is performed on different parts of the roof of the building, without semantic labeling on the detected parts. However, we assigned semantic labels to the different segments of the building's roof for clear presentation.

4.4.1 Qualitative results

According to the results related to the studied data, it can be depicted that the two-dimensional and three-dimensional region growing methods have better results than the MSAC and k-means methods. The result of the MSAC method is more suitable than the k-means method. The MSAC and k-means methods could not segment the points on the border and assign them to the proper facet. However, the 2D and 3D region growing algorithms are more sensitive to this issue and were able to assign the border points to their appropriate facades. The 2D region growing method produced five segments for the complex roof and the 3D region growing method, the MSAC method and k-means method created four segments for this type of roof with related parameters.

4.4.2 Quantitative Results

Quantitative results are tabulated in Table.1 to Table.4 for the 2D region growing, 3D region growing, MSAC, and k-means methods. In order to compare the quantitative results of the output segments, we have merged the two output facades of the 2D region growing so the results of all methods had four segments and we could use the quantitative metrics. According to Table.1, 2D region growing achieves the best segmentation results among the four segmentation methodologies. It has the highest precision, quality, recall and F1-score in comparison with other quantitative results. In contrast, although k-means appears to have identified the lowest precision from the visual inspection, it has the second rank in the quantitative segmented results. As shown in Table. 2, the k-means method gains the precision of 0.96 and the quality of 0.81. On the other hand, the precision, quality, recall and F1-score of the 3D region growing method is lower than the k-means procedure and higher than the MSAC quantitative results. According to the results in Table.4, the MSAC based segmentation presents lower precision, as expected from the visual inspection, and the border points are not segmented as points of the facet.

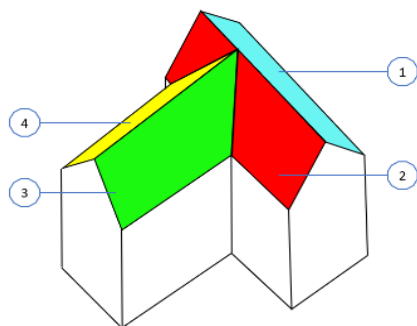


Figure 4. The four different facades

2D region growing				
	Recall	Precision	Quality	F1
Facet 1	0.8131	0.9986	0.8122	0.8964
Facet 2	0.8897	0.9864	0.8783	0.9351
Facet 3	0.8477	0.9980	0.8463	0.9168
Facet 4	0.8522	0.9895	0.8446	0.9158
mean	0.8507	0.9932	0.8452	0.9160

Table 1. Evaluation metrics for 2D region growing method.

k-means				
	Recall	Precision	Quality	F1
Facet 1	0.8363	0.9914	0.8303	0.9073
Facet 2	0.8568	0.9241	0.8004	0.8892
Facet 3	0.8350	0.9746	0.8172	0.8994
Facet 4	0.8527	0.9585	0.8224	0.9025
mean	0.8452	0.9621	0.8176	0.8996

Table 2. Evaluation metrics for k-means method.

3D region growing				
	Recall	Precision	Quality	F1
Facet 1	0.7599	0.9898	0.7540	0.8598
Facet 2	0.8035	0.9785	0.7895	0.8824
Facet 3	0.7879	0.9903	0.7819	0.8776
Facet 4	0.7976	0.9877	0.7898	0.8825
mean	0.7872	0.9866	0.7788	0.8756

Table 3. Evaluation metrics for 3D region growing method.

MSAC				
	Recall	Precision	Quality	F1
Facet 1	0.6790	0.9781	0.6688	0.8015
Facet 2	0.7074	0.9632	0.6888	0.8157
Facet 3	0.7610	0.9676	0.7421	0.8519
Facet 4	0.6654	0.9622	0.6484	0.7867
mean	0.7032	0.9678	0.6870	0.8140

Table 4. Evaluation metrics for MSAC method.

5. CONCLUSIONS

In this paper we presented a comparative study on segmenting complex roof facades from digital elevation data. The 2D region growing uses the planar coordinates for segmentation under the

condition of distance. As the input data was 2D gridded data, also according to the complex roof under studied, we have shown that the 2D region growing approach had the best performance among the segmentation approaches under studied. The four different methodologies needed initial settings as a necessary input. Moreover, the user will select the initial seed point in the 2D region growing approach, which improves the performance, while the other segmentation approaches start segmenting randomly. The roof used in this research is considered as a regularized roof with flat and hipped facades, so the 2D region growing method has obtained the best results. It is possible that noise-independent methods such as MSAC or 3D region growing have better results on point cloud.

REFERENCES

- Adam, A., et al. (2018). "H-RANSAC: A hybrid point cloud segmentation combining 2D and 3D data." *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci* **4**(2): 1-8.
- Alharthy, A. and J. Bethel (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.
- Arefi, H. and M. Hahn (2005). "A morphological reconstruction algorithm for separating off-terrain points from terrain points in laser scanning data." *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences* **36**(3/W19): 120-125.
- Dubey, S., et al. (2016). "Comparative study of various segmentation techniques with their effective parameters." *International Journal of Innovative Research in Computer and Communication Engineering* **4**(10): 17223-17228.
- Fischler, M. A. and R. C. Bolles (1981). "Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography." *Communications of the ACM* **24**(6): 381-395.
- Khaloo, A. and D. Lattanzi (2017). "Robust normal estimation and region growing segmentation of infrastructure 3D point cloud models." *Advanced Engineering Informatics* **34**: 1-16.
- Lafarge, F., et al. (2008). "Structural approach for building reconstruction from a single DSM." *IEEE Transactions on pattern analysis and machine intelligence* **32**(1): 135-147.
- Lafarge, F. and C. Mallet (2011). *Building large urban environments from unstructured point data*. 2011 International Conference on Computer Vision, IEEE.
- Louizi, M. and M. M. Gammoudi (2011). *Improved method for nose classification*. 2011 IEEE 3rd International Conference on Communication Software and Networks, IEEE.
- Rottensteiner, F. and C. Briese (2003). "Automatic generation of building models from LIDAR data and the integration of aerial images."
- Rottensteiner, F., et al. (2004). *Building detection by Dempster-Shafer fusion of LIDAR data and multispectral aerial imagery*. Proceedings of the 17th International Conference on Pattern Recognition, 2004. ICPR 2004., IEEE.

Samadzadegan, F., et al. (2010). "An agent-based method for automatic building recognition from lidar data." *Canadian Journal of Remote Sensing* **36**(3): 211-223.

Sampath, A. and J. Shan (2008). "Building roof segmentation and reconstruction from LiDAR point clouds using clustering techniques." *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* **37**(Part B3a).

Sheng, H., et al. (2016). "Feature extraction of SAR scattering centers using M-RANSAC and STFRFT-based algorithm." *EURASIP Journal on Advances in Signal Processing* **2016**(1): 1-16.

Tarsha-Kurdi, F., 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.

Vo, A.-V., et al. (2015). "Octree-based region growing for point cloud segmentation." *ISPRS Journal of Photogrammetry and Remote Sensing* **104**: 88-100.

Vosselman (2010). *Airborne and terrestrial laser scanning*. Boca Raton, CRC Press.

Vosselman, G. and S. Dijkman (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.

Zhang, W., et al. (2014). "3D building roof modeling by optimizing primitive's parameters using constraints from LiDAR data and aerial imagery." *Remote Sensing* **6**(9): 8107-8133.