

Building Extraction from LiDAR Point Clouds Based on Revised RandLA-Net

Yiru ZHANG^{1,2}, Tao WANG^{1,2*}, Xiangguo LIN³, Zihao ZHAO^{1,2}, Xiwei WANG^{1,2,4}

1. MOE Lab of 3D Spatial Data Acquisition and Application, Capital Normal University, Beijing 100048, China;

2. College of Resource Environment and Tourism, Capital Normal University, Beijing 100048, China;

3. Chinese Academy of Surveying and Mapping, Beijing 100830, China;

4. Beijing Institute of Surveying and Mapping, Beijing 100038, China

Keywords: LiDAR point clouds, RandLA-Net, building extraction, rooftop structure.

Abstract

3D building models is crucial for applications in smart cities. Automatic reconstruction of 3D buildings has been investigated based on various data sources. Point clouds from airborne LiDAR scanners can be used to extract buildings data due to its high accuracy and point density. In this paper, we present a methodology to segment buildings and corresponding rooftop structure from point clouds. First, RandLA-Net, which is an efficient and lightweight neural network for semantic segmentation of large-scale point clouds, is revised and adopted for building segmentation. By implementing local feature aggregation of each point, RandLA-Net can effectively preserve geometric details in point clouds. Besides 3D coordinates of point clouds, we incorporated point attributes including pulse intensity and return numbers into the network as additional features. Feature normalizations are applied to the input features. To achieve a better result of the local feature aggregation, hyperparameters of the network are fine-tuned according to the density of points and building size. Based on the classified building point clouds, DBSCAN clustering algorithm is implemented for segmenting individual buildings. Elevation histogram analysis is conducted to determine optimal threshold values for delineating candidate rooftop point clouds of individual buildings. For the buildings with multiple rooftops, multiple elevation threshold values are necessary to extract corresponding rooftops or walls. Then DBSCAN is employed again for segmentation of individual rooftops and denoising of point clouds of each building. Finally, Alpha-shape analysis is applied based on adaptive threshold values to build the envelope of each rooftop. Experiments show that our implementation of building segmentation using RandLA-net achieves higher mean IoU (Intersection over Union) and better classification performance in building segmentation. ISPRS benchmark data was used in our experiment and our methodology produce results with accuracy of 90.79%.

1. Introduction

With the rapid advancement of global urbanization, it is important to find solutions to improve spatial efficiency of dense urban space. Building rooftops are receiving significant attention as spaces for developing vertical greening, solar panels deploying and other applications (Mahmoud et al., 2022; Yang et al., 2023; Li et al., 2023). To effectively develop and utilize building rooftops space, detailed three-dimensional structural information is required, especially accurate information regarding the rooftops and their properties, like area, slope, orientation and structural layouts.

Airborne LiDAR (Light Detection and Ranging) is an important means of obtaining detailed three-dimensional structural data of building. Generally, there are four widely used segmentation approaches for point clouds segmentation: model fitting (Tarsha-Kurdi et al., 2007; Li et al., 2017; Adam et al., 2018), region growing (Vo et al., 2015; Zhao et al., 2021), data clustering (Zhou et al., 2016; Kim et al., 2016), and energy minimization (Sun et al., 2013; Yan et al., 2014). These segmentation methods have their specific advantages in terms of accuracy and efficiency. At the same time, they face different application constraints, such as single segmentation scale, over-segmentation, and under-segmentation. In order to improve the accuracy and efficiency of point cloud segmentation, researchers usually adopt a multi-scale, multi-level segmentation strategy, combining two or more methods to meet the needs of 3D modelling. However, in complex urban scenes

with varied terrain features, traditional methods need further optimization in practical applications.

In this work, we firstly implemented RandLA-Net (Hu et al., 2020; Hu et al., 2021) based on local feature aggregation and normalization of point attributes with fine-tuned hyperparameters for building segmentation. Then a clustering algorithm was implemented for identifying each individual building, whose rooftops were segmented based on optimal elevation threshold values. Thirdly the clustering algorithm and the alpha shape algorithm were used to delineate rooftop boundaries. In the last section, results and accuracies are to be reported.

2. Methodology

2.1 Building Point Clouds Segmentation Based on Revised RandLA-Net

To address the difficulty in setting hyper-parameters for building point cloud extraction and the misclassification between building point clouds and surrounding point clouds, this paper conducts research on the application of the deep learning-based point cloud semantic segmentation network, RandLA-Net. The study focuses on two main aspects to enhance the performance of the network model of RandLA in building point clouds extraction: enriching and optimizing the local features of point clouds input into the network and improving the network structure to increase the aggregation range of global features.

In addition to three-dimensional coordinates (X, Y, Z), this study adds three features, namely pulse intensity, return number and relative elevation as input features for the neural network. These three attribute features have distinct advantages in discriminating terrain features, and complement each other. Introduction of these three features to the three-dimensional coordinate (X, Y, Z) features will not affect the generality of the trained semantic segmentation model.

Generally, the range of values for pulse intensity information is normally either 0~255 or 0~65535. The range of values for return number is normally 1~7. And the relative elevation of ground features is typically less than 200m in most typical urbanized regions. The ranges of values for these three parameters are not consistent. This may affect the network's learning mechanism of point cloud features. Due to the significant difference in numerical scales, in the early stage of model training, the pulse intensity information component is particularly large, while the return number information component is very small. In the fully connected neural network, the feature values may be dominated by intensity information. To address this, feature normalization operation is implemented in this study for consistency of the ranges of values of the three features. This can optimize representation of features in the network model and enhance the network's adaptability to point cloud data generated by different LiDAR measurement systems. The normalization formula is as follows:

$$N_i = \frac{2(V_i - \min(V))}{\max(V) - \min(V)} - 1 \quad (1)$$

where V represents a set of input data to be normalized, V_i denotes the i^{th} value in this set of data, $\text{mean}(V)$, $\text{max}(V)$, and $\text{min}(V)$ respectively represent the mean, maximum, and minimum values of the array V . N_i represents the result after normalizing the numerical value V_i .

Each encoding layer consists of a local feature aggregation module and a random sampling operation. As shown in Figure 1, random sampling reduces the point density of the output data in each encoding layer, so the spatial range aggregated by the local feature aggregation in each encoding layer (dark solid circles in Figure 1) is larger than the spatial range aggregated by the local feature aggregation in the previous layer. The approximate influence range of a point in the encoder output data (circle in Output_Encoder in Figure 1) by input values from various layers is shown by the orange dashed circle in Figure 1. Based on the principles of local feature aggregation and sampling, the specific influence range of this point by input values from various layers can be inferred, as shown by the point cloud cluster on the right side of Figure 1.

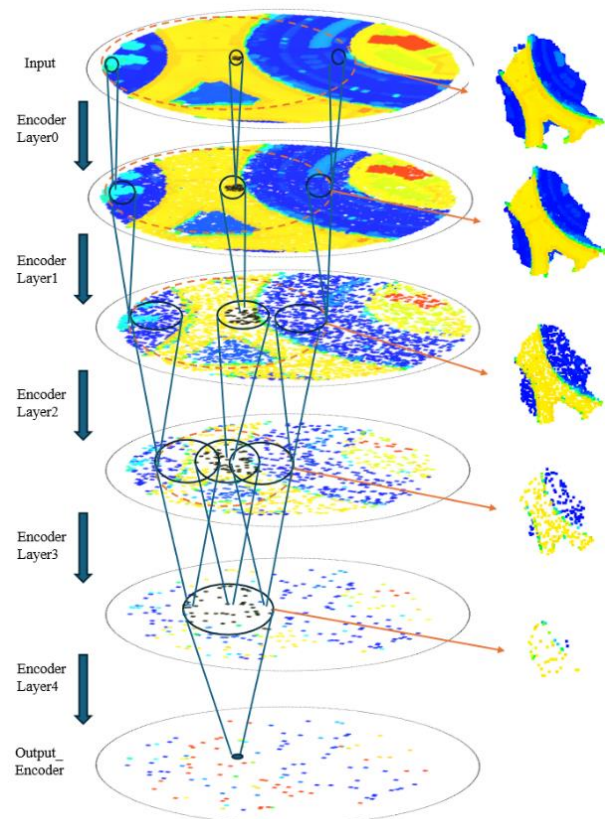


Figure 1. Aggregation of Contextual Information.

2.2 Rooftop Segmentation from Building Point Clouds

In terms of building point clouds segmented based on above RandLA-Net model, building facade points and building rooftop points (including rooftop surface points and rooftop components points) are mixed with each other and noise points exist, this paper proposes a bottom-to-up approach for rooftop point cloud segmentation.

Firstly, the elevation frequency distribution histogram is built and a peak-finding algorithm is implemented (Figure 2). Building rooftop points are to be filtered out from building facade points, while rooftops at different heights of individual buildings is achieved (Figure 3).

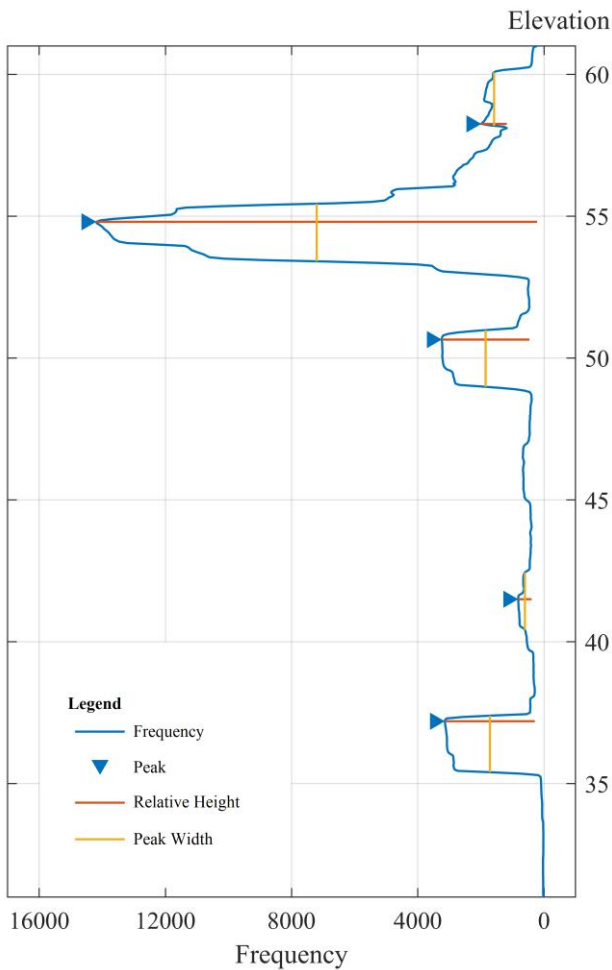


Figure 2. Peak-finding Algorithms Applied on an Elevation Frequency Distribution Histogram.

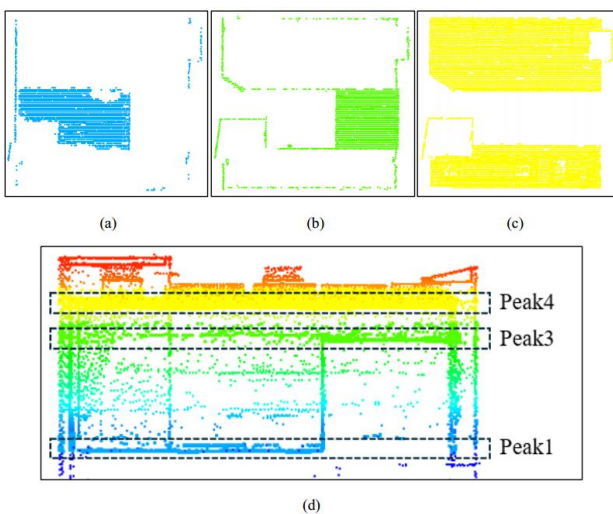


Figure 3. Extraction of Rooftops by Elevation Threshold.

Density-based clustering algorithm, DBSCAN, and building contour extraction based alpha shape algorithm are implemented. Rooftop point clouds that are not connected are extracted within the same building (Figure 4).

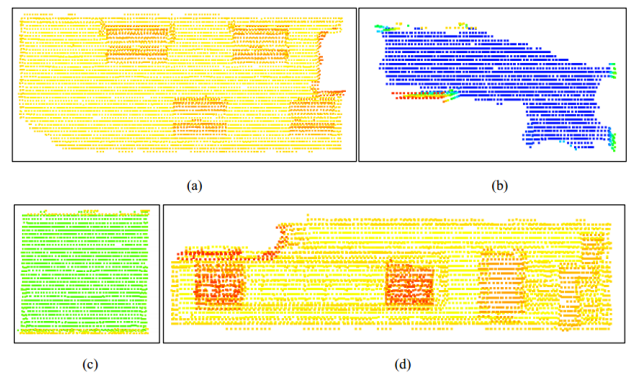


Figure 4. Extraction of Individual Rooftops.

Finally, cloth simulation filtering algorithm is implemented where rooftop surface point clouds are treated as ground point clouds, and rooftop components point clouds are treated as non-ground point clouds. Then rooftop component point clouds are segmented out (Figure 5).

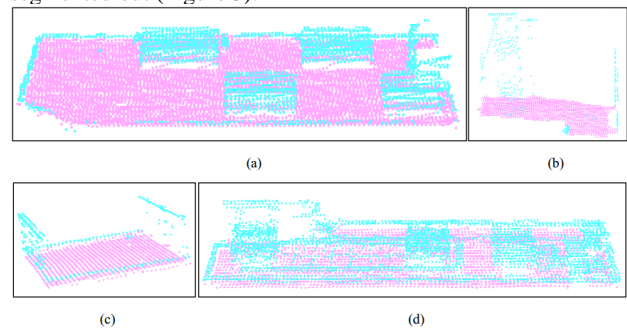


Figure 5. Segmentation of Individual Rooftops.

3. Results

3.1 Data and Environment

To evaluate the adaptability of the proposed methodology in this study, experiments were conducted on the ISPRS Toronto benchmark dataset for building point clouds extraction. This dataset also includes attribute information such as pulse intensity and return number, with pulse intensity ranging from 0 to 541 and the return number ranging from 1 to 4. The average point density of this dataset is approximately 6 points per square meter. This study conducted experiments using a deep learning workstation equipped with NVIDIA GeForce RTX 3080 GPU and utilized TensorFlow version 1.15 provided by NVIDIA as the deep learning framework. The specific software and hardware configurations are shown in Table 1.

Hardware/Software	Configuration
CPU	Intel Xeon Gold 6128 * 2
GPU	NVIDIA GeForce RTX 3080
System Memory	256GB
GPU Memory	10GB
OS	Windows 10 64bit
CUDA	11.2
Developing Environment	Python 3.7
Deep Learning Framework	NVIDIA TensorFlow 1.15

Table 1. Software and Hardware Configurations

3.2 Result of Building Point Clouds Extraction

In the experiment of building segmentation based on the revised RandLA-Net network, the Toronto dataset was denoised and relative elevations were calculated. Annotated samples were constructed. Point cloud data containing X, Y, Z, reflectance intensity, return number, and relative elevation were input into the network, and feature normalization was applied. The model achieved the best validation accuracy of 92.745% at the 68th Epoch, with an average time consumption of 235.69s per Epoch. By applying this model to classify the complete Toronto dataset and evaluating the classification results on the official validation area Area4, the confusion matrix is shown in Table 2. The experiment obtained a classification accuracy of 90.79%, a recall rate of 88.26%, and a false alarm rate of 7.55%.

Pre \ True	Building	Non-Building	Total
Building	493219	64178	557397
Non-Building	65582	785735	851317
Total	558801	849913	1408714

Table 2. Confusion matrix of the segmentation result in Area 4

According to the provided official building outlines in Area4 for accuracy evaluation, the extracted point cloud of buildings was clipped to obtain the results as shown in Figure 6. From the figure, it can be observed that a small number of buildings exhibit omission issues, primarily due to the presence of multiple rooftops at different heights, with the lowest rooftop not successfully identified, as illustrated in Figure 7. The omission of building point clouds exhibits spatial contextual features similar to ground points at terrain undulations, which may be the main reason for model misjudgement. For the buildings in the southernmost part of Area4, detailed results are shown in Figure 8, according to point cloud data and reference high-resolution optical imagery, the missed area is at the same height as the ground, not pertaining to architectural structures, this is an error in the validation sample.



Figure 6. The accuracy evaluation of building extraction in the Toronto Area 4 accuracy evaluation area.

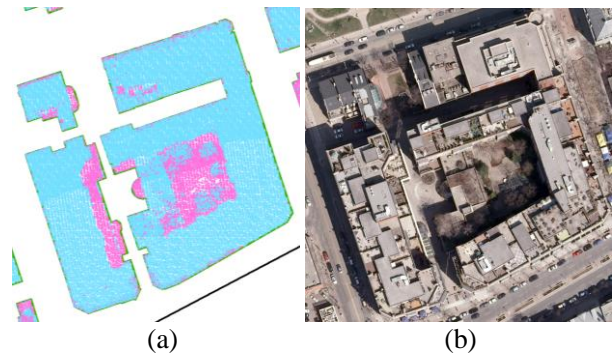


Figure 7. Example of missed points due to low roof height.

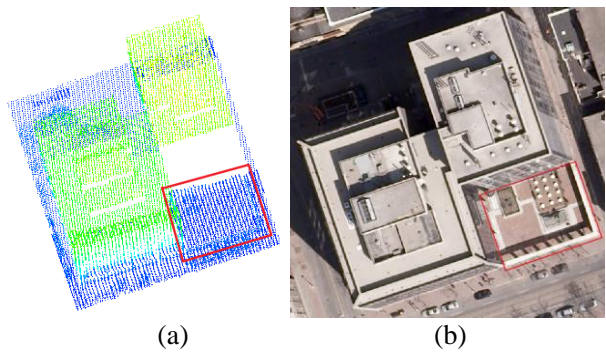


Figure 8. Example of missed scores due to incorrect labelling of validation samples.

3.3 Results of Rooftop Point Clouds Segmentation

In the rooftop component extraction, the generic parameters proposed for the study area in this paper were used to extract rooftop information from individual buildings in Toronto data. As shown in Figure 9, a building with a simple multi-roof structure is depicted, where the blue points represent rooftop surface points and the pink points represent rooftop component points, which can effectively identify items placed on the roof. For the complex multi-roof structure building shown in Figure 10, the method provides a comprehensive search for the rooftop surfaces, with only a small area in the lower left corner where some surfaces were not extracted, while the rest of the rooftop information is relatively complete, and the identification of rooftop components is accurate. For the building with a large number of components on the rooftop as shown in Figure 11, this method can effectively identify the position and shape of the components and can also identify rooftop railing points. In conclusion, the generalization ability of the rooftop information extraction method proposed in this study is good, and it can demonstrate good performance in different study areas and data sources when using generic parameters.

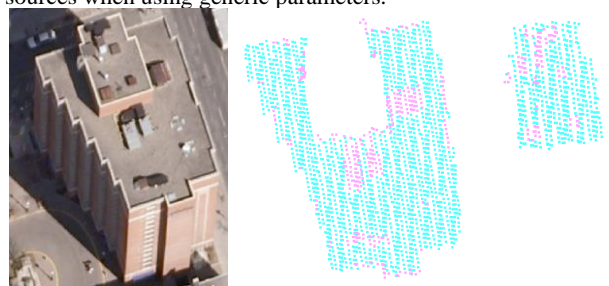


Figure 9. Simple multi-roof structure building.

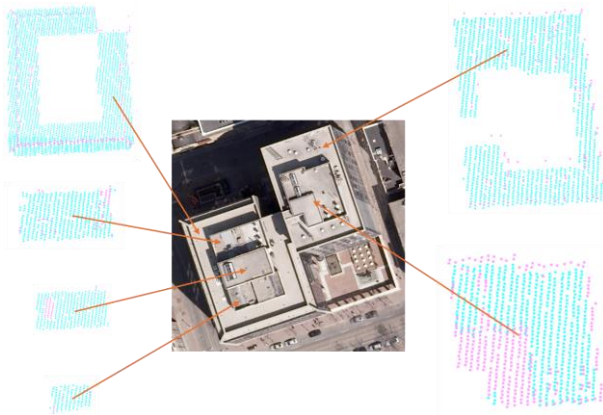


Figure 10. Complex multi-roof structure building.

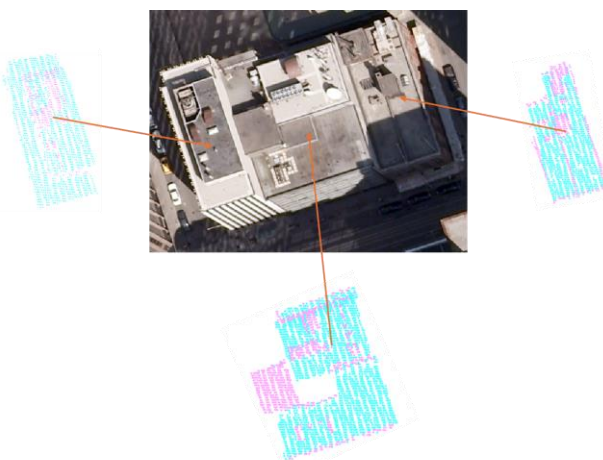


Figure 11. Building with a large number of components on the rooftop.

4. Conclusion

In this work, we proposed a methodology to segment building and rooftop component structures from dense 3D point clouds. RandLA-Net is used, and the input data are enriched by introducing extra features of point clouds. Improved performance on efficiency and accuracy is achieved. Based on DBSCAN algorithm, point clouds of individual buildings are delineated which are further segmented into rooftop by optimal threshold values acquired from elevation histogram analysis. A filtering algorithm and the alpha shape detection algorithm are implemented for rooftop structure components deification. ISPRS benchmark data was used in experiment which prove the proposed approach is effective concerning accuracy of building point clouds segmentation.

In the future we are going to test the methodology in other typical urban regions.

Acknowledgements

The authors would like to acknowledge the provision of the Downtown Toronto data set by Optech Inc., First Base Solutions Inc., GeoICT Lab at York University, and ISPRS WG III/4.

This research was partly funded by the National Natural Science Foundation of China, grant number 41671403.

References

- Adam, A., Chatzilari, E., Nikolopoulos, S., Kompatsiaris, I., 2018. H-RANSAC: A Hybrid Point Cloud Segmentation Combining 2D and 3D Data. *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.*, IV-2, 1-8. doi.org/10.5194/isprs-annals-IV-2-1-2018.
- Hu, Q., Yang, B., Xie, L., Rosa, S., Guo, Y., Wang, Z., Trigoni, N., Markham, A., 2020. RandLA-Net: Efficient Semantic Segmentation of Large-Scale Point Clouds. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 11105-11114. doi.org/10.1109/CVPR42600.2020.01112.
- Hu, Q., Yang, B., Xie, L., Rosa, S., Guo, Y., Wang, Z., Trigoni, N., Markham, A., 2021. Learning Semantic Segmentation of Large-Scale Point Clouds with Random Sampling. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 44(11): 8338-8354. doi.org/10.1109/TPAMI.2021.3083288.
- Kim, C., Habib, A., Pyeon, M., Kwon, G., Jung, J., Heo, J., 2016. Segmentation of Planar Surfaces from Laser Scanning Data Using the Magnitude of Normal Position Vector for Adaptive Neighborhoods. *Sensors*, 16(2), 140. doi.org/10.3390/s16020140.
- Li, L., Yang, F., Zhu, H., Li, D., Li, Y., Tang, L., 2017. An Improved RANSAC for 3D Point Cloud Plane Segmentation Based on Normal Distribution Transformation Cells. *Remote Sensing*, 9(5), 433. doi.org/10.3390/rs9050433.
- Li, N., Zhao, F., Chen, L., 2023. Review of rooftop solar photovoltaic electrical potential estimation: approaches and models. *Acta Ecologica Sinica*, 43(10), 4284-4293. doi.org/10.5846/stxb202205041230.
- Mahmoud, A. S., 2022. Overview of Green Roof Technology as a Prospective Energy Preservation Technique in Arid Regions. *Engineering, Technology & Applied Science Research*, 12(4), 8982-8989. doi.org/10.48084/etasr.4919.
- Sun, S., Salvaggio, C., 2013. Aerial 3D Building Detection and Modeling From Airborne LiDAR Point Clouds. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 6(3), 1440-1449. doi.org/10.1109/JSTARS.2013.2251457.
- Tarsha-Kurdi, F., Landes, T., Grussenmeyer, P. 2007. Hough-Transform and Extended RANSAC Algorithms for Automatic Detection of 3D Building Roof Planes from Lidar Data. *ISPRS Workshop on Laser Scanning 2007 and SilviLaser 2007*, 36, 407-412.
- Vo, A., Truong-Hong, L., Laefer, D. F., Bertolotto, M., 2015. Octree-based region growing for point cloud segmentation. *Isprs Journal of Photogrammetry and Remote Sensing*, 104, 88-100. doi.org/10.1016/j.isprsjprs.2015.01.011.
- Yan, J., Shan, J., Jiang, W., 2014. A global optimization approach to roof segmentation from airborne lidar point clouds. *ISPRS Journal of Photogrammetry and Remote Sensing*, 94, 183-193. doi.org/10.1016/j.isprsjprs.2014.04.022.

Yang, S., Wen, P., Jiang, F., Xu, J., Xie, Z., Zhang, G., 2023. Urban flood runoff and waterlogging characteristics-based study on optimization of green roofs spatial layout. *Water Resources and Hydropower Engineering*, 54(04), 22-36. doi.org/10.13928/j.cnki.wrahe.2023.04.002.

Zhao, C., Guo, H., Lu, J., Yu, D., Lin, Y., Jiang, H., 2021. Roof segmentation from airborne LiDAR by combining region growing with random sample consensus. *Acta Geodaetica et Cartographica Sinica*, 50(5), 621-633. doi.org/10.11947/j.AGCS.2021.20200270.

Zhou, G., Cao, S., Zhou, J., 2016. Planar Segmentation Using Range Images From Terrestrial Laser Scanning. *IEEE Geoscience and Remote Sensing Letters*, 13(2), 257-261. doi.org/10.1109/LGRS.2015.2508505.