DOI 10.34658/jacs.2018.26.2.99-106

JOURNAL OF APPLIED
COMPUTER SCIENCE
Vol. 26 No. 2 (2018), pp. 99-105

A Review on Point Cloud Semantic Segmentation Methods

Jagoda Lazarek, Michał Pryczek

Lodz University of Technology
Institute of Information Technology
Wólczańska 215, 90-924 Lodz, Poland
jagoda.lazarek@p.lodz.pl

Abstract. *Semantic segmentation of 3D point clouds is an open research problem and remains crucial for autonomous driving, robot navigation, human-computer interaction, 3D reconstruction and many others. The large scale of the data and lack of regular data organization make it a very complex task. Research in this field focuses on point cloud representation (e.g., 2D images, 3D voxels grid, graph) and segmentation techniques. In the paper, state-of-the-art approaches related to these tasks are presented.*
Keywords: *Point Clouds, Segmentation, Classification, Laser Scanning.*

1. Introduction

Point clouds are created as a result of 3D scanning space indoor or outdoor with LiDARs (Light Detection And Ranging). Such scans are characterized by tremendous accuracy of mapping, however, this is associated with a very large amount of generated data (millions of points), which must then be processed. Among known, publicly available datasets Semantic3D [1] and S3DIS [2] should be listed, which contain a large labelled 3D point cloud data set of natural scenes.

To extract useful information from such data particular objects need to be segmented, then recognized accordingly to the context. Segmentation of largescale point clouds is challenging. There are two main obstacles: the scale of the data – millions of points to process, the lack of regular data arrangement (clear structure) as a grid in images. Because of that is impossible to use proven and well-known image segmentation techniques directly. It is a reason why methods dedicated to point cloud segmentation are developed. In the paper, we present state-of-the-art solutions.

2. Point Cloud Representation

First, a very important step is a choice of point cloud representation, which has an influence on the result of segmentation. Several representations have been proposed:

- collections of 2D images, e.g., set of virtual 2D RGBD snapshots [3], spherical images [4];
- regular 3D voxel grid [5, 6];
- superpoint graph (SPG) [7];
- 3D point cloud (directly) [8, 9, 10, 11, 12].

Two first representations convert 3D point cloud to 2D form, while third and fourth ones preserve 3D structure. A transformation from 3D into 2D simplify problem but at the same time cause some problems:

- loss of information,
- eliminates details,
- requires surface reconstruction.

3. State-of-the-art methods of Point Cloud Segmentation

A possibility of using specific segmentation methods depends on point cloud representation directly. As was mentioned in the previous section, following approaches may be distinguished:

- 2D point cloud representation – SnapNet [3], PointSeg [4];

- unprocessed 3D point cloud – SPLATNet [9], PointNet [8], PointNet++ [10], PointSIFT [11], 3P-RNN [12];
- 3D voxel grid – SEGCloud [5], VGS and SVGS [6].

Methods based on 2D point cloud representation

SnapNet [3] is based on processing 2D representation of point cloud. In first step two types of 2D views (snapshots) are generated – RGB images and depth composite view (with geometric features), then 2D views are segmented separately using deep segmentation network (CNN). Finally obtained results are projected on the original data.

PointSeg [4] solution starts from transforming point cloud into spherical images, which become inputs of the convolutional neural networks (CNNs). CNNs predict the point-wise semantic mask. The method is based on SqueezeNet [13]. Segmentation results obtained on spherical images are back-projected on point cloud. Authors tested solution on the road-object segmentation task but didn't achieve good results in segmenting small objects – like pedestrians.

Methods based on 3D voxel grid point cloud representation

SEGCloud [5] is a solution works on a regular voxel grid, which is processed by a 3D fully convolutional neural network (3D-FCNN). They use trilinear interpolation (TI) and fully connected Conditional Random Fields (FC-CRF) to project obtained segmentation results on the original data (point cloud).

VGS (voxel- and graph-based segmentation) and SVGS (supervoxel- and graph-based segmentation) [6] are methods which work on voxel grid as well. Authors use octree-based voxelization to create 3D grid point cloud representation. Next, they calculate the geometric cues between voxels or supervoxels accordingly, which are used for weighting edges in the graph model. Finally, the graph-based clustering is used to segment voxels or supervoxels into homogeneous segments.

The solution proposed in [14] is based on a 2D voxel representation of a point cloud. Authors use two different 2D CNN architectures to process the voxel representation – 2D FCNN (2D Fully Convolutional Neural Network) and 2D U-Net to compute per-voxel class scores, which are projected then to the 3D point cloud. Finally, per-point class scores are obtained.

Methods based on 3D point cloud directly

SPLATNet (SParse LATtice Network) [9] processes 3D point cloud directly, without any preprocessing. Hierarchical and spatially-aware features are computed with the use of sparse and lattice filters. Authors point moreover that their solution gives a possibility to map 2D views to 3D and vice-versa.

PointNet [8] is a solution based on deep learning techniques, which provides good results in point cloud analysis, which takes point cloud as direct input. Method is based on deep net architecture and may be used for 3D shape classification, shape part segmentation and scene semantic parsing tasks. They use two networks – classification and segmentation network. At first, input and feature transformations are applied, then point features are aggregated by max pooling. To provide segmentation results, they use segmentation network. Usage of PointNet is limited by the size of point cloud which may be processed.

PointNet++ [10] is an extended version of PointNet [8] method, based on hierarchical structure. They introduced "set abstraction" levels, each consecutive level reduces the number of elements in a set.

PointSIFT [11] is based on SIFT concept [15] known from image segmentation domain and uses parametric deep learning to provide results. Method encodes information of eight orientations (orientation-encoding unit) and is adaptive to scale of shape. They use stack of several orientation-encoding units to achieve multi-scale representation. Solution is based on network with a encode-decode (downsample-upsample) framework.

3P-RNN [12] is a solution which takes into account the inherent contextual features during a process of semantic segmentation. It is built based on two modules – the pointwise pyramid pooling (3P) module and the two-direction hierarchical recurrent neural networks (RNNs). Such framework lets extract local structures at various densities (multi-scale neighbourhood analysis) and long-range spatial dependencies.

Methods based on lossless 3D point cloud representation

Landrieu and Simonovsky [7] proposed a semantic segmentation method which works on 3D structure – a superpoint graph (SPG). SPG is a result of partitioning point cloud into geometrically homogeneous elements (simple shapes), which become nodes (superpoints). Edges represent relationships between superpoints (nodes). SPG is processed by a graph convolutional network. Authors point many advantages of that solution:

- whole object parts are classified (instead of points or voxels),
- detailed relationships between adjacent nodes are introduced,
- size of SPG depends on the number of simple shapes instead of the number of points.

4. Conclusions

Among point cloud segmentation methods three directions in point cloud representation may be distinguished – 2D, 3D representation or whole, unprocessed point cloud. Choice of representation determines segmentation approaches which may be used. Methods based on 2D or 3D representations tend to be faster but their accuracy is limited because of omitting part of point dependencies. On the other hand, usage of methods which process point cloud directly provide better segmentation results but it is limited by the size of point cloud which may be processed at once.

One of the most interesting solutions, which is based on processing SPG representation of point cloud, eliminates most of the listed obstacles and sets an interesting direction for the development of point cloud segmentation methods.

References

[1] Hackel, T., Savinov, N., Ladicky, L., Wegner, J. D., Schindler, K., and Pollefeys, M., *SEMANTIC3D.NET: A new large-scale point cloud classification benchmark*, In: ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. IV-1-W1, 2017, pp. 91–98.

[2] Armeni, I., Sener, O., Zamir, A. R., Jiang, H., Brilakis, I., Fischer, M., and Savarese, S., *3D Semantic Parsing of Large-Scale Indoor Spaces*, In: Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition, 2016.

[3] Boulch, A., Saux, B. L., and Audebert, N., *Unstructured Point Cloud Semantic Labeling Using Deep Segmentation Networks*, In: Eurographics Workshop on 3D Object Retrieval, edited by I. Pratikakis, F. Dupont, and M. Ovsjanikov, The Eurographics Association, 2017.

[4] Wang, Y., Shi, T., Yun, P., Tai, L., and Liu, M., *PointSeg: Real-Time Semantic Segmentation Based on 3D LiDAR Point Cloud*, CoRR, Vol. abs/1807.06288, 2018.

[5] Tchapmi, L. P., Choy, C. B., Armeni, I., Gwak, J., and Savarese, S., *SEGCloud: Semantic Segmentation of 3D Point Clouds*, CoRR, Vol. abs/1710.07563, 2017.

[6] Xu, Y., Hoegner, L., Tuttas, S., and Stilla, U., *Voxel- and Graph-Based Point Cloud Segmentation of 3D Scenes Using Perceptual Grouping Laws*, ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. IV-1/W1, 05 2017, pp. 43–50.

[7] Landrieu, L. and Simonovsky, M., *Large-scale Point Cloud Semantic Segmentation with Superpoint Graphs*, CoRR, Vol. abs/1711.09869, 2017.

[8] Qi, C. R., Su, H., Mo, K., and Guibas, L. J., *PointNet: Deep Learning on Point Sets for 3D Classification and Segmentation*, CoRR, Vol. abs/1612.00593, 2016.

[9] Su, H., Jampani, V., Sun, D., Maji, S., Kalogerakis, E., Yang, M.-H., and Kautz, J., *SPLATNet: Sparse Lattice Networks for Point Cloud Processing*, In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018, pp. 2530–2539.

[10] Qi, C. R., Yi, L., Su, H., and Guibas, L. J., *PointNet++: Deep Hierarchical Feature Learning on Point Sets in a Metric Space*, CoRR, Vol. abs/1706.02413, 2017.

[11] Jiang, M., Wu, Y., and Lu, C., *PointSIFT: A SIFT-like Network Module for 3D Point Cloud Semantic Segmentation*, CoRR, Vol. abs/1807.00652, 2018.

[12] Ye, X., Li, J., Huang, H., Du, L., and Zhang, X., *3D Recurrent Neural Networks with Context Fusion for Point Cloud Semantic Segmentation*, In: The European Conference on Computer Vision (ECCV), September 2018.

[13] Iandola, F. N., Moskewicz, M. W., Ashraf, K., Han, S., Dally, W. J., and Keutzer, K., *SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and <1MB model size*, CoRR, Vol. abs/1602.07360, 2016.

[14] Zhang, C., Luo, W., and Urtasun, R., *Efficient Convolutions for Real-Time Semantic Segmentation of 3D Point Clouds*, In: 2018 International Conference on 3D Vision, 3DV 2018, Verona, Italy, September 5-8, 2018, 2018, pp. 399–408.

[15] Lowe, D. G., *Distinctive Image Features from Scale-Invariant Keypoints*, Int. J. Comput. Vision, Vol. 60, No. 2, nov 2004, pp. 91–110.