# A BOUNDARY-ENHANCED SUPERVOXEL METHOD FOR 3D POINT CLOUDS

*Zhengchuan Sha[1], Qing Zhu[1], Yiping Chen[1*], Cheng Wang[1], Abdul Nurunnabi[2], Jonathan Li[1*,3]*

[1]Fujian Key Laboratory of Sensing and Computing for Smart Cities, School of Informatics
Xiamen University, Xiamen, Fujian 361005, China
[2]Faculty of Science, Technology and Communication,
University of Luxembourg, Belval Campus,
2 avenue de l'Université, L-4365 Esch-sur-Alzette, Luxembourg
[3]Departments of Geography and Environmental Management and Systems Design Engineering,
University of Waterloo, Waterloo, Ontario N2L 3G1, Canada
*Corresponding authors: chenyiping@xmu.edu.cn; junli@xmu.edu.cn (J. Li)

IGARSS 2020 - 2020 IEEE International Geoscience and Remote Sensing Symposium | 978-1-7281-6374-1/20/$31.00 ©2020 IEEE | DOI: 10.1109/IGARSS39084.2020.9323330

## ABSTRACT

This paper presents a boundary-enhanced supervoxel method to solve over-segmentation problems in supervoxel generation of Voxel Cloud Connectivity Segmentation (VCCS). First, we use different searching methods to obtain the neighborhood of each point. Second, three variants of neighbor points are clustered by the local *k*-means clustering method on points directly instead of on voxels. Finally, a scale metric is used to measure the difference between two points that considers underlying 3D spatial structure of the points. Our proposed is tested on two publicly available benchmark point cloud datasets acquired by mobile laser scanning (MLS) and terrestrial laser scanning (TLS) systems, respectively. Results of the experiments show that the boundary recall approximately 7 and 4 times higher than VCCS for the best results, which our proposed methods are effective, and the cost time is feasible and effective.

***Index Terms***—Clustering, 3D point cloud, nearest neighbor, over-segmentation, scale metric.

## 1. INTRODUCTION

Precise processing of large-scale point clouds plays an essential role in 3D computer vision. In 2D image processing, clustering pixels from images into different regions, called superpixels. Points clustering is one of the most significant use in point cloud processing like segmentation and classification. Over-segmentation methods split point clouds into several supervoxels. Similar to superpixels, the use of supervoxels in 3D point clouds can greatly reduce the number of points and consequence is reducing time for respective point cloud processing. Currently, many researchers focus on grouping voxels by using methods such as object detection, scene segmentation and saliency estimation [4].

However, the supervoxel problem can result into the over-segmentation problem. Different from images, it is difficult to over-segment point clouds due to its large volume and underlying data complexity, they are not structured as images, and usually point clouds contain various complex objects of different shapes. So it is significant to design a novel supervoxel method to solve the over-segmentation problem.

Over the past decades, some studies focused on supervoxels to solve the over-segmentation problem. The authors in [1] first presented the supervoxel method called Voxel Cloud Connectivity Segmentation (VCCS) by octree based searching and local *k*-means clustering method [2]. And the VCCS method is the advanced algorithm to over-segment the points. However, the number of points that changes with the supervoxel resolutions have a detrimental effect on the results of VCCS and cannot preserve the high boundary recall value in large outdoor 3D point cloud scene. In [3], a boundary-enhanced supervoxel segmentation (BESS) method was proposed to solve the over-segmentation problem by enhancing boundary information. The BESS can detect the outdoor point cloud scene, but it needs ordered points in one direction that makes its application limited. In [4], another proposed algorithm considered supervoxel segmentation as a subset selection problem to preserve the boundary better, but it costs too much time. In [5], supervoxels were obtained by generating facet to extract the road boundary. However, the method in [5] needed to select a seed point and lost some boundary information. In order to solve the over-segmentation problem better, in this paper, our contributions are to use different nearest neighbor search methods to enhance the boundary information and modify the VCCS performed on points instead of on voxels.

The rest of the paper is organized as follows. Section 2 describes our two-step method in detail. Section 3 presents and discusses the experimental results. Section 4 concludes the paper.

## 2. METHOD

Our goal is to improve the traditional supervoxel method in two steps. First, we use three different search methods to obtain the neighbors of an interest point. Second, we modify the VCCS supervoxel method performed on points rather than voxel with a scale metric.

### 2.1. Nearest neighbor search

Without considering neighbor information of each point, traditional VCCS is calculating normal vector of each representative point based on the voxels directly. We improve the VCCS method by using nearest search methods to consider more neighbor information.

Considering the accuracy of neighbor information has an important effect on over-segmentation results. Using different search methods can obtain different types of neighbor information.

For a given set of 3D points $P = \{p_1, p_2, \cdots, p_N\}$ and a query point $q$, we find the nearest neighbor points' set that is closest to the query point $q$ w.r.t. the Euclidean distance:

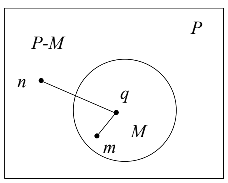



**Fig. 1** Diagram of the nearest neighbor search

$$NN(q, P) = \underset{p_i \in P}{\operatorname{argmin}}\, d(q, p_i) \quad (1)$$

where the $NN(\cdot)$ denotes finding the nearest neighbor point operation. $d(\cdot)$ is the Euclidean distance. The nearest neighbor search method has three regular ways: $k$-nearest neighbor search (kNN), radius nearest neighbor search (RNN) and radius $k$-nearest neighbor search (RkNN)[6]. The principle of nearest neighbor search is shown in Fig.1.

The kNN search is to obtain the closest $k$ points from the query point $q$. The definition of this search is written as:

$$kNN(q, k, P) = M \quad (2)$$

where $M$ is a point set. The cardinal of the $M$ is $k$ and $M$ is the subset of $P$. At the same time, for arbitrary $m \in M, n \in P - M$, $d(q, m) < d(q, n)$.

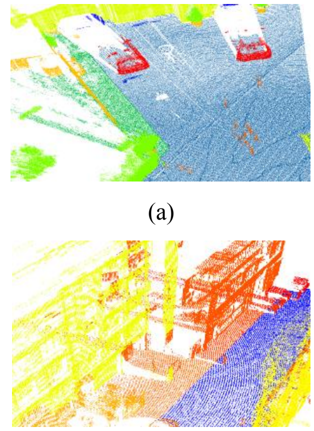
The RNN search is to find all the points located closer than the given radius $R$. The definition of RNN is defined as follows:

$$RNN(q, R, P) = \{p \in P, d(q, p) < R\} \quad (3)$$

The R$k$NN search is mainly a combination of $k$NN and RNN. We define the R$k$NN search that satisfies the condition:

$$RkNN(q, k, R, P) = M \quad (4)$$

where $|M| \le k, M \subseteq P$. For $\forall\, m \in M, n \in P - M$, we have $d(q, m) < R$ and $d(q, m) < d(q, n)$.



(a)

(b)

**Fig. 2** Our test benchmark dataset; (a) Untermaederbrunnen station from Semantic3D dataset [8], and (b) Cassette GT from IQTM dataset [9].

### 2.2. Supervoxel over-segmentation algorithm

After we obtain the nearest points for each point by aforementioned three search methods, we use PCA method [7] to calculate the normal vector of each neighbor point for next clustering based on VCCS.

The VCCS is an advanced over-segmentation algorithm that generates supervoxel by using a variant of $k$-means clustering [2]. Different from using the projection and depth information directly, the VCCS considered the 3D geometric information between two points. In [1], the VCCS method first voxelized the 3D points using octree and then partitioned the supervoxels evenly. So it is based on voxel directly. In order to consider more neighbor information in detail, we modify the VCCS so it can perform on points directly, and the above three nearest neighbor search methods (VCCS-kNN, VCCS-RNN and VCCS-RkNN) are used to obtain neighbors for further clustering points.

### 2.3. Metric method

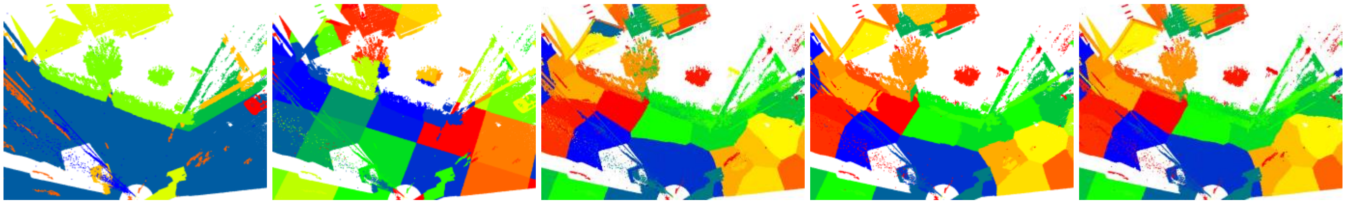
For VCCS-kNN, VCCS-RNN and VCCS-RkNN supervoxel, we use the same measure metric as defined in [1, 7]:

$$D(p, q) = 1 - \left|\overrightarrow{n_p} \cdot \overrightarrow{n_q}\right| + 0.4\frac{\|p-q\|}{R} \quad (5)$$

where the $\overrightarrow{n_p}$ and $\overrightarrow{n_q}$ are the normal vectors of two points $p$ and $q$, respectively. $|\cdot|$ denotes the inner product operation. $\|\cdot\|$ denotes the Euclidean distance. $R$ is the resolution of supervoxels. The measure metric considers the geometric information and location between two 3D points.

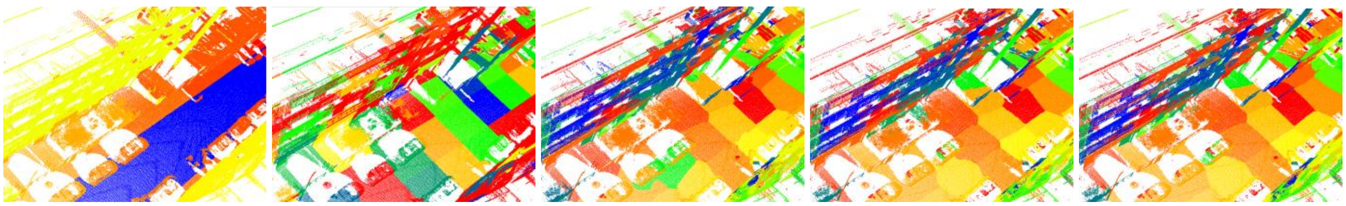
## 3. RESULTS AND DISCUSSION

As shown in Fig. 2, our experiments include two public outdoor 3D point cloud benchmark datasets, one is Semantic 3D (Untermaederbrunnen-station) dataset [8], and the other



(a) Ground-truth (b) VCCS (c) VCCS-kNN (d) VCCS-RNN (e) VCCS-RkNN

**Fig. 3** Visual representation of supervoxel results on Semantic 3D dataset.



(a) Ground-truth (b) VCCS (c) VCCS-kNN (d) VCCS-RNN (e) VCCS-RkNN

**Fig. 4** Visual representation of supervoxel results on IQTM dataset.

Table 1. The information of our tested scenes

| Scene | Numbers of points | Mean of $\bar{r}$ (cm) |
|---|---|---|
| Untermaederbrunnen_station | 1377290 | 1.98 |
| Cassette_GT | 1747830 | 2.49 |

is IQmulus and TerraMobilita (IQTM) (Cassette-GT) dataset [9]. Semantic 3D dataset is a large-scale point cloud classification benchmark collected by static terrestrial laser scanners. The IQTM dataset is acquired by a mobile laser scanning system. It has 12 millions of manually labeled points in Paris with 200m street of 15 labeled points. We down-sampled the dataset so that the numbers of points is approximately 1 million. Necessary information of our test scenes is described in Table 1.

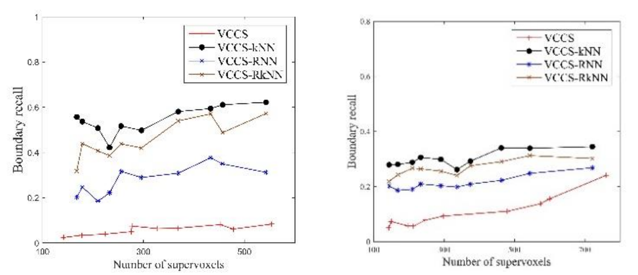
$\bar{r}$ is the average resolution of each scene [10]. It is calculated as the average distance between two adjacent points. In this paper, we set the parameters empirically, the $k$ in kNN and RkNN is 20, and the radius in RNN and RkNN is $10\bar{r}$. For three search methods, we use the KD-tree data structure to accelerate relevant processing. During the experiments, we evaluate the boundary recall and time for each scene by our methods. The boundary recall is defined in [11, 12, 13] as:

$$BR_{gtb}(\mathcal{S}) = \frac{\sum_{p\in gtb} \mathbb{I}\left(\min_{q\in svb}\|p-q\|<\varepsilon\right)}{gtb} \tag{6}$$

where $gtb$ and $svb$ denote the ground-truth boundaries and supervoxel boundaries in the form of binary. $\mathrm{I}(\cdot)$ is an indicator function, which equals to 0 when the dependent variable is 0, and equals to 1 in any other conditions [14]. The $\mathrm{I}(\cdot)$ ensures that whether the ground-truth boundary point is covered by obtained supervoxel boundaries. The threshold $\varepsilon$ is set to 0.03 empirically [4]. A high boundary recall value indicates that the supervoxels follow the boundary of objects at the groundtruth labels. Our method achieve the high boundary-recall (BR) values and preserve the boundary better than the traditional VCCS method. The pictorial presentation of the results from our proposed three methods, VCCS and ground-truth on Semantic 3D dataset are shown in Fig. 3. Meanwhile, visual representation of VCCS, VCCS-kNN, VCCS-RNN, VCCS-R$k$NN and ground-truth are in Fig. 4.

### 3.1. Boundary recall (BR)

As shown in Fig. 5(a) and (b), the proposed three methods have higher boundary recall than modified VCCS in different numbers of supervoxel $N$ on both Semantic3D and IQTM dataset. We can obtain that the proposed three methods can achieve a best BR value for all different numbers of supervoxel. Hence, our proposed methods can enhance the boundary more effective than traditional VCCS method on both datasets due to high BR values in the whole range of $N$.



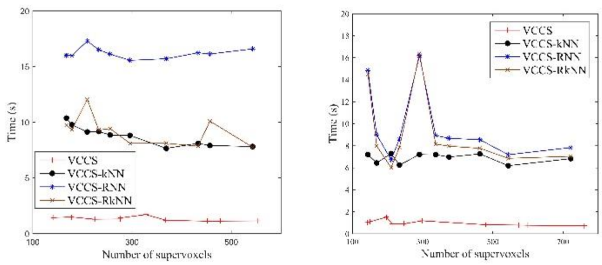

(a) (b)

**Fig. 5** Evaluation of four supervoxel methods on boundary recall metric. (a) Line diagrams for Semantic 3D benchmark dataset with the number of supervoxel $N \in [100, 600]$, (b) Line diagrams for

IQTM benchmark with the number of supervoxel $N \in [100,700]$.

### 3.2. Time performance

Fig. 6(a) shows cost time in four methods on Semantic 3D dataset. Although times to perform our methods are more than the modified VCCS, it is feasible and effective for processing a large dataset of 1.3 million points. Fig. 6(b) shows the comparison in running time of the four methods on IQTM dataset. Since our methods operate the points directly to consider more information of scenes, thus the methods are slower than modified VCCS. Results are similar for Semantic 3D dataset. But it is efficient and reasonable to process for the reason that the huge number of points.




(a) (b)

**Fig. 6** Evaluation of four supervoxel methods on time metric. (a) Line diagrams for Semantic 3D benchmark dataset with the number of supervoxel $N \in [100, 600]$, (b) Line diagrams for IQTM benchmark with the number of supervoxel $N \in [100,700]$.

## 4. CONCLUSION

In this paper, we have presented three variants of supervoxels for searching neighbor points based on improved VCCS. Different from the traditional method, our proposed methods considered three types of nearest neighbor search algorithms. The results from two publicly available 3D point cloud benchmark datasets demonstrated proposed methods are feasible and have high BR values. The BR values of VCCS-kNN, VCCS-RNN and VCCS-RkNN are around 50%, 45%, 20%, which is higher than VCCS when the number of supervoxels is about 600 on Semantic 3D. And on the IQTM dataset, the BR values of our methods are about 4, 3 and 3 times higher than the VCCS, respectively. We can observe that the BR values of VCCS-kNN are higher than those of VCCS, VCCS-RNN and VCCS-RkNN. Hence, our methods are effective. The reason might be that the scenes higher than ground belong to the same classes and kNN search can greatly obtain the neighborhood information. We intend to optimize the supervoxel methods further in future.

## ACKNOWLEDGEMENTS

This work was supported in part by the National Natural Science Foundation of China under Grants 41871380, 41471379, 61471371 and 61602499.

## REFERENCES


[1] P. Jeremie, A. Alexey, S. Markus, and W. Florentin, “Voxel cloud connectivity segmentation-supervoxels for point clouds,” in *Proc. CVPR*, 2013, pp. 2027-2034.

[2] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. Susstrunk, “Slice superpixels compared to state-of-theart superpixel methods,” *IEEE Trans. Pattern Anal. Machine Intell,* vol. 34, no. 11, pp. 2274-2282, 2012.

[3] S. Song, H. Lee, and S. Jo, “Boundary-enhanced super-voxel segmentation for sparse outdoor LiDAR data,” *Electron Lett,* vol. 50, no. 25, pp. 1917-1919, 2014.

[4] Y. Lin, C. Wang, D. Zhai, W. Li, and J. Li, “Toward better boundary preserved supervoxel segmentation for 3d point clouds,” *ISPRS J. Photogramm. Remote Sens.*, vol. 143, pp. 39-47, 2018.

[5] D. Zai, J. Li, Y. Guo, M. Cheng, Y. Lin, H. Luo, and C. Wang, “3-D road boundary extraction from mobile laser scanning data via supervoxels and graph cuts,” *IEEE Trans. Intell. Transport. Syst.*, vol. 19, no. 3, pp. 802-813, 2018.

[6] M. Muja and D. Lowe, “Scalable nearest neighbor algorithms for high dimensional data,” *IEEE Trans. Pattern Anal. Machine Intell*, vol. 36, no. 11, pp. 2227-2240, 2014.

[7] Y. Lin, C. Wang, B. Chen, D. Zai, and J. Li, “Facet segmentation-based line segment extraction for large-scale point clouds,” *IEEE Trans. Geosci. Remote Sens*., vol. 55, no. 9, pp. 4839-4854, 2017.

[8] T. Hackel, N. Savinov, L. Ladicky, J. Wegner, K. Schindler, and M. Pollefeys, “Semantic3D.net: A new large-scale point cloud classification benchmark,” *ISPRS Annuals*, vol. IV-1/W1, pp. 91-98, 2017.

[9] B. Vallet, M. Br´edif, A. Serna, B. Marcotegui, and N. Paparoditis, “TerraMobilita/iQmulus urban point cloud analysis benchmark,” *Comput. Graph.,* vol. 49, pp. 126-133, 2015.

[10] D. Zai, J. Li, Y. Guo, M. Cheng, P. Huang, X. Cao, and C. Wang, “Pairwise registration of TLS point clouds using covariance descriptors and a non-cooperative game,” *ISPRS J. Photogramm. Remote Sens.,* vol. 134, pp. 15-29, 2017.

[11] M. Liu, O. Tuzel, S. Ramalingam, and R. Chellappa, “Entropy rate superpixel segmentation,” in *Proc. IEEE CVPR*, 2011, pp. 2097-2104.

[12] Z. Ye, R. Yi, M. Yu, Y. Liu, and Y. He, “Fast computation of content-sensitive superpixels and supervoxels using q-distances,” in *Proc. ICCV*, 2019, pp. 3770-3779.

[13] Y. Shabat, T. Avraham, M. Lindenbaum, A. Fischer, “Graph based over-segmentation methods for 3d point clouds,” *Comput. Vis. Image Underst*, vol. 174. 11, pp. 12-23, 2018.

[14] E. Elhamifar, G. Sapiro, S. S. Sastry, “Dissimilarity-Based Sparse Subset Selection,” *IEEE Trans. Pattern Anal. Machine Intell*, vol. 38, no. 11, pp. 2182-2196, 2016.