A Supervoxel Approach to Road Boundary Enhancement From 3-D LiDAR Point Clouds

Zhengchuan Sha, Yiping Chen, *Senior Member, IEEE*, Yangbin Lin, *Member, IEEE*, Cheng Wang, *Senior Member, IEEE*, José Marcato, Jr., *Member, IEEE*, and Jonathan Li, *Senior Member, IEEE*

***Abstract*— Rapid and accurate enhancement of road boundaries from terrestrial laser scanning (TLS) 3-D point clouds has been a challenging task in road infrastructure inventory. To address the challenge with a lack of ability to enhance object boundaries when the supervoxel number is less, this letter proposes a novel supervoxel segmentation algorithm framework for enhancing road boundaries from 3-D point clouds. First, we utilize radius k nearest-neighbor search method to obtain the neighborhood information after partitioning points on octrees with seed points. Second, the iterative weighted least square algorithm and spatial structure judgment are used to segment point clouds based on seed points. Finally, an update method to adjust the supervoxel centroids is applied with surrounding information in the first part. To verify the excellent performance, we tested the proposed method on two publicly large-scale point clouds benchmarks—IQmulus and TerraMobilita (IQTM) and Semantic 3-D. The experimental results demonstrate that our approach achieved approximately 48.98% and 68.41% boundary recall higher than two existing classical methods in the street scene, and our running time is feasible and effective.**

***Index Terms*— Boundary-enhanced, over-segmentation, point clouds, spatial structure, supervoxel.**

I. Introduction

SUPERVOXEL segmentation algorithms from point clouds group voxels into meaningful areas that maximize the boundary of targets. In the 3-D point cloud information processing, it is efficient to operate on representative points regions rather than scattered points for large-scale scenes. Supervoxel segmentation methods have been used in many fields, such as line extraction [1], [2], semantic labeling [3], point registration [4], and object detection [5].

Supervoxel segmentation methods are designed for 3-D point clouds scenes. There are many over-segmentation approaches to generate supervoxels. According to [6], great supervoxels have three natures.

1) The boundaries of over-segmentation adhere to ground-truth boundaries as many as possible.
2) The running time of the over-segmentation method is expected to be efficient.
3) The regions of supervoxels are regular even in sparse and dense regions.

Voxel cloud connectivity segmentation (VCCS) [7] is an advanced over-segmentation method to generate supervoxels based on neighbor relationships of voxels. VCCS uses spatial and geometric information to make supervoxels confirm to ground-truth boundaries. A boundary-enhanced supervoxel segmentation (BESS) to enhance boundaries is presented in [8]. The method consists of estimating the discontinuity frame by frame and clustering points on the graph to segment points into supervoxels. A facet segmentation method to segment points into facets by region growing is presented in [2]. The modification of the facet segmentation method to extract road boundaries by evenly partitioning 3-D space with a fixed size in [1]. The graph method for oversegmenting point clouds by a point embedding network and a graph-structured loss function is first used in [9]. Context-sensitive based on the graph distance to obtain superpixels and supervoxels is proposed in [10]. For most over-segmentation methods, the most essential drawback is that they cannot enhance road boundaries when the supervoxel number is less because they do not consider the global scene information and neighborhood information. Our letter focuses on the problem of road boundaries enhancing by considering the spatial structures of scenes. Our method can expand on different tasks and we verify it on road scenes. The main contributions of this letter are as follows:

1) We consider more surrounding points to express local features compactly and utilize octrees to voxelize points for the generation of uniform supervoxels.
2) We propose the spatial structure with angle and height judgments for the road boundaries enhancement more effective and avoid over-segmentation invalid.
3) Experiments on IQmulus and TerraMobilita (IQTM) [11] and Semantics 3-D [12] data sets show that proposed novel supervoxel method outperforms state-of-the-art methods on urban scenes.

The remainder of this letter is organized as follows. Section II describes the 3-D road boundary extraction method. Section III presents the experimental results and discussion. Section IV concludes the letter.

Manuscript received June 15, 2020; revised October 15, 2020; accepted November 5, 2020. Date of publication November 24, 2020; date of current version December 29, 2021. This work was supported in part by the National Natural Science Foundation of China under Grant 41871380. *(Corresponding authors: Yiping Chen; Jonathan Li.)*

Zhengchuan Sha, Yiping Chen, and Cheng Wang are with the Fujian Key Laboratory of Sensing and Computing for Smart City, Xiamen University, Xiamen 361005, China, and also with the School of Informatics, Xiamen University, Xiamen 361005, China (e-mail: shazhengchuan@163.com; chenyiping@xmu.edu.cn; cwang@xmu.edu.cn).

Yangbin Lin is with the Faculty of Computer Engineering College, Jimei University, Xiamen 361021, China (e-mail: yblin@jmu.edu.cn).

José Marcato, Jr. is with the Faculty of Engineering, Architecture and Urbanism and Geography, Federal University of Mato Grosso do Sul, Campo Grande 79070-900, Brazil (e-mail: jose.marcato@ufms.br).

Jonathan Li is with the Fujian Key Laboratory of Sensing and Computing for Smart City, Xiamen University, Xiamen 361005, China, also with the School of Informatics, Xiamen University, Xiamen 361005, China, and also with the Department of Geography and Environmental Management, University of Waterloo, Waterloo, ON N2L 3G1, Canada (e-mail: junli@xmu.edu.cn).

Digital Object Identifier 10.1109/LGRS.2020.3037484

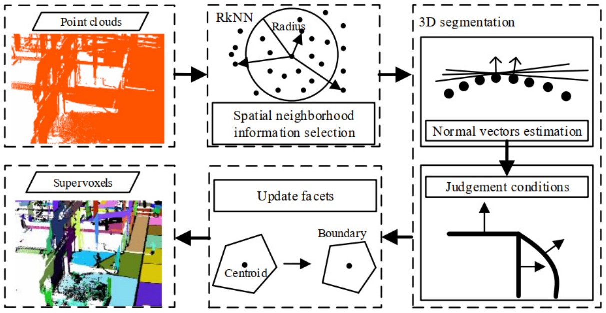


Fig. 1. Workflow of our proposed over-segmentation method.

II. Method

In this letter, we present a three-step framework to over-segment supervoxels from point clouds data. We first use the nearest neighbor method to consider spatial neighborhood information, and then we use the normal vectors and the judgment condition for 3-D segmentation. Finally, supervoxels are updated for boundaries regulation. The flowchart of the process is illustrated in Fig. 1.

A. Spatial Neighborhood Information Selection

Traditional VCCS method [7] is applied on points directly, which does not consider the neighborhood information. We use the radius k nearest-neighbor (RkNN) search method to find each point of neighborhood information given the fixed radius r and the number of k nearest neighbor, after initiating the supervoxel seeds based on the octree. If the point number is smaller than the given radius, the RkNN returns the true points in the search range. We set k equals 20. For the fixed radius r, we synthesize the point scene and neighborhood information. The r is set as 50 times of point resolution $D(p,q)$ of the whole scene, which is defined as

$$D(p,q) = \frac{1}{m}\sum_{q\in Q} \| \text{p-q} \| \tag{1}$$

where p and q are two points. Q is the k nearest neighborhood point set of point p. $\| \cdot \|$ represents the Euclidean distance between two points. m is 2 for two nearest points. The point resolution $\overline{r}$ reflects the adjacent information of the given point. A scene has only one $\overline{r}$ similar to [13], which can be considered as the inherent quality. Thus, we obtain the supervoxel number because the seed number is equal to the supervoxel number. Then we utilize neighborhood point sets of each point to calculate normal vectors for the next updating supervoxels by using the principal component analysis (PCA) method.

B. 3-D Segmentation

For the supervoxel generation, the purpose of 3-D segmentation is to label the analogous characters as the same label. Inspired by the study [2], we consider each supervoxel as a facet, which is considered as the point coordinate and its corresponding normal vector.

For each point, we apply the iterative weighted least square algorithm to obtain its corresponding normal vector. In this letter, we use the same weight to generate the tangent plane. For a given point p, the weight is calculated as in the k neighborhood range of p

$$w(p) = \begin{cases} (1-(\dfrac{d(p,T^n(p))}{\varepsilon})), & \text{if } d(p,T^n(p))<\varepsilon \\ 0, & \text{otherwise} \end{cases} \tag{2}$$

Algorithm 1 Calculate Normal Vector

```
Input: $\lambda_1$, $\lambda_2$, $\lambda_3$, $\vec{v_1}$, $\vec{v_2}$ and $\vec{v_3}$
Output: $\vec{v}$
 if $\lambda_1 = \lambda_2 = \lambda_3$ then
   $\vec{v} = (0,0,1)$
 else
   if $\vec{v_3} = (0,0,0)$ and $\vec{v} = \vec{v_1} \times \vec{v_2}$ then
     if $|\vec{v}| = 0$ then
       $\vec{v} = (0,0,1)$
     else
       $\vec{v} = \vec{v_1} \times \vec{v_2}$
     end if
   else
     $\vec{v} = \vec{v_3}$
   end if
 end if
```

where ε is a threshold of distance to judge if the point is on the plane. It only depends on the accuracy of laser point cloud equipment. Hence, the density of point cloud does not influence on this threshold. ε relates to the supervoxel resolution (R) [2]. n is the number of iterations. $d(\cdot)$ is the Euclidean distance operation. Then the tangent plane $T(p)$ is obtained by solving an iterative equation.

Different from the facet segmentation method [1], we calculate the normal vector based on each point with its corresponding k neighborhood points by the weighted PCA method. For the 3-D weighted PCA method, we obtain three eigenvalues λ_1, λ_2, and λ_3 $(\lambda_1 \geq \lambda_2 \geq \lambda_3)$ and three corresponding eigenvectors $\vec{v_1}$, $\vec{v_2}$, and $\vec{v_3}$. The normal vector is calculated as below.

After calculating all normal vectors, we initiate the number of facets equal to the number of seed points. Then we obtain the tangent plane of neighborhood points based on each seed point. When we assign the points to facets, distance $D(p,f)$ from ith point p to the facet f is computed as

$$D(p,f) = w_n(1-|\vec{n_p}\cdot\vec{n_f}|) + w_s\frac{\| p - f_p \|}{SR} \tag{3}$$

where w_n and w_s are the normal weight and spatial weight, respectively. $\vec{n_p}$ and $\vec{n_f}$ are the normal vector of p and the normal vector of facet, respectively. f_p is the point part of the facet. SR is the seed resolution value.

Each seed point constructs each initial facet. Then distances of neighbor points based on each seed point to their corresponding facets are calculated. This is considered for maintaining the generated supervoxel number the same as the seed number. Comparing the distance with the given threshold (in this letter, the threshold is set to infinity), the points in a smaller range of threshold are further labeled [14].

We applied judgment conditions to consider the spatial structure of the whole over-segmentation scene, which are summarized as

$$\text{OnPlane(NT)} \ \& \ z_{\text{NP}} < \text{TH} \| \ z_{\text{NP}} > \text{TH} \tag{4}$$

where OnPlane determines whether a plane is coplanar with standard XOY-plane or not. NT represents neighbor tangent plane. In OnPlane function, this judgment condition is that if the angle between the normal of the input plane and $z = (0,0,1)$ is higher than an angle threshold. In this letter,

TABLE I
TLS DATA SETS OF THE TESTED MODELS

Model	Numbers of points	$\overline{r}$(cm)	z_{min}(m)
IQTM	1258014	2.286	35.615
Station in Semantic 3D	3373270	0.912	-3.315
Street in Semantic 3D	3590904	1.414	-2.597

we set the angle threshold θ as 22.5°, similar to [2]. z_{NP} and TH mean the z value of the neighbor point and heights threshold, respectively. According to (4), we cluster the points into the meaningful regions by considering the relationships between normal angle differences and heights. After that, the segmentation processing is accomplished.

C. Update Facets

Similar to [7], we update facets (supervoxels). For the same label points in the generated regions, the centroid coordinates are computed for each given region as the new cluster center. Normal vectors calculated in Section II-A are used to obtain the new normal vectors as the facets after the normalization operation of them. Comparing the distances between points in the segmentation region with its corresponding centroid, nearest points indexes are acquired and the facets are updated according to new points and normal vectors.

III. RESULTS AND DISCUSSION

In this section, several experiments are conducted to assess the proposed method. We compare it with VCCS [7], VCCS-RNN, Lin's work [6], and the modified Lin's work.

A. Data Sets

We used two data sets with various characteristics. The first is the IQTM data set collected by a Stereopolis II [11] 3-D mobile laser system (MLS), which is a dense urban scene in Paris, France. It contains a fully manually annotated street of 200 m long with 12 million points. Second is Semantic 3-D data set [12] consisting of urban and rural point clouds with static terrestrial laser scanning (TLS), as shown in Figs. 5(a) and 7(a). The large-scale point cloud classification benchmark has 1 billion points with eight classes manually annotated. We choose rural (Station) and urban (Street) outdoor scenes. Due to the limitation of our memory, we down-sample the data sets to decrease both the computation time and memory. For both data sets, scenes lowest z coordinate ($z_{\min}$) is calculated as

$$z_{\min} = \min_{p_i \in P} z_{p_i} \tag{5}$$

where P is the whole point set in the scene. The information about the data sets is summarized in Table I.

B. Evolution Metrics

Boundary recall (BR) [6] and boundary precision (BP) [9] measure whether the supervoxels stick to the ground-truth boundary and minimize overlap. We use the same conception similar to [6]. A high BR and BP indicate that supervoxels properly follow objects by ground-truth labeled. Under-segmentation error (UE) measures the amount of supervoxels leakage across ground-truth boundaries [7], [15]. We adopt the same criterion to evaluate the performance. A low UE value means fewer supervoxels crossover object boundaries.

TABLE II
PARAMETERS SETTING IN OUR LETTER

ε in Eq.2	w_n in Eq.3	w_s in Eq.3	θ in Eq.4
$0.5R$	1	4	22.5^o

TABLE III
BR OF PROPOSED METHODS ON IQTM ($N = 296$)

(k, t)	Proposed-0.5	Proposed-3	(k, t)	Proposed-0.5	Proposed-3
(15, 50)	0.3418	0.4247	(20, 40)	0.3402	0.4643
(20, 50)	**0.3466**	0.4673	(20, 50)	**0.3466**	0.4673
(25, 50)	0.3405	**0.5177**	(20, 60)	0.3449	**0.4677**

TABLE IV
BR OF PROPOSED METHODS ON THE STREET SCENE ($N = 331$)

(k, t)	Proposed-0.5	Proposed-3	(k, t)	Proposed-0.5	Proposed-3
(15, 50)	0.3318	0.5913	(20, 40)	0.3461	**0.5937**
(20, 50)	**0.3463**	**0.5937**	(20, 50)	**0.3463**	**0.5937**
(25, 50)	0.3311	0.5909	(20, 60)	**0.3463**	**0.5937**

C. Parameters Discussion

Table II shows parameters are set in our letter. We set $w_n = 1$, and $w_s = 4$ in (3) (only considering the geometric information), which have been demonstrated as the most adequate weight values [16] for considering both spatial extents and normal of seed points distance measure in a feature space. For the k and radius (times (t) of $\overline{r}$) in RkNN, we empirically choose both randomly by considering the scene size and complexity of scene structure. Then we use several (k, t) with different supervoxel numbers (N) in different scenes as shown in Tables III and IV. (k, t) selection has little effect on BR results. Hence, our method is general and suitable for many kinds of tasks.

D. Scene Performance

In our experiment, we refer to the VCCS neighborhood information obtained from the Radius Nearest Neighbor (RNN) method as VCCS-RNN. For comparison, the radius of VCCS-RNN is set to $50\overline{r}$, which is the same as our method. The TH in (4) influences the enhanced results. To assess the proposed method, we set two TH values based on the scene lowest z coordinate TH $= z_{\min} + \delta$. We set δ equals to 0.5 and 3 m, which are donated as Proposed-0.5 and Proposed-3, respectively. Meanwhile, we add our proposed condition in (4) to improve Lin's method, which the TH value is set to 0.5 m higher than $z_{\min}$. Different from [6], we set the initial value λ_0 in Lin's work as the point resolution from (1) rather than the median value of the lowest dissimilarity distances because both of them are small enough. The final step is to use two different update methods. If the update method is the k medoids step, which is the same presented in [6], we donate it as m-Lin1-0.5. If the update method is the same as that found in Section II-C, we denote it as m-Lin2-0.5. Next, we compare the above seven methods with general evolution metrics in the same range of the supervoxel number (N) for comparison.

For urban scenes (IQTM and the street), road boundaries are more regular and integrated than these in rural scenes. As shown in Figs. 2(a) and 6(a), the method of Proposed-3 achieved the highest BR except in Fig. 2(a) with N smaller than 150. For three figures [Figs. 2(a), 4(a), and 6(a)], the BR of VCCS and Lin's method are mostly the lowest values, thus these two methods both lose their effectiveness.

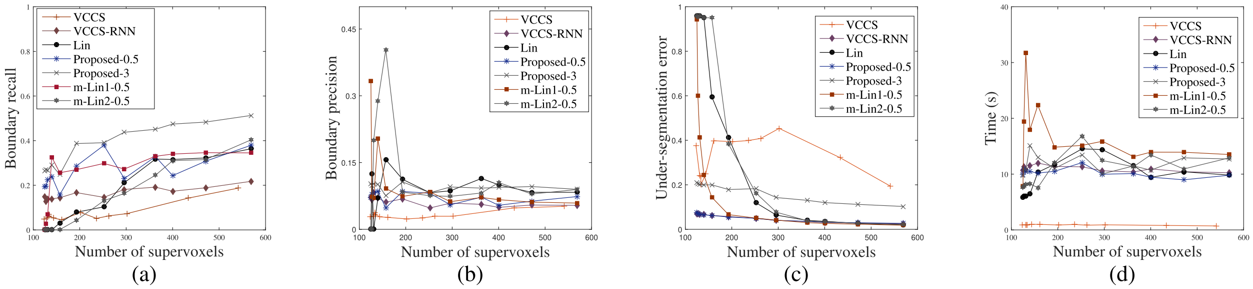


Fig. 2. (a) Evaluation of seven supervoxel methods on BR, (b) BP, (c) UE, and (d) running time for IQTM data set.

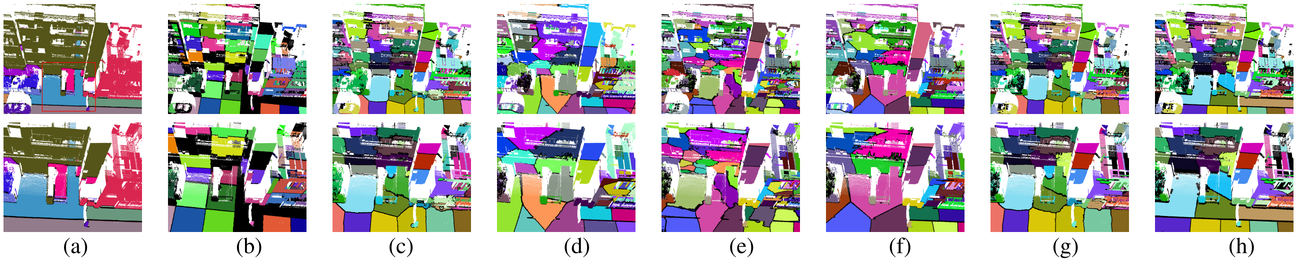


Fig. 3. Comparison of visualized results using IQTM data set (urban scene). (a) Ground-truth. (b) VCCS. (c) VCCS-RNN. (d) Lin. (e) m-Lin1-0.5. (f) m-Lin2-0.5. (g) Proposed-0.5. (h) Proposed-3.

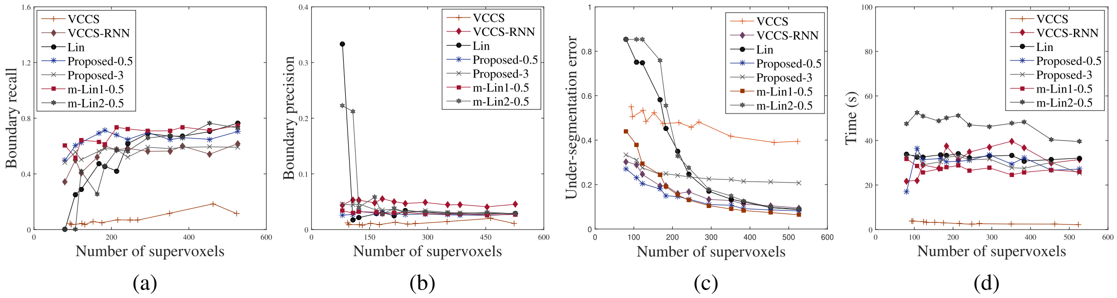


Fig. 4. (a) Evaluation of seven supervoxel methods on BR, (b) BP, (c) UE, and (d) running time for station scene of Semantic 3-D data set.

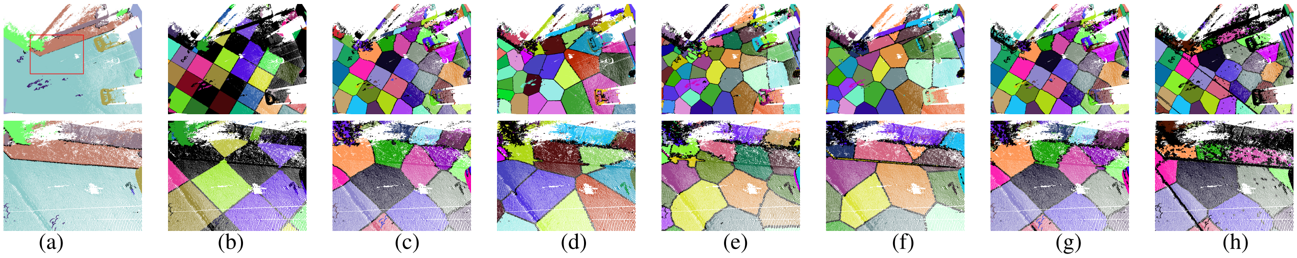
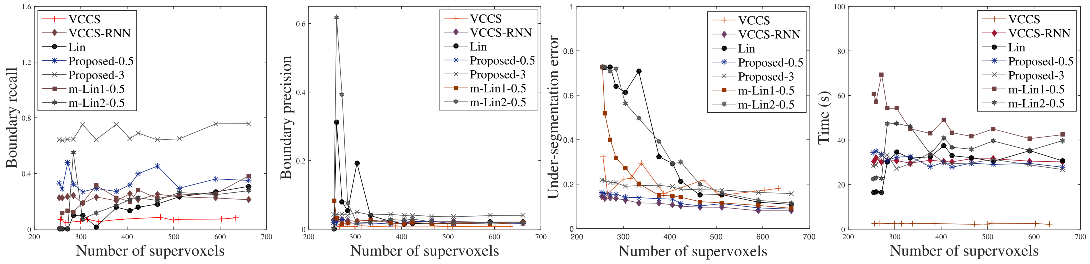


Fig. 5. Comparison of visualized results using the station scene (rural) based on Semantic 3-D data set. (a) Ground-truth. (b) VCCS. (c) VCCS-RNN. (d) Lin. (e) m-Lin1-0.5. (f) m-Lin2-0.5. (g) Proposed-0.5. (h) Proposed-3.

Meanwhile, the update method in Section II-C cannot be effective to enhance the boundaries in the same condition [Fig. 2(a) with the N lower than 200; Fig. 4(a) with the N lower than 100; Fig. 6(a) with the N lower than 300 with an exception when N is 260.]. For the evaluation of BP, it is easy to find BP values of VCCS method are lowest in three figures [Figs. 2(b), 4(b), and 6(b)]. Moreover, the UE values of VCCS are the highest with N greater than 193 in Fig. 2(c) and N greater than 214 in Fig. 4. For both Lin's method and m-Lin2-0.5, they are not effective with the N lower than 140 in Fig. 2(c) and the N lower than 107 in Fig. 2(c). For VCCS-RNN, the effect of it is always moderate except in Fig. 4(b) when N is greater than 183. The reason is VCCS-RNN only considers the neighborhood information and neighborhood metrics by using (3) without whole spatial structures when assigning the points to different regions. Therefore, our proposed method can achieve great BR with low supervoxel numbers, especially in the urban scene with whole and regular road boundaries [for Figs. 2(a) and 6(a) on Proposed-3 method].

The visual results are shown in Figs. 3, 5, and 7. In order to display our over-segmentation result better, we set segmentation boundaries as black color and different regions are used different colors generated randomly. As shown in Figs. 3(h) and 5(h), our over-segmentation results can follow ground-truth boundaries completely. It is the obvious advantage of the proposed method. As shown in Fig. 7, our proposed methods have an adaptive resolution, which means our methods can generate uniform supervoxels instead of small resolutions in dense regions vice versa. The reason is our



(a) (b) (c) (d)

Fig. 6. (a) Evaluation of seven supervoxel methods on BR, (b) BP, (c) UE, (d) and running time for street scene of Semantic 3-D data set.

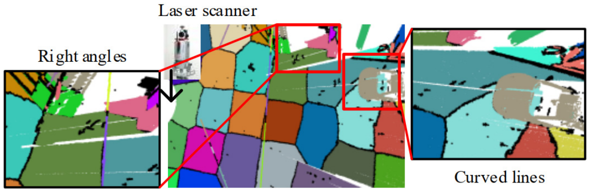


Fig. 7. Visualization results of the street with adaptive resolutions, right angles and curved lines in nonuniform density.

proposed methods initiate supervoxels on octree and partition evenly to obtain the seed points in Section II-A. Meanwhile, our method can segment curved lines and more complicated structures (e.g., right angles) effectively as shown in Fig. 7.

E. Time Performance

The proposed methods were implemented in C++ and the experiments were conducted on a PC with Ubuntu 18.04, Intel Core i5-3470 3.2 GHz CPU, and 16.0 GB memory. The running time performances are shown in Figs. 2(d), 4(d), and 6(d). The running time of VCCS is fastest because it performs on voxels directly, hence, it decreases the scale of the problem greatly. For the VCCS-RNN method, it performs on points instead of voxels so its cost time is intermediate compared with representative methods instead of VCCS as shown in Figs. 2(d), 4(d), and 6(d). As shown in Figs. 4(d) and 6(d), m-Lin2-0.5 method in station scene and m-Lin1-0.5 in street scene cost the longest time in that using our spatial structure judgment in (4) needs to calculate the normal vector of each point. Meanwhile, it is also the main reason for the proposed cost time. Due to voxelizing points on octrees, the cost time hardly changes when N does not change dramatically because voxels number do not change sharply. Hence, sacrificing running time for considering more neighborhood information can yield more effective results.

IV. Conclusion

In this letter, we present a novel over-segmentation algorithm to enhance road boundaries. We first use the 3-D partition on octrees to preserve regular supervoxels, and the neighbor nearest method (RkNN) to consider more neighborhood information. Second, the adopted iterative weighted least square algorithm and spatial structure judgment considering more cluster information are added for 3-D oversegmentation. Finally, we update facets to preserve the boundaries better.

Our proposed methods have been successfully tested on two large-scale point clouds benchmarks with several representative methods (VCCS, VCCS-RNN, Lin's method, m-Lin1-0.5, and m-Lin2-0.5). The results prove that our proposed methods have the ability to greatly enhance the road boundaries in urban scenes (with about 48.98% and 68.41% BR value higher than Lin's method and traditional VCCS) and keep regular even when the supervoxel number N is low. Meanwhile, the cost time of our proposed method is only influenced by the number of points and hardly change with N when N is small. Future works will be oriented to design various boundaries enhancement supervoxel methods and optimize the cost time.

References

[1] D. Zai *et al.*, "3-D road boundary extraction from mobile laser scanning data via supervoxels and graph cuts," *IEEE Trans. Intell. Transp. Syst.*, vol. 19, no. 3, pp. 802–813, Mar. 2018.

[2] Y. Lin, W. Cheng, 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, Sep. 2017.

[3] H. Luo *et al.*, "Patch-based semantic labeling of road scene using colorized mobile LiDAR point clouds," *IEEE Trans. Intell. Transp. Syst.*, vol. 17, no. 5, pp. 1286–1297, May 2016.

[4] W. Li, C. Wang, C. Lin, G. Xiao, C. Wen, and J. Li, "Inlier extraction for point cloud registration via supervoxel guidance and game theory optimization," *ISPRS J. Photogramm. Remote Sens.*, vol. 163, pp. 284–299, May 2020.

[5] H. Guan, Y. Yu, J. Li, and P. Liu, "Pole-like road object detection in mobile LiDAR data via supervoxel and bag-of-contextual-visual-words representation," *IEEE Geosci. Remote Sens. Lett.*, vol. 13, no. 4, pp. 520–524, Apr. 2016.

[6] 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, Sep. 2018.

[7] J. Papon, A. Abramov, M. Schoeler, and F. Worgotter, "Voxel cloud connectivity segmentation–supervoxels for point clouds," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2013, pp. 2027–2034.

[8] S. Song, S. Jo, and H. Lee, "Boundary-enhanced supervoxel segmentation for sparse outdoor LiDAR data," *Electron. Lett.*, vol. 50, no. 25, pp. 1917–1919, Dec. 2014.

[9] L. Landrieu and M. Boussaha, "Point cloud oversegmentation with graph-structured deep metric learning," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2019, pp. 7432–7441.

[10] Z. Ye, R. Yi, M. Yu, Y.-J. Liu, and Y. He, "Fast computation of content-sensitive superpixels and supervoxels using Q-Distances," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, Oct. 2019, pp. 3770–3779.

[11] B. Vallet, M. Brédif, A. Serna, B. Marcotegui, and N. Paparoditis, "TerraMobilita/iQmulus urban point cloud analysis benchmark," *Comput. Graph.*, vol. 49, pp. 126–133, Jun. 2015.

[12] T. Hackel, N. Savinov, L. Ladicky, J. D. Wegner, K. Schindler, and M. Pollefeys, "Semantic3D.net: A new large-scale point cloud classification benchmark," *ISPRS Ann. Photogramm., Remote Sens. Spatial Inf. Sci.*, vols. IV–1/W1, pp. 91–98, May 2017.

[13] D. Zai *et al.*, "Pairwise registration of TLS point clouds using covariance descriptors and a non-cooperative game," *ISPRS J. Photogramm. Remote Sens.*, vol. 134, pp. 15–29, Dec. 2017.

[14] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. Süsstrunk, "SLIC superpixels compared to state-of-the-art superpixel methods," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 11, pp. 2274–2282, Nov. 2012.

[15] A. Levinshtein, A. Stere, K. N. Kutulakos, D. J. Fleet, S. J. Dickinson, and K. Siddiqi, "TurboPixels: Fast superpixels using geometric flows," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 31, no. 12, pp. 2290–2297, Dec. 2009.

[16] S. C. Stein, M. Schoeler, J. Papon, and F. Worgotter, "Object partitioning using local convexity," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2014, pp. 304–311.