Normal Grouping Density Separation (NGDS): A Novel Object-Driven Indoor Point Cloud Partition Method

Jakub Walczak[1](✉), Grzegorz Andrzejczak[2], Rafał Scherer[3], and Adam Wojciechowski[1]

[1] Institute of Information Technology, Lodz University of Technology, Łódź, Poland
{jakub.walczak,adam.wojciechowski}@p.lodz.pl

[2] Institute of Mathematics, Lodz University of Technology, Łódź, Poland
grzegorz.andrzejczak@p.lodzd.pl

[3] Częstochowa University of Technology, Częstochowa, Poland
rafal.scherer@pcz.pl

Abstract. Precise segmentation/partition is an essential part of many point cloud processing strategies. In the state-of-the-art methods, either the number of clusters or expected supervoxel resolution needs to be carefully selected before segmentation. This makes these processes semi-supervised. The proposed Normal Grouping- Density Separation (NGDS) strategy, relying on both grouping normal vectors into cardinal directions and density-based separation, produces clusters of better (according to use quality measures) quality than current state-of-the-art methods for widely applied object-annotated indoor benchmark dataset. The method reaches, on average, lower under-segmentation error than VCCS (by 45.9pp), Lin et al. (by 14.8pp), and SSP (by 26.2pp). Another metric - achievable segmentation accuracy - yields 92.1% across the tested dataset what is higher value than VCCS (by 14pp), Lin et al. (by 3.8pp), and SSP (by 10.3pp). The experiment carried out indicates superiority of the proposed method as a partition/segmentation algorithm - a process being usually a preprocessing stage of many object detection workflows.

Keywords: Space partition · Superpoints · Point cloud segmentation · NGDS

1 Introduction

Point clouds have recently become a powerful representation of the environment due to the inherent spatial cues that they possess. Depth information, provided either by a depth camera, multiple-view interpolation, or by a laser scanner, is clue information exploited for retrieval relationships among objects in a scene [7]. This is a reason for such wide and various application of point clouds [9,16,33,35]. Depth information can be also used in other fields, such as computer

© Springer Nature Switzerland AG 2020
V. V. Krzhizhanovskaya et al. (Eds.): ICCS 2020, LNCS 12142, pp. 100–114, 2020.
https://doi.org/10.1007/978-3-030-50433-5_8

graphics, from rasterization and ray tracing algorithms [28] known for decades to modern screen-space methods of accelerating calculations without loss of image quality based on the prior creation of depth information in individual image pixels [30].

One of the key aspects of a point cloud processing is object detection [12,17,26,37] and, in general, semantic analysis [8,29,36,38]. In many aspects of a point cloud processing, the strategy relying on either ordering points or excessive partition followed by the aggregation (top-bottom-top) is frequently applied [13]. The strategy, relies on excessive space/surfaces segmentation (top-bottom, in order to minimize under-segmentation) and aggregation (bottom-top, so that over-segmentation could be decreased, keeping under segmentation error low) [5,6,11] followed by features analysis [10,26] or deep learning pipeline [2,12,17,38]. Such an approach requires a process of over-segmentation (top-bottom) to be efficient, granular on objects' edges and corners, and, what is the most important, devoid of overlapping areas between semantically different objects. In this article, the novel, intuitive and high-quality method for over-segmentation, called Normal Grouping- Density Separation (NGDS), is presented. The novelty of the presented method relies on clever application of efficient grouping algorithm in order to detect primary plane directions in a point cloud followed by histogram- and density-based separation within points belonging to a primary direction. Unlike current state-of-the-art methods, the presented strategy does not require predefined, manually selected proper number of segments to produce expected and sufficiently granular over-segmentation result suitable in the context of object detection. Most of parameters are calculated based on provided point cloud characteristic and just a few are required to be set by a user.

2 Related Works

A popular strategy of a point cloud processing, in any form, either semantic segmentation, or object detection, is to reduce the size of a problem by means of over-segmentation algorithms [32] which group points - consistent according to some criterion - into clusters whose number is usually tremendously lower than the cardinality of a point cloud. This allows applying just a few groups instead of hundreds of thousands of points. Such an strategy lies behind the idea of compression - to use approximate larger regions by means of a small representative entity [18]. However, the problem of current over-segmentation algorithms is the fact that they focus on class-driven approach rather than object-driven one. As a result, such methods produce clusters of preserved boundaries between objects of different classes whereas objects within the same class are not partitioned properly. This disables benchmark methods to be successfully applied as over-segmentation strategy for object detection purposes.

One of first benchmark partition strategy - VCCS - was introduced in [22]. The method presented therein focuses on quasi-regular clusters called supervoxels. The researchers used 39 point features to make sure that points inside a single cluster/supervoxel are consistent (intra-cluster consistency). 33 of those 39 features are traits of Fast Point Feature Histogram, the others are geometrical coordinates and color in CIELab space. VCCS relies on efficient k-means clustering to group points to clusters of predefined seeds. Though pretty efficient, the method suffers some issues. The first problem is initialization of the algorithm itself. The seed points need to be carefully chosen and the number of those seeds (hence clusters) has to be selected before. The desired method solving the issue should somehow recognize the optimal number of resulting clusters in unsupervised manner, because improper choice can make a method pointless. Too low number of clusters results in insufficient segmentation - output clusters overlap many ground-truth clusters. Too high number, in turn, wreaks longer time processing and loss of the context. Furthermore, as indicated in [14], VCCS may segment borders inaccurately.

To overcome the issue related to borders and seed points selection, Lin et al. [14] proposed extension of VCCS. In these studies, suggested cost function, which optimization provides representative points, consists of two counteractive components: the first one ensuring a representative point approximates well a collection of points; and the second one - constraining the number of expected representative points to be as close to the predefined value as possible.

Having selected all representative points, optimization of cost function is continued by assigning non-representative points to those representative points for which dissimilarity distance is the lowest.

Though improved with respect to its predecessor, the method of Lin et al. still requires the number of resulting clusters to be selected in advance what cannot be reliably done in unsupervised partition process. Moreover, the resulting groups/clusters are similar to quasi-regular grid of VCCS what produces high over-segmentation error in regions where it is redundant. The clue is to design a method which maintains the proper balanced between both under- and over-segmentation keeping both of them low.

To reduce the problem of excessive partition keeping under-segmentation error low, Landrieu and Boussaha [11] proposed the SSP method mixing deep learning approach with analytical strategies. Applying PointNet-like neural network enables the authors to extract high-level object-oriented features, called embeddings. Such embeddings are calculated for each point in a data set based on its vicinity. Based on embeddings and spatial connectivity, Generalized Minimal Partition Problem is solved with the method ℓ_0 presented in [12]. The method yields good results in class-driven approach, however, taking into account single objects, the method is not reliable. It is caused by embeddings, which themselves cannot differentiate points belonging to different objects of the same class. Because of this, the method leaks intra-class separation what is crucial element of partition oriented to objects. The last drawback of SSP method is the fact that it may requires color information to produce reliable result.

To sum up, the methods of VCCS and Lin et al. inherently take into account spatial connectivity of points what is beneficial in terms of object-oriented separation of clusters, however, they perform many redundant subdivisions, which increase over-segmentation error. On the other hand, the SSP method avoid excessive partition at the cost of class-oriented partition rather than the object-one. In addition, all those methods require an expected number of clusters to be defined prior to computation what makes these methods difficult to be applied successfully in unsupervised segmentation. Hence, there is a need to develop a method which automatically splits points into geometrically coherent sets. And it turns out that relying only on normal vector may lead to sufficiently detailed space partition which quantizes well points of objects.

3 Methodology

At first, let the point cloud $\mathbf{P}$ be of the cardinality $||\mathbf{P}|| = N$. Let the two clusterings $\mathcal{S}$ and $\mathcal{G}$ be also defined, where $\mathcal{S}$ of cardinality $|\mathcal{S}| = m$ consists of a set of m clusters: $\mathcal{S} = \{s_1, s_2, s_3, ..., s_m\}$ being the output of a method, and $\mathcal{G} = \{g_1, g_2, g_3, ..., g_n\}$, of cardinality $|\mathcal{G}| = n$, represents a set of n ground-truth (real) clusters (single objects in a scene: 1st chair, 2nd chair, 1st table, etc.). It is crucial to note that ground-truth clusters are single objects in a scene, while for algorithms' output clusters encompass usually subsets of objects and the goal of each partition method is to produce output clusters as alike to ground-truth ones as possible. Following the literature approaches for partition and segmentation validation [5,14,22], below quality measure were engaged.

3.1 Quality Measures

Under-Segmentation Error (UE). Also referred to as under-segmentation rate, indicates insufficient partition. In short, an output cluster overlaps more than one ground-truth cluster. Its value varies from 0 - if none of output clusters overlap more than one ground-truth cluster, and 1 if $|\mathcal{S}| = 1$ and $|\mathcal{G}| > 1$. In general, if $m < n$ then UE is in-between 0% and 100%. UE is expressed by (1). For visualization, see Fig. 1.

$$UE_{\mathcal{G}}(\mathcal{S}) = \frac{1}{m}[(\sum_{g_j \in \mathcal{G}} \sum_{s_i \in \mathcal{S}} < \frac{|s_i \cap g_j|}{|g_j|} > \epsilon <) - m] \tag{1}$$

where $< \cdot <$ is an Iverson bracket which takes 1 if inner condition is $True$ and 0 otherwise, ϵ is a very small value (here it is 0.1%).

Weighted Under- and Over-Segmentation Error (wUE, wOE). The formula (1) relies on binary values (sum 1 if overlapping exceeds the threshold value ϵ. But, according to [27], the measure to express UE may be weighted with the intersection part- wUE (2). In similar manner wOE may be expressed- wOE (3). Their best values are 0% which means all points within a single cluster are associated with the only one ground-truth cluster, i.e. object and vice versa.

$$wUE_{\mathcal{G}}(\mathcal{S}) = \sum_{s_i \in \mathcal{S}} (|s_i| - \max_j |s_i \cap g_j|)/N \tag{2}$$

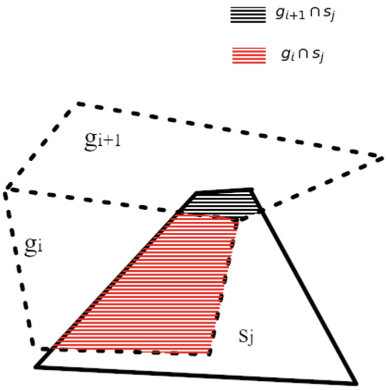

Fig. 1. Visualization of sample overlapping cases. Dashed lines represent ground- truth clusters whereas solid ones- the output cluster of a method. Black shaded region represents intersection lower than ϵ, so it is counted as 0 while red shaded region exceeds ϵ and is counted as 1. (Color figure online)

where N is the cardinality of a point cloud ($|\mathbf{P}| = \sum |g_j| = N$)

$$wOE_{\mathcal{G}}(\mathcal{S}) = \sum_{g_j \in \mathcal{G}} (|g_j| - \max_i |s_i \cap g_j|)/N \tag{3}$$

Harmonic Segmentation Error (HSE). Similarly to $F1$ score, which connected both precision and recall of classification in the form of harmonic mean, HSE may be defined as a single measure of error taking into account both weighted over- and under-segmentation errors (4).

$$HSE_{\mathcal{G}}(\mathcal{S}) = 2 \cdot \frac{wUE_{\mathcal{G}}(\mathcal{S}) \cdot wOE_{\mathcal{G}}(\mathcal{S})}{wUE_{\mathcal{G}}(\mathcal{S}) + wOE_{\mathcal{G}}(\mathcal{S})} \tag{4}$$

Achievable Segmentation Accuracy (ASA). ASA is one of quality metrics used by [15] to evaluate maximum possible accuracy in object detection task while applying proposed clusters as units. The best possible value it takes is 100%. Formally, this measure may be expressed by (5).

$$ASA_{\mathcal{G}}(\mathcal{S}) = \frac{\sum_{s_i \in \mathcal{S}} \max_j |s_i \cap g_j|}{\sum_{g_j \in \mathcal{G}} |g_j|} \tag{5}$$

Some literature studies made use of, so called, boundary recall and boundary precision [5,14]. However, indicating "boundary" points as done in [5,34] is, at least, questionable and ambiguous. Moreover, UE with low overlapping threshold ϵ and wUE directly point out if objects' borders are crossed or not. That is why boundary-based measures were skipped in the considerations presented in these studies.

3.2 Database

To make the method comparable, benchmark data sets need to be used. However, among all indoor databases widely used in studies, like: NYU RGBD v2 [21], ScanNet [4], or S3DIS [1] only the latter one distinguishes single objects. The others contained points labeled only by class what makes them useless in terms of verification of the partition method dedicated to object detection task. Therefore, S3DIS database was only selected as it is the only one available indoor database annotated by object.

S3DIS is one of the basic indoor benchmark dataset for semantic segmentation and object detection task. It was used, among others, in [5,11,12]. It contains 273 indoor-scene point sets of quite uniform densities with moderate scanning shadows present.

4 Proposed Method

In this paper, the novel partition method relying on spatial point connectivity and geometrical features - NGDS - is presented. The method consists of 6 stages:

a) normal vector estimation;
b) normal vector alignment;
c) primary directions detection;
d) level detection;
e) 2D density separation;
f) Lost points appending;

Normal Vector Estimation. Normal vector, yet trivial, is a kind of handcrafted point feature, contrary to traits learned with deep learning models [24]. Computation of a normal vector $\mathbf{n}_i$ of a point p_i is carried out by means of fitting a plane to the vicinity $\mathcal{N}$ of that point. In literature it is usually done with by eigendecomposition of the covariance matrix of neighbours' coordinates. This phase, taking into account efficient neighbours retrieving with kd-tree [3], is of time complexity: $O(n \cdot \log n) + O(|\mathcal{N}| \cdot n) + O(n) \equiv_{|\mathcal{N}|=const} O(n \cdot \log n)$

Normal Vector Alignment. Any method of calculating normal vector, which does not take into account constant reference point, cannot assure coherent orientation of normal vectors. This is the result of plane ambiguity (6).

$$\mathbf{n} \cdot p = -d \sim -\mathbf{n} \cdot p = d \tag{6}$$

Making normal vectors coherent across the same orientation simplifies to satisfying the condition (7) [23].

$$\mathbf{n}_i \cdot (v_p - p_i) > 0 \tag{7}$$

where v_p is a viewpoint, which in our studies is assumed to be in the mass center of a point cloud: $v_p = \bar{p} = \frac{1}{n} \sum_{i=1}^{n} p_i$.

In this way, parallel normal vectors get coherent orientation in symmetry with respect to v_p. This allows also distinguishing parallel planes on opposite sides of v_p (Fig. 2).

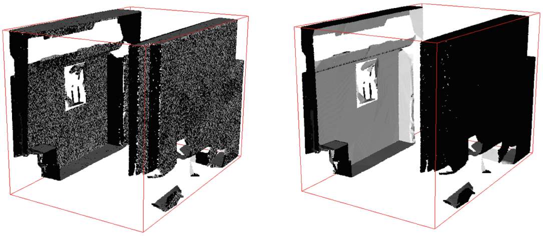

Fig. 2. Two parallel walls with associated normal vectors (brightness presents angle between a normal vector and camera optical axis). On the left- not aligned and on the right- aligned with respect to viewpoint

The stage of normal vector alignment is of linear time complexity- $O(n)$.

Primary Directions Detection (Normal Grouping). Having aligned point normal vectors, some primary directions may be identified taking a look at the distribution of normal vectors' orientations (Fig. 3a).

Such distribution seems to be dedicated for efficient k-means clustering algorithm. Though, relatively quick, it will be further accelerated if mini-batch based approach is used [25]. Though burdened with heuristics, mini-batch k-means clustering usually supplies results of sufficient accuracy with respect to optimum result in the sense of Maximum Likelihood estimation.

Widely known problem with k-means clustering concerns the proper selection of a number of resulting clusters k. In these studies, it is fixed. If normal vectors are oriented as in Fig. 3a - on unit sphere, then the allowed angle between normal vectors forms a spherical cap (Fig. 3b). Assuming expected angular tolerance to be $\Delta\theta$, the spherical cap surface associated to $\Delta\theta$ is expressed by (8).

$$P_c = 2 \cdot \pi \cdot (1 - \cos(2 \cdot \Delta\theta)) \tag{8}$$

Knowing that the unit sphere has the surface of $P_s = 4 \cdot \pi$, the expected number of clusters k may be calculated as $k = \lfloor P_s / P_c \rfloor$.

Time complexity of the mini-batch version of k-means algorithm is $O(n)$ assuming fixed number of maximum iterations and that the kd-tree for neighbours retrieving is already calculated.

Level Detection. The result of primary detection clustering provides groups of points representing a single primary direction (Fig. 4a).

Considering a single primary direction, let us define a group of points $\mathcal{D}_i \subseteq \mathbf{P}$ belonging to $i-$ th primary direction. As a one group, the points $\mathcal{D}_i$ should be

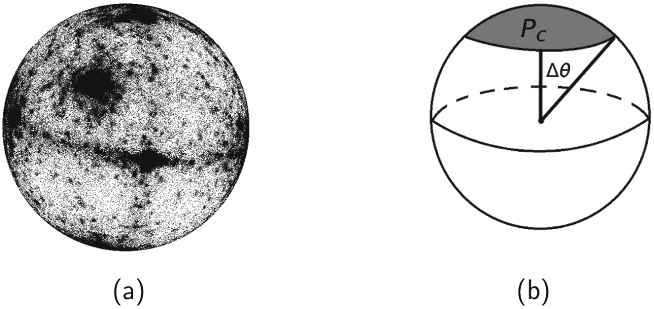

Fig. 3. (a) Unit sphere representing distribution of point normal vectors (b) spherical cap (in gray) of area P_c associated with angular tolerance $\Delta\theta$ on the unit sphere

described by a single normal vector, which may be calculated as a unit-length average normal vector $\hat{\mathbf{n}}_i$ of the group. Since they are said to share a common normal vector $\hat{\mathbf{n}}_i$, particular planar fragments may be easily distinguished by means of planes' constant factor (d, Eq. 6). For each point (and associated with it, plane fitted to its former vicinity $\mathcal{N}$), its constant factor is calculated according to the formula: $d = -p_j \cdot \hat{\mathbf{n}}_i$ (assuming $p_j \in \mathcal{D}_i$). Based on those values, a histogram may be constructed. Its peaks would clearly indicate planar fragments hung at different levels (Fig. 5). As a result, sets of co-planar points are retrieved: $\mathcal{L}_1, \mathcal{L}_1, \mathcal{L}_3, ...$ (Fig. 4b).

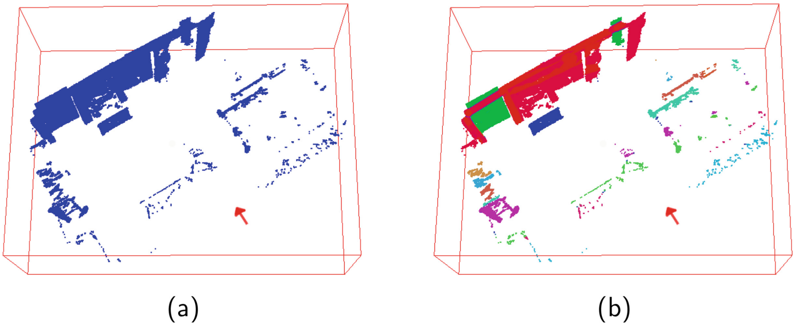

Fig. 4. (a) Groups of points around the single primary direction pointed by red arrow (b) points belonging to the same primary directions: $\mathcal{L}_1, \mathcal{L}_2, ...$, split according to associated constant factors (Color figure online)

In order to make extraction accurate, desired number of histogram bins should be carefully chosen. Intuitively, it is involved with a point cloud acquisition device tolerance, usually denoted as σ. It is involved with precision of points coordinates while noise is modelled with Gaussian distribution. The problem is,

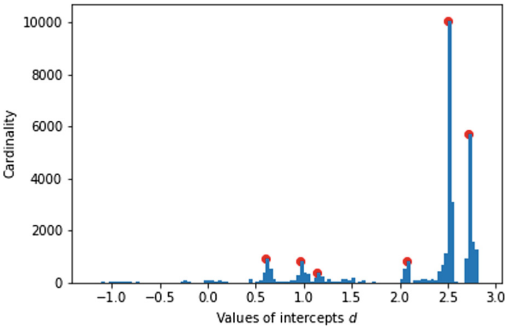

Fig. 5. Histogram of constant factors for point-planes of common primary direction

that sigma is usually not provided for indoor scans like S3DIS and it has to be somehow estimated. To do so, several random samples of all points are drawn and the distance (Euclidean or [31]) to their closest neighbours is calculated. The maximum value of this distance is saved. Across several samples, the mean maximum distance is calculated as the estimation of σ. The time complexity of this stage is linear- $O(n)$, since sigma approximation may be thought of to be of constant time complexity $O(1)$ and histogram construction takes $O(n)$ time for each primary direction.

2D Density-Based Clustering (Density Separation). Levels $\mathcal{L}_1, \mathcal{L}_3, ...$ detected in the previous stage contain points said to be co-planar. In real cases, especially when we bear in mind object detection task, co-planar points may quite often lead to insufficient partition is some areas. An example of such case is presented in Fig. 6, where two tops of two separate tables form a common group $\mathcal{L}_i$.

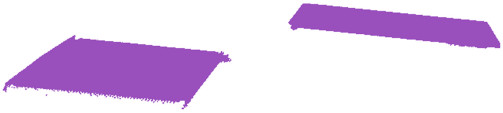

Fig. 6. Two tops of two separate tables detected as a single group $\mathcal{L}_i$

To avoid such issues, density-based HDBSCAN clustering [20] is carried out for points contained in each $\mathcal{L}_i$. Since points within $\mathcal{L}_i$ are co-planar, density separation may be reduced from 3D to 2D problem by means of Principal Component Analysis (9) so that computation time and occupied memory are both reduced. In this way, from a set $\mathcal{L}_i$ several subsets $\mathcal{L}_i = \{\mathcal{L}'_{i,1}, \mathcal{L}'_{i,2}, \mathcal{L}'_{i,3}, ...\}$ are extracted and the set of noise points $\mathcal{L}_{i,noise}$. (HDBSCAN is able to detect points deemed as noise in the context of local points density- those noise points are rejected and considered in the next stage.)

$$\mathcal{L}_i^{2D} = \mathcal{L}_i \times \mathbf{E} \tag{9}$$

where $\mathbf{E}$ is a column-wise matrix of eigenvectors

Though, HDBSCAN may yield sub-quadratic time complexity, it cannot achieve $O(n{\cdot}\log n)$, however, [19] suggests that it may approach log-linear asymptotic complexity for a number of data sets.

Lost Points Appending. During all previous stages, some points may be rejected due to rank-deficiency or density changes in HDBSCAN (noise points). To make the output point cloud conformed to the original one, these lost points need to be appended to the best matched clusters, to form $\mathcal{S}_1, \mathcal{S}_2, ...$ Assignment is done based on similarity function, defined like in [14] (10) taking as R the estimated values of σ $(R = \sigma)$.

$$D(p_i, p_j) = 1 - abs(\mathbf{n}_i \cdot \mathbf{n}_j) + 0.4 \cdot \frac{|p_i - p_q|}{R} \tag{10}$$

where R is an assumed resolution of partition, $\mathbf{n}_i$ and $\mathbf{n}_j$ are normal vectors associated to the $i-$ th and $j-$th point, $|\cdot|$ states for a norm of a vector.

The proposed method NGDS was validated on the benchmark database for indoor scenes, namely S3DIS [1].

5 Experiments

To compare the proposed NGDS method with the state-of-the-art solutions, following experiments were conducted. For benchmark dataset - S3DIS [1] partition according to VCCS [22], Lin et al. [14], and [11] methods were conducted. Ground-truth clusters $\mathcal{G} = \{g_1, g_2, g_3, ..., g_n\}$ are single objects in a scene, provided by the dataset. Experiments for VCCS and for the method of Lin et al. were carried out for all point clouds contained in the aforementioned database. To compare the state-of-the-art solutions with the proposed one, publicly available implementation of those algorithms were employed:

- Point Cloud Library implementation of VCCS was used [23] in the experiments.
- The code for the method of Lin et al. is provided by the authors on the publicly available repository [1].
- The code for SSP method is accessible in the public repository managed by the author [2].

[1] https://github.com/yblin/Supervoxel-for-3D-point-clouds.

[2] https://github.com/loicland/superpoint_graph.

6 Results

The results for the proposed NGDS method juxtaposed with the current state-of-the-art methods [11,14,22] for benchmark object database S3DIS are presented below (Table 1). For quality measures of the proposed method associated standard deviations are given.

Table 1. Comparison of quality measures for the benchmark partition methods and the proposed method NGDS for S3DIS

Method	UE [%]	wUE [%]	wOE [%]	HSE [%]	ASA [%]
VCCS [22]	46.4	21.9	89.2	35.2	78.1
Lin et al. [14]	15.3	11.8	92.7	19.6	88.3
SSP [11]	26.7	18.2	35.9	24.2	81.8
Proposed method (NGDS)	0.5 ± 0.4	4.7 ± 3.2	48.7 ± 11.5	8.6	92.1 ± 3.9

There are exemplary results of three state-of-the-art methods and NGDS presented in the Fig. 7

The results presented in Table 1 show superiority of the NGDS method over state-of-the-art solutions in terms of all presented quality measures. They prove high quality of NGDS as a partition method dedicated to object detection task. Under-segmentation error for the proposed method is lower by 99%, 97%, and 98% for VCCS, Lin et al., and SSP respectively. This confirms that clusters created by NGDS method do not tend to cross **object** boundaries, even within the same class. Lower weighted under-segmentation error, in turn, proves that only 4.7% of points are mismatched. On the other hand, weighted over-segmentation error shows that less redundant subdivisions were done with respect to the method of Lin et al. and VCCS. In comparison to SSP, wOE is slightly higher. To infer an overall trade-off of under- and over-segmentation, HSE indicator was introduced. Undoubtedly, it attains the best (lowest) value for the proposed method - respectively lower by 26.6pp, 11pp, and 15.6pp than VCCS, Lin et al., and SSP. Also the average Achievable Segmentation Accuracy clearly points out that having applied NGDS method for the task of indoor object detection, the highest accuracy may be achieved among all four methods.

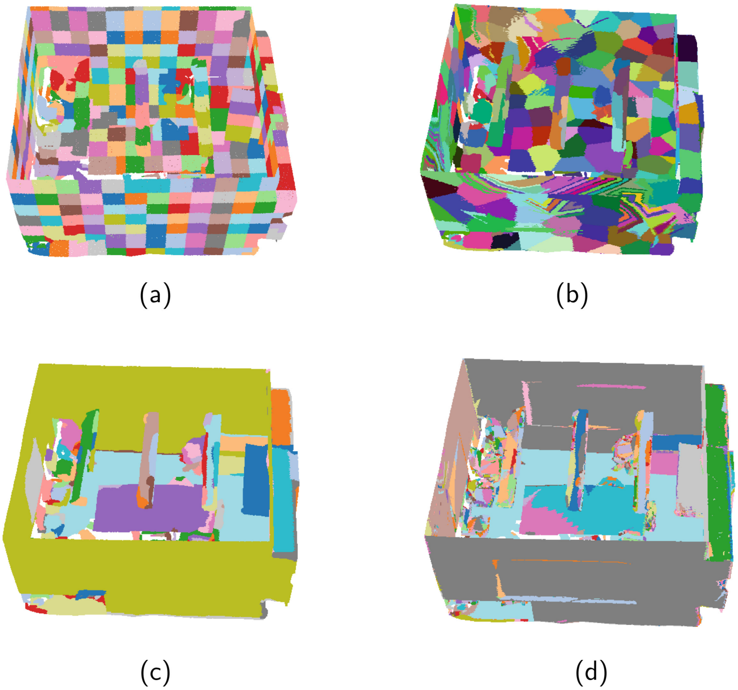

Fig. 7. Results for conferenceRoom_1 from Area_1 for (a) VCCS (b) Lin et al., (c) SSP, and (d) for NGDS

7 Conclusions

Based on performed evaluation, it may be clearly noted that the proposed method yields better results for indoor scenes than state-of-the-art partitioning algorithms. NGDS provides partition result less over-segmented than VCCS or the method of Lin et al. keeping under-segmentation ratio at the very low level (lower than competitive methods). The limitation of the method is the fact that in case of a point cloud of extremely uneven density, the over-segmentation ratio deteriorates significantly, keeping over-segmentation rate at similar, low level. Further research will focus on applying the proposed NGDS method for indoor object detection.

References

1. Armeni, I., Sax, A., Zamir, A.R., Savarese, S.: Joint 2D–3D-semantic data for indoor scene understanding. ArXiv e-prints, February 2017
2. Ben-Shabat, Y., Lindenbaum, M., Fischer, A.: Nesti-Net: normal estimation for unstructured 3d point clouds using convolutional neural networks (2018). http://arxiv.org/abs/1812.00709

3. Bentley, J.L.: Multidimensional binary search trees used for associative searching. Comm. ACM **18**(9), 509–517 (1975). https://doi.org/10.1145/361002.361007
4. Dai, A., Chang, A.X., Savva, M., Halber, M., Funkhouser, T., Nießner, M.: ScanNet: richly-annotated 3D reconstructions of indoor scenes. In: Proceedings of the Computer Vision and Pattern Recognition (CVPR). IEEE (2017)
5. Dong, Z., Yang, B., Hu, P., Scherer, S.: An efficient global energy optimization approach for robust 3D plane segmentation of point clouds. ISPRS J. Phot. Rem. Sens. **137**, 112–133 (2018). https://doi.org/10.1016/j.isprsjprs.2018.01.013
6. El-Sayed, E., Abdel-Kader, R.F., Nashaat, H., Marei, M.: Plane detection in 3D point cloud using octree-balanced density down-sampling and iterative adaptive plane extraction. IET Image Proc. **12**(9), 1595–1605 (2018). https://doi.org/10.1049/iet-ipr.2017.1076
7. Forczmański, P., Nowosielski, A.: Multi-view data aggregation for behaviour analysis in video surveillance systems. In: Chmielewski, L.J., Datta, A., Kozera, R., Wojciechowski, K. (eds.) ICCVG 2016. LNCS, vol. 9972, pp. 462–473. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-46418-3_41
8. Guerrero, P., Kleiman, Y., Ovsjanikov, M., Mitra, N.J.: PCPNet learning local shape properties from raw point clouds. In: Computer Graphics Forum, vol. 37, no. 2, pp. 75–85 (2018). https://doi.org/10.1111/cgf.13343
9. Kasaei, S.H., Tomé, A.M., Seabra Lopes, L., Oliveira, M.: GOOD: a global orthographic object descriptor for 3D object recognition and manipulation. Pattern Recogn. Lett. **83**, 312–320 (2016). https://doi.org/10.1016/j.patrec.2016.07.006
10. Kumar, A., Anders, K., Winiwarter, L., Höfle, B.: Feature relevance analysis for 3D point cloud classification using deep learning. ISPRS Ann. Phot. Rem. Sens. Spat. Inf. Sci. **4** (2019). https://doi.org/10.5194/isprs-annals-IV-2-W5-373-2019
11. Landrieu, L., Boussaha, M.: Supervized Segmentation with Graph-Structured Deep Metric Learning. IEEE Geosci. Remote Sens. Mag. (2019). http://arxiv.org/abs/1905.04014
12. Landrieu, L., Simonovsky, M.: Large-scale point cloud semantic segmentation with superpoint graphs. In: Proceedings of the IEEE CVPR Conference, pp. 4558–4567 (2018)
13. Lazarek, J., Pryczek, M.: A review on point cloud semantic segmentation methods. J. Appl. Comput. Sci. **26**(2), 99–106 (2018)
14. Lin, Y., Wang, C., Zhai, D., Li, W., Li, J.: Toward better boundary preserved supervoxel segmentation for 3D point clouds. ISPRS J. Phot. Rem. Sens. **143**, 39–47 (2018). https://doi.org/10.1016/j.isprsjprs.2018.05.004
15. Liu, M.Y., Tuzel, O., Ramalingam, S., Chellappa, R.: Entropy rate superpixel segmentation. In: CVPR 2011, pp. 2097–2104. IEEE (2011)
16. Lu, R., Brilakis, I., Middleton, C.R.: Detection of structural components in point clouds of existing RC bridges. Comput.-Aided Civil Infrastruct. Eng. **34**(3), 191–212 (2019). https://doi.org/10.1111/mice.12407
17. Ma, Y., Guo, Y., Lei, Y., Lu, M., Zhang, J.: 3DMAX-NET: a multi-scale spatial contextual network for 3D point cloud semantic segmentation. In: 2018 24th International Conference on Pattern Recognition (ICPR), pp. 1560–1566. IEEE (2018)
18. Maleika, W., Forczmański, P.: Adaptive modeling and compression of bathymetric data with variable density. IEEE J. Oceanic Eng. (2019). https://doi.org/10.1109/JOE.2019.2941120

19. McInnes, L., Healy, J.: Accelerated hierarchical density clustering. arXiv preprint arXiv:1705.07321 (2017)
20. McInnes, L., Healy, J., Astels, S.: HDBSCAN: hierarchical density based clustering. J. Open Source Softw. **2**(11), 205 (2017)
21. Silberman, N., Hoiem, D., Kohli, P., Fergus, R.: Indoor segmentation and support inference from RGBD images. In: Fitzgibbon, A., Lazebnik, S., Perona, P., Sato, Y., Schmid, C. (eds.) ECCV 2012. LNCS, vol. 7576, pp. 746–760. Springer, Heidelberg (2012). https://doi.org/10.1007/978-3-642-33715-4_54
22. Papon, J., Abramov, A., Schoeler, M., Worgotter, F.: Voxel cloud connectivity segmentation-supervoxels for point clouds. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 2027–2034 (2013)
23. Rusu, R.B., Cousins, S.: 3D is here: point cloud library. Point Cloud Library. http://pointclouds.org/. Accessed 15 2017
24. Scherer, R.: Feature detection. In: Scherer, R. (ed.) Computer Vision Methods for Fast Image Classification and Retrieval. SCI, vol. 821, pp. 7–32. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-12195-2_2
25. Sculley, D.: Web-scale k-means clustering. In: Proceedings of the 19th International Conference on WWW WWW 2010, pp. 1177–1178. ACM, New York (2010). https://doi.org/10.1145/1772690.1772862
26. Thomas, H., Goulette, F., Deschaud, J.E., Marcotegui, B.: Semantic classification of 3d point clouds with multiscale spherical neighborhoods. In: 2018 International Conference on 3D Vision (3DV), pp. 390–398. IEEE (2018)
27. Walczak, J., Wojciechowski, A.: Clustering quality measures for point cloud segmentation tasks. In: Chmielewski, L.J., Kozera, R., Orłowski, A., Wojciechowski, K., Bruckstein, A.M., Petkov, N. (eds.) ICCVG 2018. LNCS, vol. 11114, pp. 173–186. Springer, Cham (2018). https://doi.org/10.1007/978-3-030-00692-1_16
28. Walewski, P., Gałaj, T., Szajerman, D.: Heuristic based real-time hybrid rendering with the use of rasterization and ray tracing method. Open Phys. **17**(1), 527–544 (2019). https://doi.org/10.1515/phys-2019-0055
29. Wang, X., Liu, S., Shen, X., Shen, C., Jia, J.: Associatively segmenting instances and semantics in point clouds (2019)
30. Wawrzonowski, M., Szajerman, D.: Optimization of screen-space directional occlusion algorithms. Open Phys. **17**(1), 519–526 (2019). https://doi.org/10.1515/phys-2019-0054
31. Wróblewski, A., Andrzejczak, J.: Wave propagation time optimization for geodesic distances calculation using the heat method. Open Phys. **17**(1), 263–275 (2019)
32. Xie, Y., Tian, J., Zhu, X.X.: A review of point cloud semantic segmentation (2019). http://arxiv.org/abs/1908.08854
33. Xu, Y., Yao, W., Hoegner, L., Stilla, U.: Segmentation of building roofs from airborne LiDAR point clouds using robust voxel-based region growing. Remote Sens. Lett. **8**(11), 1062–1071 (2017). https://doi.org/10.1080/2150704X.2017.1349961
34. Yan, J., Shan, J., Jiang, W.: A global optimization approach to roof segmentation from airborne lidar point clouds. ISPRS J. Phot. Rem. Sens. **94**, 183–193 (2014). https://doi.org/10.1016/j.isprsjprs.2014.04.022
35. Yang, F., et al.: Automatic indoor reconstruction from point clouds in multi-room environments with curved walls. Sensors **19**(17), 3798 (2019). https://doi.org/10.3390/s19173798
36. Yin, Z., Liu, Z., Zhou, L., Zhang, F., Fu, K., Kong, X.: Superpixel based continuous conditional random field neural network for semantic segmentation. Neurocomputing **340**, 196–210 (2019). https://doi.org/10.1016/j.neucom.2019.01.016

37. Yousefhussien, M., Kelbe, D.J., Ientilucci, E.J., Salvaggio, C.: A multi-scale fully convolutional network for semantic labeling of 3D point clouds. ISPRS J. Phot. Rem. Sens. **143**, 191–204 (2018). https://doi.org/10.1016/j.isprsjprs.2018.03.018
38. Zhao, J., Liu, C., Zhang, B.: PLSTMNet: a new neural network for segmentation of point cloud. In: 2018 11th International Workshop on Human Friendly Robotics (HFR), pp. 42–47 (2019). https://doi.org/10.1109/hfr.2018.8633482