Article

An Accurate and Efficient Supervoxel Re-Segmentation Approach for Large-Scale Point Clouds Using Plane Constraints

Baokang Lai [1,2,†], Yingtao Yuan [1,2,3,†], Yueqiang Zhang [1,2,*], Biao Hu [1,2] and Qifeng Yu [1,2,3]

1   Shenzhen Key Laboratory of Intelligent Optical Measurement and Detection, Shenzhen University, Shenzhen 518060, China
2   College of Physics and Optoelectronic Engineering, Shenzhen University, Shenzhen 518060, China
3   School of Aeronautics, Northwestern Polytechnical University, Xi'an 710072, China
*   Correspondence: yueqiang.zhang@szu.edu.cn
†   These authors contributed equally to this work.

Abstract: The accurate and efficient segmentation of large-scale urban point clouds is crucial for many higher-level tasks, such as boundary line extraction, point cloud registration, and deformation measurement. In this paper, we propose a novel supervoxel segmentation approach to address the problem of under-segmentation in local regions of point clouds at various resolutions. Our approach introduces distance constraints from boundary points to supervoxel planes in the merging stage to enhance boundary segmentation accuracy between non-coplanar supervoxels. Additionally, supervoxels with roughness above a threshold are re-segmented using random sample consensus (RANSAC) to address multi-planar coupling within local areas of the point clouds. We tested the proposed method on two publicly available large-scale point cloud datasets. The results show that the new method outperforms two classical methods in terms of boundary recall, under-segmentation error, and average entropy in urban scenes.

Keywords: point cloud segmentation; under-segmentation; supervoxel; boundary enhanced; random sample consensus; roughness

1. Introduction

In recent years, significant advances in laser measurement technology and the introduction of innovative scanning equipment [1,2] have increased the accuracy and reliability of point clouds. The redundancy, rotational invariance, and lack of topological relationships between points in point clouds make it a powerful and flexible medium for representing the surface morphology of objects. These advantages have led to progress in various areas, such as building feature extraction [3], complex terrain analyses [4,5], forest investigations [6], deformation measurement [7], and structural health monitoring and assessment [8] using point clouds.

The core of point cloud processing is to infer geometric information from 3D discrete points [9]. Large-scale point clouds are characterised by non-uniform density and high levels of noise. It is common to segment point clouds into lower-order entities to extract the required information based on their spatial location, geometric features, textures, and other properties [10]. This process is known as point cloud segmentation. Traditional point cloud segmentation methods include region-growing methods [11–13], clustering-based methods, such as K-means clustering [14] and density clustering (DBSCAN) [15], as well as classic model-fitting-based methods, i.e., Hough Transform (HT) [16–18] and random sample consensus (RANSAC) [19–22].

However, region-growing-based segmentation methods rely on selecting seed points [23], and it would be challenging to obtain accurate results for large-scale and non-uniform dense point clouds. Clustering-based methods do not rely on the initialisation of seed points, but pre-defined parameters heavily influence their results. Similarly, the main

drawback of model-fitting-based methods is their sensitivity to pre-defined parameters and high time complexity. Additionally, degeneracy could occur when applying the standard RANSAC to large-scale point clouds [9].

Supervoxels are often used as a pre-processing step in segmentation algorithms. Similar to superpixels in 2D images [24], supervoxels are small regions of perceptually similar voxels and are usually used to improve the definition of object boundaries. Furthermore, supervoxels can significantly reduce the data volume of raw point clouds and the adverse impact of noise with minimal information loss [25]. It plays an important role in large-scale point cloud processing methods.

The boundaries of supervoxels should be as close as possible to object boundaries [10]. One classical method for point cloud supervoxel segmentation is voxel cloud connectivity segmentation (VCCS) [26], which uses an octree to voxelize the point cloud and then applies the K-means clustering algorithm to achieve supervoxel segmentation. However, VCCS uses a fixed resolution and relies on the initialisation of the seed, which makes it challenging to achieve high-quality segmentation boundaries in non-uniform density point clouds. Song et al. [27] proposed a boundary-enhanced supervoxel segmentation (BESS) method to address this issue. Although the method performs well in outdoor scenes with non-uniform density, it assumes that the boundary points are sequentially aligned in one direction. Therefore, it is normally restricted to single settings and unsuitable for complex, large-scale point clouds. Lin et al. [23] proposed a boundary-preserved supervoxel segmentation (BPSS) method that formulates the segmentation problem as a subset selection problem and addresses this problem by minimising an energy function. Their approach does not necessitate the initialisation of cluster seeds and is reported to generate supervoxels with varying resolutions to conform boundaries more effectively. Puligandla and Lončarić [28] proposed a supervoxel segmentation method based on the k-means algorithm with dynamic cluster seed initialisation. In recent times, the field of point cloud segmentation has seen advancements through deep learning [29–31]. Particularly, supervoxel segmentation has also seen advancements through applying deep learning techniques. Hui et al. [32] developed a deep iterative clustering network for end-to-end supervoxel generation from point clouds. Landrieu and Boussaha [33] presented the first supervised framework for over-segmenting 3D point clouds.

For most supervoxel algorithms, the number of supervoxels is related to the resolution value, with more supervoxels providing greater detail at lower resolutions. Ni and Niu [34] proposed a supervoxel segmentation method using local allocation, which preserves instance boundaries and enforces local minimisation through a new cost function. Xiao et al. [35] devised a framework for generating supervoxels that resolves an energy optimisation issue with a specified number of supervoxels. To address the problem of under-segmentation when the number of supervoxels is low, Sha et al. [36] optimised the road boundaries based on BPSS. Su et al. [37] proposed improving the selection of representative points. Instead of random sampling in BPSS, their method selects representative points within non-boundary regions. However, the issue of under-segmentation at a wide variety of resolutions is still worth investigating.

In this paper, we propose an accurate and efficient supervoxel re-segmentation method to achieve high-precision boundary preservation for large-scale point clouds with various resolutions. Our algorithm initially optimises the boundary points of BPSS supervoxels by introducing point-to-plane distance constraints. Then, to solve multi-planar coupling issues, RANSAC is applied to re-segment supervoxels with roughness over a threshold. Ultimately, our proposed method enhances the precision of point cloud boundary segmentation across various resolutions.

This research focuses on addressing the problem of under-segmented planes in supervoxels in 3D point clouds. Our specific contributions are as follows:

- We analysed two cases of under-segmentation in the current supervoxel generation results: (1) failure to segment non-coplanar supervoxel boundaries and (2) the multi-planar coupling problem caused by resolution and target size mismatch.

- We developed a novel algorithm that improves the accuracy of boundary segmentation between non-coplanar supervoxels. By introducing distance constraints during the merging stage and utilising RANSAC for re-segmentation, our algorithm effectively addresses the multi-planar coupling problem in local point cloud regions.
- We present a novel idea in supervoxel segmentation tasks: re-segment under-segmented supervoxels to improve the final segmentation result.

This paper is structured as follows. Section 1 introduces background information and related works for the supervoxel. In Sections 2.1 and 2.2, we provide an overview of the BPSS method and analyse the reasons for its under-segmentation problems. Section 2.3 presents our proposed re-segmentation method in detail, with Section 2.3.1 introducing the redefinition of the boundary exchange condition for supervoxels. Section 2.3.2 explains the re-segmentation of under-segmented supervoxels based on roughness filtering. In Section 3, we present the experimental results and compare our method with others. In Section 4, we analyse the results of the experiments and discuss the limitations of the algorithm in some cases. Finally, in Section 5, we summarise the paper and provide an outlook for future work.

2. Methods

2.1. BPSS

The algorithm proposed in this paper is an improvement upon the BPSS [23]. In BPSS, a heuristic approach for supervoxel segmentation is proposed by minimising the energy function shown in the following equation.

$$
\begin{aligned}
\min \quad & E(Z) = \sum_{i=1}^{N} \sum_{j=1}^{N} z_{ij} D(p_i, p_j) + \lambda |C(Z) - K| \\
\text{s.t.} \quad & z_{ij} = \{0, 1\}, \quad \forall i, j; \quad \sum_{i=1}^{N} z_{ij} = 1, \quad \forall j.
\end{aligned}
\tag{1}
$$

According to Equation (1), $E(Z)$ consists of two terms. The first term describes the geometric properties, where z_{ij} is a binary number of 0 or 1. If $z_{ij} = 1$, it means that p_i is a supervoxel representative point and p_j is subordinate to p_i. $D(p_i, p_j)$ is a distance metric to measure the dissimilarity between two points, which will be explained in Section 2.3.1. The second term is the constraint term, which constrains the number of supervoxel representative points $C(Z)$ to converge to K. λ is the regularisation parameter used to set the trade-off between two terms.

BPSS minimises the energy function through fusion and exchange operations. The exchange-based minimisation operation redefines boundary points based on the dissimilarity distance between the boundary point and the representative point of each of the adjacent supervoxels. This distance metric measures the weighted difference from the normal vectors and the relative position of two points, which ensures that each supervoxel has a relatively complete boundary.

2.2. Under-Segmentation Issues

Although BPSS addresses the problem of non-uniform point cloud density to some extent, its resolution setting still makes the method unable to segment regions with large local variations accurately. In supervoxel segmentation, the resolution of supervoxels corresponds to the granularity of segmentation, and supervoxel structures with different granularities can describe details at different levels [10]. Furthermore, suppose the resolution of supervoxels is set too low. In that case, the number of points within each supervoxel may be insufficient (e.g., around 10 in Figure 1g), which could negatively impact subsequent calculations [38]. Therefore, supervoxel resolution should be set to be similar to, or slightly smaller than, the expected segmentation target size in the point cloud while avoiding minimal values. As shown in Figure 1, the enlarged portion is a step-like protrusion on the façade of the point cloud. Even if the resolution of BPSS is

set according to the size of the step-like protrusion, the issue of under-segmentation still occurs, as illustrated in Figure 1e–g.
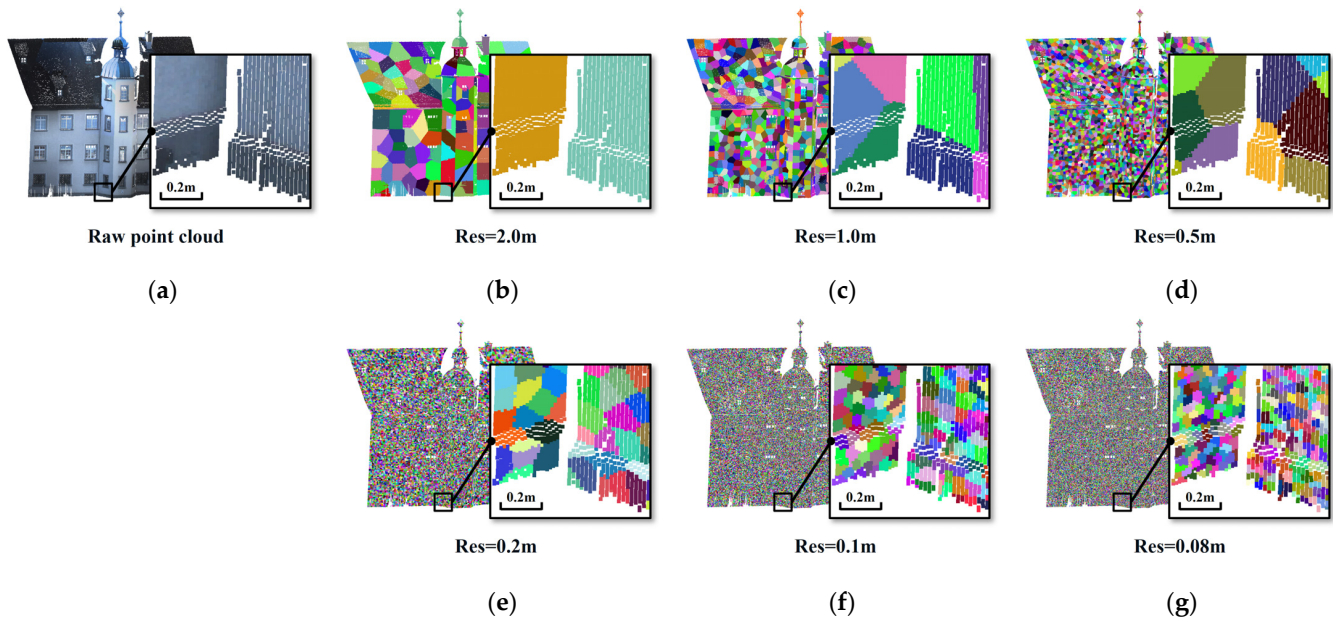

Figure 1. Segmentation results of BPSS at different resolutions. (**a**) Raw point clouds; (**b**–**g**) the segmentation results with resolution values are 2.0 m, 1.0 m, 0.5 m, 0.2 m, 0.1 m, and 0.08 m, respectively.

In addition, the results of BPSS in Figure 2 illustrate two main cases of segmentation errors in local regions. The first case, shown in Figure 2a, is the boundary point error caused by the limitations of the similarity distance judgement, which ignores the connection between points and supervoxel planes. This issue becomes even more severe when the planar normal vectors of two non-coplanar supervoxels differ by a small degree. The second case is the under-segmentation of the supervoxel, as shown in Figure 2b, where several small planes within a supervoxel are coupled due to the supervoxel resolution setting. Both cases limit the segmentation effectiveness of BPSS. Section 2.3 will discuss these errors in detail and provide corresponding solutions.
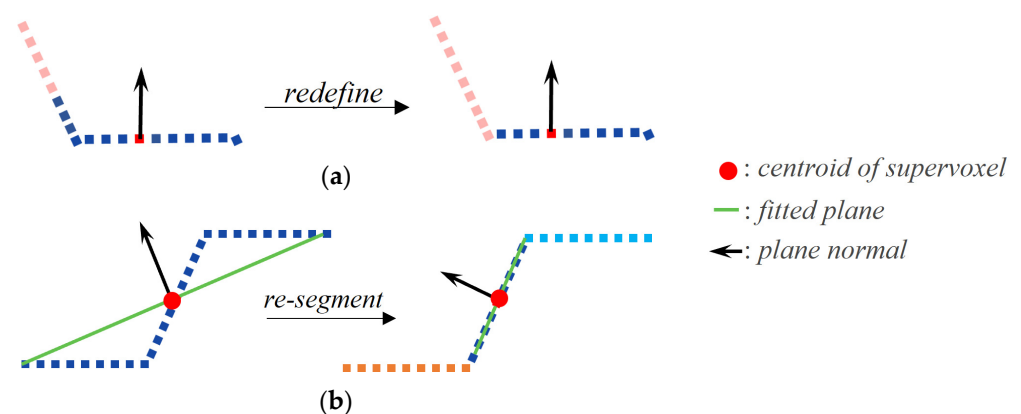

Figure 2. Possible incorrect segmentation cases, where different colors indicate distinct supervoxel planes. (**a**) Boundary point segmentation error; (**b**) under-segmentation of local regions.

2.3. Proposed Algorithm

This paper proposes a robust supervoxel segmentation algorithm to address the issues of BPSS segmentation errors in local regions. The algorithm preserves boundaries at various resolutions and consists of two steps, as shown in Figure 3. After using statistical outlier removal (SOR) [39] to remove scattered points in the point cloud, we introduced a point-to-

plane distance constraint in the initial segmentation stage of the exchange phase to redefine boundary points in the first step. In the second step, we re-segmented supervoxels with roughness above the threshold using the RANSAC algorithm.
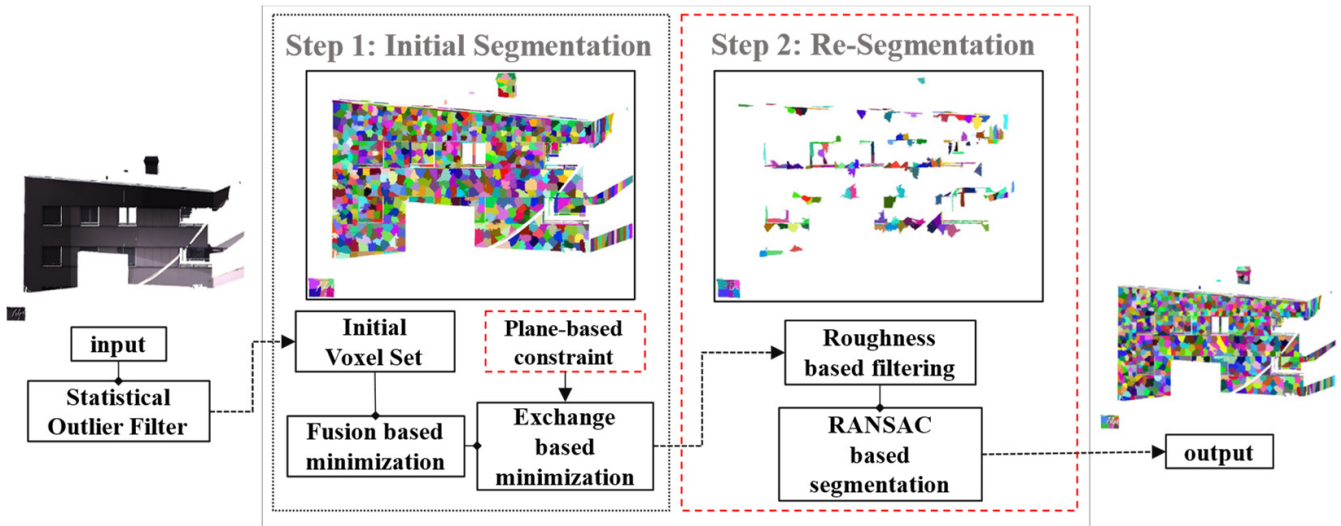

Figure 3. The proposed algorithm workflow.

2.3.1. Boundary Point Redefinition

As presented in the BPSS, Equation (2) calculates the dissimilarity distance from the boundary point p to the adjacent supervoxel representative points q_i to determine whether the point is exchanged. The dissimilarity distance is calculated by the weighted sum of the normal vector projection and the Euclidean distance,

$$D(p, q_i) = 1 - \left| n_p \cdot n_{q_i} \right| + 0.4 \frac{r(p, q_i)}{R} \tag{2}$$

where n_p and n_{q_i} are the normal vectors of p and q_i, respectively, obtained by using principal component analysis (PCA) [40], $r(p, q_i)$ denotes the Euclidean distance between points p and q_i, and R is a user-defined parameter that represents scale resolution value. However, due to resolution value, non-uniform density distribution, noise, and outliers, there are some errors in the boundary point normal vectors obtained through PCA in the boundary region between non-coplanar supervoxels. At the same time, Equation (2) is directly applied to the point-to-point relationship without considering the connection between the point and the supervoxel plane, which is the underlying cause of the segmentation errors.

In order to solve the problem of incorrect boundary segmentation of the non-coplanar supervoxels shown in Figure 2a, this paper refines the exchange minimisation in BPSS by introducing point-to-plane constraints. The algorithm flow is shown in Algorithm 1. Specifically, for a supervoxel V_i and its adjacent supervoxel V_j, consider it as plane S_i and plane S_j determined by the coordinates of the supervoxel centroid and the centroid normal vector, respectively. Calculate the Euclidean distances $r(p_b, S_i)$ and $r(p_b, S_j)$ from the boundary point p_b to the planes S_i and S_j, respectively. If it satisfies both $D(p_b, p_i) > D(p_b, p_j)$ and $r(p_b, S_i) > r(p_b, S_j)$, this indicates that the boundary point is more suitable for the supervoxel V_j; then, p_b is exchanged to complete the classification of this boundary point.

Algorithm 1: The stage of boundary exchange.

Input: Initialize a queue, Q, for points, i.e.,
while　$Q \neq \varnothing$ do
Remove the front point p_b from Q
for all $p_b \in V_i$ do
if　$D(p_b, p_i) > D\left(p_b, p_j\right) \& r(p_b, S_i) > r\left(p_i, S_j\right)$
exchange p_b to p_j
if　$p_b \notin Q$ then
Add p_b to the back of Q
end if
end if
end for
end while

The effectiveness of our proposed method in enhancing the segmentation of non-coplanar supervoxel boundaries is demonstrated in Figure 4. The figure shows the results of both the BPSS and the proposed boundary redefinition. In particular, our optimization step outperforms BPSS in segmenting the boundaries of step-like protrusions and window sills, as indicated by the red arrows in Figure 4b,c. This results in clearer and more distinct edges for non-coplanar supervoxels, thereby enhancing the algorithm's advantage in boundary segmentation of complex geometric shapes.
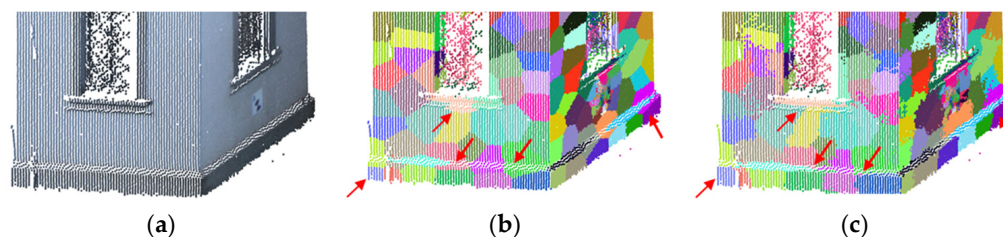

(**a**)　　　　　　　　　　　　(**b**)　　　　　　　　　　　　(**c**)

Figure 4. The segmentation results at a resolution of 0.5. (**a**) Raw point clouds; (**b**) the result of BPSS; (**c**) the result after boundary redefinition.

2.3.2. RANSAC-Based Re-Segmentation

The voxel resolution should be closely matched to the target size in order to achieve optimal segmentation outcomes. However, due to the non-uniform density distribution and complex intersection of multiple planes in large-scale point clouds, only a rough resolution range can typically be determined. Therefore, for the refined BPSS after the boundary redefinition in Section 2.3.1, we make the following two assumptions about the segmentation results of the supervoxels within the suitable resolution range:

- The first assumption is that, at a given appropriate resolution, the under-segmented supervoxels only account for a minority of the total supervoxel set.
- The second assumption is that the number of under-segmented points in an under-segmented supervoxel is in the minority.

Therefore, there is no requirement to re-segment all supervoxels.

According to the first assumption, only a proportion of supervoxels require re-segmentation. This paper uses a roughness-based statistical method to screen out the supervoxels that need to be re-segmented. For the set of supervoxels obtained from the initial segmentation, we compute the supervoxel planes using the same method as in the boundary redefinition section above and then calculate the roughness of each point within each supervoxel according to Equation (3).

$$\sigma = \sqrt{\frac{1}{N} \sum_{i=0}^{N} [r(p_i, S) - \bar{r}_S]^2} \tag{3}$$

where N is the number of points contained in the supervoxel, $r(p_i, S)$ is the Euclidean distance from the point p_i to the supervoxel plane S, and $\bar{r}_S$ is the mean distance of all points to the supervoxel plane. The mean value of the roughness represents the overall roughness of the supervoxel at each point. To reduce the impact of outliers, we exclude 5% of the extreme values when calculating the mean roughness, using only 95% of the data for analysis [41]. Figure 5 illustrates the outlier removal method proposed in this paper. Firstly, we apply SOR filtering to the entire point cloud to remove obvious outliers. Then, we use maximum suppression of roughness (MSR) to calculate the robust average roughness of the supervoxels.
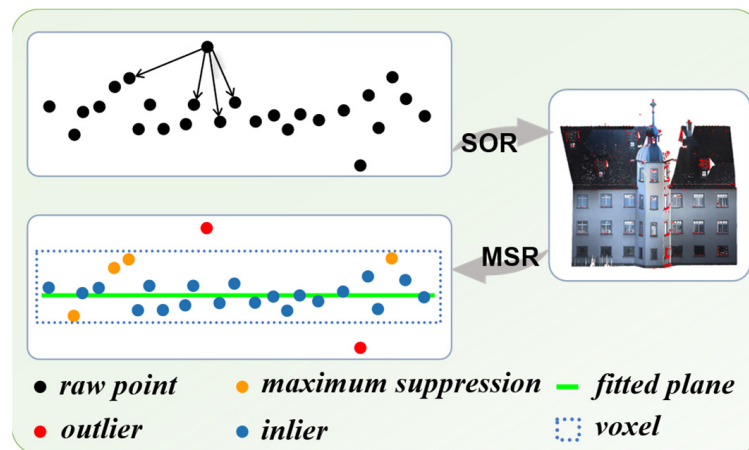

Figure 5. Statistical outlier removal and maximum suppression of roughness.

After calculating the roughness of all supervoxels, those supervoxels with roughness above the threshold are selected to be re-segmented. According to the three-sigma rule [41], the roughness threshold τ_σ is conservatively adjusted to 68% of the roughness set in ascending order. Figure 6. shows the supervoxels that were computed and selected for re-segmentation. In Figure 6a, the black dotted line corresponds to the adaptive threshold. The bar chart in Figure 6b represents the roughness distribution of the scatter plot on the left. Among them, the blue data points below the threshold account for 68% of the total, and the supervoxels represented by these points do not require re-segmentation. The remaining supervoxels, as shown in Figure 6c, need to be re-segmented. And, then, RANSAC is used to re-segment the supervoxels with roughness above the threshold value.
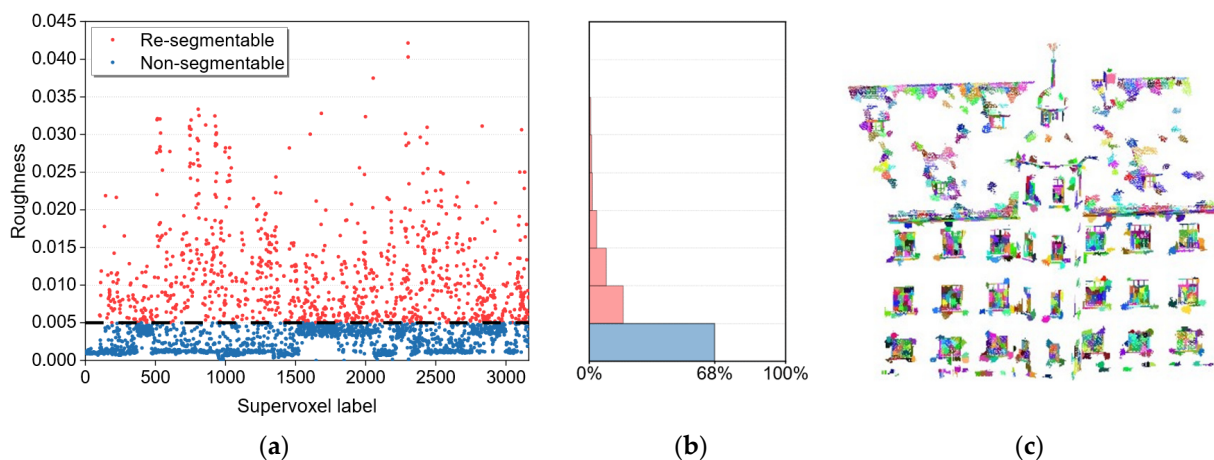

Figure 6. Roughness-filtered supervoxels for re-segmentation. (**a**) Scatter plot of roughness for each supervoxel; (**b**) percentage distribution of roughness intervals; (**c**) under-segmented supervoxels after filtering.

RANSAC is a simple but powerful tool with wide applications in outlier removal and plane detection. It robustly estimates model parameters through random sampling. In this paper, RANSAC is used to decouple multi-planar coupling supervoxels.

The distance threshold τ_d is the main parameter in the supervoxel that affects the results of RANSAC segmentation. It defines the maximum distance allowed from a point to the estimated plane. According to the second assumption, this research adaptively adjusts the distance threshold to 0.5-times the average density of each supervoxel, ensuring efficient segmentation while avoiding RANSAC degeneracy. The average density of each supervoxel is defined as the mean of the average distances between all points within the supervoxel and their eight nearest neighbours. To eliminate the impact of voids and discrete points on the computation of average density, the distance value for each point is arranged in ascending order, and only the top 90% of data is used to calculate the point cloud's average density. After using RANSAC to perform multi-plane segmentation on under-segmented supervoxels, the segmented results are re-labelled and then output. The algorithm flow for the stage of re-segmentation is shown in Algorithm 2.

Algorithm 2: The stage of re-segmentation.

Start from the voxel set V
while V_i is not processed
compute σ_i and $\bar{r}_S$ for V_i
if $\sigma_i > \tau_\sigma$
$\tau_d = 0.5 \times \bar{r}_S$
Ransac to V_i
relabel V_i
end if
end while

---

3. Results

3.1. Experimental Setup

Our proposed re-segmentation method is accurate and effective for urban scenes. To comprehensively evaluate our proposed algorithm, we implemented experiments on different publicly available datasets. We compared our results with those obtained using classical algorithms, e.g., VCCS [26], BPSS [23], and a variant of the VCCS algorithm, i.e., VCCS_KNN, implemented by Lin et al. [23]. Furthermore, for further insight into the algorithm proposed in this paper, we conducted a more in-depth analysis by separating the two-step operation into Proposed_RANSAC and Proposed_REDEFINE and comparing them with the complete algorithm. All of these algorithms were implemented with the support of the Point Clouds Library (PCL) [39]. Moreover, the code was executed in C++ on a Windows 10 PC with an Intel(R) Core(TM) i7-9750H CPU @ 2.60 GHz and 16 GB of RAM. The remainder of Section 3 will introduce the datasets used in the experiments, the specific evaluation metrics, the parameter settings, and the experimental results.

3.2. Datasets

In this paper, we evaluate our proposed algorithm on Semantic3D [42] and IQmulus & TerraMobilita (IQTM) [43]. Semantic3D is a terrestrial laser scanning (TLS) dataset published by ETH Zurich that contains a variety of scenes, such as streets, squares, and villages. IQTM is an urban street scene dataset collected by mobile laser scanning (MLS). Compared to Semantic3D, IQTM has a denser point distribution in urban scenes, but it also has relatively regular boundaries.

We selected two building point clouds from Semantic3D and intercepted a sparse road scene point cloud from IQTM to test our algorithm. Then, we manually labelled the selected point clouds and used them as ground truth to evaluate our algorithm. Figure 6 shows the point clouds and their corresponding density distributions and ground truths, where we use a colour gradient from blue–green–yellow–red to represent the change in

point cloud density distribution from sparse to dense. In Figure 7, the point is displayed in red when the number of points within its unit neighbourhood region (radius = 0.05 m) exceeds 50. Table 1 provides further details about these point clouds.
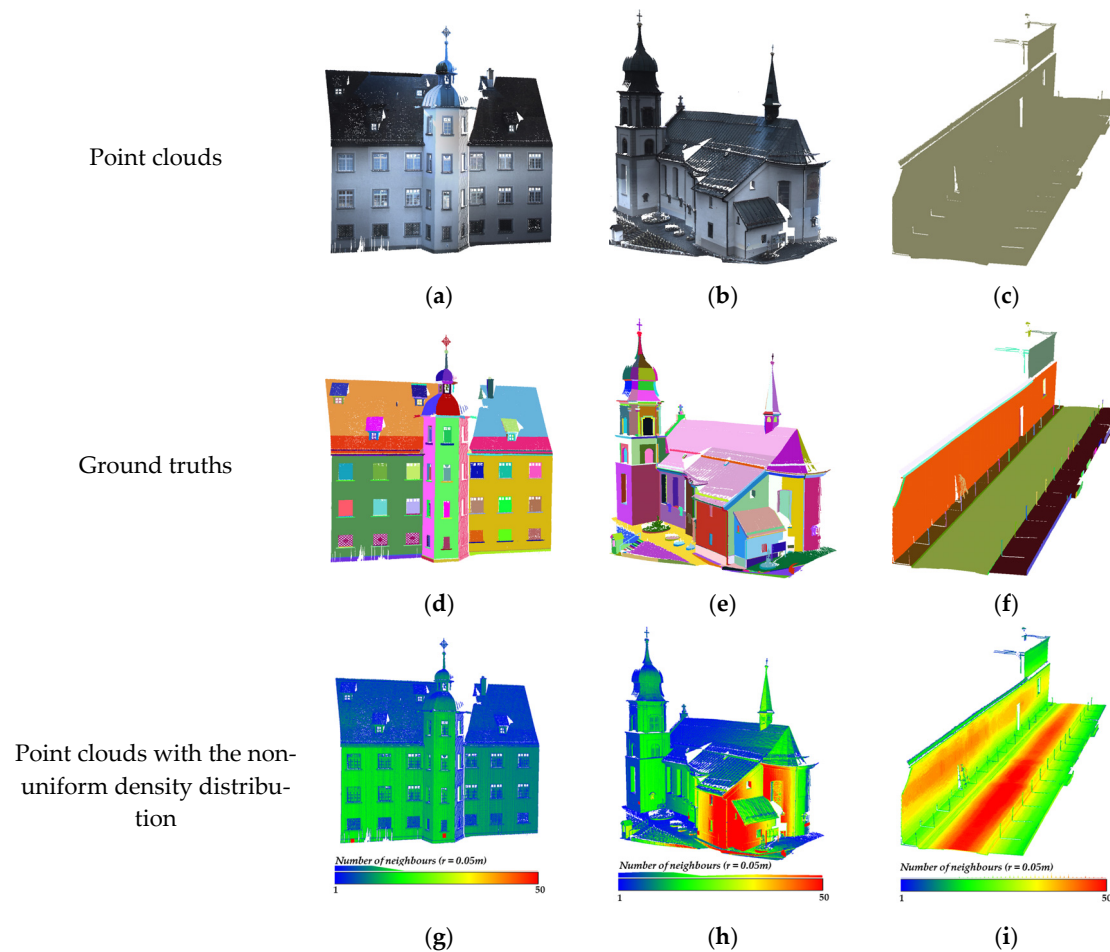
Point clouds

(**a**) (**b**) (**c**)

Ground truths

(**d**) (**e**) (**f**)

Point clouds with the non-uniform density distribution

(**g**) (**h**) (**i**)

Figure 7. The point clouds and their corresponding density distributions and ground truths. (**a**) Stgallen1; (**b**) bildstein1; (**c**) cassette; (**d–f**) corresponding ground truths; (**g–i**) corresponding density distributions.

Table 1. Introduction to point cloud data.

Data Set	Point Clouds	Point Number	Planes	Boundary Points
Semantic3D	stgallen1	620,133	142	25,930
	bildstein1	2,269,937	367	246,856
IQTM	cassette	1,189,410	44	41,737

3.3. Evaluation Metrics

To better evaluate our algorithm's performance, we use four metrics, the boundary recall (*BR*) [23], the under-segmentation error (*UE*) [26], the mean entropy of the model (*H*), and the runtime (*T*).

Boundary recall measures the overlap between the ground truth and the segmented supervoxel boundary. A high *BR* indicates that the supervoxel correctly preserves the boundaries of the plane. It is calculated using Equation (4).

$$BR = \frac{T_P}{T_P + F_N} \qquad (4)$$

In Equation (4), T_P denotes the number of truth boundary points detected; F_N denotes the number of truth boundary points not detected.

The under-segmentation error represents the leakage of the supervoxel across the truth boundary. For a truth segmentation with the region $g_1, \ldots, g_M$ of truth segmentation results and the set of supervoxels $V_1, \ldots, V_K$, the under-segmentation error is defined as,

$$UE = \frac{1}{N} \left[\sum_{i=1}^{M} \left(\sum_{V_j | V_j \cap g_i} |V_j| \right) - N \right] \tag{5}$$

where $V_j \mid V_j \cap g_i$ denotes the set of supervoxels required to cover the truth label g_i and N is the number of truth points. A lower UE value means that less supervoxels violated ground truth borders by crossing over them.

According to entropy theory [44,45], we introduce the information entropy of the segmented model as an evaluation criterion for supervoxels. Information entropy can be used to metric the level of dependency of point cloud data for a class of features. For data points in a point cloud, the entropy E_λ is calculated as follows,

$$E_\lambda = -\sum_{i-1}^{3} e_i \ln(e_i) \tag{6}$$

where $e_{i=1,2,3}$ represents the three eigenvalues obtained by calculating the local covariance matrix of the points. Then, the whole entropy of the model is obtained by averaging the entropy of the points.

$$H = \frac{1}{MN} \sum_{i=1}^{M} \sum_{V_i}^{N} E_{\lambda_j} \tag{7}$$

M is the total number of supervoxels obtained from the segmentation, and N is the number of points contained in the supervoxel. We calculated the entropy of the manually annotated ground truth. For the three-point cloud scenes stgallen1, bildstein1, and cassette, the entropy values of ground truths were 0.34, 0.31, and 0.16, respectively.

In this paper, we define the ground truth boundary as the neighbouring point between two labelled results. For any point in a plane, we traverse its eight nearest neighbours, and if a neighbour's label differs from its own, that point is considered a boundary point. This standard is also used to determine boundary points in the data to be measured and to compare them with labelled results.

3.4. Algorithms and Parameters

To comprehensively evaluate our proposed algorithm, we implemented experiments on point cloud data with various resolutions. Specifically, the proposed algorithm was compared with five other supervoxel algorithms to validate its effectiveness in processing large-scale point clouds. These algorithms include VCCS, BPSS, VCCS_KNN, and the two stages of our proposed method, i.e., Proposed_REDEFINE(P_RDF) and Proposed_RANSAC(P_RSC).

We set up a range of resolutions for the three selected point cloud datasets as parameters. This range encompassed the entire state of point cloud supervoxel segmentation, from over-segmentation to under-segmentation. The testing resolution range of different point cloud models is roughly determined based on the expected segmentation target size. Figure 8 illustrates the specific results of the BPSS and the proposed algorithm for supervoxel segmentation at different resolutions. Multiple experiments were conducted, with the number of search points for the nearest neighbours fixed at 20 for all six methods. And, in the proposed algorithm, the roughness threshold was conservatively set to filter out supervoxels beyond 68%, based on the three-sigma rule [41], while the distance threshold was set to half of the supervoxel mean density.
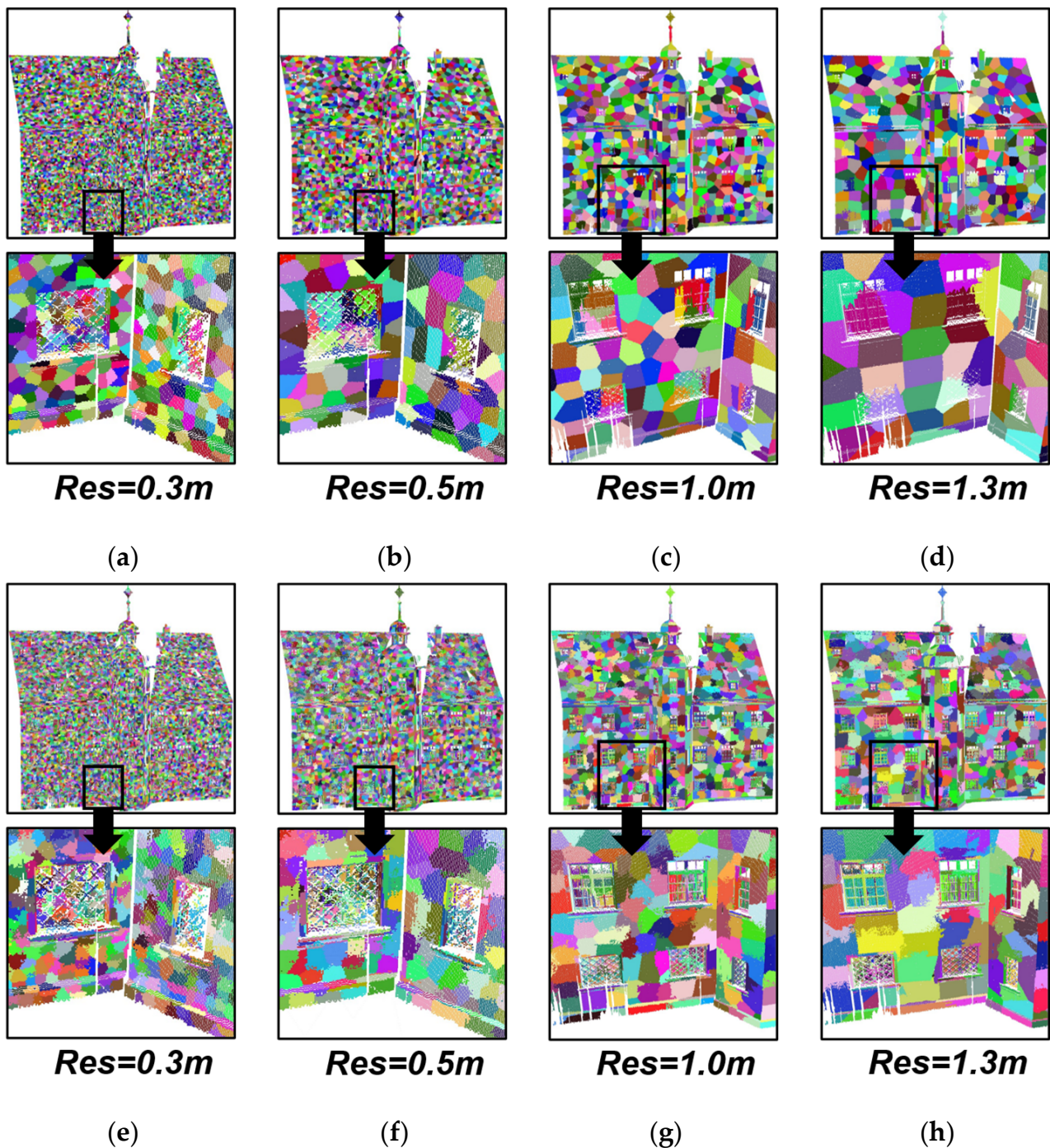
Figure 8. Segmentation results for different resolutions of supervoxels. (**a**–**d**) BPSS; (**e**–**h**) proposed algorithm.

3.5. Experimental Results

In this section, we compare our algorithm's performance with other algorithms in terms of boundary recall, under-segmentation error, information entropy, and algorithm efficiency. The reliability and validity of our method are verified through qualitative and quantitative analysis of the point clouds at different resolutions.

In the stgallen1 scene, we present segmentation results for each algorithm with a resolution set to 0.5 m. As shown in Figure 9, the results are divided into two rows with overall supervoxel display results on top and local-region zooms on the bottom. The highlighted black rectangles shows the step-like protrusions on a façade. Only our proposed

algorithm (Figure 9g) can highlight object boundaries in local enlargement, especially at the boxes. Additionally, quantitative metrics in Figure 10 show that our proposed algorithm performs excellently compared to other algorithms under evaluation metrics. Regarding boundary recall (BR), the proposed algorithm achieves the highest values at all resolutions. At the same time, the BR value of Proposed_REDEFINE is significantly better than that of BPSS, indicating that the addition of point-to-plane constraint optimisation improved the boundary recall metric. Using only RANSAC, the Proposed_RANSAC method can also achieve a high recall rate, but its performance is slightly worse than that of the proposed algorithm. It is due to the addition of boundary constraints and the optimisation of non-coplanar supervoxel boundaries. Simultaneously, the proposed algorithm also performs best in terms of under-segmentation error and entropy.
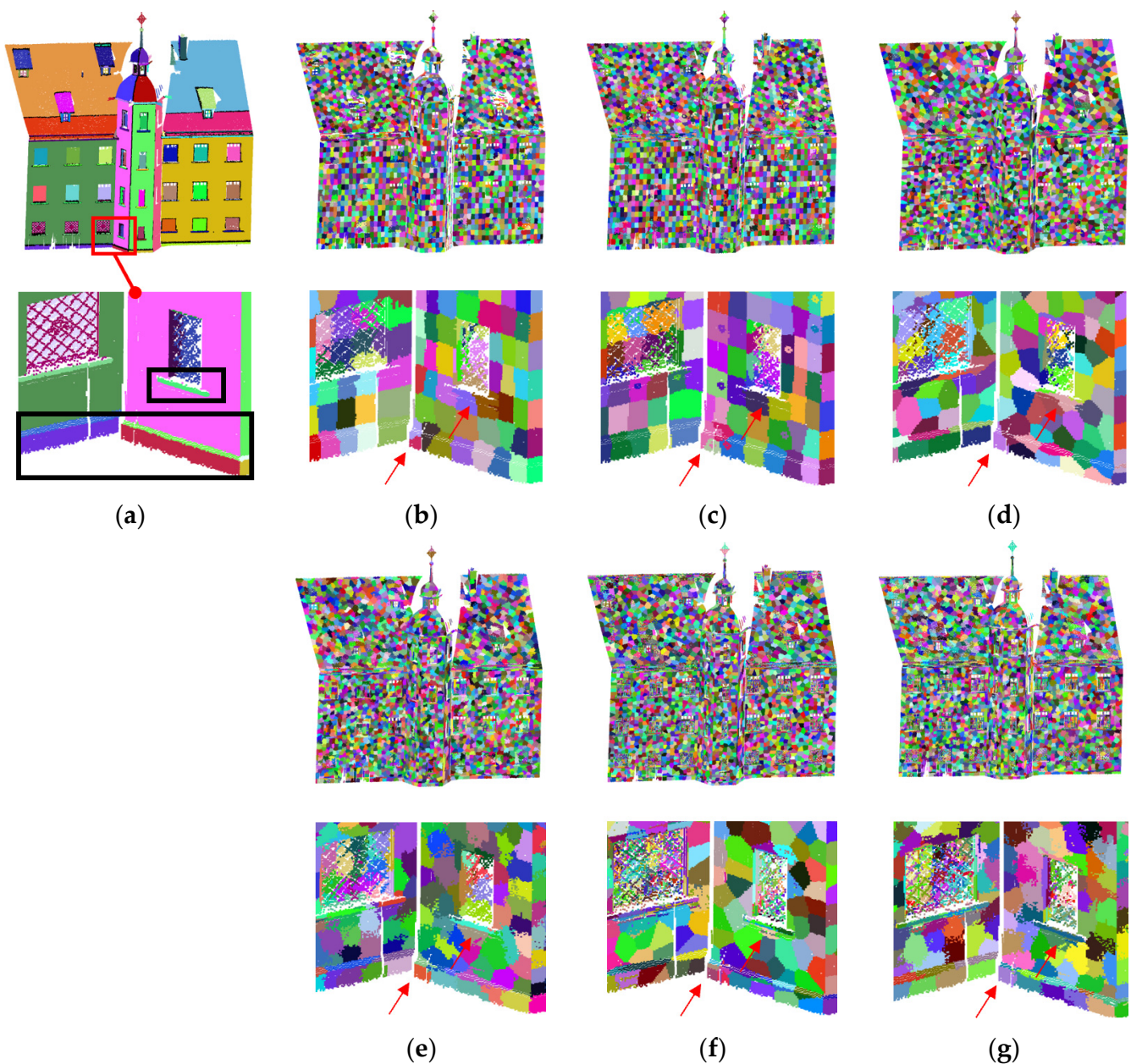

Figure 9. When R = 0.5, the stgallen1′s segmentation results in partially enlarged detail. (**a**) Ground truth; (**b**) VCCS; (**c**) VCCS_KNN; (**d**) BPSS; (**e**) Proposed_REDEFINE; (**f**) Proposed_RANSAC; (**g**) proposed.
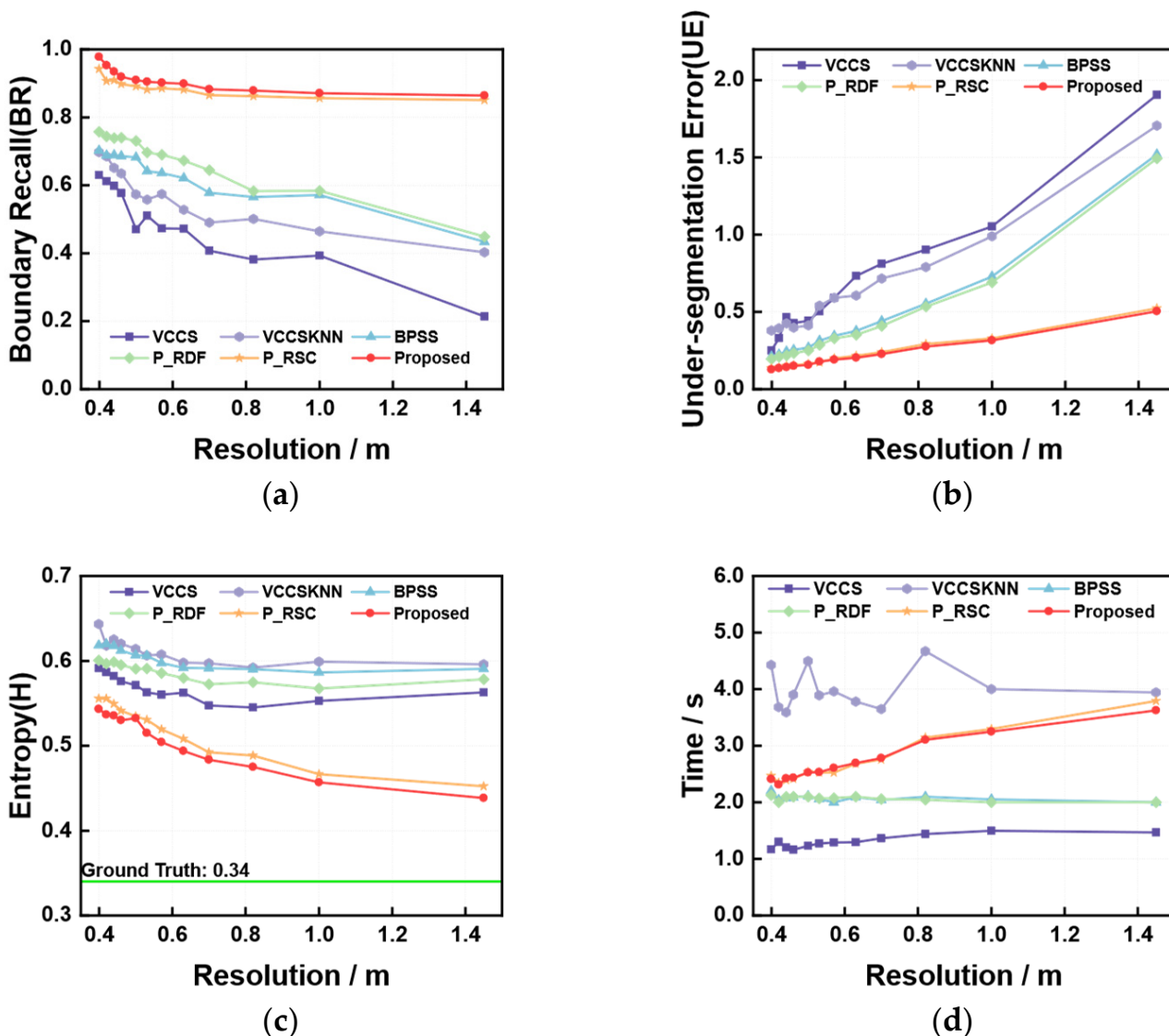
Figure 10. Evaluation metrics of algorithms on stgallen1 at different resolutions. (**a**) BR; (**b**) UE; (**c**) H; (**d**) time.

In the bildstein1 scene, the results of each algorithm at a resolution of R = 0.7 are selected for presentation in this paper. As shown in Figure 11, where local enlargement shows a window, BPSS_RANSAC (Figure 11f) and our proposed algorithm (Figure 11g) demonstrate better outcomes than other algorithms at the arrows in the presented local scene. This is reflected in their ability to segment small planes more clearly, making window features more apparent. At the same time, planes segmented by our proposed algorithm are more complete compared to only using the RANSAC method. And, it is also illustrated by boundary recall and under-segmentation metrics in quantitative metrics in Figure 12.

Figure 13 shows the segmentation results for the cassette scene for each algorithm when R = 0.7. In the local enlargement shown, VCCS and VCCS_KNN cannot segment the kerb when the resolution is too large. In contrast, as shown in Figure 13d, although the BPSS method can segment the road plane, the supervoxel region indicated by the arrow appears both above and below the step. At the same time, the BPSS method cannot accurately segment the columnar railings shown in the area marked by the black boxes. However, notably, our proposed algorithm can better address these problems, as demonstrated by its ability to segment the kerb and avoid under-segmentation of supervoxel regions accurately. The proposed method can also accurately segment small objects, such as pillar bars, with

higher accuracy and reliability. The quantitative experimental results shown in Figure 14 support the conclusions of the qualitative analysis.
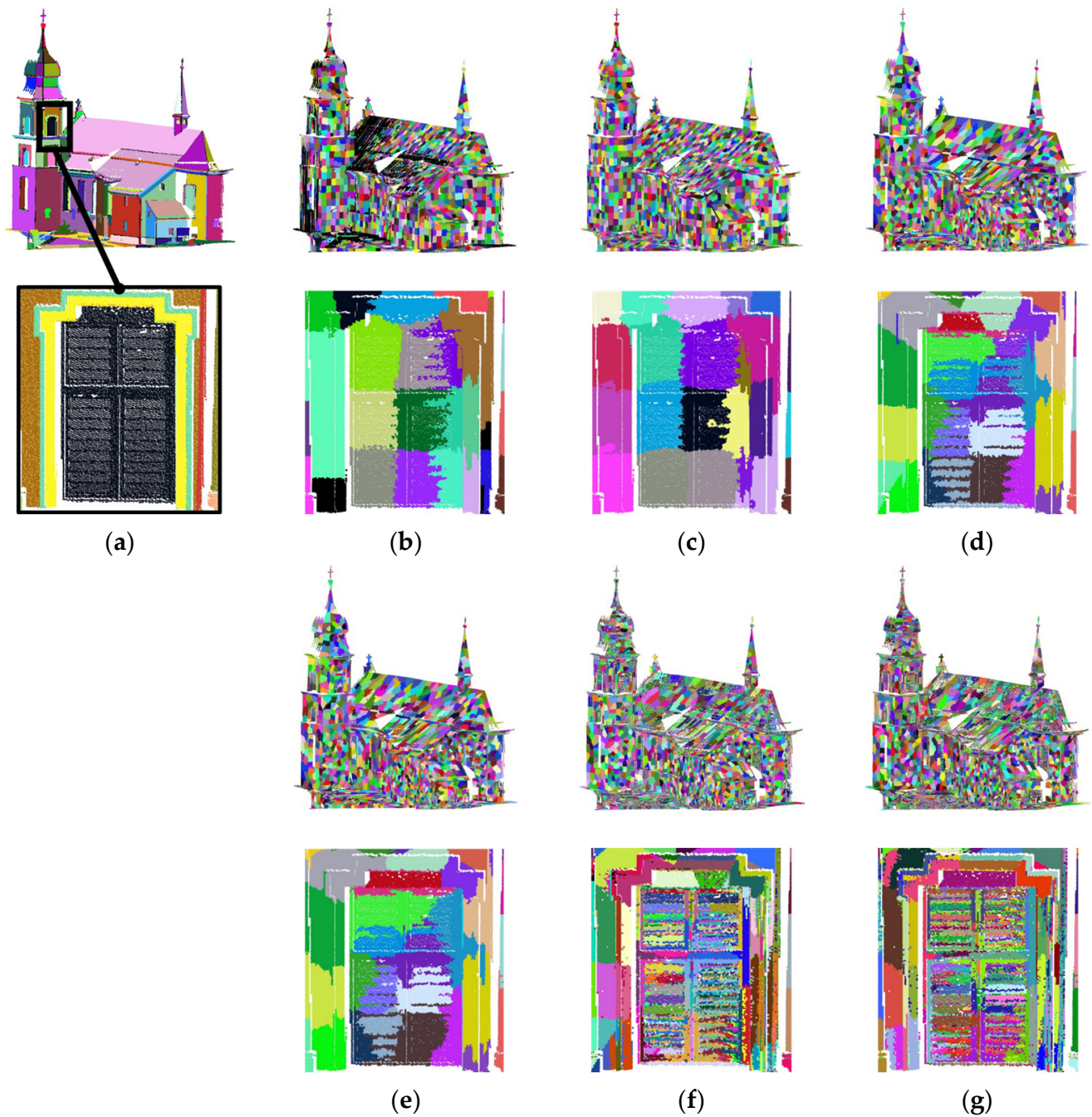

Figure 11. When R = 0.7, the bildstein1′s segmentation results in partially enlarged detail. (**a**) Ground truth; (**b**) VCCS; (**c**) VCCS_KNN; (**d**) BPSS; (**e**) Proposed_REDEFINE; (**f**) Proposed_RANSAC; (**g**) proposed.
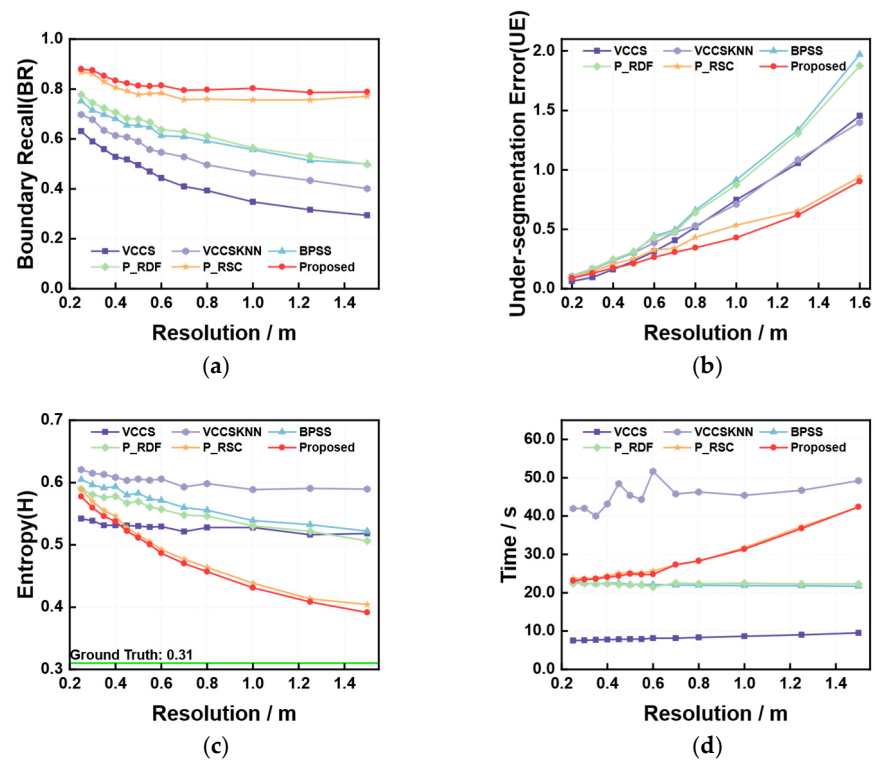
Figure 12. Evaluation metrics of algorithms on bildstein1 at different resolutions. (**a**) BR; (**b**) UE; (**c**) H; (**d**) time.
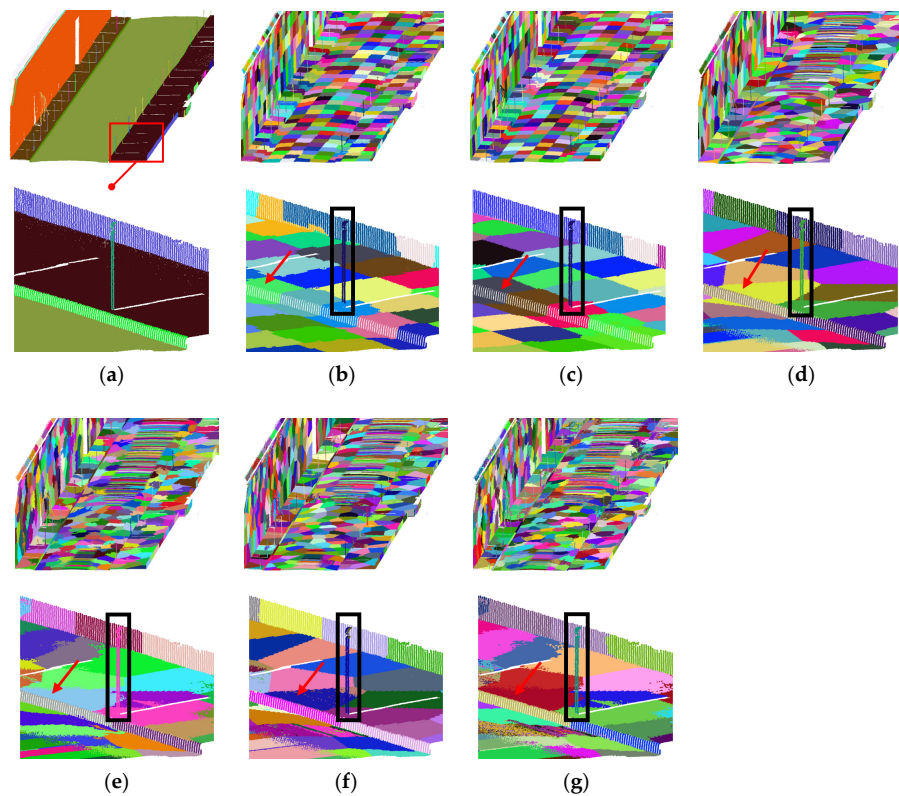

Figure 13. When R = 0.7, the cassette's segmentation results in partially enlarged detail. (**a**) Ground truth; (**b**) VCCS; (**c**) VCCS_KNN; (**d**) BPSS; (**e**) Proposed_REDEFINE; (**f**) Proposed_RANSAC; (**g**) proposed.
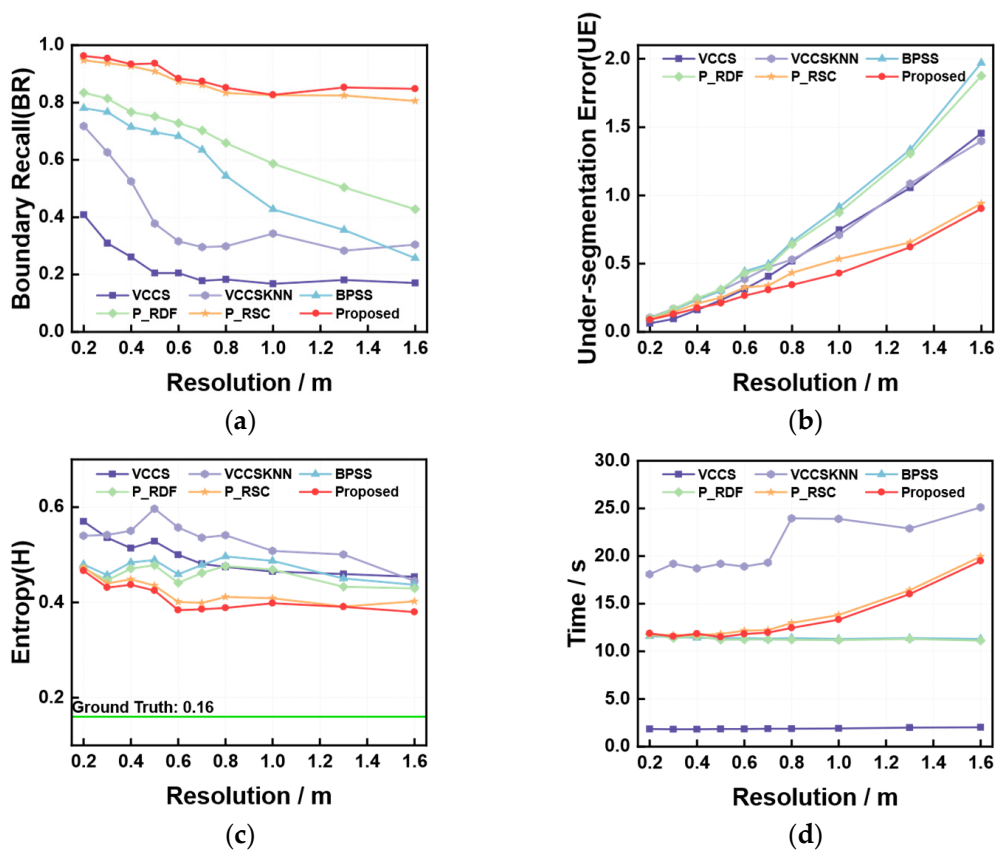
Figure 14. Evaluation metrics of algorithms on cassette at different resolutions. (**a**) BR; (**b**) UE; (**c**) H; (**d**) time.

4. Discussion

4.1. Discussion of Proposed Method

For the selected three models, the proposed algorithm shows better segmentation results compared to other algorithms in the qualitative demonstration. This is because our method introduces point-to-plane constraints to optimise the segmentation of non-coplanar supervoxel boundaries and uses the RANSAC method to re-segment supervoxels with multi-planar coupling. The quantitative analysis results correspond to the qualitative results reported, and the proposed approach achieves the best evaluation metrics at various resolutions. In terms of running time, whether RANSAC is used or not is the primary difference between the algorithm suggested in this research and BPSS. RANSAC has high time complexity, and we minimise our algorithm's runtime by using RANSAC on only a subset of supervoxels.

To optimise the segmentation outcome, it is necessary to use both optimisation operations. Specifically, boundary redefinition optimises the initial supervoxel generation stage, and then the RANSAC-based method re-segments the under-segmented supervoxels. While using only point-to-plane constraints can optimise the boundaries of adjacent non-coplanar supervoxels with minor changes in planar normal vectors, it cannot address the multi-planar coupling problem. However, on the other hand, using only RANSAC on supervoxels does not result in satisfactory boundary segmentation outcomes. This is because RANSAC's random sampling is insensitive to the boundary points of non-coplanar supervoxels when used directly for re-segmentation without boundary redefinition. Meanwhile, utilising boundary redefinition before RANSAC enhances the boundary segmentation of supervoxels below the roughness threshold. If the roughness threshold is drastically reduced or RANSAC segmentation is performed directly on all supervoxels, degeneracy

problems arise, and the running time increases significantly. As a result, both stages are complementary and contribute to improving segmentation outcomes.

Comprehensive experiments have demonstrated that the proposed method achieves good performance on high-quality TLS point clouds and low-quality MLS point clouds, making it suitable for point cloud processing tasks in urban environments. Additionally, our algorithm exhibits broad applicability and can be used in sparser ALS point clouds by simply adjusting the parameters.

The proposed method has been comprehensively validated through experiments, demonstrating exceptional performance on high-quality TLS point clouds and low-quality MLS point clouds. Furthermore, our algorithm exhibits a wide range of applicability, allowing for its utilisation in sparser ALS point clouds by adjusting parameters such as resolution.

4.2. Limitations

We also investigated two issues where the proposed method segments defectively. These two circumstances are illustrated in Figure 15.
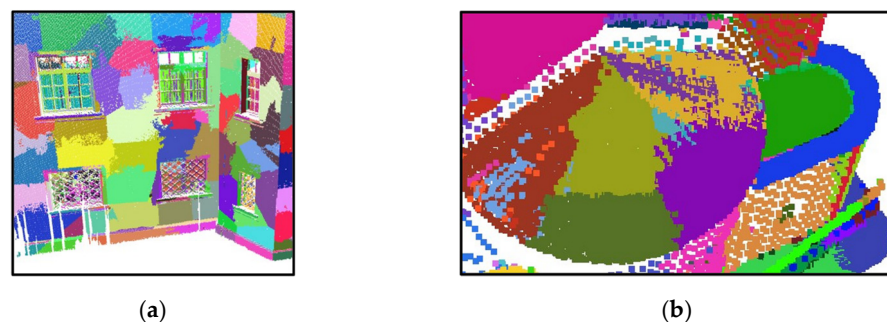

(**a**) (**b**)

Figure 15. Limitations of the proposed method: (**a**) un-compact supervoxels, and (**b**) segmentation of a curved façade into multiple small planes.

The first circumstance is the un-compact boundary points of co-planar supervoxels, as shown in Figure 15a. The redefinition of boundary points causes the limitation. However, we contend that this issue does not impede the high-level applications of super-voxels. Notably, the boundaries between hetero-plane supervoxels in Figure 15a are clearly defined, which can be advantageous for tasks such as extracting lines based on the Helmholtz theory [46,47] or other applications of supervoxels.

Due to the exclusive use of planar models, our proposed method tends to segment curved surfaces into multiple planar units, the second circumstance, as illustrated in Figure 15b. While this does not significantly affect the subsequent applications, it can be addressed by designing appropriate post-processing strategies that cater to specific use cases. To overcome this limitation, we plan to extend our approach by incorporating multiple geometric models, such as planes, spheres, and cylinders, to support the segmentation of supervoxels from point clouds.

5. Conclusions

In this research, we propose a large-scale point cloud supervoxel re-segmentation method based on point-to-plane constraints and RANSAC using roughness filtering for under-segmented supervoxels at various resolutions. Firstly, we introduce a constraint on distance judgement between boundary points and adjacent supervoxels to address boundary segmentation errors between non-coplanar supervoxels. Secondly, we use a roughness screening-based RANSAC method for re-segmentation to address multi-planar coupling problems caused by fixed-resolution mismatches in local regions. Finally, we conduct comparative experiments with classical supervoxel methods on different large-scale point clouds to validate our algorithm's efficacy. The experimental results show that our proposed algorithm can better segment boundaries between planes in large-scale

urban scenes at various resolutions without sacrificing runtime. Furthermore, the proposed method has the potential to handle point clouds from different sources.

In future work, we will further explore using more geometric models (such as cones, cylinders, spheres, etc.) to improve our algorithm while focusing on the segmentation accuracy and precision of small planes. Additionally, we will explore advanced applications of supervoxels, including 3D line extraction and point cloud segmentation.

Author Contributions: Conceptualization, B.L. and Y.Y.; methodology, B.L. and Y.Y.; software, B.L. and Y.Y.; validation, B.L. and Y.Y.; formal analysis, B.L., Y.Y. and B.H.; investigation, B.L.; resources, B.L. and Y.Y.; data curation, B.L. and Y.Y.; writing—original draft preparation, B.L. and Y.Y.; writing—review and editing, Y.Y., Y.Z. and B.H.; visualization, Y.Y.; supervision, Y.Z. and Q.Y.; project administration, Y.Z. and Q.Y.; funding acquisition, Y.Z., B.H. and Q.Y.; B.L. and Y.Y. contributed equally to this work and should be considered co-first authors. All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The two public datasets used in this paper are Semantic3D (http://semantic3d.net/, accessed on 10 March 2023) and IQmulus & TerraMobilita Contest (http://dataign.fr/benchmarks/UrbanAnalysis/#, accessed on 12 March 2023).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Li, X.; Li, Y.; Xie, X.; Xu, L. Lab-built terrestrial laser scanner self-calibration using mounting angle error correction. *Opt. Express* **2018**, *26*, 14444–14460. [CrossRef]
2. Yuan, Y.; Ge, Z.; Guo, X.; Zhang, Y.; Suo, T.; Yu, Q. Research on non-concentric spatial scanning measurement system. *Optik* **2022**, *267*, 169706. [CrossRef]
3. Lin, Y.B.; Wang, C.; Cheng, J.; Chen, B.L.; Jia, F.K.; Chen, Z.G.; Li, J. Line segment extraction for large scale unorganized point clouds. *ISPRS J. Photogramm. Remote Sens.* **2015**, *102*, 172–183. [CrossRef]
4. Li, Y.; Yong, B.; Wu, H.; An, R.; Xu, H. An improved top-hat filter with sloped brim for extracting ground points from airborne lidar point clouds. *Remote Sens.* **2014**, *6*, 12885–12908. [CrossRef]
5. Fahle, L.; Petruska, A.J.; Walton, G.; Brune, J.F.; Holley, E.A. Development and Testing of Octree-Based Intra-Voxel Statistical Inference to Enable Real-Time Geotechnical Monitoring of Large-Scale Underground Spaces with Mobile Laser Scanning Data. *Remote Sens.* **2023**, *15*, 1764. [CrossRef]
6. Liang, X.L.; Kankare, V.; Hyyppa, J.; Wang, Y.S.; Kukko, A.; Haggren, H.; Yu, X.W.; Kaartinen, H.; Jaakkola, A.; Guan, F.Y.; et al. Terrestrial laser scanning in forest inventories. *ISPRS J. Photogramm. Remote Sens.* **2016**, *115*, 63–77. [CrossRef]
7. Yuan, Y.T.; Ge, Z.D.; Lai, B.K.; Guo, X.; Zhang, Y.Q.; Liu, X.L.; Suo, T.; Yu, Q.F. Three dimensional deformation measurement method based on image guided point cloud registration. *Opt. Lasers Eng.* **2023**, *161*, 107399. [CrossRef]
8. Riveiro, B.; DeJong, M.J.; Conde, B. Automated processing of large point clouds for structural health monitoring of masonry arch bridges. *Autom. Constr.* **2016**, *72*, 258–268. [CrossRef]
9. Xia, S.B.; Chen, D.; Wang, R.S.; Li, J.; Zhang, X.C. Geometric Primitives in LiDAR Point Clouds: A Review. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2020**, *13*, 685–707. [CrossRef]
10. Xu, Y.S.; Tong, X.H.; Stilla, U. Voxel-based representation of 3D point clouds: Methods, applications, and its potential use in the construction industry. *Autom. Constr.* **2021**, *126*, 103675. [CrossRef]
11. Besl, P.J.; Jain, R.C. Segmentation through Variable-Order Surface Fitting. *IEEE Trans. Pattern Anal. Mach. Intell.* **1988**, *10*, 167–192. [CrossRef]
12. Nurunnabi, A.; Belton, D.; West, G. Robust segmentation in laser scanning 3D point cloud data. In Proceedings of the 2012 International Conference on Digital Image Computing Techniques and Applications (DICTA), Fremantle, WA, Australia, 3–5 December 2012; pp. 1–8.
13. Su, Z.H.; Gao, Z.J.; Zhou, G.Y.; Li, S.H.; Song, L.H.; Lu, X.K.; Kang, N. Building Plane Segmentation Based on Point Clouds. *Remote Sens.* **2022**, *14*, 95. [CrossRef]

14. Sampath, A.; Shan, J. Segmentation and Reconstruction of Polyhedral Building Roofs From Aerial Lidar Point Clouds. *IEEE Trans. Geosci. Remote Sens.* **2010**, *48*, 1554–1567. [CrossRef]

15. Wang, C.X.; Xiong, X.Q.; Yang, H.Q.; Liu, X.J.; Liu, L.; Sun, S.A. Application of Improved DBSCAN Clustering Method in Point Cloud Data Segmentation. In Proceedings of the 2nd International Conference on Big Data and Artificial Intelligence and Software Engineering (ICBASE), Zhuhai, China, 24–26 September 2021; pp. 140–144.

16. Duda, R.O.; Hart, P.E. Use of the Hough transformation to detect lines and curves in pictures. *Commun. ACM* **1972**, *15*, 11–15. [CrossRef]

17. Borrmann, D.; Elseberg, J.; Lingemann, K.; Nuchter, A. The 3D Hough Transform for plane detection in point clouds: A review and a new accumulator design. *3D Res.* **2011**, *2*, 3. [CrossRef]

18. Limberger, F.A.; Oliveira, M.M. Real-time detection of planar regions in unorganized point clouds. *Pattern Recognit.* **2015**, *48*, 2043–2053. [CrossRef]

19. Schnabel, R.; Wahl, R.; Klein, R. Efficient RANSAC for point-cloud shape detection. *Comput. Graph. Forum* **2007**, *26*, 214–226. [CrossRef]

20. Li, L.; Yang, F.; Zhu, H.H.; Li, D.L.; Li, Y.; Tang, L. An Improved RANSAC for 3D Point Cloud Plane Segmentation Based on Normal Distribution Transformation Cells. *Remote Sens.* **2017**, *9*, 433. [CrossRef]

21. Yang, L.N.; Li, Y.C.; Li, X.C.; Meng, Z.Q.; Luo, H.W. Efficient plane extraction using normal estimation and RANSAC from 3D point cloud. *Comput. Stand. Interfaces* **2022**, *82*, 103608. [CrossRef]

22. Chen, D.; Zhang, L.Q.; Mathiopoulos, P.T.; Huang, X.F. A Methodology for Automated Segmentation and Reconstruction of Urban 3-D Buildings from ALS Point Clouds. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2014**, *7*, 4199–4217. [CrossRef]

23. Lin, Y.B.; Wang, C.; Zhai, D.W.; Li, W.; Li, J. Toward better boundary preserved supervoxel segmentation for 3D point clouds. *ISPRS J. Photogramm. Remote Sens.* **2018**, *143*, 39–47. [CrossRef]

24. Achanta, R.; Shaji, A.; Smith, K.; Lucchi, A.; Fua, P.; Susstrunk, S. SLIC Superpixels Compared to State-of-the-Art Superpixel Methods. *IEEE Trans. Pattern Anal. Mach. Intell.* **2012**, *34*, 2274–2281. [CrossRef] [PubMed]

25. Xie, Y.X.; Tian, J.J.; Zhu, X.X. Linking Points With Labels in 3D: A Review of Point Cloud Semantic Segmentation. *IEEE Geosci. Remote Sens. Mag.* **2020**, *8*, 38–59. [CrossRef]

26. 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, Portland, OR, USA, 23–28 June 2013; pp. 2027–2034.

27. Song, S.; Lee, H.; Jo, S. Boundary-enhanced supervoxel segmentation for sparse outdoor LiDAR data. *Electron. Lett.* **2014**, *50*, 1917–1918. [CrossRef]

28. Puligandla, V.A.; Lončarić, S. A Supervoxel Segmentation Method With Adaptive Centroid Initialization for Point Clouds. *IEEE Access* **2022**, *10*, 98525–98534. [CrossRef]

29. Guo, Y.; Wang, H.; Hu, Q.; Liu, H.; Liu, L.; Bennamoun, M. Deep learning for 3D point clouds: A survey. *IEEE Trans. Pattern Anal. Mach. Intell.* **2020**, *43*, 4338–4364. [CrossRef] [PubMed]

30. Tiator, M.; Geiger, C.; Grimm, P. Point cloud segmentation with deep reinforcement learning. In *ECAI 2020*; IOS Press: Santiago de Compostela, Spain, 2020; pp. 2768–2775.

31. Wan, Z.; Hu, J.; Zhang, H.; Wang, Y. IAN: Instance-Augmented Net for 3D Instance Segmentation. *IEEE Robot. Autom. Lett.* **2023**, *8*, 4354–4361. [CrossRef]

32. Hui, L.; Yuan, J.; Cheng, M.; Xie, J.; Zhang, X.; Yang, J. Superpoint network for point cloud oversegmentation. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Montreal, QC, Canada, 11–17 October 2021; pp. 5510–5519.

33. Landrieu, L.; Boussaha, M. Point cloud oversegmentation with graph-structured deep metric learning. In Proceedings of the EEE/CVF Conference on Computer Vision and Pattern Recognition, Long Beach, CA, USA, 15–20 June 2019; pp. 7440–7449.

34. Ni, H.; Niu, X. SVLA: A compact supervoxel segmentation method based on local allocation. *ISPRS J. Photogramm. Remote Sens.* **2020**, *163*, 300–311. [CrossRef]

35. Xiao, Y.Y.; Chen, Z.G.; Lin, Z.T.; Cao, J.; Zhang, Y.J.; Lin, Y.B.; Wang, C. Merge-Swap Optimization Framework for Supervoxel Generation from Three-Dimensional Point Clouds. *Remote Sens.* **2020**, *12*, 473. [CrossRef]

36. Sha, Z.C.; Chen, Y.P.; Lin, Y.B.; Wang, C.; Marcato, J.; Li, J. A Supervoxel Approach to Road Boundary Enhancement from 3-D LiDAR Point Clouds. *IEEE Geosci. Remote Sens. Lett.* **2022**, *19*, 11. [CrossRef]

37. Su, F.; Liu, Y.; Nie, K.; Liu, Y.; Bi, J.; Zhang, R.; Zheng, G. Boundary-aware supervoxel segmentation for indoor 3D point clouds. *IEEE Access* **2023**, *11*, 73478–73489. [CrossRef]

38. Dong, Z.; Yang, B.; Hu, P.; Scherer, S. An efficient global energy optimization approach for robust 3D plane segmentation of point clouds. *ISPRS J. Photogramm. Remote Sens.* **2018**, *137*, 112–133. [CrossRef]

39. Rusu, R.B.; Cousins, S. 3D is here: Point cloud library (pcl). In Proceedings of the 2011 IEEE International Conference on Robotics and Automation, Shanghai, China, 9–13 May 2011; pp. 1–4.

40. Hoppe, H.; Derose, T.; Duchamp, T.; McDonald, J.; Stuetzle, W. Surface Reconstruction from Unorganized Points. In Proceedings of the 19th Annual Conference of the Assoc-for-Computing-Machinery: Computer Graphics and Interactive Techniques (Siggraph 92), Chicago, IL, USA, 26–31 July 1992; pp. 71–78.

41. Pukelsheim, F. The three sigma rule. *Am. Stat.* **1994**, *48*, 88–91.

42. Hackel, T.; Savinov, N.; Wegner, J.D.; Schindler, K.; Pollefeys, M. SEMANTIC3D. NET: A new large-scale point cloud classification benchmark. In Proceedings of the ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Hannover, Germany, 6–9 June 2017; pp. 91–98.
43. Vallet, B.; Brédif, M.; Serna, A.; Marcotegui, B.; Paparoditis, N. TerraMobilita/iQmulus urban point cloud analysis benchmark. *Comput. Graph.* **2015**, *49*, 126–133. [CrossRef]
44. Shannon, C.E. A mathematical theory of communication. *Bell Syst. Tech. J.* **1948**, *27*, 379–423. [CrossRef]
45. Weinmann, M.; Jutzi, B.; Hinz, S.; Mallet, C. Semantic point cloud interpretation based on optimal neighborhoods, relevant features and efficient classifiers. *ISPRS J. Photogramm. Remote Sens.* **2015**, *105*, 286–304. [CrossRef]
46. Lin, Y.; Wang, C.; Chen, B.; Zai, D.; Li, J. Facet segmentation-based line segment extraction for large-scale point clouds. *IEEE Trans. Geosci. Remote Sens.* **2017**, *55*, 4839–4854. [CrossRef]
47. Liu, L.; Xiao, J.; Wang, Y.; Lu, Z.; Wang, Y. A novel rock-mass point cloud registration method based on feature line extraction and feature point matching. *IEEE Trans. Geosci. Remote Sens.* **2021**, *60*, 1–17. [CrossRef]