



Article

Topologically Consistent Reconstruction for Complex Indoor Structures from Point Clouds

Mengchi Ai ¹, Zhixin Li ² and Jie Shan ^{2,*}

¹ College of Surveying and Geo-informatics, Tongji University, Shanghai 200092, China; aimengchi@tongji.edu.cn

² Lyles School of Civil Engineering, Purdue University, West Lafayette, IN 47907, USA; li2887@purdue.edu

* Correspondence: jshan@purdue.edu

Abstract: Indoor structures are composed of ceilings, walls and floors that need to be modeled for a variety of applications. This paper proposes an approach to reconstructing models of indoor structures in complex environments. First, semantic pre-processing, including segmentation and occlusion construction, is applied to segment the input point clouds to generate semantic patches of structural primitives with uniform density. Then, a primitives extraction method with detected boundary is introduced to approximate both the mathematical surface and the boundary of the patches. Finally, a constraint-based model reconstruction is applied to achieve the final topologically consistent structural model. Under this framework, both the geometric and structural constraints are considered in a holistic manner to assure topologic regularity. Experiments were carried out with both synthetic and real-world datasets. The accuracy of the proposed method achieved an overall reconstruction quality of approximately 4.60 cm of root mean square error (RMSE) and 94.10% Intersection over Union (IoU) of the input point cloud. The development can be applied for structural reconstruction of various complex indoor environments.



Citation: Ai, M.; Li, Z.; Shan, J. Topologically Consistent Reconstruction for Complex Indoor Structures from Point Clouds. *Remote Sens.* **2021**, *13*, 3844. <https://doi.org/10.3390/rs13193844>

Academic Editors: Lingli Zhu, Jonathan Li and Sylvie Daniel

Received: 24 August 2021
Accepted: 15 September 2021
Published: 26 September 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: indoor mapping; 3D modeling; reconstruction; point clouds; topological consistency

1. Introduction

3D reconstruction of indoor structures in complex environments has become an essential task for virtual reality, Building Information Management (BIM), emergency management, and other geoinformation-based activities [1–5]. Since reconstructing structural, watertight models with regularity and topologic consistency requires specific expertise and manual effort, developing efficient, robust and automatic reconstruction algorithms have attracted significant attention in recent years [6–10].

Structural reconstruction in complex environments poses specific challenges due to their complicated spatial relationships, irregular shapes, as well as the existence of occlusion in raw data. Existing solutions can be divided into two general categories: decomposition-based methods and primitives fitting methods [11–13]. The decomposition-based methods are designed based on strong assumptions, such as Manhattan-world or 2.5D. That means all the walls can be separated by horizontal planes and the ceilings are continuous planes without multi-level ceilings or irregular structures [14]. Although these methods can achieve good performance in environments where missing structural regions or noise points exist, they strongly depend on the prior definition of indoor structures [15–17]. In contrast, primitives fitting methods automatically detect indoor primitives, fit the point cloud into structural models and refine the estimated model through constraints [18]. However, many existing approaches focus on the local fitting without satisfying the global regularity. In addition, the reconstructed models are mostly used for visualization without establishing the topology and connectivity relationship [19]. Therefore, establishing a holistic structural model that follows topological regularity in the various indoor environments remains a challenge [20].

Specifically, the challenge of structural reconstruction based on primitive fitting can be discussed in two aspects. First, irregular structures, such as uneven ceilings or non-parallel walls, are beyond regular assumptions and it is difficult to pre-define the topological relationship between each primitive. Second, the global topological regularity issue, such as parallelism and perpendicularity, and other regularity conditions are usually ignored among the detected primitives [21].

To address these challenges, this paper proposes a topologically-consistent structural reconstruction (TC-SM) method to obtain an entire structural model from point clouds. More precisely, we reconstruct indoor structures by considering both geometric modeling fitting and topologic regularity, which taking into account topological constraints and the estimation of optimal fitting model. The contributions of the work are as follows:

(1) A primitives estimation method with detected boundary is proposed, which allows an unorganized point cloud to be modeled with a combination of semantic information and shape constraints.

(2) A topological consistent reconstruction is proposed to guarantee that the final model is topologically consistent and watertight by identifying the topology-based constraints.

The remainder of this paper is organized as follows. Section 2 briefly reviews the related literature in indoor reconstruction. Section 3 presents the proposed method in detail. Section 4 describes the experiments and results, while Section 5 concludes our work with contributions and limitations.

2. Related Work

This section reviews the literature pertaining to primitives extraction and global model reconstruction in structural reconstruction from point clouds.

Primitives extraction methods divide the indoor environment into independent structural points from raw data and estimate geometric models. They include Principal Component Analysis (PCA), Hough Transform, region-growing, Random Sample Consensus (RANSAC), and least squares-based methods [22,23]. The polygonal-structured modeling methods mainly consist of a structure detection stage and an associated parameterized (or regularized) stage. For example, the framework of GlobFit estimates the original primitives via RANSAC and optimizes the data fitting based on intrinsic and global relationships [24]. However, GlobFit's performance is not accurate when there are many noise points and outliers. With a similar framework, Jenke et al. [25] proposed an algorithm to extract a basic geometric model and contour points as the boundary contours by using an optimization function, after which surface reconstruction is implemented by dividing it into sub-problems to balance the error. The main problem for many abovementioned primitive extraction algorithms is that they can only determine the mathematical function of a primitive surface but not its boundary. In this case, a broader inter-primitive relation is required when structural primitives are irregular in shape. Indeed, without considering shape regularity, it is difficult to guarantee the fitting performance of the estimated model due to the occlusions and structural missing areas [26].

Global model reconstruction methods focus on regularization, forcing the extracted primitives to be watertight and integrated [27,28]. Normally, the regularization or refinement algorithms rely on prior knowledge or assumptions, which force planar and axis-aligned orthogonal walls. With the predefined rules and regular-shape assumption, the space-aware approaches are suitable for most common indoor environments. For example, Oesau et al. [29] detected the horizontal structures and walls for space portioning. The final model then was detected by cell partitioning into space via graph-cut. Nan and Wonka [30] obtained watertight models by minimizing the energy function, which is designed to choose the inliers' surfaces. However, an optimization algorithm cannot guarantee the basic global relationship between the primitives and the accuracy of small structural spaces. Wang et al. [31] proposed a framework that uses Conditional Generative Adversarial Network (cGAN) to deal with the structural completeness using extracted primitives; however, the algorithm cannot be applied to structures with irregular boundaries.

In summary, although the reconstruction of indoor structures has been studied in the past few years, generating a topologically-consistent semantic model in a complex environment, such as uneven ceiling, non-orthogonal walls, or irregular shape, is still a challenge [32–34]. Most published methods are based on strong structural assumptions or pre-definitions, i.e., they are limited to regularly shaped indoor structures. Moreover, while the aforementioned methods are able to extract visualized reconstruction results, most of them ignore the global topological rules and only focus on local fitting [35,36]. Although some methods pay attention to optimizing the inlier surfaces and obtaining watertight models, only a few of them focus on topologically-consistent reconstruction in complex indoor environments [37,38].

3. Methodology

The proposed topologically-consistent structural modeling (TC-SM) method is mainly comprised of three steps, as shown in Figure 1. Firstly, semantic pre-processing is performed on the raw point clouds to generate semantic patches with uniform density. Specifically, point clouds are semantically segmented into structural (walls, ceilings, and floors) points and other (non-structural) points using PointNet ++ [39]. After that, structural occlusion construction is used to reduce the holes on the structural patches. Then, the primitive extraction step estimates both the mathematical surface of the structural patches and their boundaries. Finally, the holistic reconstruction step considers the topological relationship among all the extracted primitives to simultaneously determine the structural model. This section will focus on the second step and the third step, while the semantic pre-processing will be described in Section 4.2 as part of the workflow.

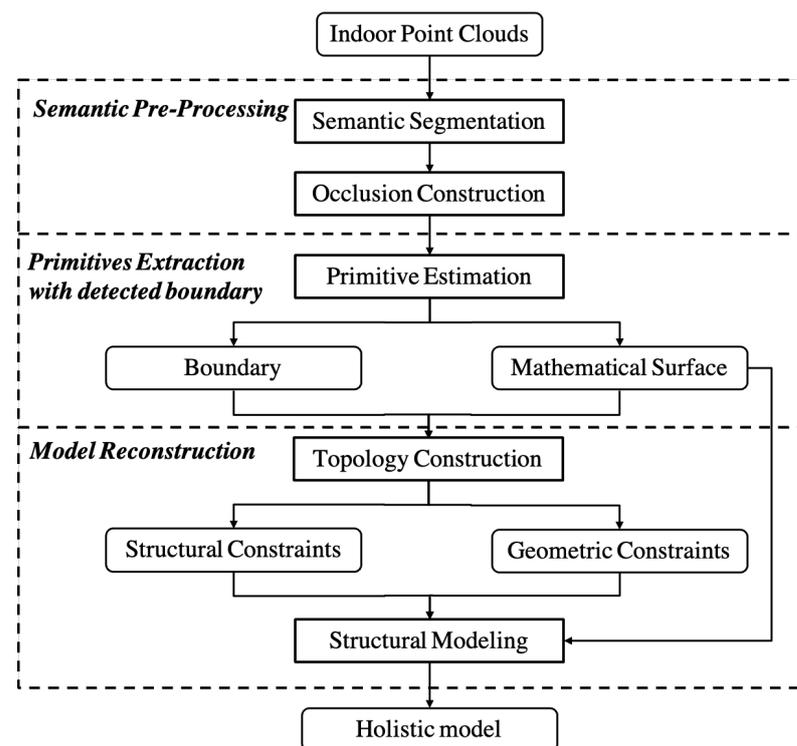


Figure 1. Flowchart of the TC-SM method.

3.1. Extraction of Primitives with Detected Boundary

The objective of this step is to find points that are on the same planar primitives and simultaneously determine the boundary of the primitives based on shape constraints [40]. This is accomplished through the following optimization framework,

$$\mathcal{H}(p, \varphi) = \mathcal{H}_d(p) + w_g \mathcal{H}_g(p) + w_\varphi \mathcal{L}(\varphi) \quad (1)$$

In this formulation, $\mathcal{H}_d(p)$, $\mathcal{H}_g(p)$, and $\mathcal{L}(\varphi)$ respectively represent the data term, the geometric term, and the shape constraints. w_g and w_φ are the corresponding weights, and φ is the number of boundary lines. p is the parameter of the mathematical surface model [41]. Equation (1) is used to determine the mathematical surface and the number of boundary lines by minimizing the energy function.

For the data term, $\mathcal{H}_d(p)$ can be expressed as Equation (2),

$$\mathcal{H}_d(p) = \sum_{s \in S} \|s - p\|^2 \quad (2)$$

where $\|s - p\|$ is the error measure between structural point $s \in S$ and estimated planar model P . Here, the squared Euclidean distance is applied as the data term.

The geometric term can be defined as,

$$\mathcal{H}_g(p) = \lambda \cdot \sum_{n,m \in \mathcal{N}} w_{mn} \cdot \delta(L_n \neq L_m), w_{mn} = \exp(-\|n, m\|^2) \quad (3)$$

where $\delta()$ is the Potts model and equals one (1) when point m and n are not belonging to the same model, otherwise equals zero (0). \mathcal{N} is the point set of neighborhoods (e.g., from the same mathematic surface or the same triangulation). Weight w_{mn} is designed as a penalty term for the pair of closer points, representing the geometric consistency. The value is higher when point m and n belong to the same model (mathematical surface or the same boundary). Here, we first use the alpha-shape method to detect the boundary points. Then, point selection is implemented based on the extracted boundary [42].

For $\mathcal{L}(\varphi)$, it is designed as the shape term by,

$$\mathcal{L}(\varphi) = |\varphi| \quad (4)$$

where φ represents the number of boundary lines. The overall objective energy function is composed of Equations (2)–(4). It is used to simultaneously extract the geometric primitives and their boundaries.

To extract each primitive, we first initialize the φ with a certain number of boundary lines ($\varphi = 10$ in this study). The setting of φ is experimental but with a maximum of 10 for each structural model. The boundary points are assigned to the closest lines and are used to re-estimate the line parameters using each group of inliers. To improve the estimation accuracy, the boundary lines with different normal vectors are considered as incorrectly estimated boundaries and deleted.

The process of primitive extraction is to minimize Equation (1) with φ and p . Starting from the initial φ , Equation (1) is minimized to estimate each p . Then, the objective function can be minimized with a sequential estimation [43]. This way we reduce the number of φ if the boundary line is not worth keeping. A worked example is shown in Figure 2 where the mathematical primitives were found with their regularized boundaries determined at the same time.

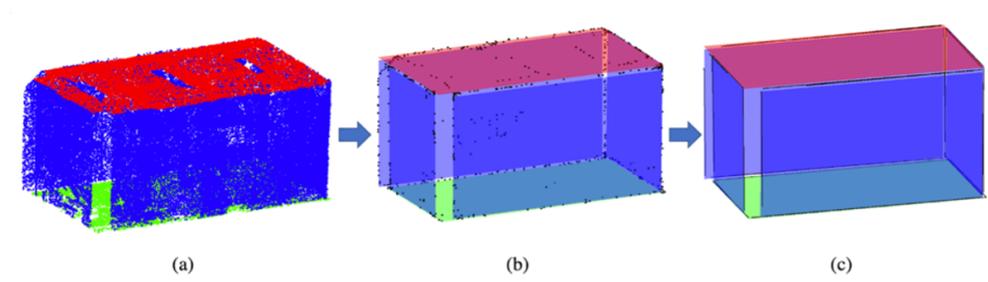


Figure 2. Example of primitive extraction with boundary: (a) semantic patches with uniform density (red: ceiling points, blue: wall points, and green: floor points); (b) geometric model and the points belonging to boundary (black points); and (c) black lines denote the boundary lines for each planar primitive.

3.2. Topology between Initial Primitives

Once the primitives are extracted along with their boundaries, the next step is to determine their topological relations. To facilitate the discussion, the topologic relations are classified into two groups: (1) the structural relations C_s are mostly relative ones such as parallelism and perpendicularity; and (2) the geometric relations C_g refer to those that directly help to determine the location of the vertices or the boundaries of the structures. Figure 3 illustrates the topological relations in a typical indoor structure.

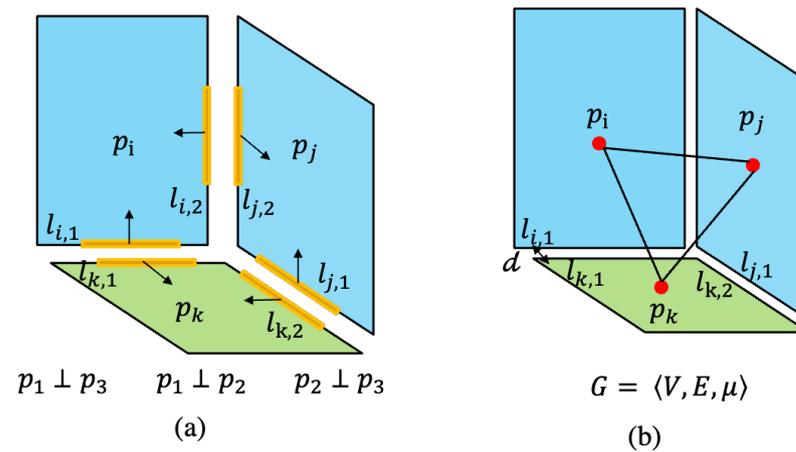


Figure 3. Two topologic relations in an indoor structure: (a) structural relations determined with the boundary lines and mathematical primitives and (b) geometric relations between estimated primitives.

These topologic relations must be inferred from the detected primitives derived in the previous step. To do so, the extended topologic graph $G = \langle V, E, \mu \rangle$ is established. As shown in Figure 3b, V represents the vertices of the primitives, E is the adjacent relationship between nodes, $E \subset \langle V, V \rangle$. The attribute set of the edges is represented by μ , where $\mu = \langle C_s, C_g \rangle$, including the structural relations and geometric relations.

For structural relations C_s , the constraint is treated as an adaptive constraint, including horizontal, parallel, or vertical relationships. The relationship is determined using Equation (5).

$$C_s = \delta(\Delta_d < T_d, \Delta_{\vec{n}} < T_\theta) \tag{5}$$

where $\delta()$ is the Potts model, with 1 for the two conditions being satisfied. Δ_d and $\Delta_{\vec{n}}$ respectively represent the distance and the degree of tilt between two primitives.

The thresholds T_d and T_θ are used to determine whether the structural constraints should be applied as the constraints in the next step. For example, the horizontal relationship can be determined when the Δ_d and $\Delta_{\vec{n}}$ are both less than the thresholds, as shown in Figure 3a.

Different from structural constraints, geometric constraint C_g is strictly enforced and designed to force a watertight polyhedron. It means that the adjacent surfaces intersect with one single corner, which will satisfy Equation (6),

$$C_g : (a_i, b_i, c_i, d_i)(X, Y, Z, 1)^T = 0 \tag{6}$$

where (X, Y, Z) is introduced as the corner point. (a_i, b_i, c_i, d_i) is the mathematical function of the structural surface. Equally important, the direction of the plumb lines is also considered as the C_g and can be represented as: $C_g : \vec{n} = (0, 0, 1, 0)$, which is in the

homogeneous coordinate. For each structural surface (a_i, b_i, c_i, d_i) , by introducing the intersected points, each primitive with C_g can be represented as Equation (7),

$$\begin{bmatrix} a_i \\ b_i \\ c_i \\ d_i \\ X \\ Y \\ Z \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} a_i \\ b_i \\ c_i \\ d_i \\ X \\ Y \\ Z \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}^T \begin{bmatrix} a_i \\ b_i \\ c_i \\ d_i \\ X \\ Y \\ Z \end{bmatrix} = 0 \quad (7)$$

Figure 4 illustrates an example for the topologic relationship based on graph and the representation of constraints. Semantic primitives with a basic adjacency graph and structural relations are shown in Figure 4a (left). The positions of the nodes (red dots) are set as the centers of the initial primitives; and the edges (black lines) indicate the adjacent relations. As for the topological relationship, we take the p_1 as an example. The corresponding relations are shown in Figure 4a (right). The structural relationships including vertical and parallel are represented respectively as blue and green lines. Then, the mathematical matrix is represented with +1 and -1, which represent the parallel and vertical relationships, respectively. Figure 4 only shows the part of the matrix that is related to p_1 , while others can be determined in a similar way

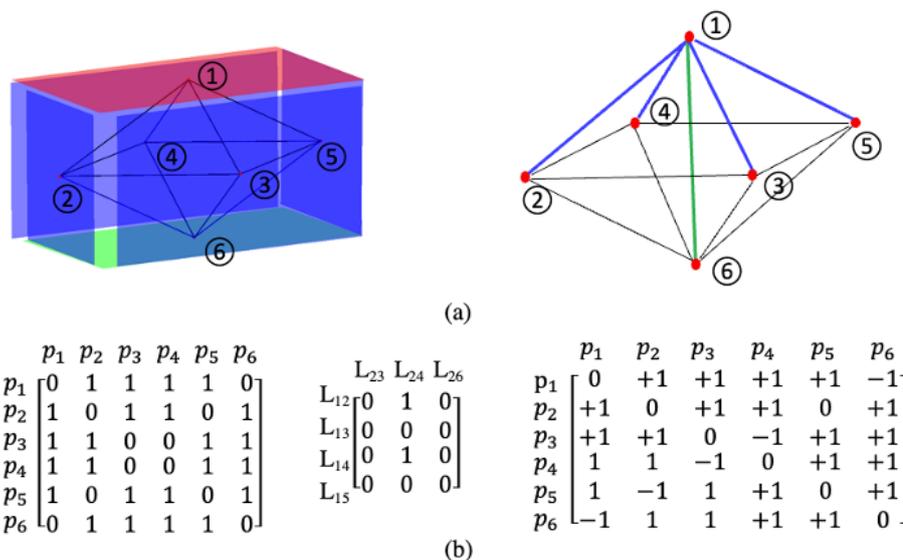


Figure 4. Set of extracted primitives (with boundary) and the corresponding relations: (a) graph-based representation and (b) matrix representation. p_i stands for plane i , L_{ij} for the intersection of p_i and p_j . The value in the matrix denotes the relationship, with +1 and -1 representing the parallel and vertical relationships. 0 is the plane itself.

3.3. Model Reconstruction under Topologic Constraints

After topologic relations C among the primitives are found, they will be used as constraints to fit the raw point clouds to a structural model. Theoretically, the fitting objective function $J(p_i)$ that meets constraint $C_{g,s}$ should be minimized [20],

$$Q = J(p_i) + \sum_{k=1}^N \zeta_k (C_{g,s}(p_i)^2) \quad (8)$$

where

$$J(p_i) = J_1 + J_2 + \dots + J_n \quad (9)$$

where J_i represents the fitting error (e.g., the squared sum of the residuals) that must be minimized by structural model associated with the raw data and topologic relations. For each model, it can be represented by the raw point vector and normal vector p of the related surface, as $h_{ij}^T p_i = 0$, where for each raw point $h_j = (x_j, y_j, z_j, 1)$. Then, the fitting error is treated as the distance error between the point and the estimated surface. For each surface, the objective function can be written as Equation (10), where M_i is the number of points belonging to surface i .

$$J_i = \sum_{j=0}^{M_i} \left(h_{ij}^T p_i \right)^2 \quad (10)$$

Under the above formulation, the constraints function and error term are included in the same objective function. The next step is to minimize the objective function subject to the topologic constraints. Combining the determined constraints, the problem can be solved by considering all the constraints in Equation (8), with N as the number of the structural constraints [44].

The topologic constraints discussed in Section 3.2 are used as the penalty function. As shown in Equation (11), the mathematical representation of the structural constraints is expressed by using a function with the cross-product terms. The solution can be performed using a constrained least-squares fitting method [45], where the optimized structural models should be subject to the topological constraints. The Jacobi matrix (the first-order derivatives) and Hessian matrix (the second-order derivatives) are calculated by using A , B and p . Table 1 lists the mathematical representation of the structural constraints and the corresponding parameter matrices [46,47].

$$C_{g,s}(p) = p^T A p + B^T p + C \quad (11)$$

where A and B are the square matrix and vector, C is scalar.

Table 1. Structural constraints and representation.

Constraints	Mathematic Representation	Parameter Matrices
Geometric Constraints		
Ceiling-Wall (Intersection)	$p_i \cdot p_n \gg 0$	$A, B, C \in \mathbb{R}$
Floor-Wall (Orthogonal)	$p_i \cdot p_n = 0$	$B = 0, C = 0$
Structural Constraints		
Horizontal	$p_i \cdot p_n = 1$	$A = 0, C = -1$
Vertical	$p_i \cdot p_n = 0$	$A = 0, C = 0$
Parallel	$p_i \times p_n = 0$	$A = 0, \text{ or } B = 0, C = 0$

4. Reconstruction Results

4.1. Datasets

Four public point cloud datasets were used for experiments. The first two datasets were the "Area1" and "Area6" of the 2D-3D-S dataset [48], containing 44 rooms and 52 rooms with trivial shapes and uneven heights. They were used to evaluate the performance under large data occlusion and on a large scale. The third dataset was selected from Synth 3 with slanted walls to test the capability of the approach to deal with the irregular structure. The fourth dataset was selected from Cottages with slanted ceilings. Synth 3 and Cottage were published by the Visualization and Multimedia Lab at the University of Zurich [49]. Table 2 shows the number of points, scans, and rooms as well as the area and the complexity for each dataset.

Table 2. Specifications of the test datasets.

Name	Area 1/2D-3D-S	Area 6/2D-3D-S	Synth 3	Cottage
# Scan Positions	-	-	25	5
Accuracy	High	High	Median	High
Complexity	Noisy	Noisy	Slant Walls	Slant Ceilings
#Room	44	52	9	3
#Points (Million)	44	42	69.4	4.5
Area [m ²]	965	935	260	32

4.2. Workflow

4.2.1. Semantic Segmentation

Before structural reconstruction, semantic patches were segmented from the raw point clouds and considered as unit for reconstruction [50]. To segment the structural patches and furniture patches, a DNN-based network was used to automatically segment the input point clouds into different patches with pre-defined labels. To achieve this goal, we integrated a semantic segmentation model based on PointNet++ into our pipeline [39]. The whole process had two stages. The first stage was to segment the input point clouds into structure points and furniture points, while the second stage improved the accuracy of segmentation. A geometric filter was used between the two stages to filter the segmented points into the different elements; and the output was a semantic label for each point. The results of the quantitative evaluation of each dataset are shown in Table 3.

Table 3. Accuracy of semantic segmentation.

Class/Dataset	Area 1/2D-3D-S	Area 6/2D-3D-S	Synth 3	Cottage
Wall (%)	75.46	68.34	81.25	88.18
Floor (%)	88.96	80.89	95.19	96.72
Ceiling (%)	91.74	87.05	94.44	93.54
Overall accuracy				
OA (%)	80.07	85.69	81.13	84.01
mAcc (%)	75.25	82.70	76.65	75.08

During the training process, some parameters were set, of which the most important were the bench size and sampling size because they affected the accuracy and density of the labeled results. Based on our experiments, the bench size and sampling size were set as 16 and 4096, respectively, larger than the default values to improve the density of semantic patches. The other parameter settings of epochs, learning rate, momentum, and radius were 200, 0.01, 0.9, and 2, respectively. All the training experiments were implemented on a single GPU with Tensorflow 1.70.

4.2.2. Structural Occlusion Construction

After the semantic labeling process, two major difficulties remained: point occlusion and uneven point density. To address these difficulties, structural occlusion construction was implemented in three steps. First, a pseudo point was introduced as the scanner's location in the indoor environment. Next, the initial structural model was estimated using the RANSAC algorithm and existing structural points [23]. Finally, to fill in the structural holes, the non-structural patches were projected onto the missing area in the structure, which was similar to the operation of the laser scanner. The projection direction was determined by the pseudo scanner and points in each non-structural patch. A real example is demonstrated in Figure 5. The raw point clouds with non-structural points and occlusion (red frame) are shown in Figure 5a. Structural patches with uneven density and missing

area were obtained after semantic segmentation, as shown in Figure 5b. After structural occlusion construction, the holes were filled in, as shown in Figure 5c.

Even after the structural occlusion construction process, some holes remained unfilled due to the imperfect pseudo scanner position. A further manual process was not required since the subsequent model reconstruction procedure could fill these holes.

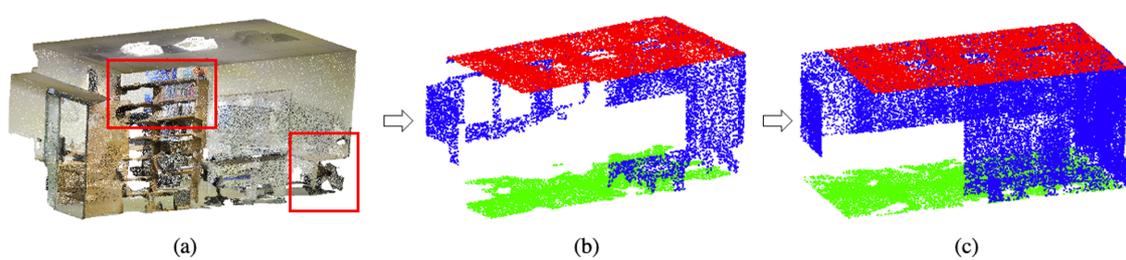


Figure 5. A detailed example of structural occlusion construction: (a) original point clouds of a complex indoor environment, where the red frames indicate the occlusion caused by furniture and limited field of view of the scanner; (b) the semantic patches with a missing region; and (c) semantic patches with uniform density.

4.2.3. Parameter Settings

This section discusses parameter selection for structural reconstruction. In general, the parameters of primitive estimation and refinement are related to prior knowledge and experimental study. Since the datasets in our experiments were from two typical sources, the values of the parameters were manually changed according to the input data.

For the process of primitive extraction with boundary, the parameter of alpha shape was important for the selection of boundary points. If α was too small, detailed boundary lines would be ignored, whereas a larger α could multiply the effect of outliers. The tests showed α between 0.5–0.8 could detect enough boundary lines with sufficient details. In addition, the initial number of boundary lines was set as $\phi = 10$. As for the energy function, the two weights were set to 0.4 and 0.6 to balance the effects between the geometric term and data term. As for the process of the final reconstruction, T_d and T_θ controlled the structural constraints. We set T_θ and T_d as 15° and 10 cm, respectively. That is to say, the structural constraints were introduced when the angle and distance were less than the thresholds.

4.3. Model Reconstruction Results

We evaluated the model reconstruction results from the point clouds with different complexity and variable noise levels. Figure 6 shows the input point cloud, surface model, and boundary lines for each dataset. In particular, Figure 6a shows the reconstructed results for Area 1/2D-3D-S, which was composed of several types of rooms (conference room, offices, and hallway). Figure 6b shows the reconstructed results for Area 6/2D-3D-S. However, rooms (hallway 1, hallway 2, open space 1, and lounge 1) cannot be reconstructed due to the large area of missing occlusion. The reconstructed results for Synth 3 with high accuracy are shown in Figure 6c. In Figure 6d, Cottage was a suitable dataset to evaluate the proposed structural constraints since it was impossible to pre-define the global relationship and regularities contained in the irregular shape of its indoor structure.

In general, the TC-SM algorithm produced structural modelings with consistent topology, guaranteeing the watertight polyhedral model under the scenario of complex indoor environments with point occlusion, irregular shapes, and noise. With its constraint-based reconstruction approach, the TC-SM method was able to detect the various shapes of each target room without any other supporting information. It is worth noting that all the reconstructed models followed the basic global structure relationship, such as the parallel relationship between the walls. For example, as shown in Figure 6c, the ceiling was not parallel to the floor, but the plumb lines should be perpendicular. Unlike other common building interiors, it was difficult to pre-define the prior assumption that is used to project

the ceiling into 2D plane. The TC-SM method was able to fit the raw data with the irregular shape and established the correct topology relationship.

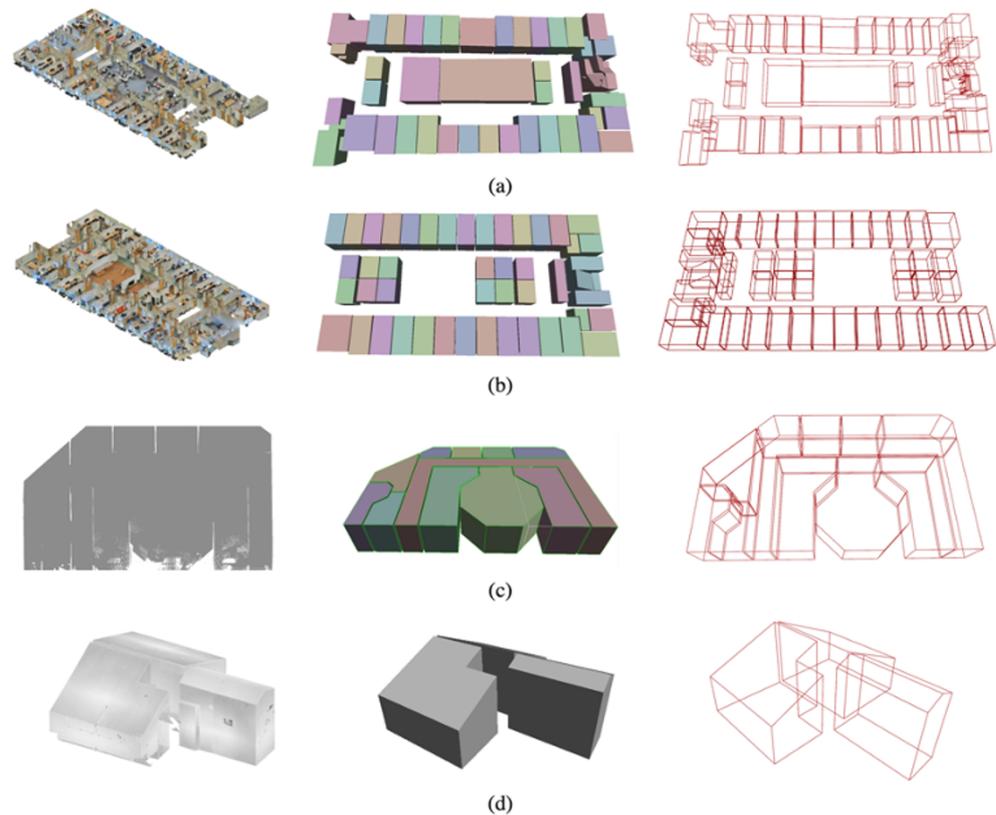


Figure 6. Reconstruction results for the four datasets (a–d) with the raw data (left), surface model (middle), and line framework (right).

To show more details, some specific rooms in Area 1/2D-3D-S and Area 6/2D-3D-S with irregular shapes and small domain structures were used to show the ability and performance of the TC-SM method. As shown in Figure 7a, the room contains several walls with irregular shapes. By estimating its regularized boundary, each primitive was fitted individually. Moreover, by determining the spatial relationship, the symmetry and integrity also were established. Similarly, Figure 7b,c showed TC-SM method performed very well for an indoor structure with a multi-level ceiling since different parts of the ceiling were enforced to be parallel. Moreover, the topologic rules made it possible to enforce the structural constraints to the model reconstruction, considering both the data fitting and the topologic regularity. Different from previous approaches based on similar strong assumptions, the TC-SM method correctly and robustly reconstructed a multi-level ceiling.

Similarly, specific rooms (office 8, hallway 3 and hallway 6) in Area 6 showed detailed results in Figure 8. As shown in Figure 8a, the room (office 8) contained uneven ceiling height with small structure. The TC-SM method could detect the lower ceiling and higher ceilings individually, and then establish the parallel relationship between them. In Figure 8b, indoor structure with irregular can be detected very well with some point occlusion. In Figure 8c, TC-SM method estimated the indoor structure almost correctly. However, the accuracy of this room is not pretty well since the stairs cannot be detected in the TC-SM method.

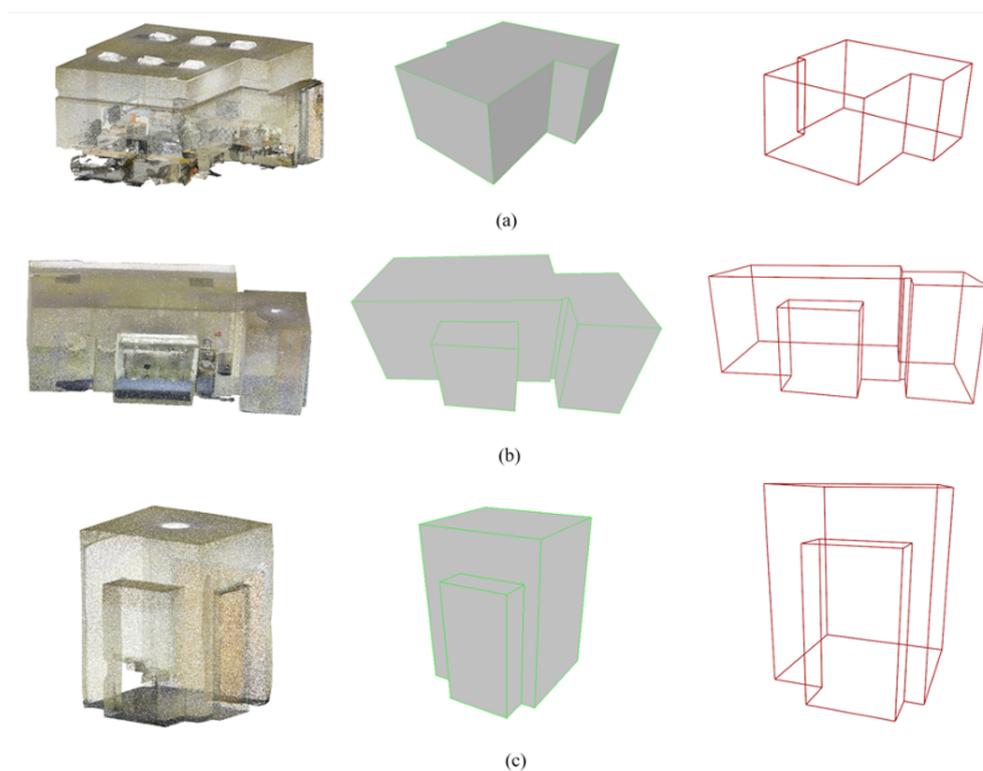


Figure 7. Zoom-in view of irregular shape indoor environment in Area 1/2D-3D-S: (a) Office 16, which is irregular in shape; (b) WC 1, which has an uneven ceiling height; and (c) the small structural in Hallway 1.

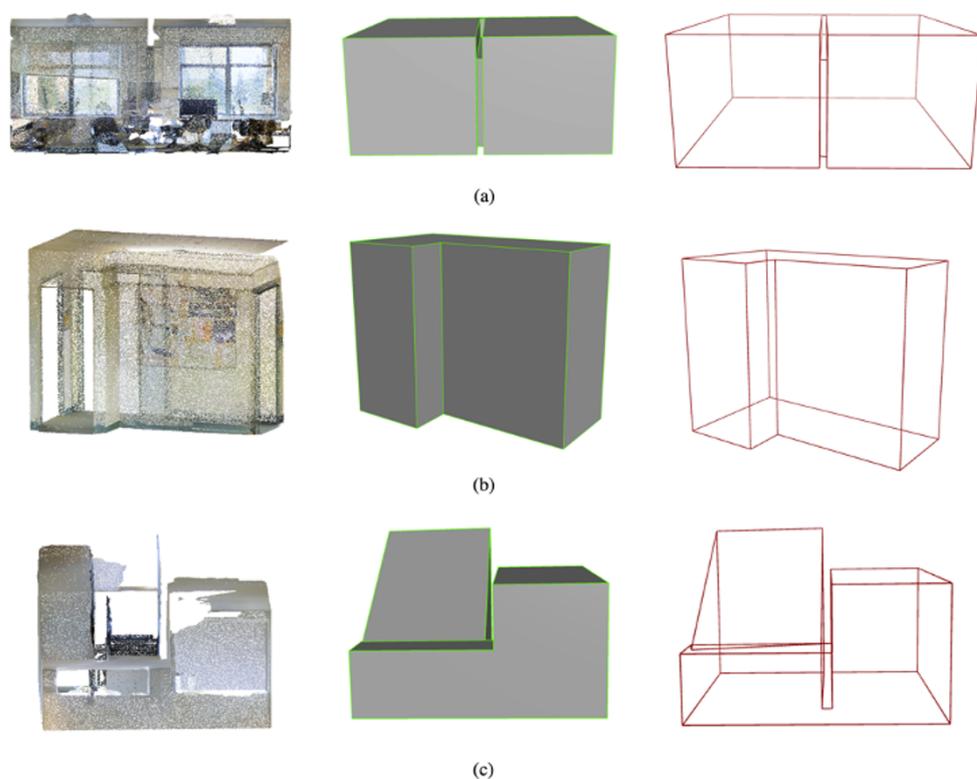


Figure 8. Zoom-in view of irregular shape indoor environment in Area 6/2D-3D-S: (a) Office 8, which has a small-structured uneven ceiling height; (b) hallway 3, which is in irregular shape; and (c) slanted walls in Hallway 6.

4.4. Quantitative Evaluation

Several quality metrics were used to evaluate the performance of the TC-SM method. Given the input points P with the total number n , and the estimated model M_i , the RMSE and IoU can be defined as Equations (12) and (13), where $\|p - M_i\|$ illustrates the Euclidean distance between the estimated model and the points. Moreover, the mean error, median error, and standard deviations were also calculated for each dataset.

$$RMSE = \sqrt{\frac{\sum_{i=1}^n \sum_{p \in P_i} \|p - M_i\|^2}{n}} \quad (12)$$

$$IoU = \frac{\text{Intersected Area}}{\text{Union Area}} \quad (13)$$

Table 4 shows the quantitative evaluation of the TC-SM method for all four datasets. In general, indoor structures with different shapes and uneven height under complex environments were successfully constructed, thereby demonstrating the robustness of the proposed method. With average medium errors of 0.38 cm and 0.30 cm, the TC-SM method obtained better results when the environment contained less furniture and the density was higher. The median distance error of Area 6/2D-3D-S was larger than the other datasets since it contained multiple small structures and multi-level stairs, leading to larger uncertainty.

Table 4. Quantitative evaluation of the reconstruction results.

Accuracy/Dataset	Area 1/2D-3D-S	Area 6/2D-3D-S	Synth 3	Cottage
RMSE (cm)	7.38	8.66	1.88	0.46
IoU (%)	92.38%	90.24%	97.64%	96.15%
Medium Error (cm)	0.69	4.69	0.38	0.30
Standard deviations (cm)	7.16	9.64	0.49	0.41

Figure 9 shows the frequency histogram of the mean error for all rooms for each dataset. For Area 1/2D-3D-S, the mean error for each room was quite stable since the office rooms were similar in shape and layout. For Hallway 7 and Hallway 8 in Area 1/2D-3D-S, there were missing areas located on the wall (see Figure 10), which the TC-SM method failed to fill and construct, resulting in large errors of 0.47 m and 0.13 m. For Hallway 8 in Figure 10a, a small structure also could not be detected since the corner was labeled as a door, resulting in a large error. Additionally, as shown in Figure 10b, TC-SM method did not produce good results for the slanted ceiling since it could not determine the connected boundary. The maximum error was around 1 m, which resulted in a large mean error in Area 1/2D-3D-S. In addition, the TC-SM method produced incomplete results when the environment contained stairs, which are not considered as structural patches in the TC-SM method (as shown in Figure 10b). Compared to Area 1/2D-3D-S and Area 6/2D-3D-S (as shown in Figure 9a,b), Synth 3 and Cottage obtained higher accuracy due to their simpler environments and yielded mean errors of 0.38 cm and 0.30 m, respectively (as shown in Figure 9c,d).

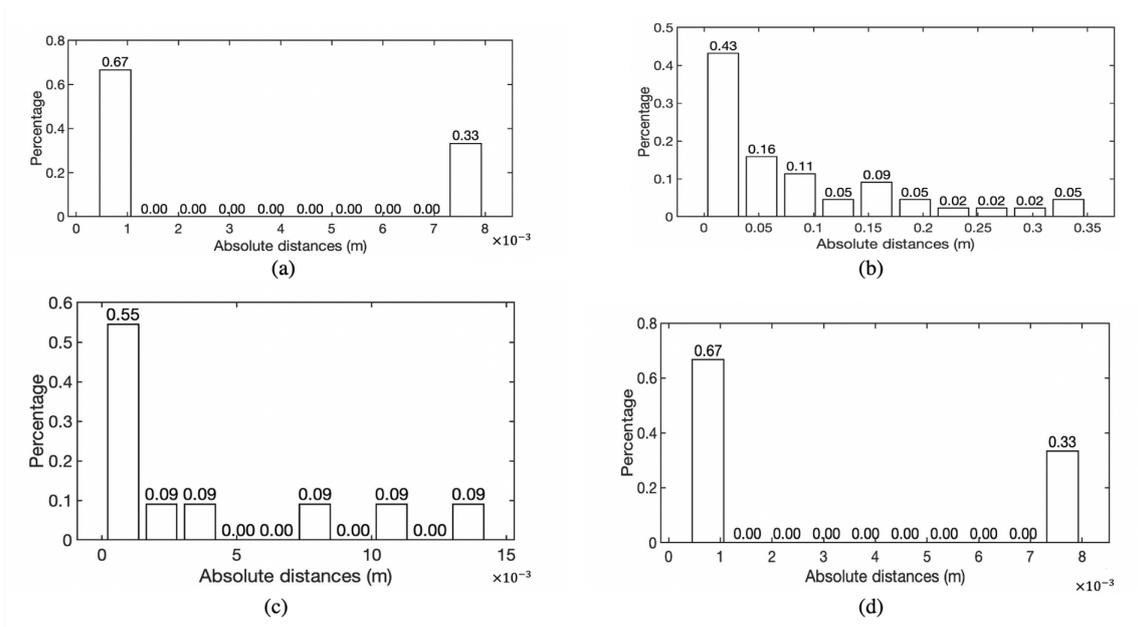


Figure 9. Comparative results of mean error in terms of individual rooms for (a) Area 1/2D-3D-S; (b) Area 6/2D-3D-S; (c) Synth 3; (d) Cottage.

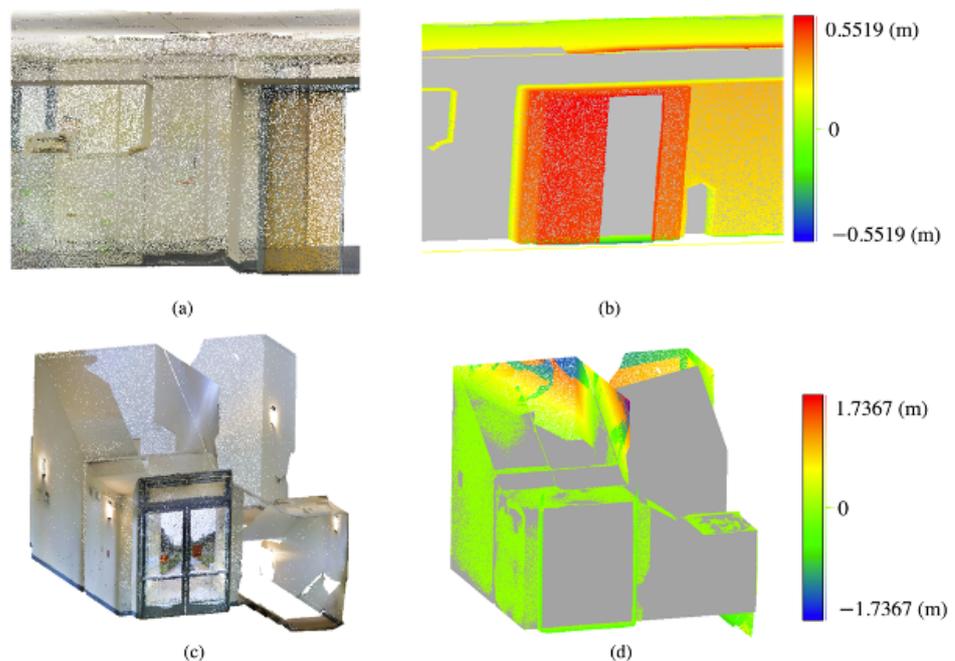


Figure 10. Large fitting errors (in meters) in hallways of Dataset 1: (a) raw data of Hallway 7; (b) color-coded maximum fitting error for the visible model; (c) raw data of Hallway 8; (d) color-coded maximum fitting error for the visible model.

Figure 11 shows the distribution of absolute distances for the structural points of three individual rooms (Office 16, WC 1, and Hallway 1, all in Area 1/2D-3D-S), which were considered as irregular-shape environments. The raw data and reconstructed results are shown in Figure 7. To evaluate the performance, the absolute distance between structural points and estimated model was calculated. The distribution of absolute distance shows that TC-SM method produced good results for irregular environments, leading to a standard deviation of 0.88 cm, 1.81 cm, and 6.92 cm, respectively. Additionally, the RMS (Root Mean Square) of each room was also calculated and slightly lower than 0.07 m. This means

that on average, in an irregular shaped environment, the structural model was reconstructed with a precision of about 0.07 m.

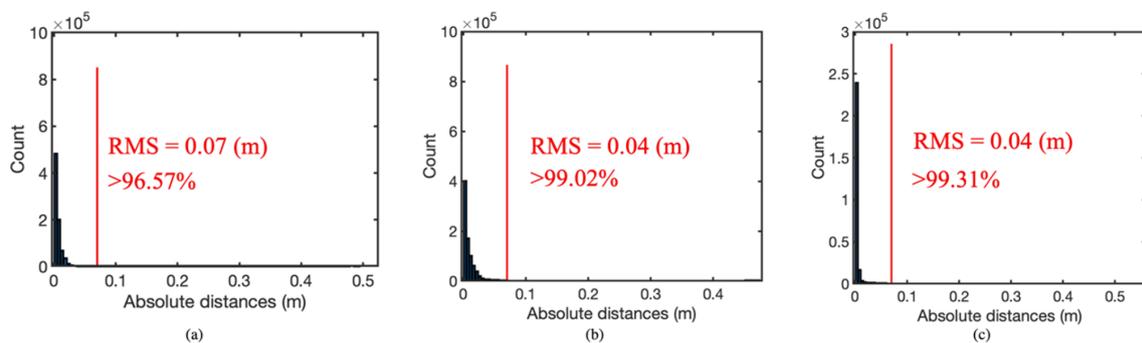


Figure 11. Comparative result of individual rooms with irregular shape; (a–c) show the mean error distribution of points for Office 16, WC 1, and Hallway 1 in Area 1/2D-3D-S.

Figure 12 shows the absolute distances of individual rooms in Area 6/2D-3D-S (Office 8, Hallway 3, and Hallway 6). The raw data and reconstructed results are shown in Figure 8. As shown in Figure 12a,b, the reconstructed results perform very well for the irregular-shape environments, with the RMS is 0.03 m and 0.04 m. However, in Figure 12c, the missing structural hole and stairs improved the larger uncertainty.

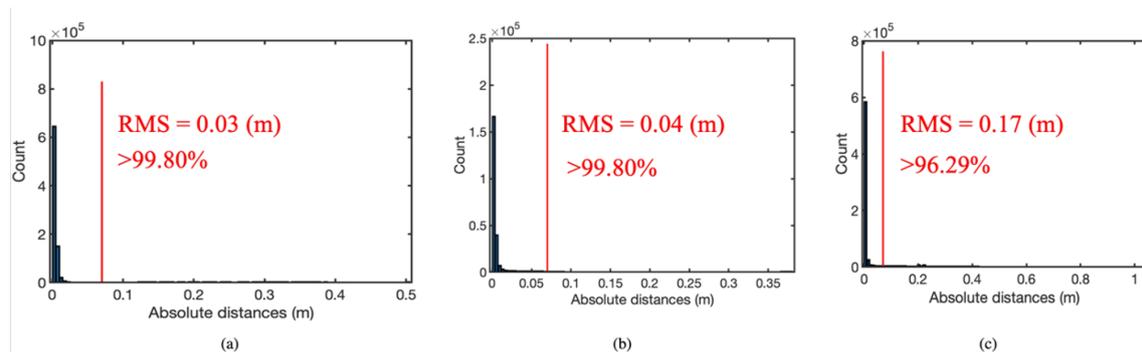


Figure 12. Comparative result of individual rooms with irregular shape; (a–c) show the mean error distribution of points for Office 8, Hallway 3, and Hallway 6 in Area 6/2D-3D-S.

4.5. Compared to Other Baseline

This section compares the proposed method with some state-of-the-art ones, including Polyfit and a parameter estimation algorithm [30,51]. Both published methods utilize a primitive-based algorithm that works well for surface reconstruction. An individual room (raw point clouds) with irregular shapes and small structures was used for comparison. The parameter estimation algorithm [51] generated stable reconstruction results by estimating the geometric parameters with the prior knowledge from a defined library of models. However, when the target environment was unknown or is irregular in shape, this algorithm could not obtain modeling results. Figure 13 represents the result for a conference room with an irregular shape from “Area 5” in 2D-3D-S [48]. The advantage of the parameter estimation method was that its algorithm was robust and highly accurate, with the mean error of 3 cm. However, small structures could not be detected without pre-definition in both mentioned state-of-art methods (as shown in Figure 13b,c).

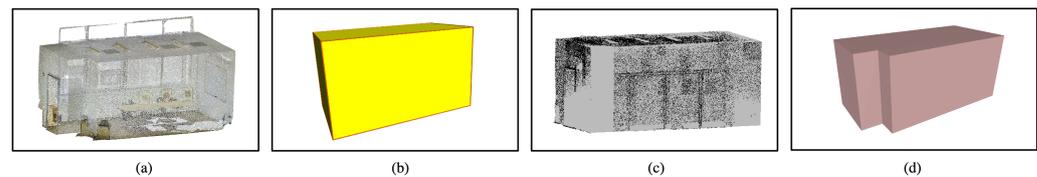


Figure 13. Reconstructed results generated from the proposed method and state-of-art ones. (a) raw input point clouds, (b) mesh surface generated from Polyfit [30], (c) the restricted results and inliers points generated from [51], (d) mesh surface generated from TC-SM method.

Polyfit [30] focuses on obtaining the entire, closed mesh of an area by intersecting the primitive extensions and selecting the inlier surface. The vertices and surfaces are estimated by minimizing the energy function using a process that is similar to the TC-SM method. However, considering the topologic properties, such as the topological relationship between structural surface, the differences were particularly notable between Polyfit and the TC-SM method. First, the TC-SM method utilized stronger regularization and topology of walls according to indoor scenarios, i.e., the generation of plumb lines, which can be explained by the fact that the TC-SM method used the fundamental constraints in the stage of closure optimization. Second, the performance of Polyfit also strongly depended on the selection of weights parameters, which must be manually selected for each different target environment. As shown in Figures 13b and 14b, the reconstructed model from Polyfit with default parameters produced wrong structure.

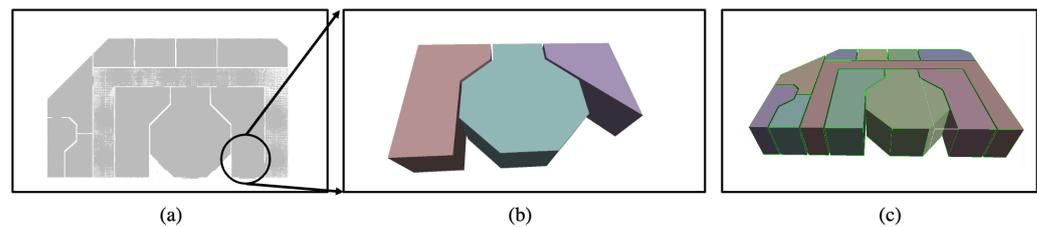


Figure 14. Raw point clouds [49] (a) and reconstructed results from Polyfit [30] (b) and the TC-SM (c). It shows that Polyfit failed to model right structural information (in middle circle), while the TC-SM method generates correct results, respectively.

5. Discussion and Conclusions

This paper proposes a novel 3D modeling method, namely topologically consistent structural modeling (TC-SM), based on consistent topology rules, to realize the structural modeling of complex real indoor scenes. Meanwhile, the constructed model can comprehensively represent the topology relationship, structural information, and semantic information. While improving the accuracy, the connection and topological relationship of the model are established. The main contribution of TC-SM is its ability to detect the topologic constraints of each structural element and thereby enforce the topologic regularity of the primitives in the process of model reconstruction. Our experimental results show that TC-SM can correctly depict the topologic relationships and closeness of a global structure without pre-defining its shape or other assumptions about its environment. Compared to other methods, experiments show that the TC-SM method has good robustness and higher accuracy in complex indoor environments.

The proposed method mainly consists of three main steps: considering the point cloud void, structural point cloud missing, and uneven point cloud density, semantic pre-processing, including semantic segmentation and structural occlusion construction is implemented to obtain structural patches with uniformed density. Then, a primitive extraction method with detected boundary is proposed to estimate both mathematical surface and boundary while considering shape constraints. Additionally, the possible topological relationship is estimated based on the initial primitives. Finally, model reconstruction clarifies the spatial topology regularity, transforms the topological rules into geometric

constraints and structural constraints that can be spatially constrained, to achieve the watertight and topological correctness of the structural model.

The innovation is to propose a watertight structural model with correct semantic information and topological information. Compared with the existing methods that focus on improving the accuracy of surface fitting, the TC-SM method can detect the real structure of indoor scenes very well and pay more attention to the topological regularity in real scenes to meet the requirements of indoor environments. Meanwhile, the TC-SM does not require spatial assumptions on the structure or shape of indoor environments to perform 3D modeling, which can better correspond to multiple types of indoor scenes. In addition, combined with the process of semantic processing, the TC-SM method can detect accurate semantic information at the same time, and better realize the rapid construction of the geographic model.

Our experiments also reveal certain limitations of the approach. In the segmentation step, small boundaries (structures) may be discarded due to the lack of structural points. Wrong reconstruction will occur when there are fewer points located on the structural boundary, since the lines are discarded and the points are considered as outliers of a mathematical surface. Another shortcoming is that the performance of reconstruction will be compromised by the mistakes in the labeling of semantic segmentation. For example, the stairs are not reconstructed but are treated as noisy data.

Author Contributions: Conceptualization, M.A. and J.S.; methodology, M.A. and J.S.; writing—original draft preparation, M.A., Z.L. and J.S.; writing—review and editing, M.A. and J.S.; visualization, M.A.; supervision J.S. All authors have read and agreed to the published version of the manuscript.

Funding: We would like to thank the China Scholarship Council for sponsoring M.A.

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

Abbreviations

The following abbreviations are used in this manuscript:

TC-SM	Topologically Consistent Structural Modeling
RMSE	Root mean square error
IoU	Intersection over Union
3D modeling	Three Dimensional Modeling

References

1. Biljecki, F.; Ledoux, H.; Stoter, J.; Vosselman, G. The variants of an LOD of a 3D building model and their influence on spatial analyses. *ISPRS J. Photogramm. Remote Sens.* **2016**, *116*, 42–54. [[CrossRef](#)]
2. Rajput, A.; Funk, E.; Börner, A.; Hellwich, O. A regularized volumetric fusion framework for large-scale 3D reconstruction. *ISPRS J. Photogramm. Remote Sens.* **2018**, *141*, 124–136. [[CrossRef](#)]
3. Tang, L.; Ying, S.; Li, L.; Biljecki, F.; Zhu, H.; Zhu, Y.; Yang, F.; Su, F. An application-driven LOD modeling paradigm for 3D building models. *ISPRS J. Photogramm. Remote Sens.* **2020**, *161*, 194–207. [[CrossRef](#)]
4. Thomson, C.P.H. From Point Cloud to Building Information Model: Capturing and Processing Survey Data towards Automation for High Quality 3D Models to Aid a BIM Process. Ph.D. Thesis, UCL (University College London), London, UK, 2016.
5. Zhou, B.; Ma, W.; Li, Q.; El-Sheimy, N.; Mao, Q.; Li, Y.; Gu, F.; Huang, L.; Zhu, J. Crowdsourcing-based indoor mapping using smartphones: A survey. *ISPRS J. Photogramm. Remote Sens.* **2021**, *177*, 131–146. [[CrossRef](#)]
6. Kang, Z.; Yang, J.; Yang, Z.; Cheng, S. A review of techniques for 3d reconstruction of indoor environments. *ISPRS Int. J. Geo-Inf.* **2020**, *9*, 330. [[CrossRef](#)]
7. Poux, F.; Neuville, R.; Nys, G.A.; Billen, R. 3D Point Cloud semantic modelling: Integrated framework for indoor spaces and furniture. *Remote Sens.* **2018**, *10*, 1412. [[CrossRef](#)]
8. Sequeira, V.; Ng, K.; Wolfart, E.; Gonçalves, J.G.; Hogg, D. Automated reconstruction of 3D models from real environments. *ISPRS J. Photogramm. Remote Sens.* **1999**, *54*, 1–22. [[CrossRef](#)]
9. Ma, Z.; Liu, S. A review of 3D reconstruction techniques in civil engineering and their applications. *Adv. Eng. Inform.* **2018**, *37*, 163–174. [[CrossRef](#)]
10. Fang, H.; Pan, C.; Huang, H. Structure-aware indoor scene reconstruction via two levels of abstraction. *ISPRS J. Photogramm. Remote Sens.* **2021**, *178*, 155–170. [[CrossRef](#)]

11. Fang, H.; Lafarge, F. Pyramid scene parsing network in 3D: Improving semantic segmentation of point clouds with multi-scale contextual information. *ISPRS J. Photogramm. Remote Sens.* **2019**, *154*, 246–258. [[CrossRef](#)]
12. Nikoohemat, S.; Diakit , A.; Zlatanova, S.; Vosselman, G. Indoor 3d modeling and flexible space subdivision from point clouds. *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.* **2019**, 285–292. [[CrossRef](#)]
13. Lim, G.; Doh, N. Automatic Reconstruction of Multi-Level Indoor Spaces from Point Cloud and Trajectory. *Sensors* **2021**, *21*, 3493. [[CrossRef](#)]
14. Wang, R.; Xie, L.; Chen, D. Modeling indoor spaces using decomposition and reconstruction of structural elements. *Photogramm. Eng. Remote Sens.* **2017**, *83*, 827–841. [[CrossRef](#)]
15. Horna, S.; Meneveaux, D.; Damiand, G.; Bertrand, Y. Consistency constraints and 3D building reconstruction. *Comput.-Aided Des.* **2009**, *41*, 13–27. [[CrossRef](#)]
16. Lim, G.; Oh, Y.; Kim, D.; Jun, C.; Kang, J.; Doh, N. Modeling of Architectural Components for Large-Scale Indoor Spaces From Point Cloud Measurements. *IEEE Robot. Autom. Lett.* **2020**, *5*, 3830–3837. [[CrossRef](#)]
17. Nikoohemat, S.; Diakit , A.A.; Zlatanova, S.; Vosselman, G. Indoor 3D reconstruction from point clouds for optimal routing in complex buildings to support disaster management. *Autom. Constr.* **2020**, *113*, 103109. [[CrossRef](#)]
18. Wu, H.; Yue, H.; Han, J.; Zhang, S. Indoor Map Boundary Correction Based on Normalized Total Least Squares of Condition Equation. *J. Surv. Eng.* **2021**, *147*, 04021015. [[CrossRef](#)]
19. Tran, H.; Khoshelham, K.; Kealy, A.; D az-Vilari o, L. Shape grammar approach to 3D modeling of indoor environments using point clouds. *J. Comput. Civ. Eng.* **2019**, *33*, 04018055. [[CrossRef](#)]
20. Xiong, B.; Jancosek, M.; Elberink, S.O.; Vosselman, G. Flexible building primitives for 3D building modeling. *ISPRS J. Photogramm. Remote Sens.* **2015**, *101*, 275–290. [[CrossRef](#)]
21. Ikehata, S.; Yang, H.; Furukawa, Y. Structured indoor modeling. In Proceedings of the IEEE International Conference on Computer Vision, Santiago, Chile, 11–18 December 2015; pp. 1323–1331.
22. Xia, S.; Chen, D.; Wang, R.; Li, J.; Zhang, X. Geometric primitives in LiDAR point clouds: A review. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2020**, *13*, 685–707. [[CrossRef](#)]
23. Yang, M.Y.; F rstner, W. Plane detection in point cloud data. In Proceedings of the 2nd International Conference on Machine Control & Guidance, Bonn, Germany, 9–11 March 2010; Volume 1, pp. 95–104.
24. Li, Y.; Wu, X.; Chrysathou, Y.; Sharf, A.; Cohen-Or, D.; Mitra, N.J. Globfit: Consistently fitting primitives by discovering global relations. *ACM Trans. Graph.* **2011**, *30*, 1–12.
25. Jenke, P.; Kr uckeberg, B.; Stra er, W. Surface Reconstruction from Fitted Shape Primitives. In Proceedings of the Vision, Modeling, and Visualization Conference 2008 (VMV 2008), Konstanz, Germany, 8–10 October 2008; pp. 31–40.
26. Becker, S.; Peter, M.; Fritsch, D. Grammar-supported 3d Indoor Reconstruction from Point Clouds for “as-built” BIM. *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.* **2015**, *2*, 17–24. [[CrossRef](#)]
27. Cui, Y.; Li, Q.; Yang, B.; Xiao, W.; Chen, C.; Dong, Z. Automatic 3-D reconstruction of indoor environment with mobile laser scanning point clouds. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2019**, *12*, 3117–3130. [[CrossRef](#)]
28. Shi, W.; Ahmed, W.; Li, N.; Fan, W.; Xiang, H.; Wang, M. Semantic geometric modelling of unstructured indoor point cloud. *ISPRS Int. J. Geo-Inf.* **2019**, *8*, 9. [[CrossRef](#)]
29. Oesau, S.; Lafarge, F.; Alliez, P. Indoor scene reconstruction using feature sensitive primitive extraction and graph-cut. *ISPRS J. Photogramm. Remote Sens.* **2014**, *90*, 68–82. [[CrossRef](#)]
30. Nan, L.; Wonka, P. Polyfit: Polygonal surface reconstruction from point clouds. In Proceedings of the IEEE International Conference on Computer Vision, Venice, Italy, 22–29 October 2017; pp. 2353–2361.
31. Wang, C.; Hou, S.; Wen, C.; Gong, Z.; Li, Q.; Sun, X.; Li, J. Semantic line framework-based indoor building modeling using backpacked laser scanning point cloud. *ISPRS J. Photogramm. Remote Sens.* **2018**, *143*, 150–166. [[CrossRef](#)]
32. Nurunnabi, A.; Belton, D.; West, G. Robust statistical approaches for local planar surface fitting in 3D laser scanning data. *ISPRS J. Photogramm. Remote Sens.* **2014**, *96*, 106–122. [[CrossRef](#)]
33. Pintore, G.; Mura, C.; Ganovelli, F.; Fuentes-Perez, L.; Pajarola, R.; Gobbetti, E. State-of-the-art in Automatic 3D Reconstruction of Structured Indoor Environments. *Comput. Graph. Forum* **2020**, *39*, 667–699. [[CrossRef](#)]
34. Yang, J.; Kang, Z.; Zeng, L.; Akwensi, P.H.; Sester, M. Semantics-guided reconstruction of indoor navigation elements from 3D colorized points. *ISPRS J. Photogramm. Remote Sens.* **2021**, *173*, 238–261. [[CrossRef](#)]
35. Krijnen, T.; Beetz, J. An IFC schema extension and binary serialization format to efficiently integrate point cloud data into building models. *Adv. Eng. Inform.* **2017**, *33*, 473–490. [[CrossRef](#)]
36. Macher, H.; Landes, T.; Grussenmeyer, P. From point clouds to building information models: 3D semi-automatic reconstruction of indoors of existing buildings. *Appl. Sci.* **2017**, *7*, 1030. [[CrossRef](#)]
37. Bassier, M.; Vergauwen, M. Topology reconstruction of BIM wall objects from point cloud data. *Remote Sens.* **2020**, *12*, 1800. [[CrossRef](#)]
38. Han, J.; Rong, M.; Jiang, H.; Liu, H.; Shen, S. Vectorized indoor surface reconstruction from 3D point cloud with multistep 2D optimization. *ISPRS J. Photogramm. Remote Sens.* **2021**, *177*, 57–74. [[CrossRef](#)]
39. Qi, C.R.; Yi, L.; Su, H.; Guibas, L.J. Pointnet++: Deep hierarchical feature learning on point sets in a metric space. *arXiv* **2017**, arXiv:1706.02413.

40. Gallo, O.; Manduchi, R.; Rafii, A. CC-RANSAC: Fitting planes in the presence of multiple surfaces in range data. *Pattern Recognit. Lett.* **2011**, *32*, 403–410. [[CrossRef](#)]
41. Zuliani, M.; Kenney, C.S.; Manjunath, B. The multiransac algorithm and its application to detect planar homographies. In Proceedings of the IEEE International Conference on Image Processing 2005, Genova, Italy, 14 September 2005; Volume 3, pp. 111–153.
42. Gardiner, J.D.; Behnsen, J.; Brassey, C.A. Alpha shapes: Determining 3D shape complexity across morphologically diverse structures. *BMC Evol. Biol.* **2018**, *18*, 184. [[CrossRef](#)]
43. Delong, A.; Osokin, A.; Isack, H.N.; Boykov, Y. Fast approximate energy minimization with label costs. *Int. J. Comput. Vis.* **2012**, *96*, 1–27. [[CrossRef](#)]
44. Fletcher, R. *Practical Methods of Optimization*; John Wiley & Sons: New York, NY, USA, 2013.
45. Jung, J.; Hong, S.; Yoon, S.; Kim, J.; Heo, J. Automated 3D wireframe modeling of indoor structures from point clouds using constrained least-squares adjustment for as-built BIM. *J. Comput. Civ. Eng.* **2016**, *30*, 04015074. [[CrossRef](#)]
46. Werghi, N.; Fisher, R.; Robertson, C.; Ashbrook, A. Object reconstruction by incorporating geometric constraints in reverse engineering. *Comput.-Aided Des.* **1999**, *31*, 363–399. [[CrossRef](#)]
47. Fiacco, A.V.; McCormick, G.P. *Nonlinear Programming: Sequential Unconstrained Minimization Techniques*; SIAM: Philadelphia, PA, USA, 1990.
48. Pham, Q.H.; Nguyen, T.; Hua, B.S.; Roig, G.; Yeung, S.K. Jsis3d: Joint semantic-instance segmentation of 3d point clouds with multi-task pointwise networks and multi-value conditional random fields. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Long Beach, CA, USA, 15–20 June 2019; pp. 8827–8836.
49. Mura, C.; Mattausch, O.; Villanueva, A.J.; Gobbetti, E.; Pajarola, R. Automatic room detection and reconstruction in cluttered indoor environments with complex room layouts. *Comput. Graph.* **2014**, *44*, 20–32. [[CrossRef](#)]
50. Xie, Y.; Tian, 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](#)]
51. Li, Z.; Zhang, W.; Shan, J. Holistic Parametric Reconstruction of Building Models from Point Clouds. *ISPRS-Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2020**, *43*, 689–695. [[CrossRef](#)]