NORMAL CLASSIFICATION OF 3D OCCUPANCY GRIDS FOR VOXEL-BASED INDOOR RECONSTRUCTION FROM POINT CLOUDS

P. Hübner,[*] S. Wursthorn, M. Weinmann

Institute of Photogrammetry and Remote Sensing, Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany
(patrick.huebner, sven.wursthorn, martin.weinmann)@kit.edu

Commission IV, WG IV/5

KEY WORDS: Indoor Reconstruction, Voxel, Building Model, Normal Vector, Point Cloud, Triangle Mesh

ABSTRACT:

In this paper, we present an automated method for classification of binary voxel occupancy grids of discretized indoor mapping data such as point clouds or triangle meshes according to normal vector directions. Filled voxels get assigned normal class labels distinguishing between horizontal and vertical building structures. The horizontal building structures are further differentiated into those with normal directions pointing upwards or downwards with respect to the building interior. The derived normal grids can be deployed in the context of an existing voxel-based indoor reconstruction pipeline, which so far was only applicable to indoor mapping triangle meshes that already contain normal vectors consistently oriented with respect to the building interior. By means of quantitative evaluation against reference data, we demonstrate the performance of the proposed method and its applicability in the context of voxel-based indoor reconstruction from indoor mapping point clouds without normal vectors. The code of our implementation is made available to the public at `https://github.com/huepat/voxir`.

1. INTRODUCTION

In recent years, digital models of indoor building environments (Borrmann et al., 2018) have experienced an ever increasing surge in importance (Ghaffarianhoseini et al., 2017; Sacks et al., 2020) in different fields of application, such as construction (Jafari et al., 2021), facility management (Gao and Pishdad-Bozorgi, 2019), energy efficiency (Jin et al., 2019) or cultural heritage (Solla et al., 2020). In this context, methods for the efficient creation of building models for existing building structures (Volk et al., 2014) have come to the focus of current research efforts (Lehtola et al., 2020; Weinmann et al., 2021) in the fields of indoor mapping (Otero et al., 2020) (i.e. the efficient acquisition of 3D indoor building geometry by means of mobile sensor systems) and indoor reconstruction (Kang et al., 2020; Pintore et al., 2020) (i.e. the automated generation of building models from indoor mapping data).

Information hinting on a distinction between interior and exterior with respect to building surfaces represented in indoor mapping geometry can offer a valuable guidance for automated indoor reconstruction approaches. This information can be provided by normal vector directions, if they are consistently oriented with respect to the building interior. While normal vectors can be efficiently determined for points of a point cloud by analyzing the point distribution in the neighbourhood of each respective point (Yu et al., 2019; Sanchez et al., 2020a), consistently determining the absolute orientation (i.e. pointing fowards or backwards along the determined direction) can be more challenging (Ochmann and Klein, 2019).

When an indoor mapping system provides information about the position of the sensor at the time of recording of each respective point, the normal vector can be oriented towards the sensor. Often, however, only the point coordinates are provided. Other indoor mapping systems provide output in the form of triangle meshes as derived product generated in a black-box process from the primary sensor measurements like e.g. the Microsoft HoloLens (Khoshelham et al., 2019; Hübner et al., 2020a) or the Matterport system (Chang et al., 2017). These triangle meshes were found to be comparable with point clouds in regard of their aptitude towards classification and segmentation tasks while representing a significantly more compact form of data (Bassier et al., 2020; Weinmann et al., 2020).

Furthermore, these triangle meshes provide consistently oriented normal vectors. A recently published voxel-based indoor reconstruction approach (VoxIR) (Hübner et al., 2020b, 2021a) with publicly availabe code is tailored towards such indoor mapping triangle meshes and depends on their normal vectors as input data. However, it is not straightforward to apply this approach to indoor mapping point clouds, where consistently oriented normal vectors are not available.

Thus, we provide in this paper the following contributions:

- We present a novel method for the automated classification of 3D occupancy voxel grids of indoor building environments according to their normal direction in vertical and horizontal structures. Voxels with vertical normal direction are further distinguished in those, whose normal direction is directed upwards, downwards or both with respect to the building interior.

- We apply the proposed normal classification method in the context of voxel-based indoor reconstruction, extending an established approach so far only applicable to triangle meshes with normal vectors to indoor mapping point clouds without normal information.

- We present a thorough qualitative and quantitative evaluation on two different publicly available benchmark datasets for indoor reconstruction.

In the following, Section 2 gives a brief overview on related work. Afterwards, the proposed methodology is explained in Section 3 before Section 4 presents both qualitative and quantitative evaluation results which are further discussed in Section 5. Finally,

---

[*]Corresponding author

the paper closes with a concluding summary and suggestions for future research in Section 6.

2. RELATED WORK

Automated reconstruction of building environments from indoor mapping data such as point clouds is a wide and active field of research (Kang et al., 2020; Pintore et al., 2020). The various proposed approaches differ significantly in the amount of assumptions that are made with respect to the building structure to be reconstructed and thus in their flexibility towards challenging building environments, ranging from single room scenarios (Li et al., 2020; Sanchez et al., 2020b), Manhattan World structures where all surfaces are orthogonal to the coordinate axes (Ryu et al., 2020; Kim et al., 2020) to diagonal (Shi et al., 2019; Tran and Khoshelham, 2020) or even curved walls (Yang et al., 2019; Wu et al., 2020) and slanted ceilings (Nikoohemat et al., 2020; Lim and Doh, 2021).

The available methods also differ in their general approaches. Some follow a bottom-up strategy, where small local plane patches are detected in the point cloud data and assembled to constitute rooms (Xie et al., 2019; Shi et al., 2020; Oh et al., 2021). Others follow a more top-down approach by first detecting the dominant global planes in the dataset and intersecting them with one another. The cells of the resulting cell complex are then partitioned into building interior and outside space. Storey-wise 2D cell complexes (Li et al., 2018; Tran and Khoshelham, 2020) as well as fully three-dimensional cell complexes can be used (Coudron et al., 2018; Ochmann et al., 2019). Other reconstruction methods make use of trajectory information of the mobile mapping system if available (Cui et al., 2019; Nikoohemat et al., 2020; Lim and Doh, 2021) or operate in a discretized voxel grid (Fichtner et al., 2017; Flikweert et al., 2019; Gorte et al., 2019). Recently, reconstruction methods relying on deep learning methods are gaining in prevalence (Kim et al., 2020; Gankhuyag and Han, 2021; Yang et al., 2021).

Regarding the task of determining normal directions with a consistent orientation with respect to the building interior as a preprocessing step to indoor reconstruction (i.e. not deriving the information about the normal orientation from the reconstructed indoor models), only few works are available that are dedicated specifically to this topic (Ochmann and Klein, 2019). The task of determining floor and ceiling layers from indoor mapping point clouds, on the other hand, is addressed more frequently in the context of indoor reconstruction (Macher et al., 2017; Fichtner et al., 2017; Elseicy et al., 2018; Li et al., 2018; Leoni et al., 2019; Romero-Jarén and Arranz, 2021). Usually, however, the ceilings and floors are assumed to occur in structure of fixed storey levels globally over the whole building and are often assumed to be planar.

3. METHOD

In the following, we present a novel methodology to classify the filled voxels of 3D occupancy grids derived from indoor mapping data according to their main normal direction with respect to the represented building structures. First, Section 3.1 describes the proposed approach. Afterwards, Section 3.2 describes its application in the context of voxel-based indoor reconstruction.

3.1 Normal Classification of Occupancy Grids

The aim of the approach presented here is to classify the filled voxels of a 3D occupancy grid representing indoor building environments into vertical structures (of horizontal normal direction class N_H) and horizontal structures (of vertical normal direction
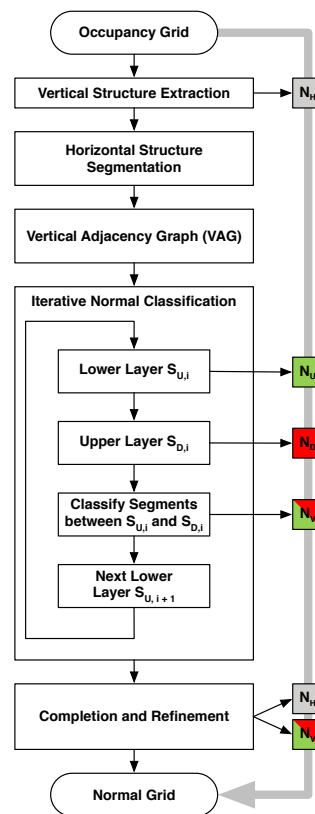

Figure 1. Schematic workflow of the proposed normal classification approach for occupancy voxel grids.

N_V). The N_V voxels are further subdivided into the normal classes N_U (normal up), N_D (normal down) and N_{UD} (normal up and down) with regard to their normal direction pointing towards the room interior. Here, the normal class N_{UD} considers voxels vertically subdividing rooms of different storeys, i.e. where a ceiling surface and the next floor surface above are covered by one and the same voxel.

As is the case with the indoor reconstruction procedure VoxIR (Hübner et al., 2020b, 2021a) in which the proposed normal classification method is to be deployed, we do not assume room surfaces to be planar or subject to common restrictions like the Manhattan World assumption (Hübner et al., 2021b). In particular, we do not assume that rooms are necessarily structured in distinct storeys but account for floors and ceilings to be on different height levels over different rooms as well as within one room.

The proposed normal classification approach is further detailed in the following sections. A graphical overview is presented in Figure 1, while exemplary results for a section of the dataset 'Office' of (Hübner et al., 2021a) are depicted in Figure 2.

3.1.1 Vertical Structure Extraction In a first step, clearly vertical structures such as wall surfaces are detected in the input voxel occupancy grid and labeled as N_H. To this aim, continuous vertical columns of filled voxels of at least $0.5\,\mathrm{m}$ of height are detected. This parameter choice was established as producing satisfactory results over a range of different datasets. Exemplary results are depicted in Figure 2(a).

3.1.2 Horizontal Structure Segmentation The remaining voxels not labeled as N_H are preliminarily assumed to belong to horizontal structures (i.e. N_V) and are segmented into horizontal voxel segments S_V depicted in Figure 2(b). As in VoxIR, a 2.5D
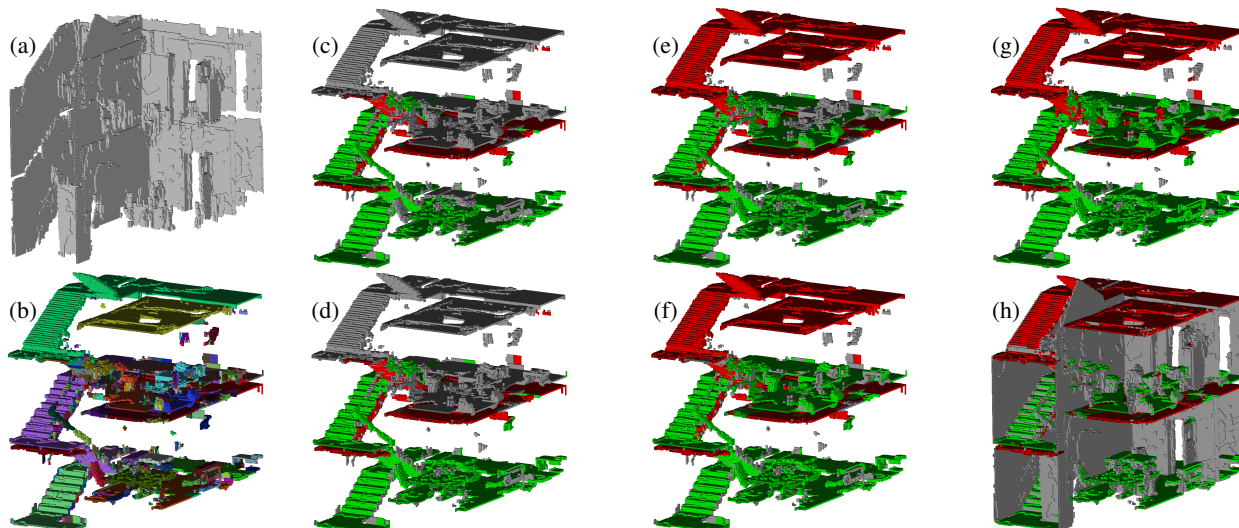
Figure 2. Exemplary results of the normal classification approach presented in Section 3.1 for a section of the dataset 'office' from (Hübner et al., 2021a). (a) N_H voxels. (b) S_V segments. (c) After classification of S_{U_0} and S_{D_0}. (d) After classification of S_V between S_{U_0} and S_{D_0}. (e) After classification of S_{U_1} and S_{D_1}. (f) After classification of S_V between S_{U_1} and S_{D_1}. (g) After completion of classification of S_V. (h) Final normal grid after refinement.

region growing with a threshold on height differences between positionally neighbouring voxels is used, in order to allow S_V to stretch over small height offsets such as stair steps. This can however lead to S_V under-segmenting ceiling surfaces together with the floor surface of the next room above. To prevent this, an additional stopping condition is introduced for the region growing process by not only regarding height differences in the voxel segment itself but also differences in the vertical distance above the respective voxels until the next voxels not yet classified as N_H (or the upper border of the voxel grid) are encountered.

3.1.3 Vertical Adjacency Graph The resulting horizontal segments S_V are assembled in a Vertical Adjacency Graph (VAG) where the voxel segments are the nodes and the edges represent a vertical adjacency (i.e. above/below) relation between them, weighted with the respective area of coverage and mean vertical distance. Besides the S_V, the upper and lower outside O_D and O_U are included as two additional nodes in the graph as neighbour for segments that to not have another segment above/below them. The resulting graph serves as input to the subsequent iterative step of normal classification.

3.1.4 Iterative Normal Classification of Horizontal Segments Some of the S_V as nodes of the VAG are assigned normal classes N_U, N_D and N_{UD} in an iterative process aiming at identifying main layers of segments S_U and S_D which are clearly definable as deliminating rooms vertically from below and above, respectively. While these layers pass vertically upwards through the voxel grid during the iterative process, they are not necessarily restricted to a single height level at a time (i.e. storey-wise).

The iterative process starts from below as we assume the lower surfaces of a room (i.e. floor and horizontal furniture surfaces acquired from above by the respective indoor mapping system) to be more complete than the upper surfaces. The process can however be inverted starting from above if this is more suited for a given indoor mapping system. All examples and results presented in this work are processed from bottom to top.

First, the first lower layer S_{U_0} is initialized by detecting S_V whose largest neighbour below with respect to coverage is the lower outside node O_U of the VAG. Starting from here, each lower layer

S_{U_i} is classified as N_U and a corresponding upper layer S_{D_i} is determined as those S_V whose largest neighbour below is one of the S_{U_i} segments. Here, however, a minimum threshold of $1.5\,\mathrm{m}$ on the mean vertical distance between S_V is applied, as we associate S_{U_i} and S_{D_i} with lower and upper surfaces of rooms and assume rooms to have a certain minimum height. The detected S_{D_i} are assigned the N_D label.

In each iteration, further S_V between the detected main layer segments can be classified as N_U or N_D if its largest upper neighbour is among the current S_{D_i} and its largest lower neighbour is among the S_{U_i} and is also a lower neighbour of the respective upper S_{D_i} neighbour. For the classification of an intermediate segment detected in this manner, its lateral borders are examined in the height range between its S_{U_i} and S_{D_i} neighbours. If there are voxels indicating a lateral vertical surface below/above the S_V segment, it is assumed to represent the upper/lower surface of a piece of furniture (e.g. table surface / lower surface of a lamp) and assigned N_U or N_D accordingly. If there are no lateral surfaces to be found, the decision is made based upon height above the lower S_{U_i} neighbour, assuming that horizontal surfaces below $1.5\,\mathrm{m}$ above the floor are likely to be captured from above and thus assigned N_U.

Based on the current S_{D_i}, the lower main layer of the next iteration $S_{U_{i+1}}$ is detected, again as those S_V, whose largest lower neighbour is among the S_{D_i}. Here, however, it needs to be considered that depending on voxel resolution and width of building structures, ceiling surfaces and the next floor surfaces above could be covered by the same horizontal layer of voxels. Thus, if the mean vertical distance from a S_{D_i} segment to its $S_{U_{i+1}}$ neighbour above is more than $1.5\,\mathrm{m}$, the respective S_{D_i} segment gets assigned an additional N_U label and is itself part of $S_{U_{i+1}}$ instead of its upper neighbour.

This process terminates, when no $S_{U_{i+1}}$ can be found anymore. Intermediate results of the iterations are exemplarily depicted in Figure 2(c) to (f).

3.1.5 Completion and Refinement S_V segments that are so far not yet assigned a normal class value (e.g. because due to occlusion or incomplete acquisition of building geometry they

do not have consistent upper and lower neighbours among the main layers of the same iteration) are finally considered in this step. All direct voxel contacts of an unclassified segment with other voxels that are already assigned a value (N_H as well as N_V) are considered and counted. A segment is assigned to the class with the largest contact count. If it does not directly contact any classified voxels, it is assigned N_H. The final classification result of the S_V segments is depicted in Figure 2(g).

Lastly some refinement steps are conducted on the resulting normal grid. For instance, vertical pillars of multiple N_V voxels are resolved by leaving only the topmost (for N_U) or lowest (for N_D) voxel N_V and turning the others to N_H as they clearly form a vertical structure. The final normal grid is exemplarily depicted in Figure 2(h).

3.2 Voxel-based Indoor Reconstruction from Point Clouds without Normal Vectors

The voxel-based indoor reconstruction method VoxIR presented in (Hübner et al., 2020b, 2021a) is tailored towards indoor mapping triangle meshes as acquired for instance with the Microsoft HoloLens (Hübner et al., 2020a) or the Matterport system (Chang et al., 2017) as these provide normal vectors consistently oriented with respect to inside/outside of the building structure. These oriented normals are further discretized to a normal grid of arbitrary resolution with the values N_H, N_U and N_D as described above which is the input to the voxel-based indoor reconstruction process.

In the scope of this work, we extent VoxIR to also be applicable to point clouds without normal vectors by means of the normal classification approach on occupancy grids presented in Section 3.1. Furthermore, we extend VoxIR to consider N_{UD} voxels representing a ceiling surface as well as the next floor above. The code of the extended version of VoxIR adapted to indoor mapping point clouds such as those of the ISPRS Benchmark on Indoor Modeling (Khoshelham et al., 2017) including our implementation of the normal classification procedure described above is released at https://github.com/huepat/voxir.

4. EVALUATION AND RESULTS

In order to quantitatively evaluate our proposed approach, we use the metrics of completeness, correctness and accuracy as proposed in (Khoshelham et al., 2018) and used in the context of evaluating the contributions to the ISPRS Challenge on Indoor Modeling (Khoshelham et al., 2020, 2021). However, these metrics expect reference and reconstructed building geometry to be given in continuous, Euclidean space, considering the fraction of reconstructed building structures within a given buffer distance around the reference building structure. As our method operates in discrete voxel space, we also use discretized versions of the evaluation metrics

$$\text{Completeness}(b) = \frac{\#\text{Vox}_{TP}(b)}{\#\text{Vox}_{Ref}} \quad (1)$$

and

$$\text{Correctness}(b) = \frac{\#\text{Vox}_{TP}(b)}{\#\text{Vox}_{Rec}} \quad (2)$$

with $\#\text{Vox}_{Ref}$ denoting the number of voxels representing building geometry in the discretized reference, $\#\text{Vox}_{Rec}$ the same value for the reconstruction voxel grid and $\#\text{Vox}_{TP}(b)$ being the number of true positive reconstructed voxels with respect to a buffer level b. For $b = 1$, we only consider a voxel as a true positive, if the corresponding voxel position in the reference grid

Table 1. Evaluation results of the normal classification of the HoloLens triangle meshes published in (Hübner et al., 2021a) with $5\,\text{cm}$ voxel resolution.

Dataset	Completeness	Correctness
Office	0.88	0.87
Attic	0.85	0.84
Basement	0.90	0.89
Residential House	0.82	0.80

reports building structure as well. For $b > 1$, we also consider voxels as true positives, where reference building structure is found within a (26-)neighbourhood of the respective voxel position. For $b = 2$, we consider the direct neighbours of the voxel position and for $b = 3$ also the neighbours of these neighbours, and so on. This can be regarded as a disrectized version of the buffer distance used in (Khoshelham et al., 2018) as well as as a generalization of the concept of 'neighbourhood precision/recall' in (Hübner et al., 2021a) where only $b = 2$ is considered.

Furthermore, we present accuracy values in analogy with (Khoshelham et al., 2018), quantifying the distance d of reconstructed building geometry to corresponding reference geometry depending on the buffer level b:

$$\text{Accuracy}(b) = \text{Mean}(d(b)) \quad (3)$$

In our case, we use as distance d the distance between the respective voxel center coordinates. Furthermore, instead of using the median, we chose to use the arithmetic mean, as possible distances in our case are restricted to a few possible values due to the discretization of the voxel grid.

While in (Hübner et al., 2021a) quantitative evaluation results are presented for different semantic classes, here, we consider only the building structure without any further semantic distinction as proposed in (Khoshelham et al., 2018). In doing so, we consider the semantic classes of 'Wall', 'Floor' and 'Ceiling' as belonging to the building structure, while 'Interior Object' representing furniture and clutter is disregarded as well as 'Empty Interior' and 'Wall Opening'. For all presented experiments, a fixed voxel resolution of $5\,\text{cm}$ is used. Investigating the impact of this parameter exceeds the scope of this work and is left to future research.

4.1 Normal Classification of Occupancy Grids

First, we evaluate completeness and correctness of the normal classification procedure presented in Section 3.1 on indoor mapping triangle meshes, where normal vectors and reference values are available. To this aim, we use the four publicly available datasets of HoloLens triangle meshes of different indoor environments presented in (Hübner et al., 2021a), where further details on the datasets can be found. The respective datasets are discretized to binary occupancy grids and, for each filled voxel, the normal class is determined. As reference data, the voxel grid with normal classification as derived by VoxIR from the triangle meshes and their normal vectors is used. The results are presented in Table 1. Here, we regard a voxel as a true positive, if it has the same normal class label as the corresponding voxel in the reference grid, while only considering buffer level $b = 1$.

4.2 Impact on VoxIR

The four datasets of triangle meshes used in Section 4.1 are furthermore used to investigate the impact of determining the normal classification from binary occupancy grids instead of from the

Table 2. Evaluation of the impact of the normal grid on VoxIR reconstruction results on the HoloLens triangle meshes published in (Hübner et al., 2021a) with 5 cm voxel resolution. The presented values refer to results achieved when using normal grids determined with the method presented in Section 3.1. The values in parentheses refer to results achieved with normal grids determined directly from the triangle meshes.

Dataset	Buffer Level	Completeness	Correctness	Accuracy [m]
Office	1	0.86 (0.87)	0.69 (0.75)	0.00 (0.00)
	2	1.13 (1.11)	0.91 (0.95)	0.02 (0.02)
	3	1.19 (1.14)	0.96 (0.98)	0.03 (0.02)
Attic	1	0.80 (0.81)	0.65 (0.70)	0.00 (0.00)
	2	1.05 (1.04)	0.86 (0.90)	0.02 (0.02)
	3	1.11 (1.08)	0.91 (0.94)	0.03 (0.02)
Basement	1	0.83 (0.87)	0.71 (0.72)	0.00 (0.00)
	2	1.07 (1.11)	0.91 (0.92)	0.02 (0.02)
	3	1.12 (1.16)	0.95 (0.96)	0.03 (0.02)
Residential House	1	0.72 (0.83)	0.58 (0.71)	0.00 (0.00)
	2	0.97 (1.05)	0.78 (0.90)	0.02 (0.02)
	3	1.04 (1.09)	0.84 (0.94)	0.03 (0.02)

normal vectors of the triangles on the reconstruction results of VoxIR. Table 2 presents an evaluation of reconstruction results on the four datasets for three buffer levels. The reported values refer to reconstruction results achieved when using a normal grid determined by the method from Section 3.1 from occupancy grids. In parantheses, the respective values are given, if instead of a mere occupancy grid, the normal grid derived directly from the triangle mesh is used. As can be seen, using discretized indoor mapping data without normal vectors only slightly diminishes the reconstruction quality of VoxIR on most datasets, with the exception of dataset 'Residential House' showing a decrease in completeness and correctness of around 10 %.

Note that the reported values for completeness reach values above 1. This is due to the fact, that VoxIR tends to reconstruct room surfaces with a thickness of several voxels while they are usually represented as surfaces with a thickness of one voxel in the discretized reference data. Thus, there are much more reconstructed voxels representing building geometry than reference voxels which can lead to there being more true positive voxels than reference voxels when using a buffer level $b > 1$.

4.3 VoxIR Applied to Point Clouds

Finally, the normal classification method presented in this work allows to evaluate the voxel-based indoor reconstruction VoxIR also on indoor mapping point clouds like those of the ISPRS benchmark on indoor modelling (Khoshelham et al., 2017). To this aim, the reference IFC models were converted to triangle meshes using IfcConvert 0.6.0[1]. All triangle meshes belonging to building structures not represented in the point clouds (i.e. the outer surfaces of the volumetric geometries) were removed in a semi-manual process using Blender 2.9.1[2]. Furthermore, the point clouds were manually cleaned as well, as they contain parts of the building environments that are not represented by the reference models. The removed parts are marked in red in the occupancy grids of the discretized point clouds depicted in the first column of Figure 3. From the resulting occupancy grids, normal grids were determined with the method presented in Section 3.1. These were then used as input for VoxIR, with the resulting reconstruction voxel grids being depicted in the right column of Figure 3 (with voxels of semantic classes 'Wall Opening' and 'Interior Object' omitted as they are not considered in the evaluation). The reference triangle

---
[1] http://ifcopenshell.org/ifcconvert
[2] https://www.blender.org/

Table 3. Evaluation of VoxIR reconstruction results on the point clouds published in (Khoshelham et al., 2017, 2020) with 5 cm voxel resolution.

Dataset	Buffer Level	Completeness	Correctness	Accuracy [m]
Case Study 1	1	0.65	0.34	0.00
	2	1.45	0.77	0.05
	3	1.67	0.88	0.06
Case Study 2	1	0.37	0.23	0.00
	2	1.02	0.63	0.05
	3	1.32	0.82	0.08
Case Study 3	1	0.46	0.42	0.00
	2	0.82	0.74	0.04
	3	0.89	0.81	0.05
Case Study 4	1	0.43	0.27	0.00
	2	1.10	0.69	0.05
	3	1.34	0.84	0.07
Case Study 5	1	0.49	0.30	0.00
	2	1.14	0.71	0.05
	3	1.30	0.80	0.06
Case Study 6	1	0.32	0.21	0.00
	2	0.84	0.56	0.05
	3	1.11	0.74	0.08

meshes manually derived from the IFC models were discretized as well and they are depicted in the middle column of Figure 3 to serve as reference data for quantitative evaluation. The results are presented in Table 3.

5. DISCUSSION

The results of the evaluation of the normal classification procedure presented in Table 1 hint on a good performance, and Table 2 shows that using these normal grids determined from occupancy grids does not significantly reduce reconstruction quality when fed as input into the VoxIR pipeline. Still, the reconstruction results on the ISPRS benchmark presented in Table 3 leave room for improvement, especially when regarding only the first buffer level, i.e. only considering exact same voxel positions in result and reference grid. However, the quality rapidly improves when considering higher buffer levels.

Overall, the results presented in the right part of Figure 3 on the right are visually also quite close to the discretized reference building structures depicted in the middle column. Among noticeable deviations from the reference, the inner walls in Figure 3(c) are partly missing or false, as are the big hexagonal column and the inner wall in Figure 3(d). These missing walls are reconstructed as interior objects (i.e. piece of furniture) and thus omitted in the visualization in Figure 3.

The room in the front right corner in Figure 3(c) is only partly reconstructed as well. Generally, as VoxIR is quite generic with respect to its assumptions on building structures, partly scanned rooms are reconstructed close to the shape of the parts included in the input data. Better reconstruction results are possible when applying more restricting assumptions on building geometry. For instance, the mentioned room in Figure 3(c) could potentially be restored better, when assuming that rooms are rectangular.

On the other hand, this flexibility of VoxIR towards possible building structures can also be regarded as its strength. For instance, the protrusions of the ceiling of the large room in Figure 3(e) are reconstructed quite well, as are the unusual room shapes in Figure 3(f). Here, in Figure 3(f), the inner yard is also reconstructed as a room with its ceiling height determined from the height of the surrounding ceilings.

(a) Case Study 1.

(b) Case Study 2.

(c) Case Study 3.

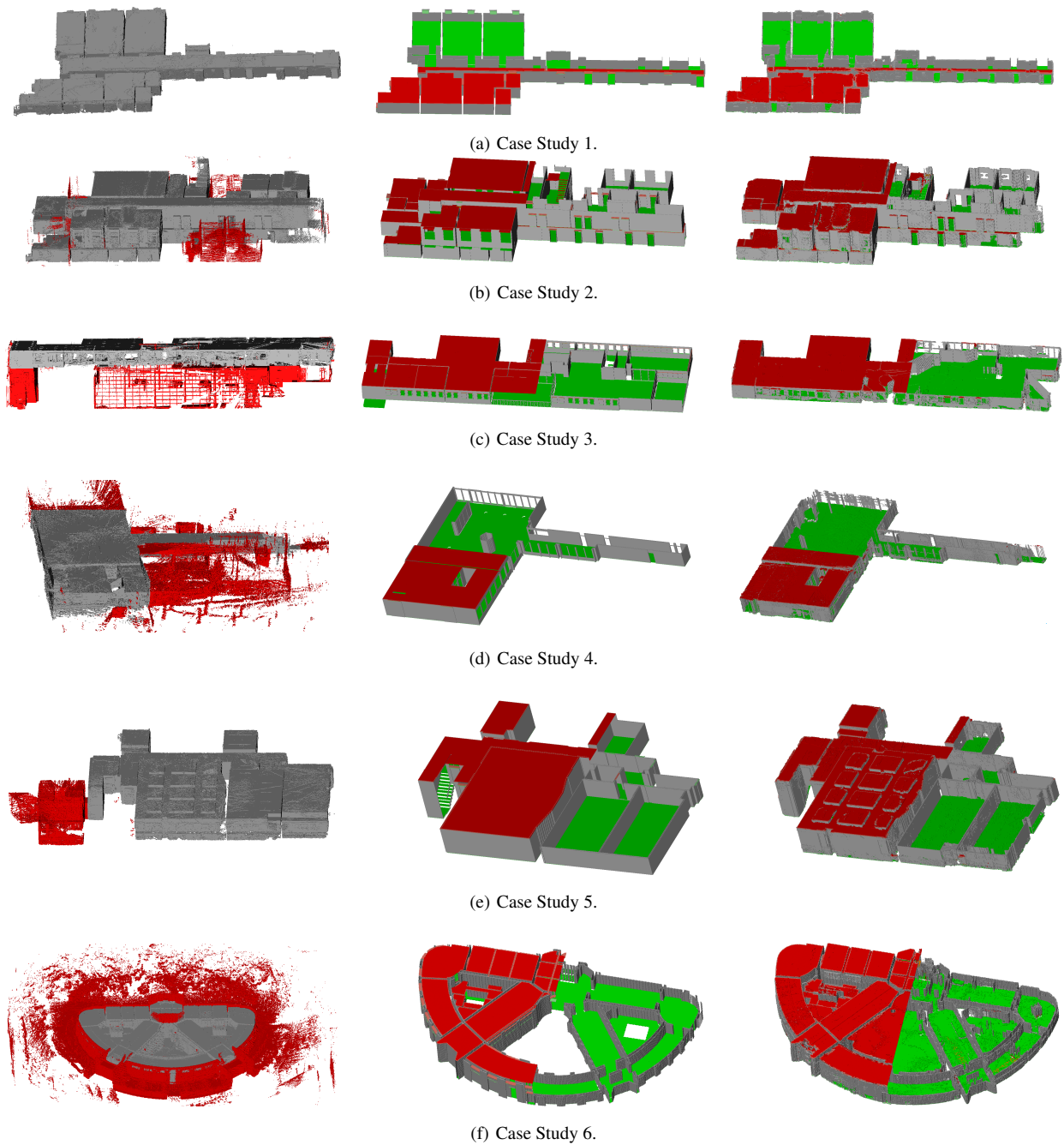(d) Case Study 4.

(e) Case Study 5.

(f) Case Study 6.

Figure 3. Visualization of input occupancy grids (manually removed parts depicted in red) on the left, discretized ground truth data in the middle and VoxIR reconstruction results on the right (both with $5\,\mathrm{cm}$ voxel resolution, parts of the ceilings are removed for better visibility) for the point clouds of the ISPRS benchmark on indoor modeling (Khoshelham et al., 2017, 2020). In the middle and right parts, ceiling is depicted in red, while floor and walls are depicted in green and grey, respectively.
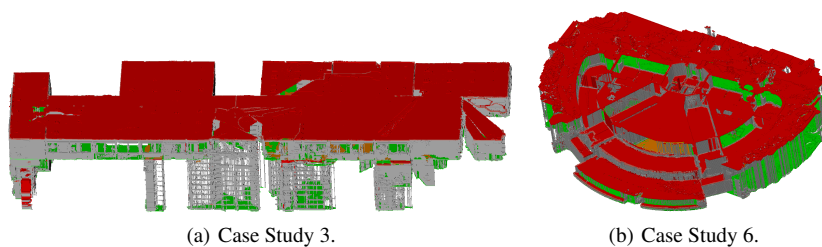

(a) Case Study 3.

(b) Case Study 6.

Figure 4. Reconstruction results for two of the datasets depicted in Figure 3, when the red parts of the occupancy grids on the left in Figure 3 are included in the input occupancy grids.

As exemplified in Figure 4, the reconstruction results are also more or less reasonable when including the removed parts of the input occupancy grids depicted in red on the left-hand side of Figure 3 which are not considered in the reference IFC models. For instance, the lower floor of Figure 4(a) is partly reconstructed where there are parts of its floor or ceiling available in the input data. Also, the outdoor terrain around the building in Figure 4(b) is included in the reconstruction, with its height being determined by the height of the trees in the rearward part behind the building. In this case, the height of the ceilings reconstructed above the inner yards gets set to the tree height as well. Generally indoor/outdoor transitions and their consideration in modeling and automated reconstruction of buildings is still an interesting topic for future research (Previtali et al., 2014; Koch et al., 2016).

6. CONCLUSION

In this paper, we presented a novel method for normal classification of voxel occupancy grids of discretized indoor mapping data along with qualitative and quantitative evaluation results. Furthermore, the proposed method was used to extend an existing voxel-based indoor reconstruction pipeline to be applicable to input data in the form of point clouds. This enabled us to evaluate the indoor reconstruction approach, which so far was only applicable to triangle meshes, on the point clouds of the well-known ISPRS Benchmark for Indoor Modeling.

While the presented results are promising, there is still ample opportunity for further research left for future work. For instance, the evaluation presented here focuses exclusively on the reconstruction of building structures without regard for further semantics or room partitioning which are also considered in the evaluated voxel-based indoor reconstruction approach. Furthermore, the discussed method (for normal classification as well as for indoor reconstruction) is still restricted to the discretized voxel space. Conversion of the results towards actual surface or even volume geometries in Euclidean space would be a valuable extension, providing the means for the automated generation of actual BIM models from indoor mapping data and enabling better comparability of the results with those of other reconstruction methods. Still, we believe that voxels hold great potential for building-related analysis tasks (Gorte et al., 2019; Song et al., 2019; Wang et al., 2020).

References

Bassier, M., Vergauwen, M. and Poux, F., 2020. Point Cloud vs. Mesh Features for Building Interior Classification. Remote Sens. 12(14), pp. 2224:1–26.

Borrmann, A., König, M., Koch, C. and Beetz, J., 2018. Building Information Modeling: Why? What? How? In: Building Information Modeling, Springer, pp. 1–24.

Chang, A., Dai, A., Funkhouser, T., Halber, M., Nießner, M., Savva, M., Song, S., Zeng, A. and Zhang, Y., 2017. Matterport3D: Learning from RGB-D Data in Indoor Environments. In: International Conference on 3D Vision (3DV), pp. 667–676.

Coudron, I., Puttemans, S. and Goedemé, T., 2018. Polygonal Reconstruction of Building Interiors from Cluttered Pointclouds. In: European Conference on Computer Vision (ECCV), pp. 459–472.

Cui, Y., Li, Q. and Dong, Z., 2019. Structural 3D Reconstruction of Indoor Space for 5G Signal Simulation with Mobile Laser Scanning Point Clouds. Remote Sens. 11(19), pp. 2262:1–32.

Elseicy, A., Nikoohemat, S., Peter, M. and Oude Elberink, S., 2018. Space Subdivision of Indoor Mobile Laser Scanning Data Based on the Scanner Trajectory. Remote Sens. 10(11), pp. 1815:1–26.

Fichtner, F. W., Diakité, A. A., Zlatanova, S. and Voûte, R., 2017. Semantic Enrichment of Octree Structured Point Clouds for Multi-Story 3D Pathfinding. Trans. GIS 22, pp. 233–248.

Flikweert, P., Peters, R., Díaz-Vilariño, L., Voûte, R. and Staats, B., 2019. Automatic Extraction of a Navigation Graph Intended for IndoorGML from an Indoor Point Cloud. ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci. IV-2/W5, pp. 271–278.

Gankhuyag, U. and Han, J.-H., 2021. Automatic BIM Indoor Modelling from Unstructured Point Clouds Using a Convolutional Neural Network. Intelligent Automation & Soft Computing 28(1), pp. 133–152.

Gao, X. and Pishdad-Bozorgi, P., 2019. BIM-Enabled Facilities Operation and Maintenance: A Review. Adv. Eng. Inform. 39, pp. 227–247.

Ghaffarianhoseini, A., Tookey, J., Ghaffarianhoseini, A., Naismith, N., Azhard, S., Efimova, O. and Raahemifarb, K., 2017. Building Information Modelling (BIM) Uptake: Clear Benefits, Understanding its Implementation, Risks and Challenges. Renew. Sust. Energ. Rev. 75, pp. 1046–1053.

Gorte, B., Zlatanova, S. and Fadli, F., 2019. Navigation in Indoor Voxel Models. ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci. IV-2/W5, pp. 279–283.

Hübner, P., Clintworth, K., Liu, Q., Weinmann, M. and Wursthorn, S., 2020a. Evaluation of HoloLens Tracking and Depth Sensing for Indoor Mapping Applications. Sensors 20(4), pp. 1021:1–24.

Hübner, P., Weinmann, M. and Wursthorn, S., 2020b. Voxel-Based Indoor Reconstruction from HoloLens Triangle Meshes. ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci. V-4-2020, pp. 79–86.

Hübner, P., Weinmann, M., Wursthorn, S. and Hinz, S., 2021a. Automatic Voxel-based 3D Indoor Reconstruction and Room Partitioning from Triangle Meshes. ISPRS J. Photogramm. Remote Sens. 181, pp. 254–278.

Hübner, P., Weinmann, M., Wursthorn, S. and Hinz, S., 2021b. Pose Normalization of Indoor Mapping Datasets Partially Compliant with the Manhattan World Assumption. Remote Sens. 13(23), pp. 4765:1–34.

Jafari, K. G., Sharyatpanahi, N. S. G. and Noorzai, E., 2021. BIM-Based Integrated Solution for Analysis and Management of Mismatches During Construction. J. Eng. Des. Technol. 19(1), pp. 81–102.

Jin, R., Zhong, B., Ma, L., Hashemi, A. and Ding, L., 2019. Integrating BIM with Building Performance Analysis in Project Life-Cycle. Autom. Constr. 106, pp. 102861.

Kang, Z., Yang, J., Yang, Z. and Cheng, S., 2020. A Review of Techniques for 3D Reconstruction of Indoor Environments. ISPRS Int. J. Geo-Inf. 9(5), pp. 330:1–31.

Khoshelham, K., Díaz Vilariño, L., Peter, M., Kang, Z. and Acharya, D., 2017. The ISPRS Benchmark on Indoor Modelling. Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. XLII-2/W7, pp. 367–372.

Khoshelham, K., Tran, H., Acharya, D., Díaz Vilariño, L., Kang, Z. and Dalyot, S., 2020. The ISPRS Benchmark on Indoor Modelling – Preliminary Results. Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. XLIII-B5-2020, pp. 207–211.

Khoshelham, K., Tran, H., Acharya, D., Díaz Vilariño, L., Kang, Z. and Dalyot, S., 2021. Results of the ISPRS Benchmark on Indoor Modelling. ISPRS Open Journal of Photogrammetry and Remote Sensing 2, pp. 100008:1–13.

Khoshelham, K., Tran, H. and Acharya, D., 2019. Indoor Mapping Eyewear: Geometric Evaluation of Spatial Mapping Capability of HoloLens. Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. XLII-2/W13, pp. 805–810.

Khoshelham, K., Tran, H., Díaz-Vilariño, L., Peter, M., Kang, Z. and Acharya, D., 2018. An Evaluation Framework for Benchmarking Indoor Modelling Methods. Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. XLII-4, pp. 297–302.

Kim, T., Cho, W., Matono, A. and Kim, K.-S., 2020. PinSout: Automatic 3D Indoor Space Construction from Point Clouds with Deep Learning. In: SIGSPATIAL '20: Proceedings of the 28th International Conference on Advances in Geographic Information Systems, pp. 211–214.

Koch, T., Körner, M. and Fraundorfer, F., 2016. Automatic Alignment of Indoor and Outdoor Building Models Using 3D Line Segments. In: IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), pp. 689–697.

Lehtola, V. V., Nikoohemat, S. and Nüchter, A., 2020. Indoor 3D: Overview on Scanning and Reconstruction Methods. In: Handbook of Big Geospatial Data, Springer Nature Switzerland AG, chapter 3, pp. 55–97.

Leoni, C., Ferrarese, S., Wahbeh, W. and Nardinocchi, C., 2019. Extraction of Main Levels of a Building from a Large Point Cloud. Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. XLII-5/W2, pp. 41–47.

Li, L., Su, F., Yang, F., Zhu, H., Li, D., Zuo, X., Li, F., Liu, Y. and Ying, S., 2018. Reconstruction of Three-Dimensional (3D) Indoor Interiors with Multiple Stories via Comprehensive Segmentation. Remote Sens. 10(8), pp. 1281:1–30.

Li, Y., Li, W., Tang, S., Darwish, W., Hu, Y. and Chen, W., 2020. Automatic Indoor as-Built Building Information Models Generation by Using Low-Cost RGB-D Sensors. Sensors 20(1), pp. 293:1–21.

Lim, G. and Doh, N., 2021. Automatic Reconstruction of Multi-Level Indoor Spaces from Point Cloud and Trajectory. Sensors 21(10), pp. 3493:1–16.

Macher, H., Landes, T. and Grussenmeyer, P., 2017. From Point Clouds to Building Information Models: 3D Semi-Automatic Reconstruction of Indoors of Existing Buildings. Appl. Sc. 7(10), pp. 1030:1–30.

Nikoohemat, S., Diakité, A. A., Zlatanova, S. and Vosselman, G., 2020. Indoor 3D Reconstruction from Point Clouds for Optimal Routing in Complex Buildings to Support Disaster Management. Autom. Constr. 113, pp. 103109:1–18.

Ochmann, S. and Klein, R., 2019. Automatic Normal Orientation in Point Clouds of Building Interiors. In: Computer Graphics International Conference, pp. 556–563.

Ochmann, S., Vock, R. and Klein, R., 2019. Automatic Reconstruction of Fully Volumetric 3D Building Models from Point Clouds. ISPRS J. Photogramm. Remote Sens. 151, pp. 251–262.

Oh, S., Lee, D., Kim, M., Kim, T. and Cho, H., 2021. Building Component Detection on Unstructured 3D Indoor Point Clouds Using RANSAC-Based Region Growing . Remote Sens. 13(2), pp. 161:1–20.

Otero, R., Lagüela, S., Garrido, I. and Arias, P., 2020. Mobile Indoor Mapping Technologies: A Review. Autom. Constr. 120, pp. 103399.

Pintore, G., Mura, C., Ganovelli, F., Fuentes-Perez, L., Pajarola, R. and Gobbetti, E., 2020. State-of-the-Art in Automatic 3D Reconstruction of Structured Indoor Environments. In: R. Mantiuk and V. Sundstedt (eds), EUROGRAPHICS 2020, Vol. 39, pp. 667–699.

Previtali, M., Scaioni, M., Barazzetti, L. and Brumana, R., 2014. A Flexible Methodology for Outdoor/Indoor Building Reconstruction from Occluded Point Clouds. ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci. II-3, pp. 119–126.

Romero-Jarén, R. and Arranz, J. J., 2021. Automatic Segmentation and Classification of BIM Elements from Point Clouds. Autom. Constr. 124, pp. 103576:1–17.

Ryu, M. W., Oh, S. M., Kim, M. J., Cho1, H. H., Son, C. B. and Kim, T. H., 2020. Algorithm for Generating 3D Geometric Representation Based on Indoor Point Cloud Data. Appl. Sc. 10(22), pp. 8073:1–13.

Sacks, R., Girolami, M. and Brilakis, I., 2020. Building Information Modelling, Artificial Intelligence and Construction Tech. Developments in the Built Environment 4, pp. 100011:1–24.

Sanchez, J., Denis, F., Coeurjolly, D., Dupont, F., Trassoudaine, L. and Checchin, P., 2020a. Robust Normal Vector Estimation in 3D Point Clouds through Iterative Principal Component Analysis. ISPRS J. Photogramm. Remote Sens. 163, pp. 18–35.

Sanchez, J., Denis, F., Dupont, F., Trassoudaine, L. and Checchin, P., 2020b. Data-Driven Modeling of Building Interiors from LiDAR Point Clouds. ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci. V-2-2020, pp. 395–402.

Shi, P., Ye, Q. and Zeng, L., 2020. A Novel Indoor Structure Extraction Based on Dense Point Cloud. ISPRS Int. J. Geo-Inf. 9(11), pp. 660:1–25.

Shi, W., Ahmed, W., Li, N., Fan, W., Xiang, H. and Wang, M., 2019. Semantic Geometric Modelling of Unstructured Indoor Point Cloud. ISPRS Int. J. Geo-Inf. 8(1), pp. 9:1–20.

Solla, M., Gonçalves, L. M. S., Gonçalves, G., Francisco, C., Puente, I., Providência, P., Gaspar, F. and Rodrigues, H., 2020. A Building Information Modeling Approach to Integrate Geomatic Data for the Documentation and Preservation of Cultural Heritage. Remote Sens. 12(24), pp. 4028:1–24.

Song, Y., Niu, L. and Li, Y., 2019. Combinatorial Spatial Data Model for Building Fire Simulation and Analysis . ISPRS Int. J. Geo-Inf. 8(9), pp. 408:1–21.

Tran, H. and Khoshelham, K., 2020. Procedural Reconstruction of 3D Indoor Models from Lidar Data Using Reversible Jump Markov Chain Monte Carlo. Remote Sens. 12(5), pp. 838:1–26.

Volk, R., Stengel, J. and Schultmann, F., 2014. Building Information Modeling (BIM) for Existing Buildings – Literature Review and Future Needs. Autom. Constr. 38, pp. 109–127.

Wang, Q., Zuo, W., Guo, Z. and Li, Q., 2020. BIM Voxelization Method Supporting Cell-Based Creation of a Path-Planning Environment. J. Constr. Eng. Manag. 146(7), pp. 04020080:1–16.

Weinmann, M., Jäger, M. A., Wursthorn, S., Jutzi, B., Weinmann, M. and Hübner, P., 2020. 3D Indoor Mapping with the Microsoft HoloLens: Qualitative and Quantitative Evaluation by Means of Geometric Features. ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci. V-1-2020, pp. 165–172.

Weinmann, M., Wursthorn, S., Weinmann, M. and Hübner, P., 2021. Efficient 3D Mapping and Modelling of Indoor Scenes with the Microsoft HoloLens: A Survey. PFG – J. Photogramm. Rem. 89, pp. 319–333.

Wu, K., Shi, W. and Ahmed, W., 2020. Structural Elements Detection and Reconstruction (SEDR): A Hybrid Approach for Modeling Complex Indoor Structures. ISPRS Int. J. Geo-Inf. 9(12), pp. 760:1–15.

Xie, L., Wang, R., Ming, Z. and Chen, D., 2019. A Layer-Wise Strategy for Indoor As-Built Modeling Using Point Clouds. Appl. Sc. 9(14), pp. 2904:1–27.

Yang, F., Zhou, G., Su, F., Zuo, X., Tang, L., Liang, Y., Zhu, H. and Li, L., 2019. Automatic Indoor Reconstruction from Point Clouds in Multi-Room Environments with Curved Walls. Sensors 19(17), pp. 3798:1–19.

Yang, J., Kang, Z., Zeng, L., Akwensi, P. H. and Sester, M., 2021. Semantics-Guided Reconstruction of Indoor Navigation Elements from 3D Colorized Points. ISPRS J. Photogramm. Remote Sens. 173, pp. 238–261.

Yu, Z., Wang, T., Guo, T., Li, H. and Dong, J., 2019. Robust Point Cloud Normal Estimation via Neighborhood Reconstruction. Adv. Mech. Eng. 11(4), pp. 1–19.