DOI: 10.58895/ksp/1000150865 -16 erschienen in:
Forum Bildverarbeitung 2022 | Image Pocessing Forum 2022
DOI: 10.5445/KSP/1000150865 | https://www.ksp.kit.edu/site/books/m/10.58895/ksp/1000150865/

Indoor floorplan estimation from 3D point clouds for *Scan-to-BIM*

Oscar H. Ramírez-Agudelo[1], Antje Alex[2], Lena Schreiber[1], Norman Niemann[1], Edoardo Milana[1], and Christof Hammer[1]

[1] German Aerospace Center (DLR), Institute for the Protection of Terrestrial Infrastructures, Rathausallee 12, 53757 Sankt Augustin, Germany
[2] German Aerospace Center (DLR), Institute for the Protection of Maritime Infrastructures, Fischkai 1, 27572 Bremerhaven, Germany

Abstract Societies depend on the unrestricted availability of their infrastructures. Events such as (natural) disasters, emergencies, or even attacks, could threaten their safety and security. Indoors models provide relevant information that could help in this regard. Their floorplans contain key information such as their location, design, and layout. The architecture, engineering, and construction (*AEC*) community work together to create the respective indoor models within the Building Information Modelling (*BIM*) framework. *BIM* modelling has recently gotten the attention in the computer vision domain. The 1st international *Scan-to-BIM* challenge, organised within the *CVPR 2021* conference, helped to establish research interest and common goals between the *AEC* and computer vision community. In this paper, we introduce a method to estimate floorplans from 3D point cloud data by using the *Scan-to-BIM* dataset. Our work has been developed by using image processing techniques. It does not aim to replace state-of-the-art approaches, which are more elaborate and robust. Instead, it constitutes a non CPU intensive alternative that fairly estimates floorplans for the *Scan-to-BIM* dataset.

Keywords Floorplan estimation, 3D point clouds, *Scan-to-BIM*, data and image processing

1 Introduction

Modern societies depend on the unrestricted availability of their critical infrastructures [1], where buildings constitute main terrestrial infrastructures. They impact our quality of life in many of the same ways as other infrastructures. To protect them from dangers is essential for prosperity and social stability. Events such as (natural) disasters, emergencies, or even attacks, could threaten their safety and security [2]. Therefore, it is important to gather detailed information as well as to provide indoor models [3]. The Institutes for the Protection of Terrestrial and Maritime Infrastructures, subscribed to the German Aerospace Center (DLR), are dedicated to develop concepts and technologies to help to improve the safety and security of critical maritime and terrestrial infrastructures.

The floorplan of buildings becomes a relevant representation of their interiors. In the architecture, engineering, and construction (*AEC*) community, it is standard that such models are done manually, being prone to human errors. Additionally, due to renovation and maintenance, floorplans are often outdated. Moreover, there are other cases, where the floorplans do not even exist [4]. In recent years, with 3D point cloud scanning and technologies such as building information model (*BIM*), the modelling has become a common practice [3]. Although it still encounters computational challenges such as data diversity, accurate geometry, large-scale input, etc. [5], it is currently an active area of research.

Computer vision has already made progress in the detection of walls from buildings [6]. Deep learning has shown promising potential in object detection [7] or in room layout reconstruction tasks such as segmentation and parsing geometry [8–12]. Deep neuronal networks have also been applied to floorplan reconstruction [13–15] (see Section 2).

In this paper, we propose an automatic and light alternative to estimate the 2D floorplan from 3D point cloud data by implementing an image processing approach. This work is structured as follows. Section 2 reviews relevant literature. Section 3 describes the methodology adopted in this work. The results are presented in Section 4 and discussed in Section 5. The conclusion is presented in Section 6.

2 Relevant works

Computer vision and deep learning tasks have made an effort to reconstruct indoor floorplan environments. In computer vision, some representative works are for instance [14] and [16]. The authors generate the 2D floorplan by using line detection algorithms such as *CANNY* [17] and *RANSAC* [18]. In the latter case, the output model is provided within the *BIM* format. It is important to note, however, that first the reconstruction is only based on walls, i.e. excluding information such as doors and stairs. Second, the reconstruction follows the Manhattan-layout assumption, i.e. the orientation of the floorplan can only be horizontal or vertical.

In deep learning, *Floor-Net* [13] and *FloorPP-Net* [15] are representative frameworks to reconstruct floorplans from 3D point clouds. By using the *Scan-to-BIM* dataset [19], *FloorPP-Net* converts it into point pillars. Then the network learns to predict the corners and the edges, generating the desired floorplan output model. Again the final model is only based on walls. Computer vision and deep learning approaches are currently working to include the information of doors and stairs in their future models. However, due to class imbalance (i.e. $data(wall) \gg data(door)$ or $data(stair)$), this aim is a relatively difficult to accomplish. Besides, due to data pre-processing and algorithm implementation (line detection for computer vision or neuronal networks for deep learning), these frameworks could take up to several minutes to compute ($\sim$5 minutes) and require special graphical processing unit (GPU); making them computing time intensive.

3 Methodology

3.1 Dataset

In this paper, we introduce a method to estimate floorplans from 3D point cloud data by using the *Scan-to-BIM* dataset. The dataset has been obtained from the 1^{st} International *Scan-to-BIM* Challenge [19]. It was published at the workshop *Computer Vision in the Built Environment* (*CVBE*) as part of the *Computer Vision Pattern Recognition* (*CVPR*) conference in 2021 [20][3]. The dataset includes a wide variety of constructions such as libraries, office labs, short-, medium- and large-offices as well as parking sites. The sample contains a total of 31 buildings with multiple floors each and dozens of rooms on each floor. For 20 buildings it also contains floorplan ground truths. The labels range from wall, door, stair, etc.

3.2 Framework

The methodology developed in this work is based on image processing techniques. The framework is implemented in two-stages.

Algorithm 1

The first-stage consists of the construction of a 2D histogram where all data-points are projected to the $x - y$ plane with the bin size as parameter (*bins*). The histogram returns the x and y edges of the grid (i.e. x_{edges} and y_{edges}) as well as the number of data points per the bi-dimensional bin (H) computed in log-scale.

Algorithm 2

The second-stage computes the floorplan estimation in the following way:

1. The consideration of the output of x_{edges}, y_{edges}, and H computed in the first-stage.
2. The values are normalised with respect to the bin-size to make our method independent of the dimension of the input point cloud.

[3] The CVPR 2022 hosted the 2^{nd} version of the *CVBE*-workshop, where the same dataset has been made available.

3. The ground truth, provided for 20 buildings by the *Scan-to-BIM* dataset, supply the annotation of the labels by segments. A segment is defined by two coordinates; i.e. (x_1, y_1) and (x_2, y_2). For each segment, the class-category (e.g., wall) is assigned to *bins* which distance to that segment is smaller than a *Criterion* (*Crit*) and if the content of that bin, namely H, is larger than the *threshold* (Thr).

The proposed methodology has been applied to several cases. The numeric values of the parameters are: $bins = 1000$, $Crit = 25$, and $Thr \geq 0$. The parameters have been selected after a grid search. They optimise our results without incurring in over-fitting.

3.3 Metrics

We aim to evaluate the position and length of the detected features (e.g. wall) by using the *precision* and *recall*. Based on true-positive (TP), false-positive (FP), and false-negative (FN), we calculate the recall as follows:

$$Precision = \frac{TP}{FP + TP} \tag{1}$$

$$Recall = \frac{TP}{FN + TP} \tag{2}$$

where:

- TP refers to the area of the detected feature (e.g. wall) that is that feature (e.g. wall) in the ground truth.
- FP refers to the area of the detected feature that is not a feature in the ground truth.
- FN is the area that is the feature in the ground truth but is not detected as a wall by the proposed algorithm.

Finally, the Structural Similarity Index (*SSIM*) has also been calculated following the equation 13 of [21]. This is an image quality assessment to compare two images for structural information ranging from 0 (no similarity) to 1 (similar). More details can be found in [21].

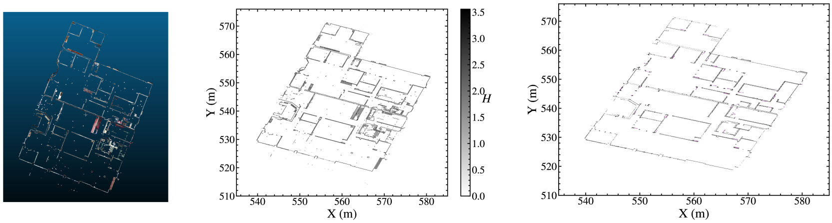

Figure 1: *Small building* of the *Scan-to-BIM* Challenge (see Section 3.2). *Left-panel:* Point Cloud 12_SmallBuilding_02_F1. *Middle-panel:* 2D histogram. Outcome from Sect. Algorithm 1. *Right-panel:* Floorplan estimation. Outcome from Sect. Algorithm 2 (see Sect. 3.2). Labels: walls (black), doors (purple), and stairs (gold).

4 Results

The proposed methodology presented in Sect. 3 has been applied to two different cases. They belong to the training set of *Scan-to-BIM* dataset. Both have ground truth annotations with three categories: wall, door and stair.

4.1 Small building

Figure 1 presents the first experiment. *Left-panel* shows the point cloud of the first floor of a small building with about 17 million data points. First of all, note that this point cloud does not follow the Manhattan layout, i.e. the orientation of the walls of the building does not follow a horizontal or vertical orientation [6]. Second, the data points do not have information of the ceiling or floor. Third, the content of clutter or noise is minimal. Therefore, this becomes an excellent study case.

Middle-panel is the outcome of applying the steps described in Algorithm 1 to the left-panel. It shows the 2D histogram where the maximum value of H is about $H_{max} = 3.5$. *Right-panel* has been constructed by applying the steps described in Algorithm 2 (see Sect. 3.2). There, the floorplan estimation of the *Small building* has been obtained with label annotations. Doors and stairs are rather difficult to retrieve.

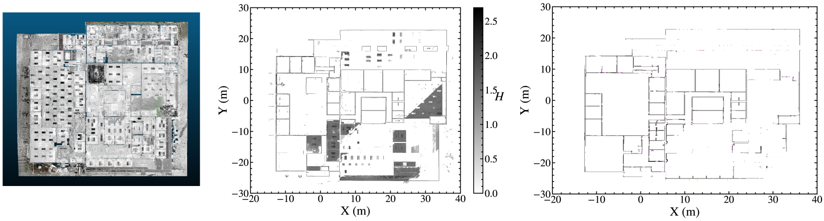

Figure 2: *Office Lab* of the *Scan-to-BIM* Challenge (see Section 3.2). *Left-panel:* Point Cloud 12_SmallBuilding_02_F1. *Middle-panel:* 2D histogram. Outcome from Sect. Algorithm 1. *Right-panel:* Floorplan estimation. Outcome from Sect. Algorithm 2 (see Sect. 3.2) Labels: walls (black), doors (purple), and stairs (gold).

Besides, considering the definition of *precision* and *recall* (see Sect. 3.3), for the feature wall then: *precision* = 1 and *recall* = 0.54 ($TP = 11527$, $FN = 9723$ and $FP = 0$).

4.2 Office Lab

Figure 2 presents the second case. Panel *a)* shows the Office Lab with about 120 million points. This case follows the Manhattan layout. However, it has information of the ceiling. This information needs to be removed. Therefore, this experiment constitutes a much more complex case to study.

Following the work of [16][4], we first proceed with an analysis of height to take out the ceiling as well as the clutter (see sections 3.1 and 3.2 f the mentioned paper). The point cloud is reduced to about six million points. Afterwards, we continue with the implementation of our framework. Panel *b)* shows the 2D histogram. The maximum value of H is about $H_{max} = 2.5$. Panel *c)* shows the floorplan estimation accounting for the label-categories. Once again doors and stairs are rather difficult to retrieve.

As for the metrics defined in Sect. 3.3, for the feature wall then: *precision* = 1 and *recall* = 0.43 ($TP = 16158$, $FN = 21853$ and $FP = 0$).

[4] Repository available in [22].

5 Discussion

5.1 Floorplan: Estimation vs. reconstruction

Figures 3 and 4 compare the floorplan estimation for the *Small building* and *Office Lab*, presented in Figs. 1 and 2, with the ground truth. Table 1 provides statistical insight to our findings. For the *Small building*, the *ratio* between the number of points of the estimated feature divided by the total number of points of the ground truth of that feature, i.e. wall, door and stair are: 54%, 40%, and 6%, respectively (see values in Table 1). Note that the ground truth presents a room around the coordinate $(X, Y) = (578, 556)$ (see Fig. 3 right-panel) that is not present at all in the original point cloud data (see left-panel of Fig. 3). This inconsistency is intrinsic to the dataset. Although it contributes to the discrepancy in our results, it does not explain the difference altogether.

The *ratio* for the *Office Lab* are: 43%, 26% and 4%, respectively. Due to a class imbalance[5], the identification of doors and stairs is limited. This is a well known issue in the literature, where it is common to provide floorplans purely based on walls, e.g. [15,16] (see also Sect. 5.2).

In our approach, the *FPs* are zero for both cases (see Sect. 3). Thus, the *ratio* and *recall* have the same values. The *SSIM* for the *Small building* is 0.91 and for the *Office Lab* is 0.86 (see Table 1), indicating that the floorplan estimation and ground truth, at least for walls, are similar.

5.2 Comparison to other methods

State-of-the-art (*SOTA*) approaches (i.e. computer vision or deep learning) make use of metrics such as *Intersection over Union* (*IoU*), *recall* and *precision*. For instance, the computer vision work of [16] presents great results in their experiments with a *precision* and *recall* over 90%. Similarly, the deep learning work *FloorPP-Net* [15] by using the *Scan-to-BIM* dataset reported a *precision* of 7%, *recall* of 39% and an *IoU* of 12% in the floorplan only based on walls (i.e. without

[5] $data(wall) \gg data(door)$ or $data(stair)$.

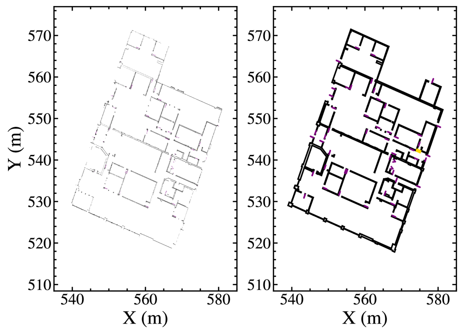

Figure 3: Floorplan estimation (left-panel) vs. the ground truth (right-panel) for the *Small building* presented in Sect. 4.1 where the walls (black), doors (purple) and stairs (gold) are shown.

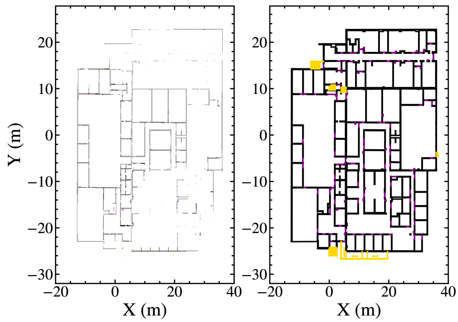

Figure 4: Floorplan estimation (left-panel) vs. the ground truth (right-panel) of the Office Lab presented in Sect. 4.2 where the walls (black), doors (purple) and stairs (gold) are shown.

including information of any other feature such as door or stair).

Computer vision and deep learning are still improving not only in the automatic detection of walls but also in the detection of doors and stairs. However, it is important to note, the calculation could take up to several minutes to compute and often require a special graphical processing unit (GPU).

Table 1: Columns 1-4: Number (#) of data points per class category for the floorplan estimation (our method) vs. Ground truth. The ratio $\left(\frac{\#points\ Estimation}{\#points\ Ground\ Truth}\right)$. Column 5: Result of the Structural Similarity Index (*SSIM*) between the estimation and ground truth.

–	**Wall** (# points)	**Door** (# points)	**Stair** (# points)	***SSIM*** [0,1]
Small building				
Estimation	11527	724	4	0.91
Ground Truth	21250	1791	71	
ratio	54 %	40 %	6 %	–
Office Lab				
Estimation	16158	818	100	0.86
Ground Truth	38011	3181	2698	
ratio	43 %	26 %	4 %	–

Comparing this work to *SOTA*, and by considering that the metric *SSIM* can be understood as a proxy of *IoU*, the results of this work compare well (see also *recall*). Besides, it can be seen as an alternative method to estimate floorplan of buildings. It constitutes a light implementation (i.e. CPU-based), which provides fast and fair floorplan estimation for the *Scan-to-BIM* dataset. By virtue of its simplicity, in the future, its implementation will be extended to other datasets .

6 Conclusion

Based on image processing techniques, we develop an alternative method to estimate floorplan of buildings in the *Scan-to-BIM* dataset. Our method does not aim to replace state-of-the-art approaches, which are more elaborate and robust. It, however, provides a fair automatic floorplan estimation, which may lead to the reconstruction of floorplans.

References

1. "Directive of the european parliament and of the council on the resilience of critical entities," https://eur-lex.europa.eu/resource.html?uri=cellar:

74d1acf7-3f94-11eb-b27b-01aa75ed71a1.0001.02/DOC_1&format=PDF, accessed: 2022-10-03.

2. "Dlr-pi," https://www.dlr.de/pi/en/desktopdefault.aspx/, accessed: 2022-10-03.
3. A. A. Aibinu and E. Papadonikolaki, "Conceptualizing and operationalizing team task interdependences: Bim implementation assessment using effort distribution analytics," *Construction Management and Economics*, vol. 38, no. 5, pp. 420–446, 2020. [Online]. Available: https://doi.org/10.1080/01446193.2019.1623409
4. C. Wang, K. Yong, and C. Kim, "Automatic bim component extraction from point clouds of existing buildings for sustainability applications," vol. 56. Autom. Constr., 2015, pp. 1–13.
5. R. Volk, J. Stengel, and F. Schultmann, "Building information modeling (bim) for existing buildings – literature review and future needs," vol. 38. Autom. Constr., 2014, pp. 109–127.
6. I. Hanagnostopoulos, V. Patraucean, I. Brilakis, and P. Vela, "Detection of walls, floors and ceilings in point cloud data," *Construction Research Congress 2016*, pp. 2302–2311, 2016.
7. Z. Ge, S. Liu, F. Wang, Z. Li, and J. Sun, "YOLOX: exceeding YOLO series in 2021," *CoRR*, vol. abs/2107.08430, 2021. [Online]. Available: https://arxiv.org/abs/2107.08430
8. C. Liu, K. Kim, J. Gu, Y. Furukawa, and J. Kautz, "Planercnn: 3d plane detection and reconstruction from a single image," *CoRR*, vol. abs/1812.04072, 2018. [Online]. Available: http://arxiv.org/abs/1812.04072
9. C. Zou, A. Colburn, Q. Shan, and D. Hoiem, "Layoutnet: Reconstructing the 3d room layout from a single RGB image," *CoRR*, vol. abs/1803.08999, 2018. [Online]. Available: http://arxiv.org/abs/1803.08999
10. C. Sun, C. Hsiao, M. Sun, and H. Chen, "Horizonnet: Learning room layout with 1d representation and pano stretch data augmentation," *CoRR*, vol. abs/1901.03861, 2019. [Online]. Available: http://arxiv.org/abs/1901.03861
11. C. Zou, J. Su, C. Peng, A. Colburn, Q. Shan, P. Wonka, H. Chu, and D. Hoiem, "3d manhattan room layout reconstruction from a single 360 image," *CoRR*, vol. abs/1910.04099, 2019. [Online]. Available: http://arxiv.org/abs/1910.04099
12. C. Yang, J. Zheng, X. Dai, R. Tang, Y. Ma, and X. Yuan, "Learning to reconstruct 3d non-cuboid room layout from a single

RGB image," *CoRR*, vol. abs/2104.07986, 2021. [Online]. Available: https://arxiv.org/abs/2104.07986

13. C. Liu, J. Wu, and Y. Furukawa, "Floornet: A unified framework for floorplan reconstruction from 3d scans," in *Computer Vision – ECCV 2018*, V. Ferrari, M. Hebert, C. Sminchisescu, and Y. Weiss, Eds. Cham: Springer International Publishing, 2018, pp. 203–219.
14. J. Han, M. Rong, H. Jiang, H. Liu, and S. Shen, "Vectorized indoor surface reconstruction from 3d point cloud with multistep 2d optimization," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 177, pp. 57–74, 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0924271621001222
15. Y. Wu and F. Xue, "Floorpp-net: Reconstructing floor plans using point pillars for scan-to-bim," 2021.
16. U. Gankhuyag and J.-H. Han, "Automatic 2d floorplan cad generation from 3d point clouds," *Applied Sciences*, vol. 10, no. 8, 2020. [Online]. Available: https://www.mdpi.com/2076-3417/10/8/2817
17. J. Canny, "A computational approach to edge detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. PAMI-8, no. 6, pp. 679–698, 1986.
18. H. Cantzler, "Random sample consensus (ransac)," *Institute for Perception, Action and Behaviour, Division of Informatics, University of Edinburgh*, 1981.
19. "1st international scan-to-bim challenge," https://cv4aec.github.io/, accessed: 2022-02-04.
20. "Cvpr virtual 2021," https://cvpr2021.thecvf.com, accessed: 2022-10-07.
21. Z. Wang, A. Bovik, H. Sheikh, and E. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600–612, 2004.
22. "Repository of: Automatic 2d floorplan cad generation from 3d point clouds," https://github.com/joyjo/to-generate-2D-floorplan-CAD-from-3D-point-clouds, accessed: 2022-02-04.