2022 The 22nd International Conference on Control, Automation and Systems (ICCAS 2022)
BEXCO, Busan, Korea, Nov. 27~Dec. 01, 2022

Parsing Indoor Manhattan Scenes Using Four-Point LiDAR on a Micro UAV

Eunju Jeong[1], Suyoung Kang[2], Daekyeong Lee[1], and Pyojin Kim[1*]

[1]Department of Mechanical Systems Engineering
[2]Department of Electronics Engineering
Sookmyung Women's University, Seoul, South Korea
{eunju0316,1913084,swbubl,pjinkim}@sookmyung.ac.kr

Abstract: We propose the first 3D mapping algorithm using four-point LiDAR for a micro unmanned aerial vehicle (UAV). Existing mapping approaches depend on 360° 2D laser scanner and RGB-D camera, which are unsuitable for micro UAV with a small payload. The proposed method builds a 3D structure map with an accumulated point cloud obtained from low-cost and lightweight four ToF sensors suitable for micro UAV in four directions: front, back, left, and right. The noise of range measurement by the low-cost ToF sensor and inaccurate 6-DoF pose estimation of Crazyflie make a noisy point cloud. We overcome these problems by utilizing the geometric constraints of the interior structures, the Manhattan world (MW), and the proposed method successfully parse the floor plan of the Manhattan scenes. We evaluate the proposed method in various MW structures and demonstrate that the proposed method produces comparable results to the ROS Gmapping algorithm, which uses a 360° 2D laser scanner.

Keywords: Sparse Sensing, Four-Point LiDAR, Micro UAV, Parsing, Manhattan World

1. INTRODUCTION

3D mapping of the indoor environment is essential in mobile robotics, such as indoor autonomous driving. Micro UAV is a suitable platform for indoor mapping due to its small size and wide range of applications. Multi-agent micro UAVs rather than one large UAV can speed up indoor mapping and navigation [1]. Previous SLAM methods use a 2D laser scanner with a wide range and consistent reliability and accuracy, but it is not available for micro UAVs of small payloads.

The low-cost and lightweight time-of-Flight (ToF) sensor is suitable for micro UAV, but has a short measuring range. Micro UAV with low-cost ToF sensors generate a noisy point cloud due to the inaccurate 6-DoF pose estimation of micro UAV and the range measurement limitation of ToF sensors. It is challenging to build a reliable map with a micro UAV with a few ToF sensors.

To the best of our knowledge, no studies have performed 3D mapping with sparse range sensing with micro UAV. We propose the first 3D mapping algorithm using an accumulated point cloud obtained by the four-point LiDAR. Since, in most indoor structures, the adjacent walls are orthogonal to each other, and the ceiling and floor are parallel, we apply Manhattan World (MW) Assumption [2] to overcome the limitation of low-cost ToF sensors and inaccurate estimation of the 6-DoF pose of a micro UAV.

We employ Crazyflie 2.1, an open-source micro quadcopter of Bitcraze, a multi-ranger deck with five ToF sensors (up, front, back, left, and right), and a flow deck with an optical flow sensor, enabling 6-DoF state estimation on the Crazyflie firmware. The proposed method uses only four ToF sensors of the multi-ranger deck to generate the point clouds: front, back, left, and right (four-point LiDAR), as shown in Fig. 1 (left). We only use the up and down ToF sensor to measure the distance between the ceiling and the floor, not to generate the point cloud.

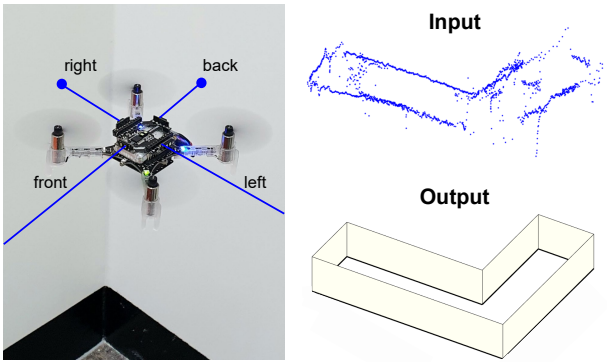

Fig. 1. A micro UAV measuring the ranges of the surrounding environment with four-point LiDAR (left). Conversion of ranges to point cloud expressed as the global frame (top right). The output of the proposed method, a 3D structure map (bottom right).

The proposed method first converts the accumulated range measurements from the four-point LiDAR into a 3D point cloud expressed in a global frame, as shown in Fig. 1 (top right), and projects the 3D point cloud into a plane. We cluster the point cloud corresponding to the wall (wall clustering) and fit lines with the line model estimated by RANSAC. Then, we perform structure alignment with the MW assumption and generate a 3D map as shown in Fig. 1 (bottom right) using the ceiling height obtained by the up and down ToF sensors.

Our method produces comparable maps to the ROS Gmapping algorithm [3], which uses the 2D laser scanner. The main contributions of this paper are as follows:

- We present the first 3D mapping algorithm that parses the Manhattan structures using low-cost four-point LiDAR suitable for micro UAV.

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korean government (MSIT) (No. NRF-2021R1F1A1061397), and the Korea Foundation for Women In Science (WISET) (No. WISET-2022-075). * Corresponding author

• We evaluate the proposed method with other algorithms on the various Manhattan world structures from room scale to building scale.

2. RELATED WORK

Most existing studies have used a 2D laser scanner or RGB-D camera for mapping, and several studies use a 2D laser scanner and RGB-D camera together. [4] produces a floor plan with a point cloud obtained by a 2D laser scanner, performs point cloud clustering, and then changes it to a straight line similar to our method. However, since [4] does not combine clusters corresponding to the same wall, it cannot represent the floor plan as a single line per a single wall. In [5] and [6], they build an interior 3D map using a 2D laser scanner, and [6] uses an intensity camera with the 2D laser scanner. Unlike [5] and [6], the 3D structure map built by our method can recover ceiling height through distance obtained by up and down ToF sensors. We build a building-scale map using low-cost, lightweight four-point LiDAR, whereas [7] uses a 2D laser scanner and RGB-D camera to construct a room-scale floor plan, and [8] takes aligned panorama RGB-D scans as input to build a building-scale floor plan. [9] uses a Manhattan structure similar to the proposed method. While [9] presents a parsing algorithm in the MW structure of an indoor scene from a single RGB-D frame, we parse the indoor MW structure using the low-cost four-point LiDAR.

Recent robotics research using micro UAV has been actively studied in [1], [10], [11], and [12]. [1] presents an indoor navigation algorithm of multiple Crazyflies, and [10] performs a swarm of 49 Crazyflies, not mapping. [10] is only possible in a space with a motion capture system, so they always have to attach markers to the Crazyflie. However, we do not have to attach markers and can build a reliable 3D structure map in an unknown indoor environment where the global navigation satellite system (GNSS) is unavailable. The most relevant research to the proposed method is [11]. [11] presents a 2D mapping algorithm of the open-source micro UAV, Crazyflie 2.0 [12] with a custom-made range deck. They use the same VL53L1x ToF sensor as ours, but they use 13 ToF sensors while we use 6 ToF sensors.

3. PROPOSED METHOD

The overall procedure of the proposed method is shown in Fig. 3. We overcome the lack of enough points due to sparse sensing, the short range of the ToF sensors, and inaccurate estimation of the position and orientation of Crazyflie by exploiting geometric constraints where most of the interior structures of buildings are orthogonal and parallel.

3.1 Converting Ranges to 3D Point Cloud

Fig. 2 shows the accuracy of the ToF sensor of the four-point LiDAR. Measurement resolution is 10cm for distances up to 3m and a 20cm step up to 5m. The difference between the distance measured with a ruler and the distance measured with the ToF sensor is about 0.59cm up to 2.5m on average, so we treat the maximum reliable range as 2.5m.

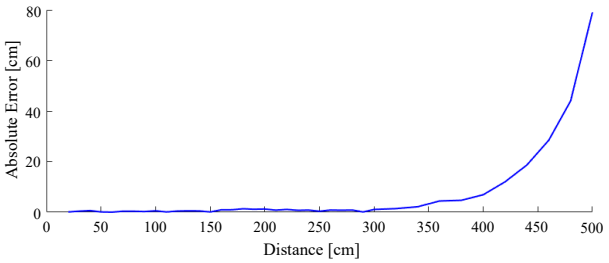

Fig. 2. The comparison of LiDAR measurement with a ruler. The average absolute error up to 2.5m is 0.59cm. We assume the reliable maximum range is 2.5m and fix the range of usable laser measurement within 2.5m.

To plot the 3D point cloud, it is essential to convert each range measurement from a body frame to a global frame. We transform the coordinate of the ranges within 2.5 m obtained by the four ToF sensors (front, back, left, and right) to Cartesian coordinates shown in Fig. 3 (a). We do not convert the up and down distances as a 3D point cloud but only use them to measure the height between the ceiling and the floor to make a 3D map in the final step of the proposed method.

$$\begin{bmatrix} X_g \\ Y_g \\ Z_g \\ 1 \end{bmatrix} = T_{gb} \begin{bmatrix} X_b \\ Y_b \\ Z_b \\ 1 \end{bmatrix} \tag{1}$$

$$T_{gb} = \begin{bmatrix} R_{gb} & \mathbf{t} \\ \mathbf{0_{1\times 3}} & 1 \end{bmatrix} \in \mathbb{R}^{4\times 4} \quad R_{gb} \in \mathbb{R}^{3\times 3} \tag{2}$$

where X_g, Y_g, and Z_g are the 3D point coordinates described in the global frame, and X_b, Y_b, and Z_b are the 3D point coordinates described in the body frame of the Crazyflie. $T_{gb} \in SE(3)$ is 4×4 rigid body transformation matrix which represents 6-DoF pose of the Crazyflie. $\boldsymbol{t} \in \mathbb{R}^3$ and $R_{gb} \in SO(3)$ are the estimated 3-DoF translational motion and 3-DoF rotational motion obtained from an optical flow sensor of flow deck, and IMU embedded to Crazyflie.

3.2 2D Projection and Eliminating Noise

We remove noisy points and remain meaningful points for successful wall clustering, as shown in Fig. 3 (b). When the Crazyflie reaches a certain altitude h from the floor, it moves only on a 2D plane and measures the distance from the surrounding obstacles. We assume that all points obtained from the flight of Crazyflie are on the same plane with the z-axis coordinates h, and we project all points to a horizontal 2D plane.

Noisy points occur due to the inaccurate range measurement of the ToF sensor and inaccurate positioning of the Crazyflie, as shown in Fig. 4 (left). These noisy points

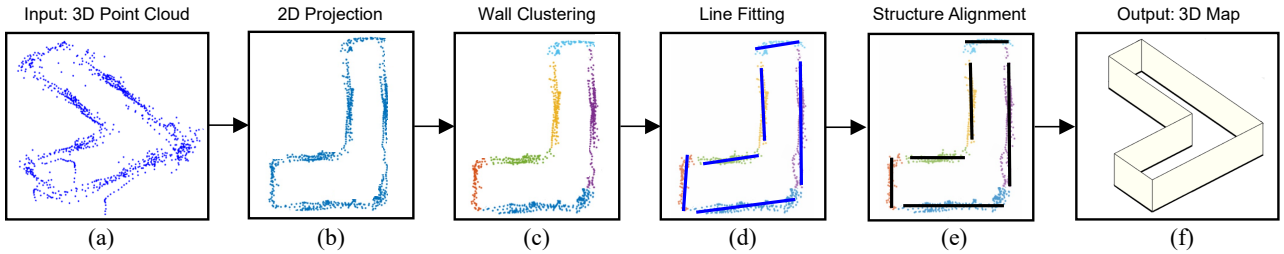

Fig. 3. The pipeline of the proposed method. Conversion of range measurement to point cloud (a), 2D projection and eliminating the noise of point cloud (b), wall clustering with Hierarchical clustering (c), line fitting of each cluster with line model estimated by RANSAC (d), alignment of lines with respect to the Manhattan structures (e), building a 3D structure map by projecting the floor plan as measured ceiling height on the z-axis (f).

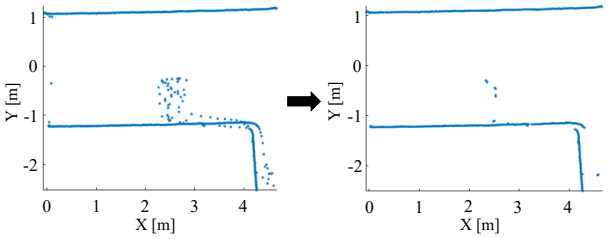

Fig. 4. The ToF sensor often measures inaccurate range, resulting in noise (left). We eliminate the noise by ignoring the points where the nearest Euclidean distance with another point is more than 5cm (right).

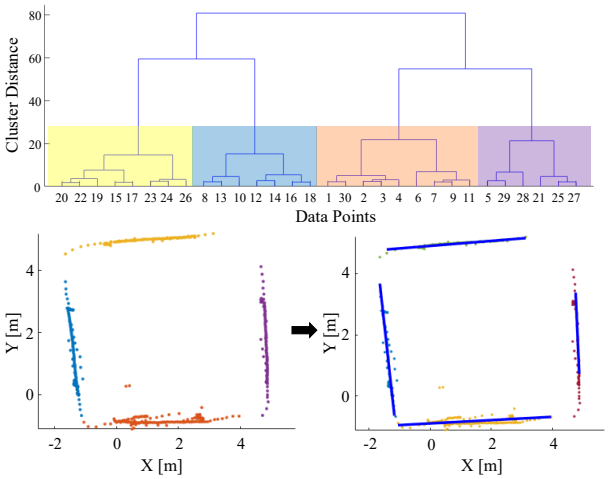

Fig. 5. Dendrogram of the clusters. The predetermined number of clusters is four (top). Cluster divided by Hierarchical clustering (bottom left), Line fitting for each cluster (bottom right).

are sparse, unlike the points generated for obstacles such as actual walls. If the distance from the nearest point in the point cloud is more than 5cm, We treat them as noise and ignore them when we build a map. The proposed method can successfully remove almost all of the noisy points, as shown in Fig. 4 (right).

3.3 Wall Clustering and Line Fitting

The criterion for point cloud clustering is the Euclidean distance between points. We select Hierarchical clustering (HC) [13] for wall clustering since HC clusters point cloud based on the distance between points. HC treats all points as different clusters and stores the distance between all clusters in advance. HC repeats two steps until all points become one cluster. First, find the two clusters with the closest distance. Second, combine the two found clusters into one cluster. Fig. 5 (top) is the hierarchical relationship between the clusters. We obtain a dominant line for each wall by determining the number of clusters, as shown in Fig. 5.

We fit lines using the line model estimated by RANSAC [14] for each cluster as shown in Fig. 3 (d). The RANSAC algorithm estimates the parameters of a model from data with high measurement noise. RANSAC is suitable for the point cloud obtained by sparse sensing due to the high measurement noise of the low-cost ToF sensor.

3.4 Structure Alignment in Manhattan World

The two adjacent walls are in an orthogonal relationship in most interior structures. We refit the slopes of lines to be orthogonal relationships using this MW geometric constraint as shown in Fig. 3 (e). The proposed method aligns the structure of the curved point cloud due to inaccurate 6-DoF pose estimation of micro UAV and sparse sensing.

We treat the slope of the cluster with the highest number of inlier points as the most reliable and the reference slope. We compare the slope of the reference with the slope of the other lines. If the angle difference between two lines is 45 degrees or more, we define them as an orthogonal relationship. If it is less than 45 degrees, we define them as a parallel relationship and refit the slope of the line.

If the offset between two parallel lines is less than 5 cm, We treat them as the same wall, combining the inlier points of the two clusters and representing them as a single line.

3.5 Completion of the Floor Plan and 3D Map

The proposed method generates a 3D map by extruding the z-axis direction by ceiling height on the floor plan. We can obtain the distance between the ceiling and the floor through the up and down ToF sensors attached to Crazyflie. Most of the ceiling and the floor are parallel, and the proposed method can build a 3D map.

If the refitted slope is less than 1, we draw a line by limiting the scope of the x-axis. If the refitted slope is greater than or equal to 1, we draw a line by limiting the

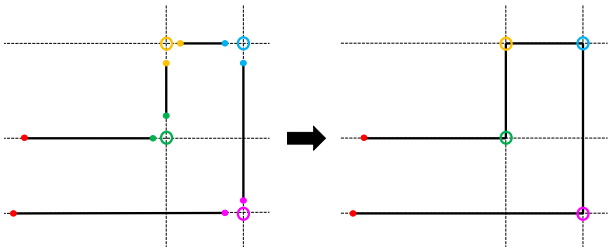

Fig. 6. The process of completing the floor plan. We represent the pairs of the nearest endpoints of each line in the same color. We extend two lines to the intersection if the relationship between the two lines corresponding to the nearest endpoints is orthogonal. If the relationship between the two lines is parallel, we do not extend them.

scope of the y-axis, as shown in Fig. 6 (left).

We represent the pairs of the nearest endpoints of each line in the same color, as shown in Fig. 6. The two lines of the orthogonal relationship corresponding to the closest endpoint pair are adjacent wall relationships. If the two lines corresponding to each nearest endpoint pair are in an orthogonal relationship, we replace the endpoints of each line with the intersecting point denoted by an unfilled circle. If the two nearest lines are parallel, we cannot obtain the intercepting point and do not change the endpoints, like the red-filled circle pair shown in Fig. 6.

4. EXPERIMENTS

We implement and test the proposed method in various real MW environments to evaluate its effectiveness. We experiment with five indoor MW structures that are easily encountered in typical indoor buildings, as shown in Fig. 9. The square room is a size of 5.6m×6m with no obstacles in the middle, and the L-shaped corridor and straight corridor are 1.76 m wide. In the straight corridor, we block the front and back of the open hallway with boxes, unlike the L-shaped corridor. Lounge 1 and lounge 2 are large spaces with no obstacles in the middle.

We use Crazyflie 2.1, a size of 9.2mm×9.2mm and a weight of 27g, and we use a multi-ranger deck with five VL53L1x ToF sensors for range measurement. To obtain the 6-DoF pose of Crazyflie, We use Flow deck v2 and an IMU sensor embedded in the Crazyflie. The flow deck has one VL53L1x ToF sensor in the down direction. We attach a multi-ranger deck at the top of the Crazyflie and a flow deck at the bottom. We control the Crazyflie with the laptop keyboard using the Python API and Crazyradio PA. We log the position and orientation of Crazyflie and six-way distances from the six ToF sensors at 10 Hz and continuously accumulate each data. We use four ToF sensors of the multi-ranger deck to create the point cloud: front, back, left, and right (four-point LiDAR). We use the up and down ToF sensor to measure the distance between the ceiling and the floor, not to generate the point cloud.

We compare the proposed method with the two other methods. Log-odds OGM, an occupancy grid mapping (OGM) algorithm proposed in [15], uses the four-point LiDAR and Crazyflie. The proposed method uses the Manhattan world assumption to correct the distorted raw point cloud, whereas Log-odds OGM does not correct the distorted structure caused by inaccurate range measurements of the ToF sensor and inaccurate 6-DoF state estimation from the flow deck. To demonstrate that the floor plans built by the proposed method using four-point LiDAR and those produced by a 2D laser scanner are comparable, we utilize the ROS Gmapping algorithm[3] using 2D laser scanner LDS-02 on TurtleBot3-Burger. Since the 2D laser scanner is not attachable to the Crazyflie due to the small payload, we apply the 2D laser scanner LDS-02 to the TurtleBot3-Burger to build a floor plan using ROS Gmapping. We control the TurtleBot3-Burger with the laptop keyboard. We use range measurement within 2.5m for a reliable map for all three algorithms.

Fig. 8 shows samples of the Manhattan scenes that the proposed method parses accurately. The black dots and lines on each map are obstacles, such as a wall that each algorithm recognizes. We use the four-way distance to make the point cloud in the square room, and the two-way distance of left and right in four other MW environments since the ToF sensor generates severe noise data in the front-rear direction in wide-open spaces. The linear and angular velocities of Crazyflie are 0.3m/s, and the linear and angular velocities of TurtleBot3-Burger are 0.05m/s and 0.1m/s, respectively.

Our method best represents the orthogonal structure of the interior wall compared to the other two algorithms. Our method removes the noisy points caused by the inaccurate range measurements of low-cost and lightweight ToF sensors, whereas Log-odds OGM fails to remove noise caught in the empty space and recognizes it as an obstacle. The inaccurate 6-DoF estimation of Crazyflie calculated by IMU embedded on Crazyflie and the optical flow sensor of the flow deck causes the distorted point cloud. The estimated trajectory of the Crazyflie tends to bend when the Crazyflie rotates or flies a long distance, such as in the L-shaped corridor and the straight corridor. The proposed method corrects the distorted structure using the MW assumption and builds an accurate indoor map, while Log-odds OGM fails to fix the distorted form of the point cloud, and the proposed method produces comparable results with ROS Gmapping using a 2D laser scanner.

We overlay the mapping results with the proposed method on the actual floor plan, as shown in Fig. 9. Red is the actual floor plan, and black is ours. We define the error metric as the subtraction of the wall length of our floor plan from the actual wall length measured directly by the ruler and obtain errors for all walls as shown in Fig. 7. For unblocked structures such as the L-shaped corridor, We calculate errors for the width and length of the corridor. As shown in Table. 1, the average length error in the five MW structures is 0.457m, and we obtain

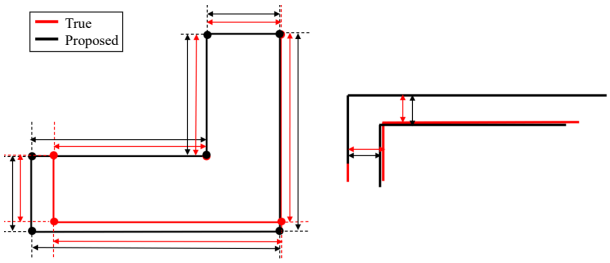

Fig. 7. Error metric applied to lounge 1 (left) and the L-shaped corridor (right). The error metric in an open corridor is the average difference in width. Filled circles are the corners of each floor plan.

Table 1. The quantitative evaluation results and flight time of Crazyflie in each of the five MW structures.

	Length Error (m)	Flight Time of Crazyflie
Square room	0.232	46 sec
L-shaped corridor	0.17	1 min 46 sec
Straight corridor	0.744	1 min 30 sec
Lounge 1	0.587	1 min 18 sec
Lounge 2	0.551	2 min 14 sec

the 6-DoF pose data of Crazyflie and distance data to the surrounding walls during an average flight time of 1 min 30 sec. Given that this is a building scale, the proposed method produces accurate and reliable floor plans despite using a point cloud obtained during a short flight time with a sparse four-point LiDAR.

5. CONCLUSION

We propose the first approach to parsing the Manhattan scenes using the four-point LiDAR on a micro UAV. Most indoor structures have adjacent walls orthogonal to each other, Manhattan World. We overcome the noise of range measurement by the low-cost ToF sensor and inaccurate 6-DoF pose estimation of Crazyflie by effectively utilizing the Manhattan world assumption. Experiments in various MW structures demonstrate the superior performance of the proposed method, and our approach produces comparable results to the ROS Gmapping algorithm, which uses a 2D laser scanner. We extend the scope of the 3D mapping platform by enabling 3D mapping, previously only possible with large and heavy 2D laser scanners, with a four-point LiDAR suitable for a micro UAV. We expect that it will be able to scale faster and broader in the future using multi-agent micro UAVs and will be helpful in SLAM using micro UAV.

REFERENCES

[1] K. McGuire, C. De Wagter, K. Tuyls, H. Kappen, and G. C. de Croon, "Minimal navigation solution for a swarm of tiny flying robots to explore an unknown environment," *Science Robotics*, vol. 4, no. 35, p. eaaw9710, 2019.

[2] J. M. Coughlan and A. L. Yuille, "Manhattan world: Compass direction from a single image by bayesian inference," in *Proceedings of the seventh IEEE international conference on computer vision*, vol. 2. IEEE, 1999, pp. 941–947.

[3] G. Grisetti, C. Stachniss, and W. Burgard, "Improved techniques for grid mapping with rao-blackwellized particle filters," *IEEE transactions on Robotics*, vol. 23, no. 1, pp. 34–46, 2007.

[4] P. Słowak and P. Kaniewski, "Lidar-based slam implementation using kalman filter," in *Radioelectronic Systems Conference 2019*, vol. 11442. SPIE, 2020, pp. 198–207.

[5] A. Adan and D. Huber, "3d reconstruction of interior wall surfaces under occlusion and clutter," in *2011 International Conference on 3D Imaging, Modeling, Processing, Visualization and Transmission*. IEEE, 2011, pp. 275–281.

[6] J. Li and R. Stevenson, "2d lidar and camera fusion using motion cues for indoor layout estimation," in *2021 IEEE 24th International Conference on Information Fusion (FUSION)*. IEEE, 2021, pp. 1–6.

[7] X. Zhang, J. Lai, D. Xu, H. Li, and M. Fu, "2d lidar-based slam and path planning for indoor rescue using mobile robots," *Journal of Advanced Transportation*, vol. 2020, 2020.

[8] J. Chen, C. Liu, J. Wu, and Y. Furukawa, "Floor-sp: Inverse cad for floorplans by sequential room-wise shortest path," in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2019, pp. 2661–2670.

[9] C. J. Taylor and A. Cowley, "Parsing indoor scenes using rgb-d imagery," in *Robotics: Science and Systems*, vol. 8, 2013, pp. 401–408.

[10] J. A. Preiss, W. Honig, G. S. Sukhatme, and N. Ayanian, "Crazyswarm: A large nano-quadcopter swarm," in *2017 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2017, pp. 3299–3304.

[11] T. Chathuranga, M. Padmal, D. Bibile, P. Jayasekara, and N. Kottege, "Sensor deck development for sparse localization and mapping for micro uavs to assist in disaster response."

[12] W. Giernacki, M. Skwierczyński, W. Witwicki, P. Wroński, and P. Kozierski, "Crazyflie 2.0 quadrotor as a platform for research and education in robotics and control engineering," in *2017 22nd International Conference on Methods and Models in Automation and Robotics (MMAR)*. IEEE, 2017, pp. 37–42.

[13] S. C. Johnson, "Hierarchical clustering schemes," *Psychometrika*, vol. 32, no. 3, pp. 241–254, 1967.

[14] M. A. Fischler and R. C. Bolles, "Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography," *Communications of the ACM*, vol. 24, no. 6, pp. 381–395, 1981.

[15] P. Anantharam. (2020, Dec.) Simultaneous localization and mapping (slam) with crazyflie. [Online]. Available: http://https://pramodatre.github.io/2020/12/21/slam-with-crazyflie/#fn:1

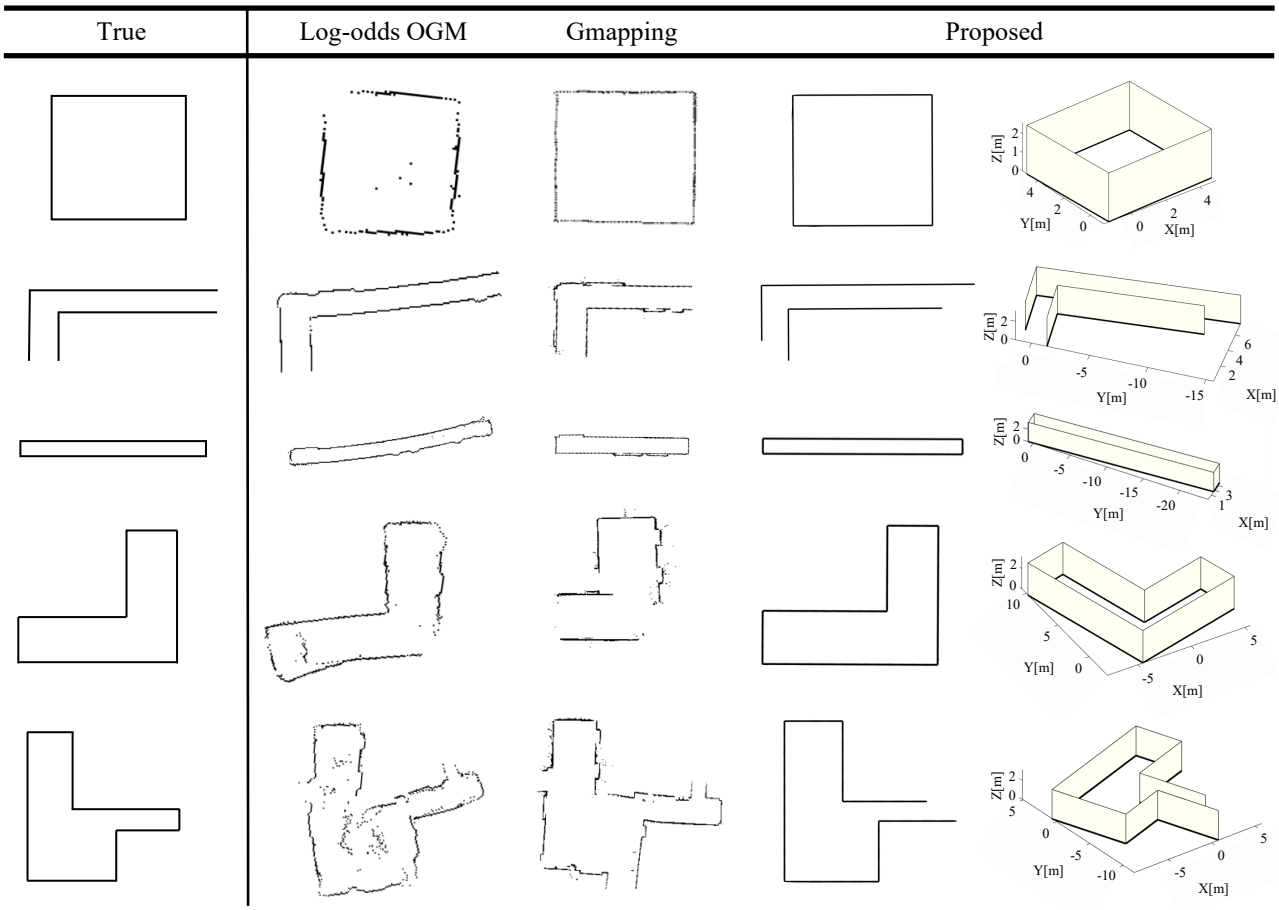

Fig. 8. Qualitative comparisons of Log-odds OGM[15], Gmapping[3], and the proposed method. The leftmost is the ground truth floor plan, and the rightmost is the floor plan and 3D structure map built by the proposed method.

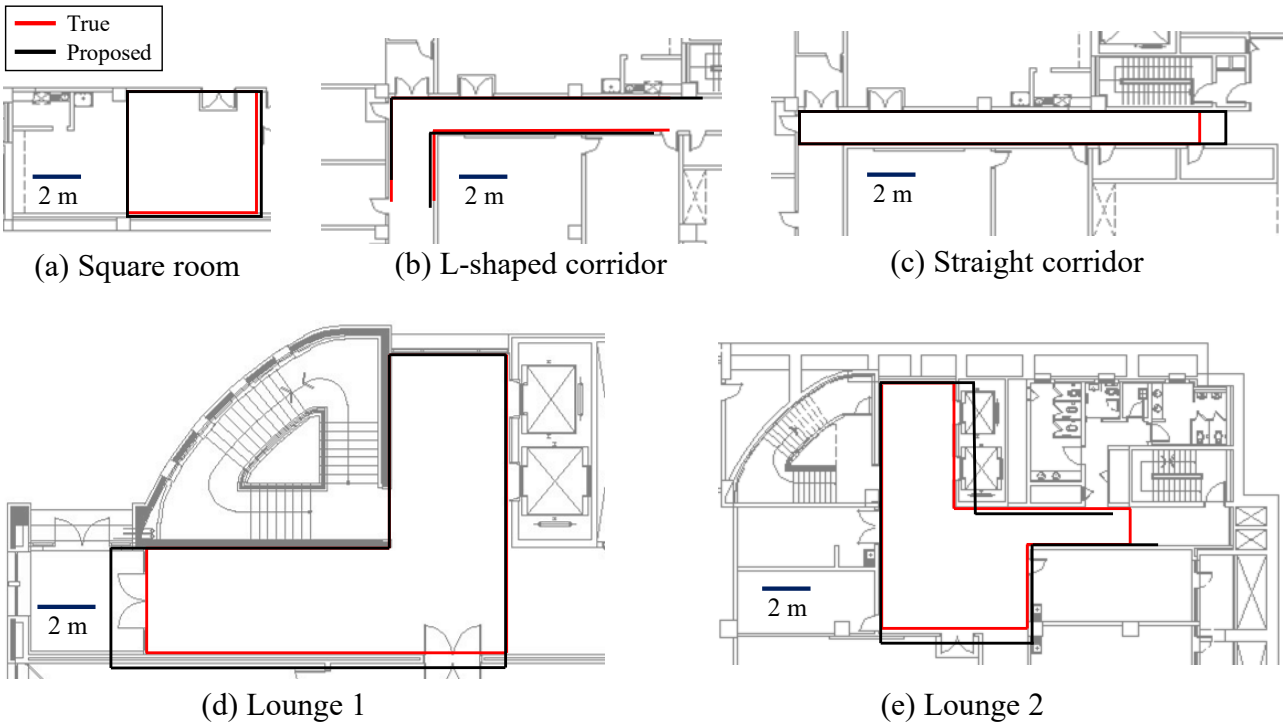

Fig. 9. Ground truth floor plan (red line) and the floor plan of the proposed method (black line). We overlap the floor plan of the proposed method with the floor plan of the buildings to check the accuracy of the proposed method.