Robot-Assisted 3D Scene Reconstruction Using Fixed FOV LiDAR-Based Depth Camera

Vincent J.L. Gan[1,2]

[1] Department of the Built Environment, National University of Singapore, Singapore
[2] Centre of 5G Digital Building Technology, National University of Singapore, Singapore
vincent.gan@nus.edu.sg (correspondence)

Abstract –

This study introduces an innovative robot-assisted 3D scene reconstruction approach featuring a fixed field-of-view (FOV) LiDAR-based depth camera. The approach integrates a quadruped robot with LiDAR-camera for depth measurement and for capturing RGB colour and texture. The RGB-D information is then used in SLAM to enable detailed 3D scene reconstruction and representation of as-built conditions. This hybrid system reduces the costs and operational limitations of traditional terrestrial laser scanning (TLS) while maintaining decent accuracy for downstream tasks such as condition monitoring and scan-to-BIM. By enhancing scanning productivity, this approach can be further integrated with automated point cloud semantic instance segmentation techniques for a streamlined workflow in high-fidelity geometric modelling, advancing robotic 3D BIM reconstruction and contributing to efficient and accurate digital construction applications.

Keywords –

Ground Robot; LiDAR; Image-based Sensing; 3D Scene Reconstruction; Scanning

1 Introduction

There is a growing momentum to harness automated, intelligent technologies to capture and reconstruct 3D digital representations of the built environment, translating the information into virtual platforms. The digital representations (such as Building Information Models) are useful in the architecture, engineering and construction (AEC) industry due to their capability to enhance communication and information exchange between project stakeholders, facilitating lifecycle management of built assets [2]. Traditionally, this process relies on accurate scanning to capture 3D point cloud data reflecting the as-built condition of the structure. This is followed by extensive manual processing of the scan point cloud data to generate the 3D scene or geometric model [3]. This workflow is not only labour-intensive and time-consuming, but also prone to human errors [4].

Several studies have explored the use of Light Detection and Ranging (LiDAR) integrated with robotic platforms, such as ground quadruped robots and wheeled robots, for 3D scene reconstruction of built environments [5-9]. For example, ground robots equipped with 3D LiDAR sensors can autonomously plan routes, navigate to specific targets, and collect point cloud data. The integration of BIM and IndoorGML were presented to optimise quadruped robot navigation for 3D scanning, achieving 70–90% coverage with high accuracy (0.006–0.021 mm), reducing human effort, and outperforming heuristic algorithms in efficiency, coverage, and scan planning [10, 11]. Likewise, an automatic 3D scaffold reconstruction using robot-dog-acquired point clouds was also investigated, achieving 90.84% F1 in semantic segmentation and generating 3D CAD models through deep learning for improved safety management [12]. For building façades [13], ground robot was also used for automated 3D scene reconstruction and visual inspection, demonstrating its growing potential in addressing unique challenges in complex environments. Researchers are actively exploring Simultaneous Localisation and Mapping (SLAM) as a viable option to 3D scene reconstruction. SLAM enables real-time mapping and localisation by allowing a robot or sensor system to build a spatial map of its environment while simultaneously determining its position within that map [1]. In one of the studies, an autonomous mobile robot using 2D SLAM and real-time 3D point cloud mapping was introduced [14] to navigate construction sites, optimise scan positions, and capture RGB-mapped point clouds, addressing occlusions and reducing time for geometric reconstruction.

Nowadays, unmanned aerial vehicles (UAVs) have been utilised for tasks such as surveying outdoor areas, which collect data more efficiently than ground robots. Collaborations between UAVs and ground robots have shown potential to overcome some limitations. Kandath et al. [15] proposed a system where UAVs enhance ground robot navigation by providing real-time obstacle location data in indoor environments. Furthermore, Kim et al. [16] introduced UAV-based 3D terrain mapping and generation. Peterson et al. [17] also developed an integrated robotic system with ground robots with drones, significantly enhancing the navigation and exploration for outdoor settings.

An example of multi-agent robotic system (MARS) in previous research is shown in Figure 1. The UAV agent is a lightweight drone equipped with an onboard camera, Wi-Fi module, and open-source controller. The UAV can capture 5MP images or 720p live video streams, transmitting data to a processing unit for tasks like path planning. It has a flight time of 8 minutes and 30 seconds, making it ideal for short inspection missions. The ground robot features a wheeled chassis with an intelligent controller and modules for human detection and avoidance. Designed for heavy payloads, the ground robot supports advanced sensing devices and extended operation times.

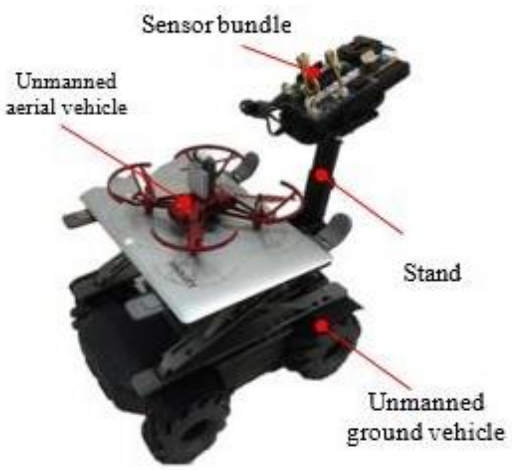


Figure 1. Ground robot equipped with UAV (configuration adapted from Reference [1])

In collaborative operation, the ground robot inspects surroundings and streams video data. As Figure 2 shows, the ground robot carries the UAV and follows a predefined path, capturing facility data and transmitting it to the mediating agent. The mediating agent directs the UAV when obstacles block the ground robot. The mediating agent calculates the optimal UAV trajectories for inaccessible areas and controls the UAV in navigating and scanning the areas that are inaccessible for the ground robot. Results validate the system's feasibility for indoor inspections in cluttered environments. With potential semantic navigation capabilities, ground robots and UAVs can further enhance their ability to interpret the built environment and navigate indoors. This advancement enables efficient and robust 3D scanning by intelligently interpreting and adapting to complex surroundings, ensuring comprehensive spatial data acquisition in challenging environment [18]. Previous research has highlighted the evolving role of ground robots and UAVs in advancing automated data capture in AEC projects.

Nevertheless, research in this field remains in progress with challenges persisting in achieving high-quality scanning outcomes [19]. In particular, significant challenges exist in achieving optimal scanning quality and cost efficiency. Traditional methods such as terrestrial laser scanning (TLS), while highly accurate, are often expensive and cumbersome, especially for large scenes where complex feature mapping requires manual intervention. Current robot-assisted systems, integrated with SLAM algorithms, show good potential for navigating complex environments while effectively capturing data. However, they often struggle to produce dense, high-quality 3D measurements. Furthermore, the use of cost-effective sensors to capture both depth and RGB information in a dense format remains underexplored, particularly without compromising depth accuracy.

This paper investigates a robot-assisted 3D scene reconstruction approach featuring a fixed field-of-view (FOV) LiDAR-based depth camera. By integrating a quadruped robot with this sensor, the approach leverages LiDAR for depth measurement and an integrated camera for capturing RGB, delivering detailed 3D representations of as-built conditions. This approach addresses the cost barriers of traditional TLS by employing an affordable sensor while preserving measurement accuracy. The quadruped robot's manoeuvrability enables it to efficiently scan diverse and complex structural areas, even in cluttered environments. By combining cost-efficient sensors with robotic mobility, this approach not only advances the field of robot-assisted 3D reconstruction but also opens new possibilities for applications such as smart city modelling, continuous monitoring, surveillance, and facilities management, providing a scalable and practical solution for the built environment.

Figure 2. Ground robot equipped with UAV for automated robotic inspection (configuration adapted from Reference [1])

2 Methodology

To achieve complete and accurate point cloud reconstruction that reflects as-built conditions, this study emphasises the integration of a ground quadruped robot with a fixed FOV LiDAR-based depth camera. This integration ensures robust spatial data acquisition while addressing scanning complex environments. To validate the proposed approach, an experiment is conducted on the NUS campus. The scan data are further analysed using two metrics: the completeness and accuracy of the point clouds.

The results provide insights into the feasibility and effectiveness of using quadruped robot-based scanning for detailed 3D scene reconstruction. The methodology and experimental setup are detailed in the subsequent sections.

2.1 Quadruped Robot-assisted Scanning

Figure 3 illustrates the setup of the quadruped robot, featuring an Intel RealSense L515 LiDAR-based depth camera mounted on top. The quadruped robot incorporates high-torque-density electric motors at each leg's knee, thigh, and hip joints, enabling precise control over its movements. This advanced actuation mechanism allows the robot to perform various dynamic actions, such as walking, jumping, and executing pitch, roll, and yaw motions. These capabilities make it well-suited for operating efficiently in confined and cluttered environments for 3D scanning.

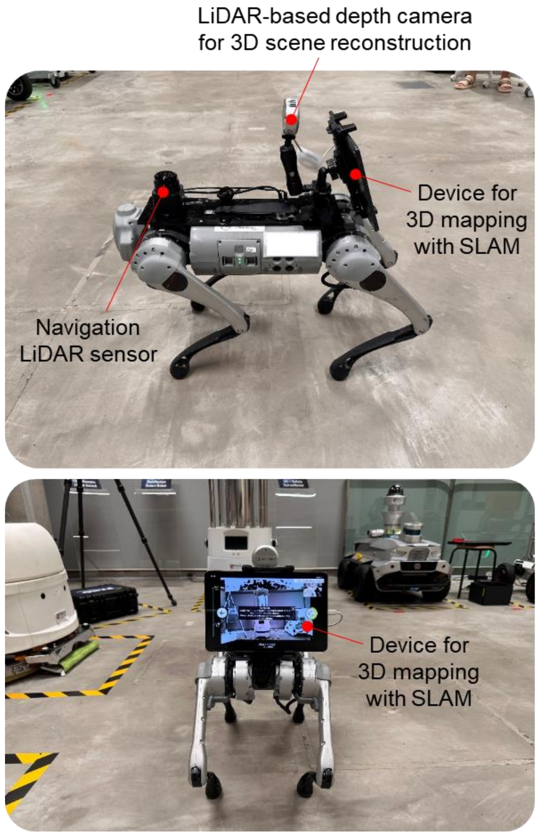


Figure 3. Ground quadruped robot equipped with fixed FOV LiDAR-based depth camera

Inspired by Reference [9], the LiDAR-based depth camera is directly connected to a device mounted on the quadruped robot for real-time SLAM-based 3D mapping. The sensor captures detailed depth maps and RGB colour and texture data, which are processed by the device for 3D scene reconstruction.

By integrating the LiDAR-based depth camera with the quadruped robot, the system utilises the robot's exceptional manoeuvrability to compensate for the fixed FOV of the LiDAR-based depth camera. As the robot moves in confined and cluttered spaces, this configuration enables efficient scanning of structural components while addressing obstacles and occlusions commonly encountered in such environments. The resulting data acquisition ensures decent 3D representations of as-built conditions, with depth measurements and rich spatial data supporting comprehensive mapping and scanning.

2.2 3D Scene Reconstruction

In this study, SLAM algorithm is utilised to map the constructed environment. SLAM operates by generating point clouds from individual scans and matching them within a global coordinate system. To overcome challenges such as obstacles and occlusions, the viewpoints for the robots are strategically identified, ensuring comprehensive scan completeness of the 3D scene. The ground robot then moves along the optimised path while scanning and capturing the RGB-D information for SLAM algorithm to generate the point clouds. This process results in the as-built 3D scenes with further refinement through loop detection.

The Intel RealSense L515 LiDAR-based depth camera is employed. This portable device is mounted on the ground robot's microprocessor and operated using Dot3D Scan software. The placement height is carefully adjusted to maximise the effective scanning range. Key specifications of the sensor include:

- **Dimensions**: 61 mm (diameter) x 26 mm (thickness)
- **Depth Range**: 0.25 to 9 meters
- **Accuracy**: ±14 mm
- **Horizontal Field of View (FOV)**: ~70°
- **Vertical Field of View (FOV)**: ~43°

LiDAR captures depth maps, while the image sensor captures RGB colour, and both are fused into SLAM for 3D scene reconstruction. This setup enables ground robot-mounted scanning to efficiently generate as-built 3D point clouds of the environment.

3 Experiment Results and Discussion

3.1 Testing Site

The effectiveness of quadruped robot-based scanning is demonstrated through experiments conducted on the NUS campus, focusing on 3D scanning and mapping within a selected laboratory in an educational building. Figure 4 illustrates the process of 3D scene reconstruction. In Figure 4(a), the robot moves to the first predefined scanning waypoints. In Figure 4(b), the robot adjusts its motion dynamically (e.g., tilting its head/sensor upward and downward) to scan both upper and lower parts of the structures. The captured RGB-D data are processed in real time for point cloud generation and 3D scene reconstruction (Figure 4(c)). In Figure 4(c), the depth sensor's FOV has a pyramid shape with a square projection at its bottom. This is a result of the sensor's line of sight on structural elements such as walls, meaning that when scanning a planar surface, the projected area should appear square-shaped rather than spherical.

L515 LiDAR sensor is fixed on the quadruped robot, meaning that as the robot moves and tilts its head upward or downward, or reorients in different directions, the sensor's viewpoint shifts accordingly. This movement allows the sensor to capture different regions of the indoor environment to enhance the completeness of 3D scene reconstruction. Using SLAM, the robot achieves dense, accurate scene mapping while travelling to subsequent viewpoints, including confined areas (Figure 4 (d, e)). The final output, shown in Figure 4(f), is a detailed, integrated 3D map of the scanned environment, seamlessly combining data from multiple scanned frames.

3.2 Visual Inspection of 3D Scenes

Figure 5 (a, b) showcases the reconstructed 3D scenes of the testing site, highlighting various structural elements. The reconstruction combines depth and RGB data captured by the LiDAR-based depth camera and processes it in real time using SLAM. The results are dense, visually enriched point clouds that represent the spatial and visual characteristics of the structure and the environment, including the orientation, shape, and

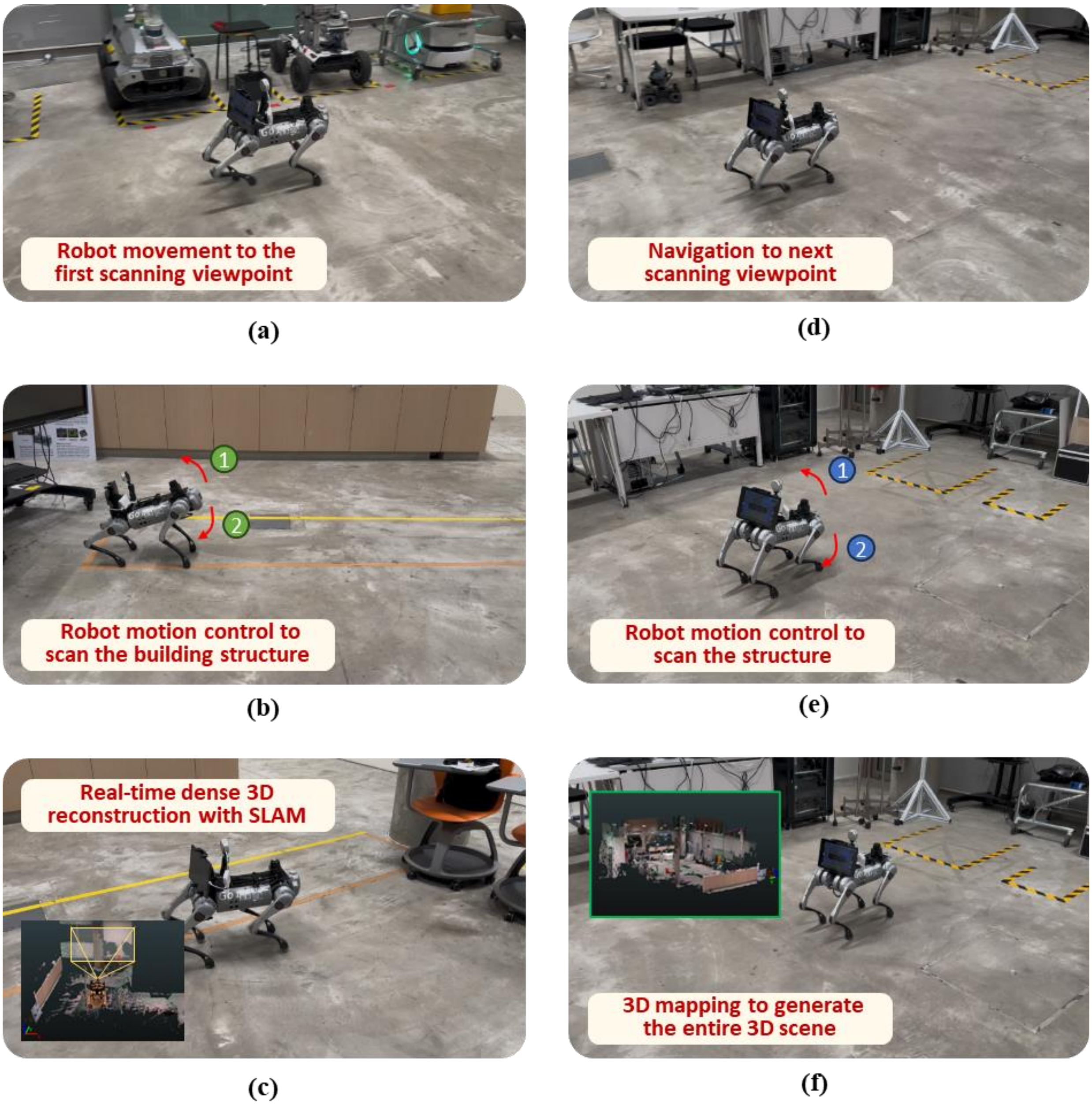
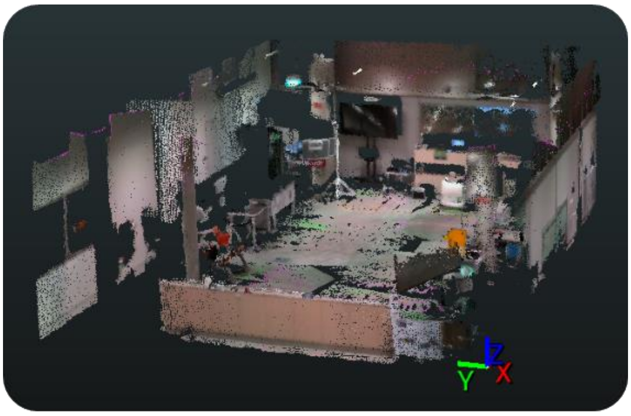


Figure 4. Automated 3D scanning using quadruped robot and fixed FOV LiDAR-based depth camera



(a)

(b)

Figure 5. Reconstructed 3D scenes for the testing site

dimensions of walls, floors, furniture, and other architectural components. The addition of colour data further enhances the 3D scenes by reflecting surface textures and material properties. This rich geometric and texture information is invaluable for applications like training AI models for semantic segmentation, enabling effective differentiation and classification of structural and non-structural elements.

The limitation of robot-assisted scanning with a fixed FOV LiDAR-based depth camera is also apparent. Due to its limited FOV, the dimensions of the theoretical measurement area are constrained. The length and width of the measurement area can be calculated using Equations (1) and (2), as follows:

$$L = 2 \cdot \tan\left(\frac{\mathrm{FOV}_{\mathrm{horizontal}}}{2}\right) \cdot \mathrm{Distance} \quad (1)$$

$$W = 2 \cdot \tan\left(\frac{\mathrm{FOV}_{\mathrm{vertical}}}{2}\right) \cdot \mathrm{Distance} \quad (2)$$

wherein L refers to the length of the measurement area, W represents the width of the measurement area, $\mathrm{FOV}_{\mathrm{horizontal}}$ and $\mathrm{FOV}_{\mathrm{vertical}}$ stand for the horizontal and vertical FOV respectively (in degrees), and Distance is the distance from the sensor to the measurement area. Even with the quadruped robot's ability to adjust its sensor orientation for scanning ceilings and floor slabs, the ceiling and upper sections of the structure, including mechanical and electrical (M&E) systems, remain difficult to capture comprehensively due to the sensor's limited measurement distance and area. Such limitations may prevent the reconstructed 3D scenes from meeting the requirements of various applications, such as digital twin creation for M&E systems.

3.3 Quantitative Analysis of Data Quality

The completeness and accuracy of the scan data are analysed using 2D projections of the XZ and YZ planes from the 3D scenes, as shown in Figures 6 and 7. Unscanned areas, highlighted in purple in Figure 6, indicate that approximately 65% of the selected planes were successfully covered. Key structural elements, including walls, windows, and equipment, are clearly identifiable, though corner areas remain partially unscanned due to occlusions and the limited FOV of the LiDAR-based depth camera. A notable fraction of the purple areas corresponds to windows, and the absence of depth data in these regions should not be considered an error but rather a natural limitation of the LiDAR sensor, as L515 struggles with reflective or transparent surfaces, making it unable to capture reliable depth data. However, this missing data can be leveraged to approximate or identify window boundaries from wall point clouds. The missing data in the upper regions of the wall is not solely due to transparent glass surfaces but may also result from sensor occlusion or range limitation. Future research could further analyse the correlation between these factors on 3D scene reconstruction.

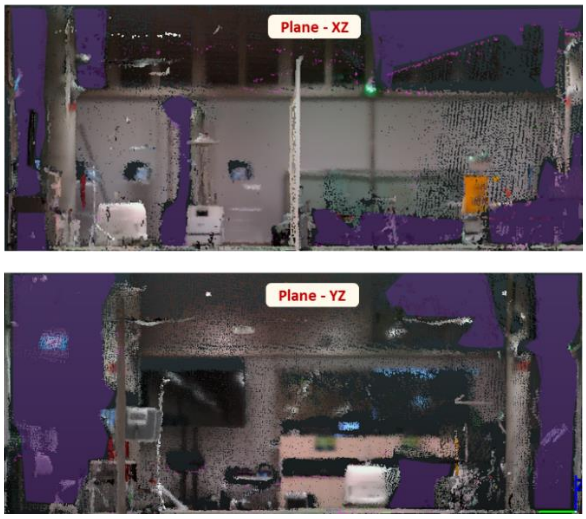


Figure 6. Completeness of scan data

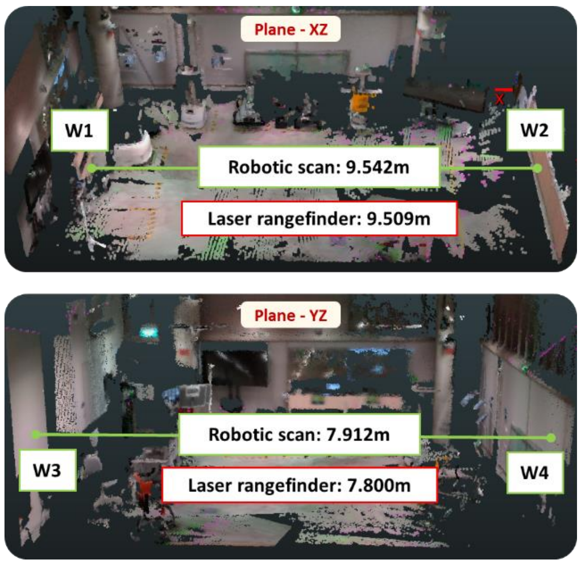


Figure 7. Accuracy of scan data

Figure 7 compares the distance measurement from robotic scanning against conventional laser rangefinders. For the XZ plane, the robotic scan records 9.542m from W1 to W2, while the laser rangefinder measures 9.509m. Similarly, for the YZ plane, the robotic scan reports 7.912m from W3 to W4, compared to the laser rangefinder of 7.800m. Figure 8 illustrates the discrepancies between the measurements obtained from robotic scan and laser rangefinder. The differences are minimal, with a deviation of only 1.4% in the XZ plane and 0.3% in the YZ plane. These differences may result from vibration during the robotic dog's movement, affecting SLAM-based 3D mapping, and the lower accuracy of L515 LiDAR-based depth camera used in the robotic scan.

Despite these constraints, the results demonstrate that robot-assisted scanning using a fixed FOV LiDAR camera is a viable approach for producing decent as-built 3D measurements. This highlights the potential of integrating robotics with LiDAR to facilitate spatial data collection, enabling downstream applications such as scan-to-BIM and facility inspection.

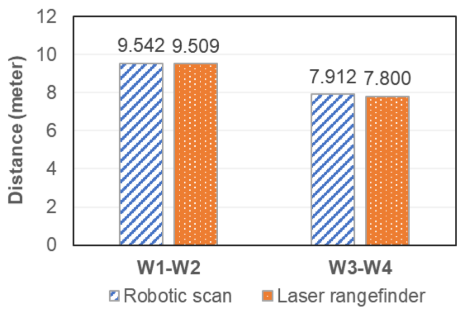


Figure 8. Comparison of distance measurement by robotic scan and laser rangefinder

4 Conclusions

This study demonstrates the potential of integrating a quadruped ground robot with a fixed FOV LiDAR-based depth camera for 3D point cloud reconstruction. The proposed system addresses challenges in capturing spatial data by leveraging the advanced manoeuvrability of a quadruped robot and the precision of a LiDAR-based depth camera. This combination enhances the robot's ability to efficiently move in confined spaces and cluttered environments, while enabling detailed 3D scanning and mapping. The research involves the validation through experimental studies conducted on NUS campus. The field tests focus on scanning an educational building to assess the system's capabilities in capturing as-built conditions. The quadruped robot, equipped with the Intel RealSense L515 LiDAR camera, demonstrates its ability to perform detailed 3D measurements. The results underscore the feasibility and effectiveness of the proposed quadruped robot-based scanning for detailed environmental mapping.

Integrating a fixed FOV LiDAR camera into robotic systems marks a significant advancement, but several areas require improvement. Firstly, the sensor's limitations in depth measurement, RGB colour, and texture highlight the need for adopting 3D LiDAR-camera. Additionally, given that the L515 sensor provides RGB and depth maps in addition to point clouds, future work of this research could consider integrating these depth maps with RGB data during the 3D mapping process using novel SLAM algorithms, such as Gaussian Splatting-based SLAM, to improve 3D scene reconstruction accuracy. Lastly, advanced data acquisition planning optimisation is essential to guide the robot effectively through complex structural environments. These algorithms should also ensure adaptability to confined spaces, broadening the system's applicability across various use cases.

Acknowledgements

This research is supported by Singapore Ministry of Education under Academic Research Fund Tier 1 (A-8002138-00-00). Any opinions, findings and conclusions or recommendations expressed in the material are those of the author and do not reflect the views of the grantor. The author would like to thank Hu Difeng whose robotic work was referenced in the literature review, and the robot dog 'Ben Ben' for four years' experiments.

References

[1] D. Hu, V. J. Gan, T. Wang, and L. Ma. Multi-agent robotic system (MARS) for UAV-UGV path planning and automatic sensory data collection in cluttered environments. Building and Environment, 221:109349, 2022.
[2] P. Tang, et al. Automatic reconstruction of as-built building information models from laser-scanned point clouds: A review of related techniques. Automation in Construction, 19(7):829–843, 2010.
[3] V. J. L. Gan, K. Li, M. Li, and L. B. E. Halfian. 3D reconstruction of building information models with weakly-supervised learning for carbon emission modelling in the built environment. Applied Energy, 377:124695, 2025.
[4] H. Son, C. Kim, and Y. Turkan. Scan-to-BIM: An overview of the current state of the art and a look ahead. In ISARC. Proceedings of the International Symposium on Automation and Robotics in Construction, pages 12–34, Citeseer, 2015.
[5] S. I. El-Halawany and D. D. Lichti. Detecting road poles from mobile terrestrial laser scanning data. GIScience & Remote Sensing, 50(6):704–722, 2013.
[6] K. S. Yen, T. A. Lasky, and B. Ravani. Cost-benefit analysis of mobile terrestrial laser scanning applications for highway infrastructure. Journal of Infrastructure Systems, 20(4):04014022, 2014.
[7] J. Gené-Mola, et al. Fruit detection in an apple orchard using a mobile terrestrial laser scanner. Biosystems Engineering, 187:171–184, 2019.
[8] J. R. Rosell-Polo, et al. Kinect v2 sensor-based mobile terrestrial laser scanner for agricultural outdoor applications. IEEE/ASME Transactions on Mechatronics, 22(6):2420–2427, 2017.
[9] D. Hu, V. J. L. Gan, and C. Yin. Robot-assisted mobile scanning for automated 3D reconstruction and point cloud semantic segmentation of building interiors. Automation in Construction, 152:104949, 2023.
[10] R. Zhai, et al. Semantic enrichment of BIM with IndoorGML for quadruped robot navigation and automated 3D scanning. Automation in Construction, 166:105605, 2024.
[11] V. J. Gan, D. Hu, R. Zhai, and Y. Wang. Automated As-built 3D reconstruction using quadruped robot and 3D LiDAR sensor. In ISARC Proceedings of the International Symposium on Automation and Robotics in Construction, IAARC Publications, 2024.
[12] J. Kim, D. Chung, Y. Kim, and H. Kim. Deep learning-based 3D reconstruction of scaffolds using a robot dog. Automation in Construction, 134:104092, 2022.
[13] T. Wang and V. J. Gan. Automated joint 3D reconstruction and visual inspection for buildings using computer vision and transfer learning. Automation in Construction, 149:104810, 2023.
[14] P. Kim, J. Chen, and Y. K. Cho. SLAM-driven robotic mapping and registration of 3D point clouds. Automation in Construction, 89:38–48, 2018.
[15] H. Kandath, T. Bera, R. Bardhan, and S. Sundaram. Autonomous navigation and sensorless obstacle avoidance for UGV with environment information from UAV. In 2018 Second IEEE International Conference on Robotic Computing (IRC), IEEE, 2018.
[16] P. Kim, J. Park, Y. K. Cho, and J. Kang. UAV-assisted autonomous mobile robot navigation for as-is 3D data collection and registration in cluttered environments. Automation in Construction, 106:102918, 2019.
[17] J. Peterson, et al. Online aerial terrain mapping for ground robot navigation. Sensors, 18(2):630, 2018.
[18] D. Hu, V. J. Gan, and R. Zhai. Automated BIM-to-scan point cloud semantic segmentation using a domain adaptation network with hybrid attention and whitening (DawNet). Automation in Construction, 164:105473, 2024.
[19] A. Aryan, F. Bosché, and P. Tang. Planning for terrestrial laser scanning in construction: A review. Automation in Construction, 125:103551, 2021.