

# Autonomous operation of a robot dog for point cloud data acquisition of scaffolds

D. Chung<sup>a</sup>, S. Paik<sup>a</sup>, J. Kim<sup>a</sup>, and H. Kim<sup>a</sup>

<sup>a</sup>Department of Civil and Environmental Engineering, Yonsei University, Korea

E-mail: [jungduho1@yonsei.ac.kr](mailto:jungduho1@yonsei.ac.kr), [tjsdnd4066@yonsei.ac.kr](mailto:tjsdnd4066@yonsei.ac.kr), [kah5125@yonsei.ac.kr](mailto:kah5125@yonsei.ac.kr), [hyoungkwan@yonsei.ac.kr](mailto:hyoungkwan@yonsei.ac.kr)

## Abstract –

Scaffolds are essential temporary structures on construction sites. Since scaffolds are frequently installed and dismantled, the inspection needs to be performed in real-time. This paper proposes a framework to automate the acquisition process of scaffold point cloud data using a robot dog. First, a Simultaneous Localization and Mapping (SLAM) algorithm (LIO-SAM) is deployed for real-time map creation based on laser-based 3D data. Scaffolds are automatically detected using the bird's eye view (BEV) projection images of the registered 3D point clouds. A scanning distance is also determined for each detected scaffold to move the robot dog to an optimal location. The robot dog can successfully scan the scaffolds on construction sites by using the proposed framework.

## Keywords –

Autonomous Operation; Mobile Laser Scanning; Robot Dog; Real-time Detection; Scaffold

## 1 Introduction

Scaffolds are an indispensable factor on construction sites, and it is one of the major risk factors for construction safety management. According to the statistics by the Korean Ministry of Employment and Labor, more than half (51.5%) of fatalities in the construction industry are falling accidents, and scaffolds are the contributing factor (19.9%) to falling fatalities [1]. Because scaffolds are frequently installed and dismantled during construction, safety management is difficult. Real-time inspection is ideal for a thorough inspection, but it rarely becomes a reality due to its labor-intensive and costly nature. Automating the inspection process using a mobile robot could be a solution for the effective monitoring of scaffolds. By adding repeatability to the labor-intensive inspection process, fast and efficient monitoring could become a reality.

Kim et al. [2] proposed a framework for automatic scaffold segmentation and 3D reconstruction based on 3D point clouds acquired by Mobile Laser Scanning

(MLS). A robot dog was used in the study for the point cloud data acquisition process, but the robot dog was teleoperated. Teleoperation can reduce human labor for data acquisition but still requires human intervention. There are some studies that use autonomous operation for data acquisition. Kim et al. [3] used an Unmanned Aerial Vehicle (UAV) to make a map of the construction site for calculating optimal scanning points. The map was given to an Unmanned Ground Vehicle (UGV) for autonomous scanning. The UGV relied on the map for its scanning process. Kim et al. [4] provided a fully automatic 3D data acquisition and registration system using a UGV. 2D SLAM was used for localization and navigation, and 3D reconstruction was performed based on the SLAM result. The study was intended to produce a general scan result of the construction site without specific target objects.

This paper proposes a new framework to automate the data acquisition process for scaffold point clouds using a robot dog. The proposed framework aims to bring a focused attention to a specific construction component - scaffolds, without the need for any prior knowledge such as scaffold location or construction site map. This would enable truly dynamic and real-time scan planning, practically applicable in scaffold inspection. To the best of the authors' knowledge, it is difficult to find previous studies in which a mobile robot was tried to automatically acquire point cloud data of specific construction objects. The overview of this framework is shown in Figure 1.

## 2 Methodology

### 2.1 System architecture

The scanning platform for this study uses a Unitree A1 robot dog, with Ouster OS1-128 Mobile LiDAR and Microstrain Inertial Measurement Unit (IMU), and an on-board computer (NVIDIA Jetson TX2), as described in Figure 2. Robot dogs can have two major advantages compared to wheeled robot for mobile laser scanning. First, robot dogs can walk stably through rough terrains and small obstacles. Unlike refined workspaces, construction sites generally have uneven surfaces, and a

wheeled robot's bump can easily affect the results of mobile laser scanning. Second, robot dog has more Degree of Freedom (DoF) in joints, which can easily increase the Field of View (FoV) of scanning without additional actuators. In our study, we controlled the roll and pitch simultaneously to increase the robot's FoV as much as possible.

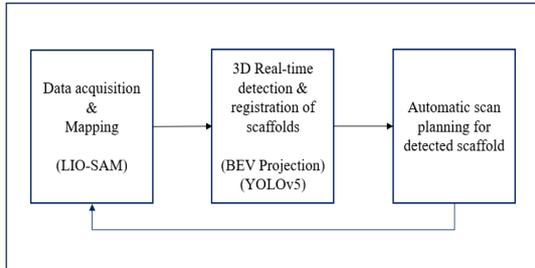


Figure 1. Overview of the proposed framework

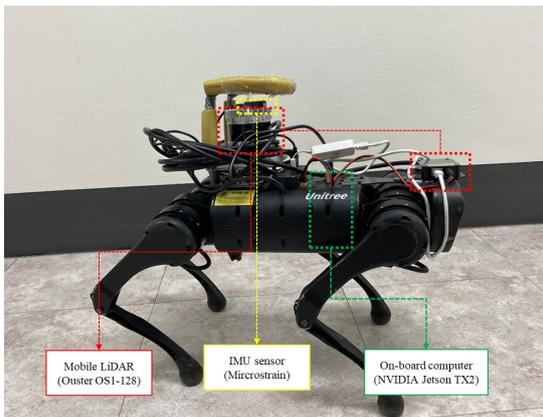


Figure 2. Scanning platform description

3D Simultaneous Localization and Mapping (SLAM) algorithm is used for the localization of the robot dog and for registration of the obtained data into a 3D point cloud map. The 3D point cloud map enables the robot dog to understand the environment, and it also works as the final product of the scanning process. LIO-SAM [5], a 3D SLAM algorithm based on sensor fusion between LiDAR and IMU, was used in this study.

LIO-SAM receives 3D point cloud and IMU sensor data as input. LiDAR odometry is initially predicted based on the motion estimated by IMU data, and the IMU bias is repeatedly corrected based on LiDAR odometry. The LiDAR odometry is sent to a path planning algorithm for the localization of the robot dog. The registered 3D point cloud map is used to perform scaffold detection and allow the robot dog to understand the location of the scaffold. Figures 3(a) and 3(b) show the scaffolds used for the experiment and its registered point clouds, respectively.

## 2.2 SLAM based BEV 3D scaffold detection

To automate the scaffold data acquisition process, the robot needs to understand the goal of scanning. In this study, deep learning-based object detection is used to detect scaffolds. Understanding the environment can be divided into two categories: image-based and point-cloud-based. Image-based methods are accurate and fast, but it lacks spatial information of the object. Point-cloud has very accurate spatial information but has sparse visual information. 3D point cloud detection also suffers from high computational costs. To reduce the computing cost, some studies attempted bird's eye view (BEV) projection-based object detection [6, 7]. The main idea is to translate 3D point cloud into 2D images by projecting the points vertically, and to apply convolutional neural network (CNN) for object detection. This idea can detect objects from the 3D point cloud in real-time, but a lack of visual information can lower detection performance.

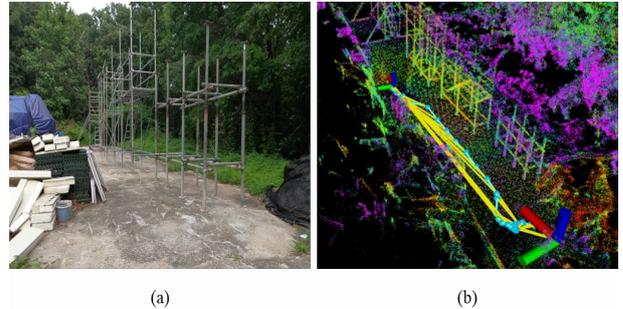


Figure 3. Scaffolds at Yonsei University; (a) photogrammetry, (b) registered point clouds

To overcome this problem, we used SLAM-based registered points instead of raw point cloud data to generate BEV images. By using registered points, the visual information becomes denser and can detect objects more accurately in real-time. In the proposed method, registered points are projected into 2D images, and each pixel value represents the height, density, and intensity features of the registered point. YOLOv5, a real-time object detection algorithm, is then applied to the generated BEV images [8] for detecting scaffolds. The scaffold detection method is shown in Figure 4.

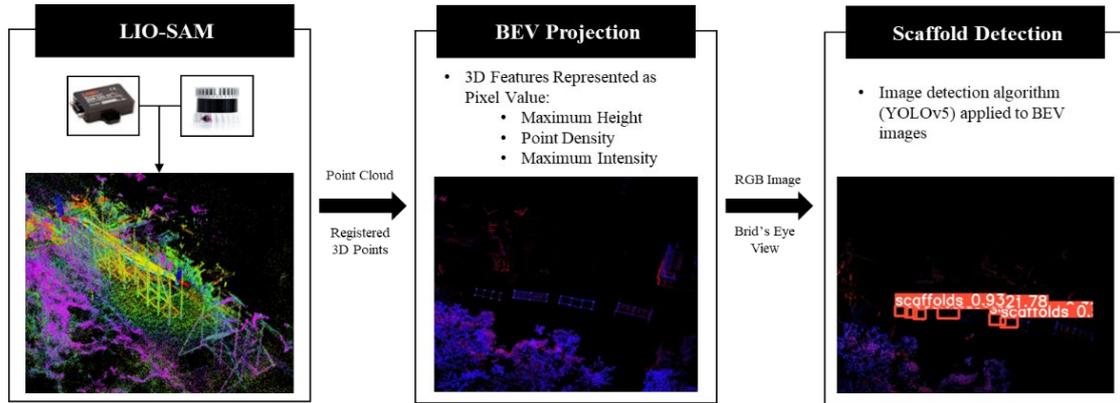


Figure 4. Scaffold detection using the SLAM data

### 2.3 Implementation

The proposed framework was implemented using a robot operating system (ROS). LIO-SAM subscribes to sensor data from LiDAR and IMU, then publishes registered 3D point cloud to the BEV projection node. The BEV projection node projects registered point cloud data to 2D images and publishes them to the detection node. The detection node detects scaffolds from the subscribed images and calculates the scaffold's Cartesian coordinates and maximum height. To prevent overlapping results from the same scaffolds, detected scaffolds are registered only when it's more than a threshold value away from each other. Maximum height is calculated by detecting the highest pixel inside the detection bounding box.

The path planning node subscribes to the scaffold data, calculates the optimal scanning distance, and publishes the command for scanning. The optimal scanning distance is decided based on the FoV of the LiDAR and the maximum control range of the robot's

pitch by Eq. (1). If there is no recognized scaffold, the robot performs a pre-defined scanning motion to get more information about the environment. After the robot recognizes the scaffold, the robot decides the closest scaffold as a goal, and moves towards the scaffold until it reaches the optimal scanning distance. Once the robot reaches the scanning distance, the robot performs the scanning motion and moves on to the next scaffold. Fig. 6 describes the flowchart of the path planning node. The control node subscribes to the published command to control the robot hardware.

$$\text{Scan distance} = \frac{\text{Scaffold height}}{\tan(\text{LiDAR FoV} + \text{Maximum Pitch})} \quad (1)$$

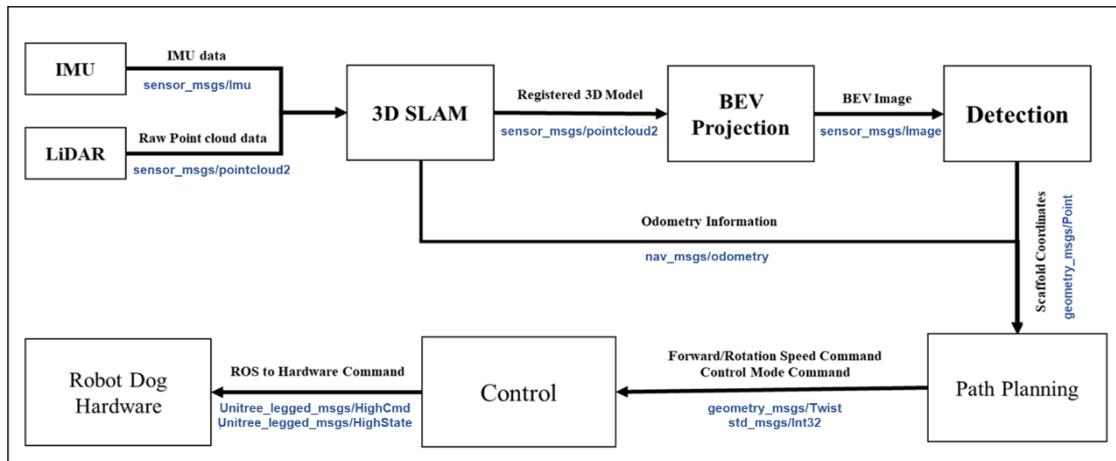


Figure 5. ROS-based implementation for the proposed method

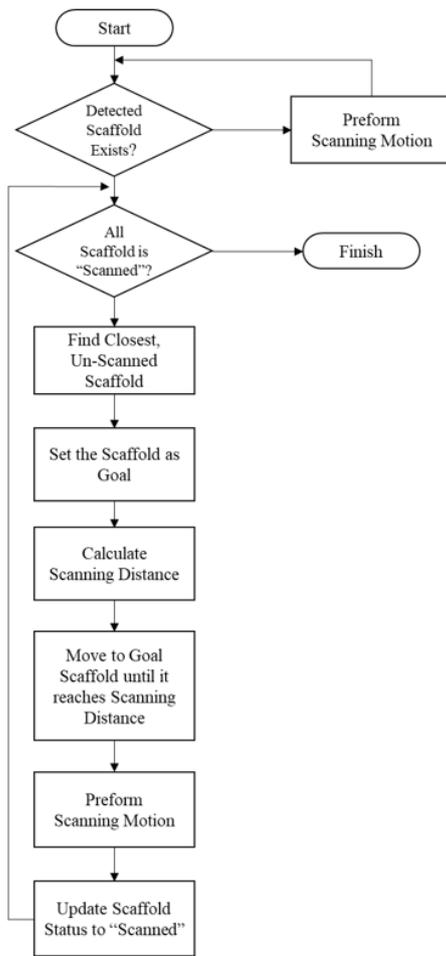


Figure 6. Flowchart for path planning node

### 3 Experiments and Results

For training the object detection model, 300 BEV images from scaffolds at Chung-Ang University were used, and 58 images from scaffolds at Yonsei University were used to test the model. All datasets were gathered by the robot dog, as conducted in [2]. Figures 7(a) and 7(b) are examples for training and testing sets, respectively.

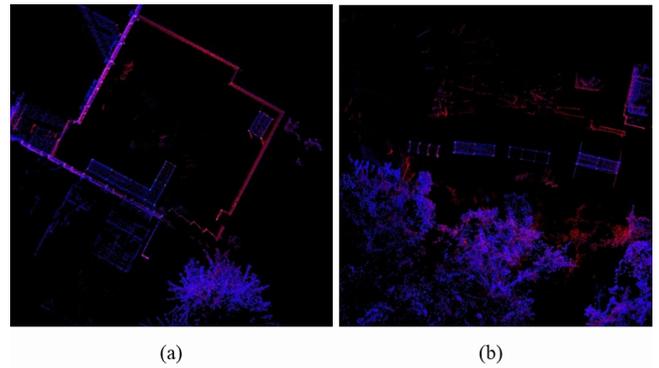


Figure 7. BEV images; (a) an example for training, (b) an example for testing

	Precision	Recall	F1-score
Scaffold	86.9%	73.5%	79.6%

Table 1. The performance of the scaffold detection model

The scaffold detection model is trained for 50 epochs, with pre-trained weights based on the COCO dataset [9]. Table 1 shows the performance of the scaffold detection model. The model achieved 86.9% precision, 73.5% recall, and 79.6% F1-score on scaffolds. Figure 8 shows an example of scaffold detection results, proving that the framework effectively detects scaffolds in real-time.



Figure 8. Scaffold detection results

For this experiment, we used a single scanning motion in which we changed the roll and pitch of the robot joint for 10 seconds. The scanning motion is shown in Figure 9. The robot dog also had a fixed region of interest of 30m x 30m square with the robot's starting point as origin.

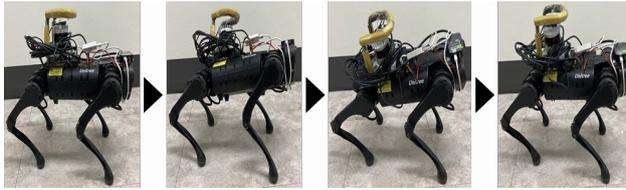


Figure 9. Scanning motion

Figure 10(a) shows an example case of the robot's trajectory and detected scaffolds, Figure 10(b) and 10(c) show the result of automatic data acquisition. As shown in Figure 10, the scaffold points have been successfully obtained. Even though there are still some limitations in the navigation algorithm, the experiment shows that the proposed framework allows the robot dog to automatically move around the site for the successful scanning of scaffolds.

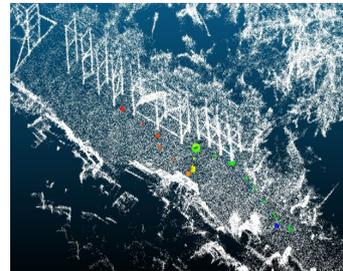
#### 4 Conclusions

This study proposed a new framework for automating the scaffold point cloud data acquisition process using a robot dog. The proposed framework with a real-time 3D scaffold detection algorithm with an SLAM-based BEV image was implemented for a robot dog, and it was tested on a real-world outdoor construction site. The experiments show that the robot dog can automatically perform the end-to-end data acquisition process without any human intervention.

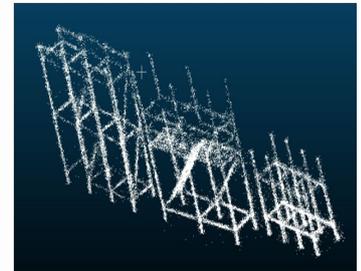
This study currently has some limitations in the navigation system. First, an obstacle avoidance system needs to be developed. Second, the path needs to be optimized. Third, the scan planning algorithm needs to be more generalized for a range of construction sites. With the improvement, the proposed method is expected to enable a fully autonomous operation of smart mobile robots designed to monitor construction sites for safety and productivity management.



(a)



(b)



(c)

Figure 10. Results; (a) robot's trajectory represented by the numbers and detected scaffolds with scanning points connected to its corresponding scaffolds by the same colors, (b) the 3D point cloud of the site, (c) the 3D point cloud of the scaffolds

#### Acknowledgment

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Ministry of Education (No. 2018R1A6A1A08025348) and the "National R&D Project for Smart Construction Technology (No.22SMIP-A156488-03) funded by the Korea Agency for Infrastructure Technology Advancement under the Ministry of Land, Infrastructure and Transport, and managed by the Korea Expressway Corporation.

#### References

- [1] Korea Ministry of Employment and Labor. Industrial Accident Analysis for 2020. Online: [https://www.moel.go.kr/news/enews/report/enews\\_View.do?news\\_seq=12149](https://www.moel.go.kr/news/enews/report/enews_View.do?news_seq=12149), Accessed:02/23/2022
- [2] Kim, J. H., Chung, D. H., Kim, Y. H., and Kim, H.

- K., Deep learning-based 3D reconstruction of scaffolds using a robot dog. *Automation in Construction*, 134: 104092, 2022.
- [3] Kim, P., Park, J., and Cho, Y. K., UAV-assisted autonomous mobile robot navigation for as-is 3D data collection and registration in cluttered environments. *Automation in Construction*, 106: 102918, 2019.
- [4] Kim, P., Chen, J., and Cho, Y. K., SLAM-driven robotic mapping and registration of 3D point clouds. *Automation in Construction*, 89: 38-48. 2018.
- [5] Shan, T., Englot, B., Meyers, D., Wang, W., Ratti, C., and Rus, D. LIO-SAM: Tightly-coupled lidar inertial odometry via smoothing and mapping. In *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 5135-5142, Las Vegas, USA, 2020.
- [6] Beltran, J., Guindel, C., Moreno, F. M., Cruzado, D., Garcia, F., & De La Escalera, A. Birdnet: a 3d object detection framework from lidar information. In *International Conference on Intelligent Transportation Systems (ITSC)*, pages 3517-3523, Maui, USA, 2018
- [7] Simony, M., Milzy, S., Amendey, K., & Gross, H. M. Complex-yolo: An euler-region-proposal for real-time 3d object detection on point clouds. In *Proceedings of the European Conference on Computer Vision (ECCV) Workshops* page 0-0, Munich, Germany, 2018
- [8] Jocher, Glenn, et al. "yolov5." Code repository <https://github.com/ultralytics/yolov5> (2020).
- [9] Lin, T. Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., ... & Zitnick, C. L. Microsoft coco: Common objects in context. In *European conference on computer vision*, page 740-755, Zurich, Switzerland. 2014