# MOVING OBJECT CLASSIFICATION USING MULTILAYER LASER SCANNING WITH SPACE SUBDIVISION FRAMEWORK

M. Nakagawa <sup>1,</sup> \*, M. Taguchi <sup>2</sup>

<sup>1</sup> Shibaura Institute of Technology, 3-7-5 Toyosu, Koto-ku, Tokyo 135-8548, Japan - mnaka@shibaura-it.ac.jp
<sup>2</sup> Azuma Kigyo Co. Ltd., 2-19-14, Kiba, Koto-ku, Tokyo 135-0042, Japan - taguchi@azumakigyo.co.jp

Commission IV, WG IV/5

KEY WORDS: Laser scanning, Multilayer laser scanner, Object classification, Object tracking, Space subdivision

### **ABSTRACT:**

In this paper, we focus on the development of intelligent construction vehicles to improve the safety of workers in construction sites. Generally, global navigation satellite system positioning is utilized to obtain the position data of workers and construction vehicles. However, construction fields in urban areas have poor satellite positioning environments. Therefore, we have developed a 3D sensing unit mounted on a construction vehicle for worker position data acquisition. The unit mainly consists of a multilayer laser scanner. We propose a real-time object measurement, classification and tracking methodology with the multilayer laser scanner. We also propose a methodology to estimate and visualize object behaviors with a spatial model based on a space subdivision framework consisting of agents, activities, resources, and modifiers. We applied the space subdivision framework with a geofencing approach using real-time object classification and tracking results estimated from temporal point clouds. Our methodology was evaluated using temporal point clouds acquired from a construction vehicle in drilling works.

## 1. INTRODUCTION

The construction field has recently focused on technical and political issues, such as construction management costs, productivity improvement, and reducing the number of accidents (Dong et al. 2018). Various actions are available to address these issues based on building information modeling (BIM). BIM is used as a visual database to visualize and manage safety operations (Guo et al. 2017) with cameras, laser scanners, global navigation satellite system (GNSS) devices, unmanned aerial vehicles (UAVs), intelligent construction vehicles (Doishita et al. 2010), virtual reality, and augmented reality. In this paper, we focus on using intelligent construction vehicles to improve the safety of workers.

Generally, GNSS positioning is applied to obtain the position data of workers and construction vehicles in construction sites. However, when position data are shared among construction vehicles and workers, instead of using GNSS devices, wireless communication systems and computing systems should be distributed to share position data between workers and construction vehicles. Thereby, the sensing cost increases and the sensing system becomes more complex. Moreover, construction fields in urban areas have poor satellite positioning environments. Thus, to address these issues, we applied 3D sensing to provide more stable worker position data acquisition and collision-avoidance sensing of construction vehicles to improve incident prediction.

UAVs and terrestrial laser scanners can acquire 3D data of static construction fields (Figure 1). However, with UAVs and terrestrial laser scanners, it is not easy to measure and represent changing objects and environments, such as moving workers, vehicles, and construction fields in real time. Therefore, we propose a methodology for real-time object measurement, classification and tracking from temporal point clouds acquired with a multilayer laser scanner. However, only using real-time 3D measurement and tracking is not sufficient to cover incident prediction. Thus, we also propose a methodology to estimate and visualize object behaviors with a spatial model based on a space subdivision framework consisting of agents, activities, resources, and modifiers (Sithole and Zlatanova, 2016). Although the framework is designed for indoor navigation, the idea of space subdivision for indoor modeling can be extended to closed spaces in outdoor environments. In this paper, we applied the space subdivision framework with a geofencing approach using the results of real-time object classification and tracking from temporal point clouds. We evaluated our methodology using temporal point clouds acquired from a fixed position in indoor and outdoor spaces. We also evaluated our methodology using temporal point clouds acquired from a construction vehicle in drilling works.



Figure 1. 3D measurement result using a TOF camera in a construction site

#### 2. METHODOLOGY

Our methodology consists of background estimation, moving object extraction, moving object tracking, object classification, and activity classification (Figure 2). First, resources are extracted from background data estimated from temporal point clouds. Second, moving objects consisting of agents and modifiers are extracted from temporal point clouds. After moving object tracking, moving objects are classified into agents and modifiers. Finally, activities are estimated and classified using agents, modifiers and resources.



Figure 2. Proposed methodology

## 2.1 Moving object extraction

We extract moving and changing parts from point clouds to use as agents and modifiers. When moving objects are classified into pedestrians and other objects, pedestrians are estimated with features and behaviors of moving objects, such as volume, actions, and moving speed, through the following moving object tracking process.

Moving objects can be extracted with a background subtraction processing, which is a conventional approach, when point clouds are acquired from a fixed position. However, when point clouds are acquired from a moving platform, the background subtraction approach is not easily applied to point cluster extraction. Therefore, we applied a point cloud segmentation approach consisting of four steps (Figure 3).

#### Point clouds

1) Temporal point cloud projection into range image
2) Point cloud labeling in range image
3) Labeled point cloud clustering
4) Moving object extraction

Moving objects

Figure 3. Moving object extraction

First, temporal point clouds are projected into temporal range images. The temporal range image is prepared as 7D spaces consisting of 3D coordinate values (X, Y, and, Z), intensity values, scanning directions (horizontal angles), scanning layers (vertical angles), and scene numbers. Second, point clouds higher than ground height are labeled in the range images. The ground height is determined using a major horizontal plane estimated with robust plane fitting. Third, labeled point clouds are clustered to generate moving object candidates with voxel segmentation processing, and we apply the region-growing methodology for the voxel segmentation. Fourth, moving objects are extracted from moving object candidates. When point clouds are acquired from a construction vehicle, the closest moving object candidate from a scanner is assumed to be a bucket, while the other moving object candidates are assumed to be workers with geometric constraints such as height and volume.

## 2.2 Moving object tracking and classification

Moving objects are tracked to be constantly fixed as candidates of moving objects during several scenes in a temporal 3D space.

When a scanner position is fixed, the nearest cluster tracking can be applied for simple object tracking. However, when the scanner translates and rotates, tracking results using acquired point clouds are unstable (Figure 4). Thus, simultaneous localization and mapping (SLAM) (Durrant-Whyte and Bailey 2006; Durrant-Whyte and Bailey 2006) is integrated to detect and track moving objects (Vu et al. 2011) to improve the stability of moving object tracking from a moving scanner. In our methodology, rotation and translation parameters are estimated with SLAM. Then, the nearest clusters are searched from rotated and translated point clouds. At the same time, spike noise and unclear points can be rejected from moving object candidates. Tracked moving objects are classified into agents and modifiers, and both agents and modifiers have object names or identified numbers, positions (gravity points of point clouds), and clock data. Status information, such as moving, stopping, and sitting, is generated from velocity and height changes estimated from temporal position data.



Figure 4. Moving object tracking and filtering

## 2.3 Activity classification

In our research, events and actions are defined as activities. Activities are estimated and classified using relative positions between agents, modifiers, and resources with basic behavior information described in the status, such as moving, stopping, sitting, or operating. The status can be estimated from each moving and changing object estimated from point clouds. When we focus on the velocity of moving objects, the statuses of moving and stopping can be distinguished. When we focus on the changing height of objects, the statuses of standing and walking can be distinguished. A geofencing approach is applied for activity recognition using the relative positions between agents, modifiers, and resources. In a geofencing approach, virtual fences are generated on a map at a radius around a store or at a point location. Virtual fences are virtual parameters for a real-world geographic area. The geofencing approach provides various services and a push-based information distribution using the virtual fences. When a user with a location-aware device enters or exits the virtual fences, the mobile device receives notification such as location-based assistance or alerts. For example, when a worker approaches a construction vehicle, an alert can be sent to an operator using a virtual fence around the construction vehicle. Estimated activities are summarized as annotations in point clouds.

In this study, virtual fences are generated from temporal point clouds based on a space subdivision, and we classified a construction space into four categories: outdoor, semi-outdoor, underground, and semi-underground (Figure 5). A safety area in a construction space is defined as the outdoor category; a construction vehicle's motion area is defined as the semioutdoor category; a drilling space is defined as the semiunderground category; and a space under the semi-underground is defined as the underground category. Here, virtual fences are generated using the semi-outdoor category to provide alerts for workers approaching construction vehicles.





Figure 5. Subdivided construction space

## 3. EXPERIMENTS

We conducted three types of experiments: the first was a preliminary experiment for object tracking and recognition with laser scanning from a fixed point to visualize pedestrian behaviors in an indoor space; the second was laser scanning from a fixed point to visualize behaviors of workers and construction vehicles in a construction environment; and the third was laser scanning from a moving construction vehicle to track activities of workers in the same construction environment. In all three experiments, we used a multilayer laser scanner (VLP-16, Velodyne) (Figure 6) for point cloud acquisition. Moreover, measured areas were subdivided into agents, activities, resources, and modifiers (Table 1).



Figure 6, VLP-16 (Velodyne)

Table 1. Subdivided spaces

			-	
		Elevator hall	Construction environment	
Resource		Chair, Elevator button	Pit, Cone, Rope	
Modifier	Name	Elevator, Elevator door	Construction vehicle	
	Status	Opened, Closed	Operated, Stopped	
Agent	Name	Pedestrian	Worker	
	Status	Moving, Stopping, Sitting	Moving, Stopping	
Activity		Elevator getting on and off	Alerted near a construction	
			vehicle	

### 3.1 Preliminary experiment of indoor laser scanning

We selected an elevator hall on our campus (Figure 7). The laser scanner was installed at a height of 1.0 m in one corner of the elevator hall. We acquired 11,000 scenes of laser scanner data for approximately 18 min (2.5 million points). The processed area (8 m  $\times$  8m) included pedestrians and three elevators. Several pedestrians existed in the measured area per minute, and their behaviors were classified into walking, calling an elevator, waiting for an elevator, and sitting. In this paper, we use 1,000 scenes (100 s) of laser scanner data. Virtual fences were created around each elevator using temporal point clouds.



Figure 7. Laser scanning in an indoor environment

### 3.2 Object recognition from a fixed point

We prepared a simulated construction environment (Figure 8). The laser scanner was installed at a height of 1.3 m on one side of the construction site. We acquired point clouds during construction work, such as excavation, piping, and filling works, for 30 min (18,000 scenes). The processed area (14 m  $\times$  5 m) included workers and construction vehicles. There were also several workers and construction vehicles present in the measured area per a minute. The workers' activities were

classified into moving and stopping. Here, we describe our results using 500 scenes (50 s). Virtual fences were created around each construction vehicle with a range of motion using temporal point clouds.



Figure 8. Simulated construction environment

## 3.3 Object recognition from a construction vehicle

Here, we used the same simulated construction environment as described in the experiment of object recognition from a fixed point (section 3.2). The laser scanner was mounted on a backhoe (Figure 9), and we acquired point clouds during construction work for 30 min (18,000 scenes). In total, we used 134,955,204 points (9,523 scenes) in all acquired point clouds for our data processing.



Figure 9. Laser scanner mounted on a construction vehicle

## 4. RESULTS

## 4.1 Laser scanning in an indoor environment

Results of indoor laser scanning in Figure 10 indicate that our methodology estimated the positions of pedestrians and elevator doors from temporal point clouds. The figure also shows that our methodology added annotations to represent an object's status, such as a door opening or closing, and a pedestrian walking, stopping, or sitting.



Figure 10. Results in indoor laser scanning

Figure 11 shows temporal plan views reconstructed from acquired point clouds. Scene numbers 383 and 437 indicate that our methodology extracted pedestrians and elevator doors. In addition, scene number 403 shows that occluded area interpolation processing estimated a pedestrian's position and

status from missing point clouds, and scene number 406 shows that occluded area interpolation processing estimated a door's position and status from missing point clouds.



Figure 11. Object recognition results (plan view)

### 4.2 Object recognition from a fixed point

Figure 12 (a, b, c, and d) shows classified objects such as workers and construction vehicles from laser scanning data acquired at a fixed point. Figure 12a and Figure 12b show object recognition results in scene number 9 in the acquired temporal point clouds. Figure 12c and Figure 12d show object recognition results in scene number 30 in the acquired temporal point clouds. Figure 12a and Figure 12c show object recognition results from a bird's eye view. Figure 12b and Figure 12d show object recognition results around recognized workers represent workers approaching a construction vehicle using virtual fences generated from a construction vehicle's motion areas. Figure 12c and Figure 12d also show that all workers were

successfully recognized, even if one worker was behind another worker and point clouds were missing.

Figure 13 shows the worker tracking results for all 500 scenes (50 s) and workers' activities such as stopping and moving during drilling works. For the visualization of a construction vehicle's motion, we added virtual fences and point clouds to scene numbers 265 and 300 in Figure 13, which shows that our proposed methodology can provide stable object tracking processing. Although there is some spike noise in the tracking results, smooth lines can be generated with spike noise filtering. When we observe workers' activities in drilling pits, we can change the position and vertical angle of the laser scanning. However, several laser scanners are required to avoid an occlusion problem in piping works.



Figure 12. Object recognition result in a construction environment



Figure 13. Worker tracking result in a construction environment (plane view)

### 4.3 Object recognition from a construction vehicle

Figure 14 shows extracted moving objects in an intensity image generated from acquired point clouds. The vertical axis indicates scanning layers extended 8.0 times with linear interpolation, and the horizontal axis indicates horizontal scanning angles with 0.25 resolutions.



Figure 14. Intensity image and extracted moving objects

Figure 15 shows our results after segmentation and clustering. Object tracking and recognition results are shown in Figure 16, and indicate that our methodology can stably trace workers. Table 2 shows the processing time for 9,523 scenes and the average processing time per scan. Our processing environment was Intel Core i7-6567U (3.30 GHz). We confirmed that object extraction was processed with a frequency of approximately 10 Hz. We also confirmed that overall processing was processed with a frequency of approximately 5 Hz from SLAM to moving object tracking.



Figure 15. Results after segmenting and clustering of point clouds



Figure 16. Worker recognition results (plan view)

Table 2. Processing time

	SLAM	Moving object extraction	Moving object tracking	Total
Processing time [sec]	1023.690	759.050	13.130	1795.870
Processing time (average) [sec]	0.107	0.080	0.001	0.187

#### 5. SUMMARY

In this paper, we propose a methodology for real-time object measurement, classification and tracking from temporal point clouds acquired with a multilayer laser scanner. We also propose a methodology to estimate and visualize object behaviors with a spatial model based on a space subdivision framework consisting of agents, activities, resources, and modifiers. We conducted an experiment to evaluate our methodology using temporal point clouds acquired from a construction vehicle in drilling works. We also verified that the space subdivision framework can be applied for construction site visualization with a geofencing approach using the results of real-time object classification and tracking from temporal point clouds. We confirmed that our methodology can extract and track objects with a multilayer laser scanner with real-time processing.

#### REFERENCES

Bailey, T., Durrant-Whyte, H., 2006. Simultaneous localization and mapping (SLAM): part II, Robotics & Automation Magazine, IEEE, Vol. 13, Issue: 3, pp. 108-117.

Doishita, K., Muramoto, E., Kouda, T., 2010. Application of ICT to construction machinery, Komatsu Technical Report, Vol. 56, No.163.

Dong, S., Li, H., Yin, Q., 2018. Building Information Modeling in Combination with Real Time Location Systems and Sensors for Safety Performance Enhancement, Safety Science, Volume 102, pp. 226-237.

Durrant-Whyte, H., Bailey, T., 2006. Simultaneous localization and mapping (SLAM): part I, Robotics & Automation Magazine, IEEE, Vol. 13, Issue: 2, pp. 99-110.

Guo, H., Yu, Y., Skitmore, M., 2017. Visualization technologybased construction safety management: A review, Automation in Construction, Vol.73, pp.135-144.

Sithole, G., Zlatanova, S., 2016. Position, Location, Place and Area: An Indoor Perspective, ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume III-4, pp.89-96.

Vu, T., Burlet, J., Aycard, J., 2011. Grid-based Localization and Local Mapping with Moving Object Detection and Tracking, Information Fusion, Volume 12, Issue 1, pp.58-69.