SEMANTIC SEGMENTATION OF POINT CLOUDS USING WATER-BORNE MMS IN URBAN RIVERS

Tetsu Yamaguchi [1], Takeshi Komori [2], Nobuaki Kubo [3], Etsuro Shimizu [4] and Masafumi Nakagawa [5]
[1]Graduate Student, Department of Civil Engineering, Shibaura Institute of Technology
3-7-5, Toyosu, Koto-ku, Tokyo 135-8548, JAPAN
Email: mh23033@shibaura-it.ac.jp
[2] Graduate Student, Graduate School of Marine Science and Technology, Tokyo University of Marine Science and Technology,
1-6, Etchujima, Koto-ku, Tokyo
Email: m225021@edu.kaiyodai.ac.jp
[3] Professor, Marine Science and Technology, Tokyo University of Marine Science and Technology,
1-6, Etchujima, Koto-ku, Tokyo
Email: nkubo@kaiyodai.ac.jp
[4] Professor, Marine Science and Technology, Tokyo University of Marine Science and Technology,
1-6, Etchujima, Koto-ku, Tokyo
Email: shimizu@kaiyodai.ac.jp
[5]Professor, Department of Civil Engineering, Shibaura Institute of Technology
3-7-5, Toyosu, Koto-ku, Tokyo 135-8548, JAPAN
Email: mnaka@shibaura-it.ac.jp

KEY WORDS: Geographic Information Systems, Point cloud, Segmentation, Simultaneous Localization and Mapping (SLAM), Rivers

ABSTRACT: In crowded and narrow urban rivers, high-resolution 3D maps can assist the navigation of autonomous boats using global navigation satellite system (GNSS) positioning and 3D scanning. In Japan, the Ministry of Land, Infrastructure, Transport and Tourism is promoting project “PLATEAU,” which is a project to develop 3D maps of urban areas. However, compared to land areas including roads and structures, it remains difficult to obtain 3D mapping of river areas, and point cloud data classification is inadequate. Therefore, this study aims to develop classified maps of urban rivers. We propose a method for the semantic segmentation of point cloud data acquired by light detection and ranging for a river as a step towards the mapping of urban rivers. The proposed methodology comprises two steps. The first step is to apply a mobile mapping system to the boat in environments where robust GNSS signals cannot be guaranteed. The second step is to apply semantic segmentation to the point clouds. In this study, the revetments and piers of urban expressways could be extracted with an accuracy of approximately 60%. Because the accuracy of attribute information depends on the results of point cloud acquisition, several technical issues remain with object-recognition information. The first issue is a lack of feature points caused by limited measurement angles and positions from the boat. The second issue is vibration caused by waves when laser scanning.

1. INTRODUCTION

In recent years, the Ministry of Land, Infrastructure, Transport and Tourism (MLIT) and other authorities have been promoting the “PLATEAU” project to digitize cities, which creates 3D city models from various existing sources of information about urban areas, adds attribute information, and maintains them as geospatial data. PLATEAU is a base map for the variety of urban data that can be used not only to improve the efficiency of urban planning work but also to simulate disasters and explore environmental issues. It is expected to be used not only to improve the efficiency of urban planning work but also to simulate disasters and address environmental issues. Wide-area urban spaces, such as road spaces and the indoor spaces of buildings, have already been developed as 3D maps. However, in the urban river space, despite the expected demand for its use during disasters and for autonomous boat navigation, data development that can meet these requirements has not progressed to date. In developing such data, it is necessary to consider that urban rivers are often too narrow or too low for boats to safely navigate and guidelines must be set according to water levels that can change over time or are linked to tide levels. This is because many cities, such as Tokyo, are in waterfront areas, and many urban rivers are located near waterfront areas. Therefore, 2D maps are not sufficient and 3D maps need to be developed. In this study, a method for generating 3D models with attribute information added by applying semantic segmentation to point clouds acquired from ships was investigated to meet the demand for their use in urban rivers during disasters and for autonomous boat navigation.

2. METHODOLOGY

The proposed method is outlined in Figure 1. In this study, we propose to generate a map model using point clouds acquired by a waterborne mobile mapping system (MMS). This study also focuses on the classification of landmarks from point clouds. The MMS point clouds used in this study are those acquired by the Centimeter-Level Augmentation Service (CLAS) and by simultaneous localization and mapping (SLAM) using two light detection and ranging (LiDAR) units(Naoto Kimura et al., 2023). First, the point cloud is partitioned into groups based on the relationship between the nearest-neighbor points at each point in the acquired point cloud. Clustering is then performed by applying the proposed semantic segmentation method to the point cloud. This creates groups of points and outputs labeled point clouds. The surface is reconstructed by using a ball-pivot method on these point clouds. The surface model reconstructed by these algorithms is fitted to the point cloud to generate polygons and then used as a vector model of the MMS point clouds.

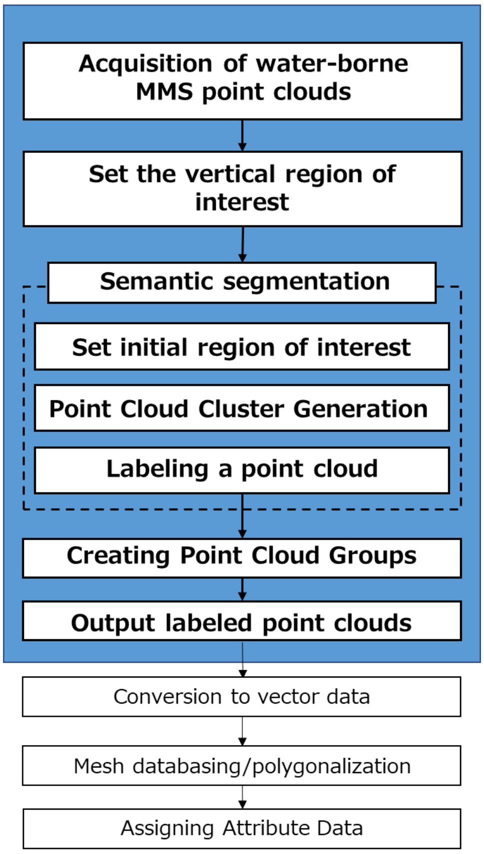

Figure 1. Flowchart of the proposed methodology processing.

2.1 Semantic segmentation

The semantic segmentation used in the proposed method estimates the attributes of geographic features in the acquired point clouds by estimating river structures and buildings based on known information such as geographic features and then applying multi-planar classification. Multi-planar classification is achieved by combining random-sample consensus (RANSAC) planar extraction with the clustering process. For the road domain, there are many datasets that can be used as training data, and methods such as deep learning have been proposed (Cheng WANG et al., 2020). Many of these methods are based on images acquired by in-vehicle MMSs (Yangzi Cong et al., 2023) and attempt to label attributes such as building facades, traffic signals, and signs(Fumiki Tonomura et al., 2013; Shi Pu et al., 2011; Takashi Michikawa et al., 2015). However, there is less training data for the river domain than for the road domain, with adequate datasets yet to be constructed. In addition, the target objects in the river domain are often revetments and piers, which are obstacles to navigation but are treated differently from those in the road domain. In addition to LiDAR, many methods use cameras for measurement in the road domain. However, for urban rivers, the degree of change in light intensity between sections with good overhead visibility and dark sections with no lighting environment, such as the lower sections of bridges and viaducts, makes it difficult to use image measurement as the primary method. Therefore, in this study, LiDAR is the primary sensor used to acquire the geometry of geographic features. However, although LiDAR is more robust against light intensity changes than a camera, the point cloud obtained by LiDAR is sparse compared to an image, making it difficult to read and recognize objects. For this reason, images captured by a camera are used as auxiliary data for attribute data assignment and grouping accuracy verification.

2.2 Attribute information assignments

In assigning attribute information, the previously acquired point clouds are grouped according to the neighboring points from each point. The grouped points are then labeled, and each label is classified according to the characteristics of the bridge or revetment to provide attribute information. For shape estimation, visual shape estimation is performed using images acquired by a 360-degree camera and a digital camera. The M-estimator sample consensus (MSAC) algorithm is used to estimate the attributes of the point cloud. However, this algorithm cannot separate many planes simultaneously. In addition, because of noise in the point cloud, even flat objects are not portrayed as perfectly flat in the point cloud(S. Xu et al., 2014). Therefore, we address these problems by using prior knowledge to refine the attributes, such as using the fact that a metropolitan highway slab will be overhead, and that revetments and bridge piers are geographic features that will be in contact with the water surface. For these knowledge-based attribute estimations, the type of geographic feature is estimated based on the relative position of the measurement target from the origin of the LiDAR position on a small boat (see Figure 2).

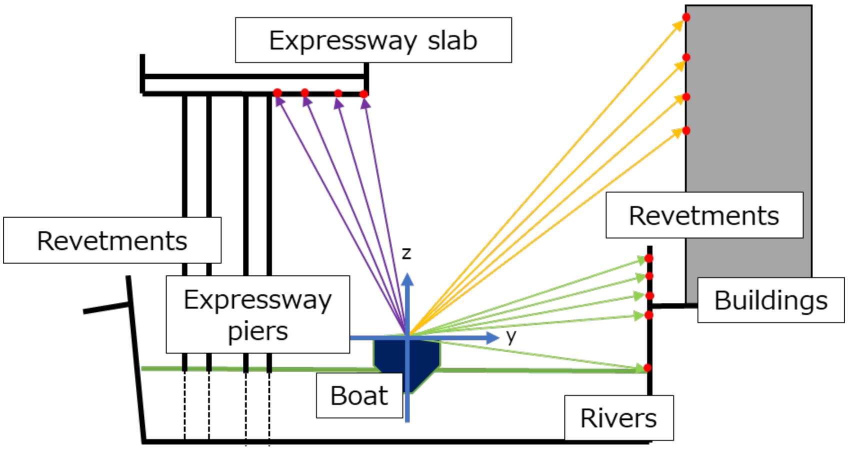

Figure 2. Objects measured using LiDAR from a boat

3. EXPERIMENTS

The canals of the Nihonbashi River, the Kanda River, and the Shiomi area in urban Tokyo were selected for the experiment (see Figure 3). On-water measurements were taken on December 3, 2021, September 28 (Kanda River), November 14 (Nihonbashi River), and December 2 (Shiomi Canal), 2022. The water-borne MMS point clouds were acquired using the battery-propelled boat “Raicho I” (see Figure 4), which was equipped with a horizontal-scan LiDAR (VLP-32C, Velodyne), diagonal-scan LiDAR (VLP-16, Velodyne), and an omnidirectional camera (Ladybug 5+, FLIR). The navigational paths were set to be the Kayababashi to Suidobashi section, which is covered by the Metropolitan Expressway (an urban expressway), the Asakusabashi to Suidobashi section, which has a clear view of the sky, and the Shiomi Canal area, where the river is relatively wide and there are few structures in the vicinity. These rivers are characterized by being narrow rivers with a width of about 10 m for almost all sections. For the Sumida River, where the river is wide, and in the Kiba area, it was not easy to acquire sufficient point clouds because they were sometimes outside the coverage area of the LiDAR system. Measurements for these sections were therefore excluded. Moreover, because many of the bridges in these rivers have low girders, which might hinder the safe navigation of ships, point clouds were not acquired during high tide. For similar reasons, no measurements were taken in the section of the Kanda River upstream of the Iida Bridge, where the water is shallow and there would be a risk of stranding. Of the measurement targets, point cloud processing was performed only for the Kanda and Nihonbashi Rivers because the point clouds created were of poor quality in the Shiomi Canal, caused by SLAM degeneration. A desktop PC (Intel Core- i7 12700, 3.6 GHz 16 GB RAM) was used for the point cloud processing that extracted and created vector models of revetments, Metropolitan Expressway slabs, and Metropolitan Expressway piers.

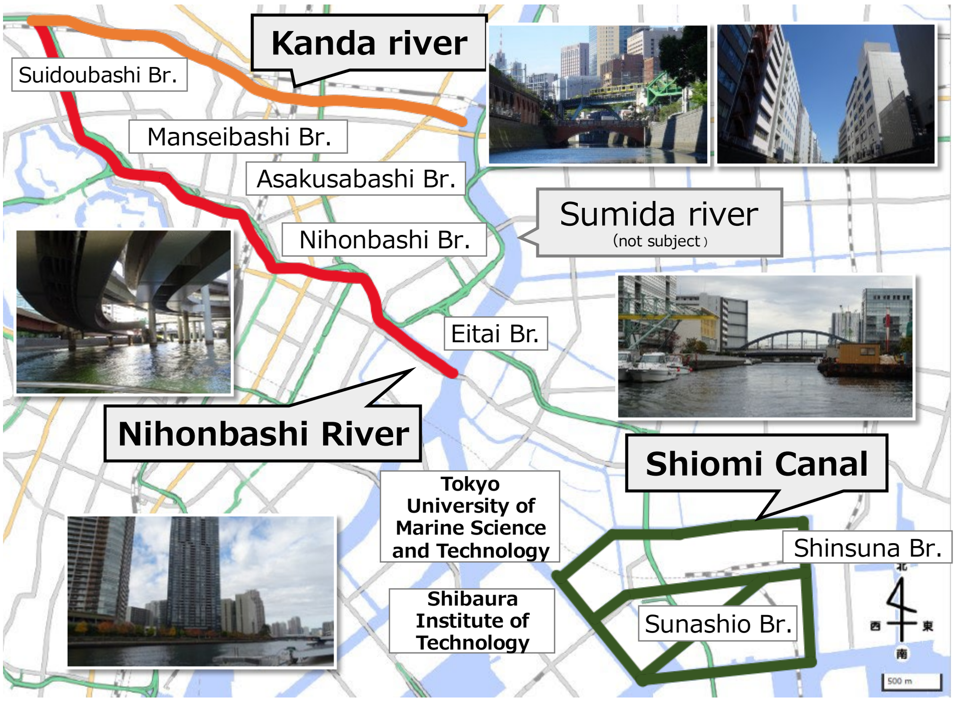

Figure 3. Study area

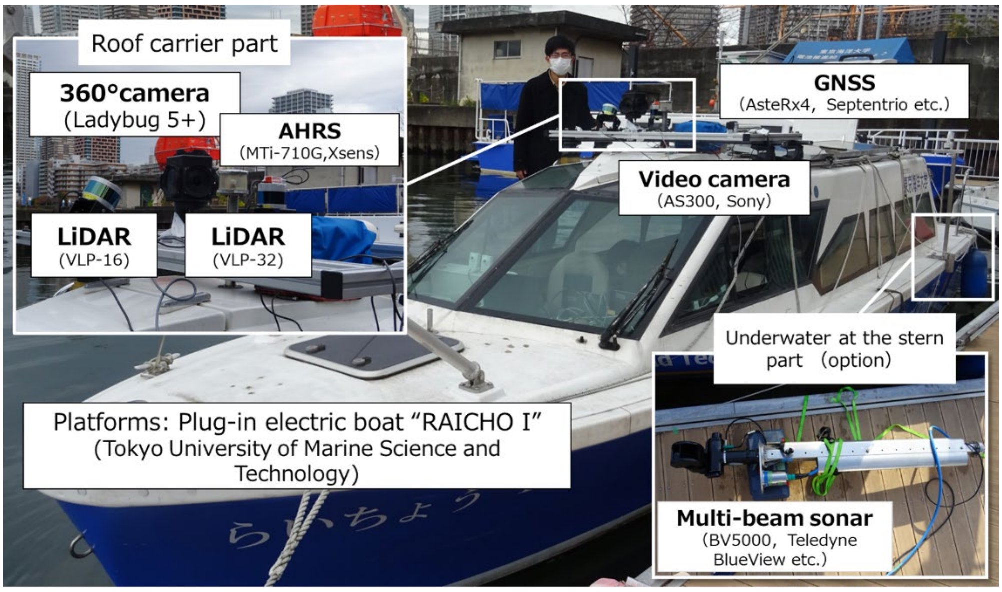

Figure 4. Sensors mounted on the waterborne MMS

4. RESULTS

Figure 5 shows part of the segmentation results, namely the results of the segmentation of the revetment, the slab of the Metropolitan Highway and the piers of the Metropolitan Highway. The results of the surface model generation and polygon generation for these waterborne MMS point clouds using the proposed method are shown in Figure 6.

4.1 Processing time of segmentation

The segmentation process was performed using the proposed method. The number of input points was 2,098,080 and the number of frames was 500. First, the process stage of constructing a 2D environmental map using only horizontal LiDAR was denoted on the environmental map construction (2D) stage. Next, the process stage of adding height information to the horizontal map by combining horizontal and oblique LiDAR was denoted on the environmental map construction (3D) stage. Vector model generation was used to generate a surface model from the point clouds created by these processes. The processing time results (s) and processing time per frame (frames/ms) are given in Table 1.

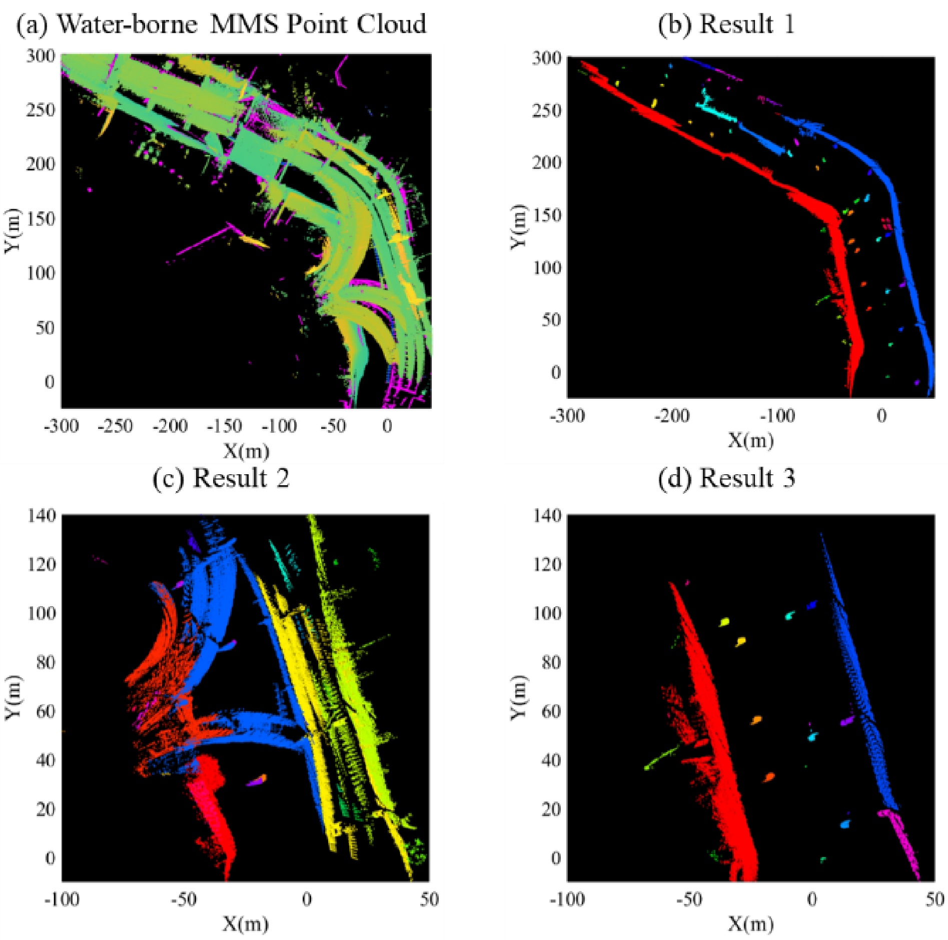

Figure 5. Point cloud and segmentation results around Nihonbashi Bridge

(The revetment, the slab of the Metropolitan Highway and the piers of the Metropolitan Highway are shown)

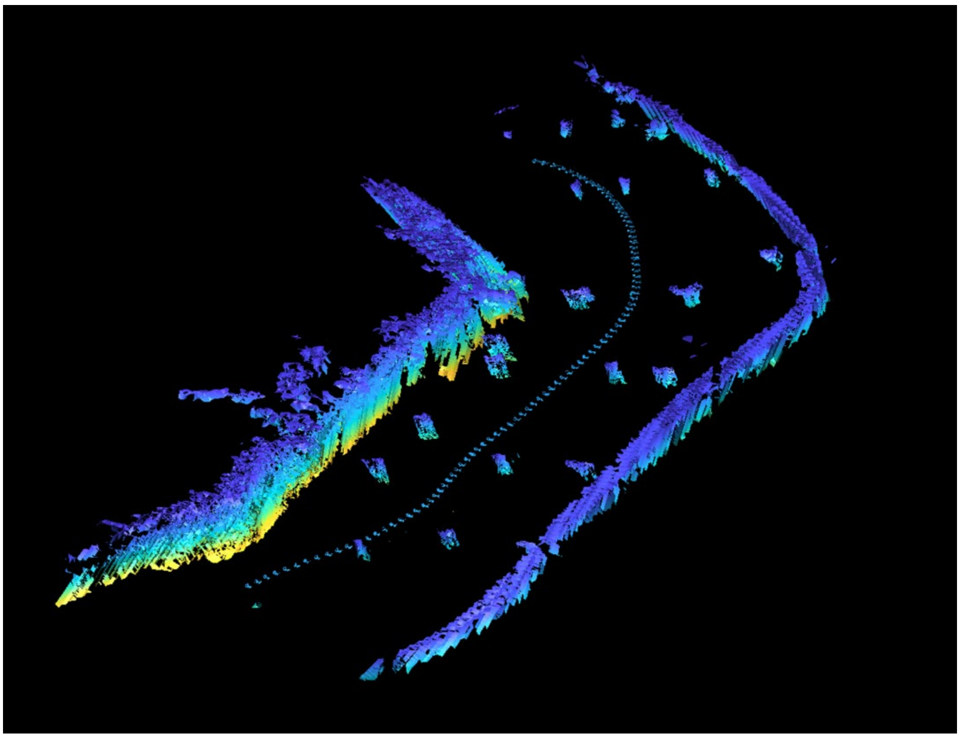

Figure 6. Result of polygon model generation around Nihonbashi bridge.

4.2 Processing results of segmentation

The numbers of revetments, piers of the Metropolitan Expressway, and slabs of the Metropolitan Expressway extracted by segmentation and the number of visual readings based on the image data acquired by the 360-degree camera within the set processing range were classified according to five evaluation aspects for each generated segment, as shown in Table 2.

Table 1. Processing time for the segmentation process

Processing	Processing time (s)	Processing time per frame (flame/ms)
Environmental map construction(2D)	121.81	243.6
Environmental map construction(3D)	36.92	73.8
Extraction of revetment	0.86	1.7
Extraction of piers	2.60	5.2
Extraction of slab	1.87	3.7
Generating Vector Models	17.13	34.2

Table 2. Extraction pattern classification in the same process

Visual readings based on the image data	Revetments	Piers	Slab	Total
a) Exact match between extracted segment and geographic object	6	12	3	21
b) Multiple segments make up a single geographic feature	1	0	3	4
c) One segment contains more than one geographic feature.	1	1	4	6
d) No segment for land objects	0	0	4	4
e) Segment does not correspond to land property	The number of segment : 9			

5. DISCUSSION

The results for the ship MMS point cloud acquisition and segmentation in an urban river showed that the revetment, the piers of the Metropolitan Expressway, and the slab of the Metropolitan Expressway could be extracted with a significant degree of accuracy.

The accuracy of the extraction varied depending on the geographic features involved. The seawall and the piers of the Metropolitan Highway were easily visible from the boat and were within the observation range of the LiDAR system. However, the floor slab for the Tokyo Metropolitan Expressway and the piers of some railroad bridges were only approximately captured because they were partially obscured by other geographic features, depending on their relative position with respect to the intersection and navigation route (see Figure 2).

Not all surfaces could be measured and recognized, and there were sections where there were no segments visible for geographic objects. The piers of a bridge on a public road were not acquired as a point cloud that included the slab and abutment, which could not be separated from the revetment when estimating the geometry of the geographic feature, making it difficult to recognize the object. Moreover, because the experiment was conducted at low tide (for safe navigation), the measurement height was low, and the point cloud of buildings obscured by the high revetment could not be acquired accurately, which may have caused some defects in the shape estimation of the revetment and the attribute estimation for buildings around the river. However, because the point cloud was acquired via one-way navigation in this experiment, there is the possibility that these problems could be improved by multiple round-trip navigations (at a cost to measurement efficiency). In the segmentation process, a large proportion of the processing time was required for SLAM processing. Because the point clouds integrated by SLAM are used to transform the data, their accuracy can have a significant impact on the segmentation accuracy. Therefore, it will be necessary to reduce the SLAM processing time if the accuracy is to be improved.

Since underwater point cloud data has not yet been acquired, it is expected that underwater structures will also be addressed by acquiring and integrating these data.

6. CONCLUSIONS

In this study, we have investigated a semantic segmentation method for creating urban river maps from waterborne MMS point clouds. We confirmed that about 60% of information about geographic features that cannot be obtained by conventional methods can be extracted by applying waterborne MMSs to urban rivers, where stable measurement is otherwise difficult. However, there are issues with processing time and accuracy in the preprocessing stages. Despite this, we believe that the 3D model generated by the proposed method can contribute to the effective use of autonomous vessels in urban rivers both in normal times and during disasters.

ACKNOWLEDGMENTS

This research was supported by the MEXT Coordination Funds for Promoting Aerospace Utilization; Grant Number JPJ21460038.

REFERENCES

Cheng, W., Chenglu, W., Yudi, D., Shangshu, Y., Minghao, L., 2020. Urban 3D modeling using mobile laser scanning: a review. *Virtual Reality & Intelligent Hardware 2020*, Vol 2, Issue 3:175-212.

Tonomura, F., Ishikawa, K., Amano, Y., Hashizume, T., 2013. Recognition of road objects from 3D point cloud of Mobile Mapping data Performance evaluation of training data made of Mobile Mapping data. *JSPE 2013*, pp. 857-859.

Infrastructure, Transport and Tourism (MLIT) , 2023. *PLATEAU*, https://www.mlit.go.jp/plateau/.

Kimura, N., Nakagawa, M., Ozeki, T., Kubo, N., Shimizu, E., 2022. Seamless Indoor-outdoor Positioning and Trajectory Interpolation using SLAM and PPP-RTK for River Mapping. *The 43rd Asian Conference on Remote Sensing.*

Shi, P., Martin, R., George V., Sander, O., 2011. Recognizing basic structures from mobile laser scanning data for road inventory studies. *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 66, pp. s28-s39.

Xu, S., Vosselman, G., Oude Elberink, S., 2014. Multiple-entity based classification of airborne laser scanning data in urban areas. *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 88, pp. 1-15.

Michikawa, T., Moriwaki, K., Yabuki, N., Fukuda, T., Hara, K., Kurimoto, S., 2015. Extraction of roadside trees from point clouds and its application to maintenance. *Environmental systems research*, pp89-95.

Yangzi, C., Chi, C., Bisheng, Y., Fuxun, L., Ruiqi, M., Fei, Z., 2023. Change-aware online 3D mapping with heterogeneous multi-beam and push-broom LiDAR point clouds. *ISPRS Journal of Photogrammetry and Remote Sensing*, pp.204-219.