Segmentation and Feature Classification of Point Clouds Acquired by LiDAR-SLAM in Urban Rivers

Tetsu Yamaguchi[1*], Naoto Kimura[1], Takeshi Komori[2], Nobuaki Kubo[2], Etsuro Shimizu[2], Masafumi Nakagawa[1]

[1]Shibaura Institute of Technology, Japan

[2]Tokyo University of Marine Science and Technology, Japan

*mh23033@shibaura-it.ac.jp

Abstract *In urban crowded and narrow rivers, high-resolution 3D maps can assist the navigation of autonomous boats with GNSS positioning and 3D scanning. In Japan, the Ministry of Land, Infrastructure, Transport and Tourism (MLIT) is promoting the "project PLATEAU," which is a project to develop 3D maps of urban river areas. However, compared with land areas such as roads and structures, the 3D mapping of river areas is still difficult to obtain data and is still insufficient in many aspects. Therefore, this study aims to develop a map of urban rivers and implement our proposed methodology on autonomous boats. In this study, we proposed a method for the segmentation of point cloud data acquired by a handheld simultaneous localization and mapping (SLAM)-light detection and ranging (LiDAR) and water-borne mobile mapping system (MMS) on a river as a stage in the mapping of urban rivers. The proposed methodology consists of two steps. The first step is point cloud acquisition using a handheld SLAM-LiDAR and water-borne MMS on a boat in urban river environments where GNSS signals are unstably received. The second step is a semantic segmentation of the acquired point clouds to be used for the geographic information system (GIS). In our experiments, we confirmed that revetments, river crossing structures, and piers of urban expressways can be measured by a handheld SLAM-LiDAR and water-borne MMS from rivers. We also confirmed that revetments and piers of urban expressways were extracted with an accuracy of approximately 60%. The 3D model generated by the proposed method can contribute to the effective use of autonomous boats in urban rivers both in peacetime and during disasters. Some in-river structures of urban expressways were not completely extracted. Because the accuracy of attribute information depends on the results of point cloud acquisition, several technical issues remained in object recognition and segmentation of point clouds. The first issue is the lack of feature points because of the limitation of measurement angle and position from the boat. The second issue was vibration caused by water waves in laser scanning. False point clouds caused by bubbles captured by LiDAR on the water surface cause errors during segmentation.*

Keywords: *Point clouds, Segmentation, Simultaneous Localization and Mapping (SLAM), Urban Rivers, Geographic Information Systems*

Introduction

In recent years, the Ministry of Land, Infrastructure, Transport and Tourism (MLIT) and other organizations in Japan have been promoting projects such as the PLATEAU and VIRTUAL SHIZUOKA to digitize cities for urban digital twinning. In these projects, 3D geometrical city models of buildings and other features are created from data obtained by survey data, and attribute information, such as buildings' names, age of building, and building uses, which is added to the models. These data are described by CityGML for multipurpose use as open data. Various urban activity data are integrated into a 3D urban model to achieve a fusion of physical and cyber spaces for urban digital twinning with the data cycle. VIRTUAL SHIZUOKA is a project in which Shizuoka Prefecture, one of the local governments, is trying to utilize 3D urban models not only for urban development, infrastructure maintenance, and disaster prevention, but also for tourism, automatic driving, entertainment, and other applications throughout society. The PLATEAU is expected to be used as a base map for a variety of urban data to improve the efficiency of urban planning, disaster prevention, and environmental management. Although the PLATEAU covers various features, such as buildings, roads, and indoor spaces, in urban spaces, in the case of an urban river space, despite the expected demand for utilization in times of disaster and autonomous vessel navigation, data development to meet these requirements has not progressed. Project PLATEAU provides data on rivers, including flood inundation zones for rivers managed by the national government and the metropolitan government, but the data from previous research by MLIT are aimed at flood control in urban areas. Therefore, urban rivers are not modeled in detail. In addition, the data do not provide accurate geometry data for urban rivers. The lack of in-channel data means that information such as water depths, river widths, and allowable heights are not included. Furthermore, although this information is designed to be superimposed on a geographic coordinate system, it is primarily intended for mapping submerged areas in the event of a disaster and cannot be used to map the river itself or to delineate navigable areas for boats. To map urban rivers, it is necessary to reproduce the exact geometry of revetments and piers. It is also necessary to ensure that rivers are wide enough and deep enough for boats to navigate safely and to consider changes in elevation conditions caused by changes in tidal levels. Therefore, 2D maps are not sufficient and 3D maps need to be developed. In this study, we focus on the classification of landmarks from point clouds measured from a boat. We propose a method of adding attribute information to point clouds by semantic segmentation. Moreover, we verify the proposed methodology in experiments using a boat in urban rivers.

Literature Review

In urban river environments, the structural elements are the revetments and slabs of the Metropolitan Expressway. Although the structures are generally built vertically, the revetments are sometimes not. The Metropolitan Expressway slabs are also sometimes not horizontal, so it is necessary to use a method that accounts for the slope. Other previous research has proposed a method that can quickly scan large urban areas by using dynamic moving platforms in larger road sections (Norbert Haala et al., 2008). The goal of this study is to quickly collect data in large areas where the measurement area includes tunnels or where buildings are densely constructed. In this study, especially in a non-GNSS environment such as a Japanese bridge, dense point coverage of the facades of the adjacent architecture is required. Therefore, these methods are improved and used as water-borne MMS. In previous studies, some indoor locations sharing the same lack of GNSS availability have used mobile laser scanning (MLS) systems to process point cloud data acquired to identify and reconstruct structural elements (walls, floors, ceilings, openings) and furniture in buildings (Shayan Nikoohemata et al., 2020). However, not all building walls and ceilings are horizontal or vertical, so these feature detections are recognized as technical issues.

Methodology

The proposed method is shown in Figure 1. We propose a methodology to generate a map model using point clouds acquired by a water-borne mobile mapping system (MMS). We also focus on the classification of landmarks from point clouds. The MMS point clouds are acquired by the GNSS and by simultaneous localization and mapping (SLAM) using two light detection and ranging (LiDAR) units (Naoto Kimura et al., 2023). SLAM is a simultaneous processing of self-position estimation and point cloud acquisition. SLAM includes visual SLAM, depth SLAM, and LiDAR-SLAM. For river measurements, we navigate boats in environments with large changes in lighting from well-illuminated open sky areas to dark areas such as under bridges and river crossing structures without lighting. Thus, the visual SLAM using a camera has a disadvantage in 3D measurement and odometry. Moreover, for river measurements, the distance from a boat to measured objects is 5 m or more. Thus, depth SLAM using a time-of-flight camera has also a disadvantage. Therefore, we select LiDAR-SLAM using LiDAR, which has the advantage of long-distance measurement in illumination-changing river environments. In this study, point clouds are acquired by a handheld SLAM-LiDAR, and point clouds are acquired by a water-borne MMS. First, point clouds are acquired from boats using LiDAR. Clustering is then performed by applying the proposed semantic segmentation method to the point cloud. The three main processes are the setting of the region of interest, point cluster generation, and point cloud labeling. 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.

By accumulating data, it is possible to determine the difference from the previous measurement. The data obtained from these processes are employed when the difference is below a certain threshold value and are used for data for river surveys and autonomous navigation.

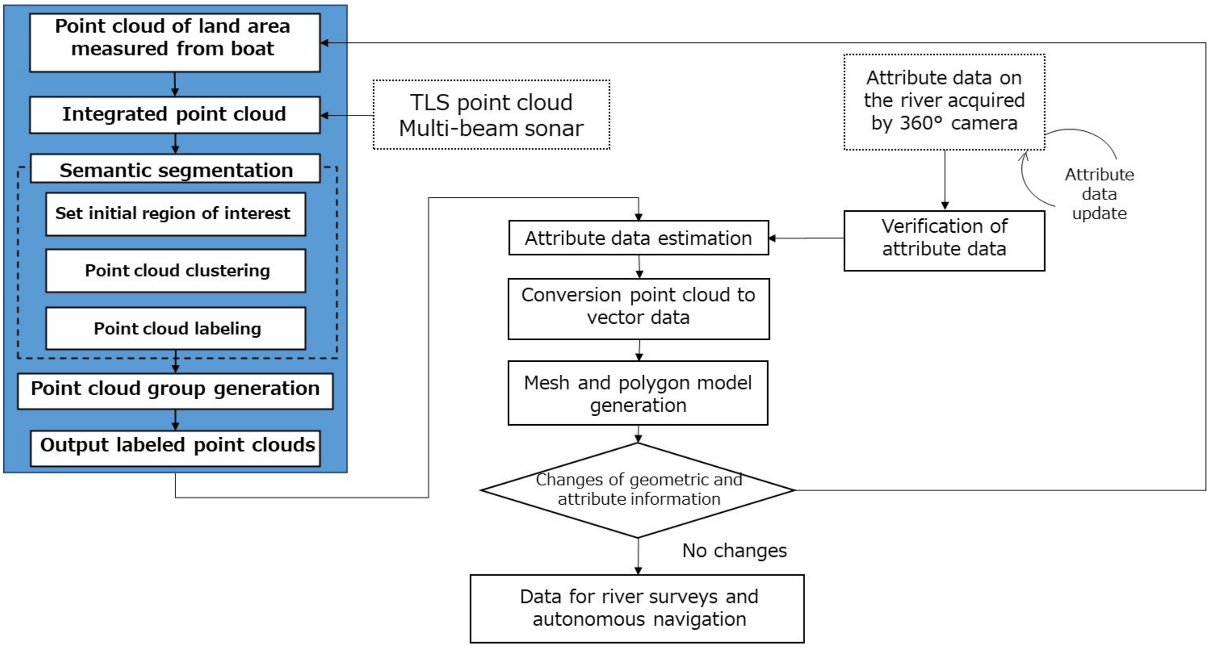

Figure 1: Proposed methodology.

a. Point Cloud Cluster Generation:

In the point cloud cluster generation process, the attributes of river structures and buildings are estimated from point clouds based on knowledge such as landform geometry and basic building shapes. Moreover, a planar extraction process is applied to the point clouds with random sample consensus (RANSAC) and a multiplanar classification process combining planar extraction and clustering processes. For the road domain, many data sets 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). On the one hand, there is less training data for the river domain than for the road domain, with adequate data sets yet to be constructed. The target objects in the river domain are often revetments and piers, and these are obstacles to navigation but are treated differently from those in the road domain. In addition, the target objects in river space are often revetments and piers that are obstacles to navigation. This is different from the objectives of the road space, where the purpose is to detect a large number of pedestrians, prevent collisions with other moving objects, and acquire peripheral information by reading road signs. Therefore, it is not desirable to apply the training data obtained in road spaces directly to urban river spaces.

Many methods use cameras for measurement in road spaces. However, in urban river spaces, the amount of change in light intensity between sections with good overhead visibility and dark sections without illumination, such as the lower sections of bridges and viaducts, is large, and methods that mainly use image measurement are not suitable for use. Although SLAM-LiDAR is more robust against light intensity changes than an RGB camera, point clouds obtained by LiDAR are sparse compared with image-based point cloud acquisition. Thus, images captured by an RGB camera are used as auxiliary data for verification of attribute data assignment.

b. Assignment of Attribute Information:

In attribute information estimation from point clouds, we group the neighboring points from each point in the point clouds of the handheld LiDAR and the water-borne MMS equipped with a boat. The point cloud clusters are labeled, and each label is classified according to the characteristics of bridges or revetments to provide attribute information. For geometrical shape estimation, visual confirmation using images acquired by a panorama camera is applied. The M-estimator sample consensus (MSAC) algorithm is used to estimate the attributes of the point clouds. This algorithm cannot separate many planes simultaneously. Therefore, the algorithm can deal with these problems by narrowing down the attributes using prior knowledge that metropolitan highway slabs are overhead and that revetments and bridge piers are geographic features that are in contact with the water surface. The problem is addressed by using a prior knowledge such as following in the knowledge-based attribute estimation using the sequence of points, scan angle, and height, and the type of geographic feature is estimated based on the relative positional relationship between the LiDAR position (external target element) and the measurement target. The following prior knowledge was used. Revetments are roughly longitudinal structures, located perpendicular to the navigation channel, concerning boats. The buildings are built vertically and are located further away from the revetments with respect to the boats. The slab covering the sky is also located directly above the boats, and its geometrical feature is approximately horizontal. A pier is a cylindrical geometrical feature, and it is located on the boat side of the river space; that is, farther from the revetments as shown in Figure 2.

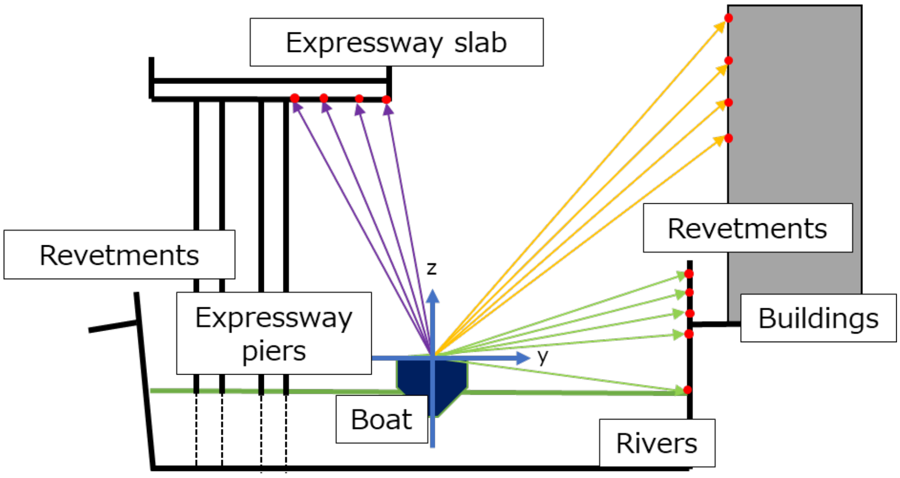

Figure 2: Objects measured using LiDAR from a boat.

The process of constructing a two-dimensional environmental map using only horizontal LiDAR is called the environmental map construction (2D) stage, and the process of adding height information to a horizontal map using both horizontal and diagonal LiDAR is called the environmental map construction (3D) stage.

c. Processing Results of Segmentation:

For evaluation, the five evaluation items for each generated segment, based on the number of revetments, piers of the Metropolitan Expressway, and slabs as shown in Figure 3 of the Metropolitan Expressway extracted by segmentation and visual verification using the panorama camera in the study area was classified into the following five patterns.

a) Correct matches: point clouds are represented as a segment

b) Separated segments: point clouds are represented as parallel segments

c) Merged segments: point clouds are represented as merged segments

d) No segments: point clouds noises or no objects

e) No point clouds: no measured point clouds

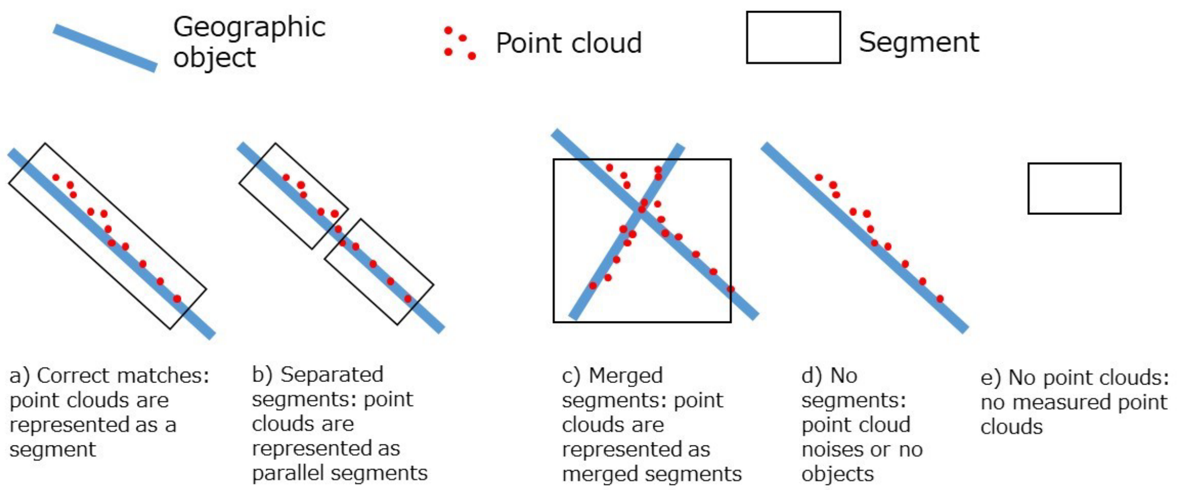

Figure 3: Polygon model generation.

Experiments

a. Measurement Area:

The Nihonbashi River and the Kanda River were selected as experimental targets as shown in Figure 3, and 3D measurements were conducted on November 14, 2022, September 13, 2023, and September 29, 2023, using the battery-propelled boat Raicho I as shown in Figure 4. We set the navigation path as from the Kayababashi Bridge to the Suidobashi Bridge at the Nihonbashi River, where the metropolitan highway slab covered the sky shown in Figure 3. We also set the navigation path from the Asakusabashi Bridge to the Suidobashi Bridge at the Kanda River, where the sky visibility was sufficient. These rivers were characterized as narrow urban rivers with approximately 10 m width and 3 m depth along almost all sections. High-rise and crowded buildings existed along the revetments. The navigation experiment routes included construction sites in the rivers. In addition, boat navigation was possible only at low tide because many of the bridges in these rivers had low girders, potentially hindering the safe navigation of boats. The section of the Kanda River upstream of the Suidobashi Bridge where the shallow water depth was rejected from our study areas. Water-borne MMS performs SLAM processing from the obtained scan data. On the other hand, handheld LiDAR uses point clouds that have already been SLAM processed.

b. Measurement Method:

We acquired point clouds with a handheld SLAM-LiDAR from the rear deck of the boat. We also acquired point clouds with a water-borne MMS consisting of two LiDAR and an omnidirectional camera mounted on the roof carrier of the boat as shown in Figure 4.

For the LiDAR, we used handheld SLAM-LiDAR (Hovermap, STX Emesent) as shown in Figure 4. We also used horizontal LiDAR (VLP-32C), diagonal LiDAR (VLP-16, Velodyne), GNSS antenna and receiver for centimeter-level augmentation services (CLAS) (AsteRx4 and mosaic-X5, Septentrio), and an omnidirectional camera (Ladybug 5+, FLIR) mounted on the boat's roof carrier as an auxiliary data.

The point cloud of a water-borne MMS equipped with a boat has a different coordinate system because the origin is set to the starting point of the measurement in post-processing. The direction of river longitude and river traverse can be estimated from the trajectory of hand-held LiDAR. Unlike boat-mounted LiDAR, handheld LiDAR requires a process to estimate the position of the origin point. Therefore, the offset between the boat-mounted LiDAR and the handheld LiDAR was measured during the experiment, and alignment was performed by post-processing.

A desktop PC (Intel Core- i7 12700, 3.6 GHz, 16 GB RAM) was used for point cloud post processing to extract and reconstruct geometrical features, such as revetments, metropolitan highway slabs, and metropolitan highway bridges piers, from point clouds obtained by the water-borne MMS.

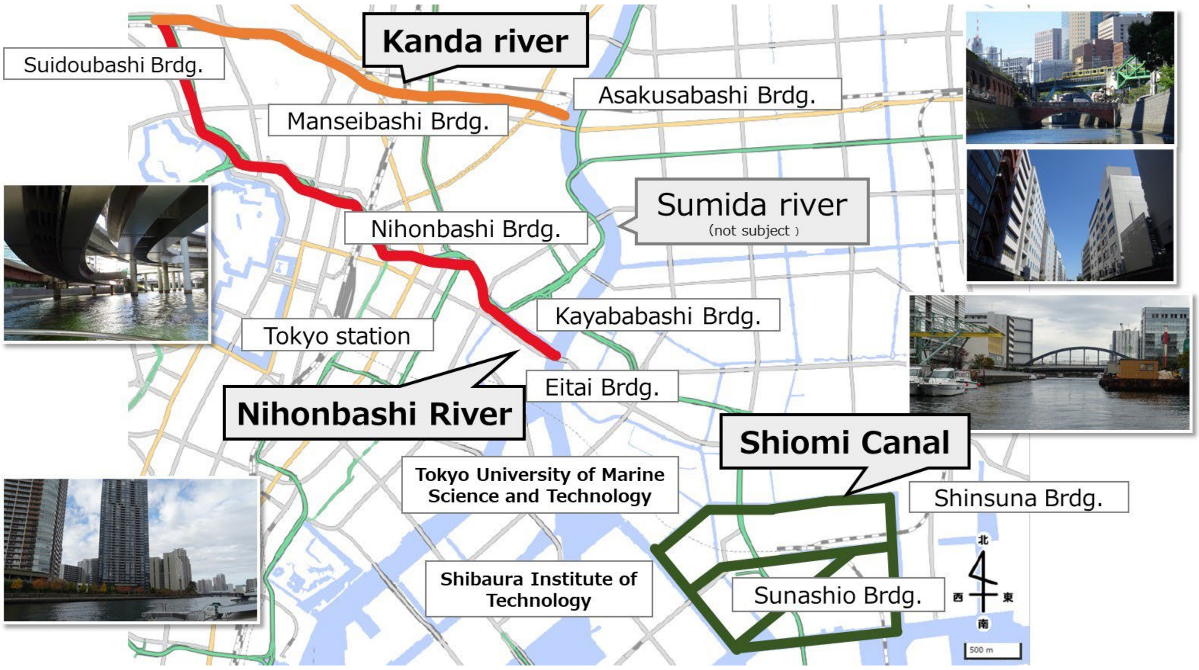

Figure 4: Study area.

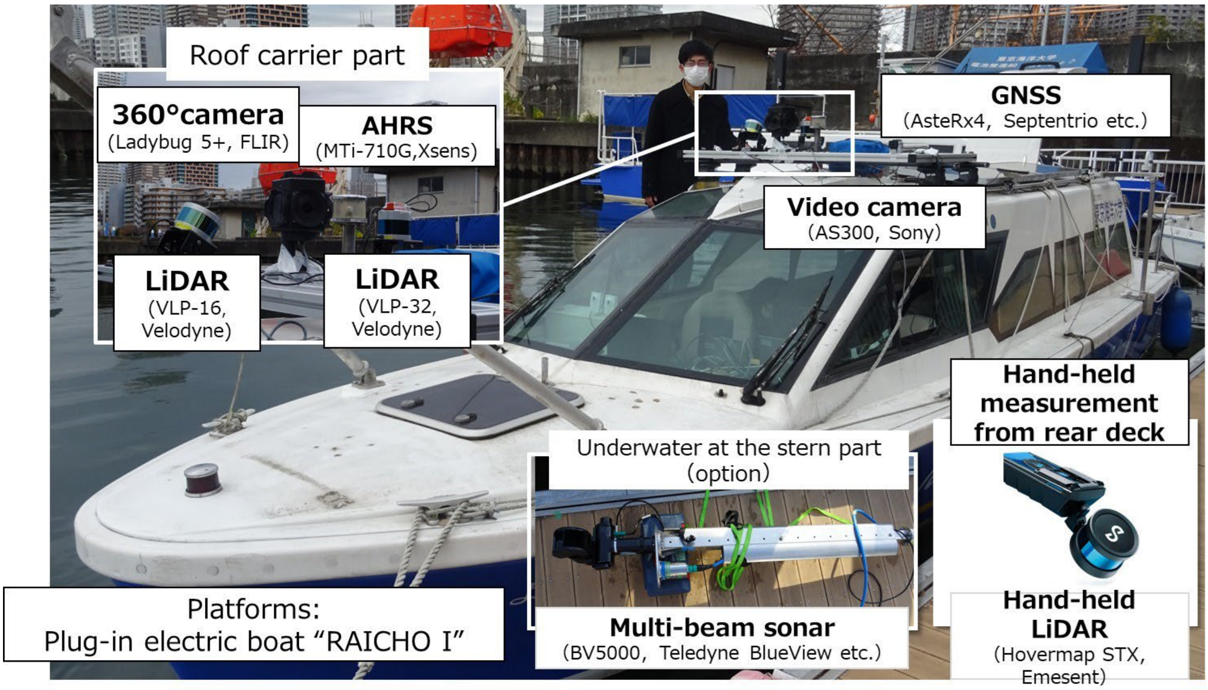

Figure 5: Sensors mounted on the water-borne MMS.

Results

a. Segmentation Results:

Figure 6 shows a part of the segmentation results of revetments and the slabs and the piers of the Metropolitan Highway generated from point clouds using the water-borne MMS. Figure 6 shows the other part of the segmentation result generated from point clouds using the hand-held LiDAR. The results of the surface model and polygon generation for these point clouds using the proposed method are shown in Figure 7 and Figure 8.

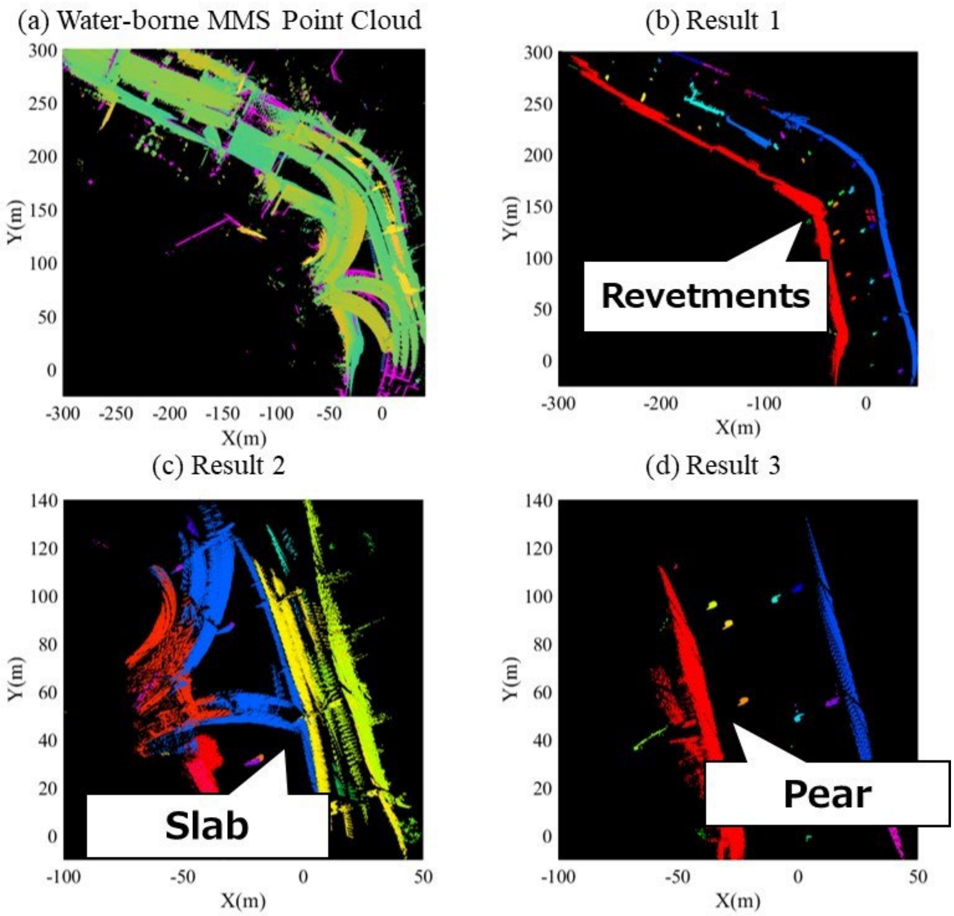

Figure 6: Point cloud and segmentation results around Nihonbashi Bridge.

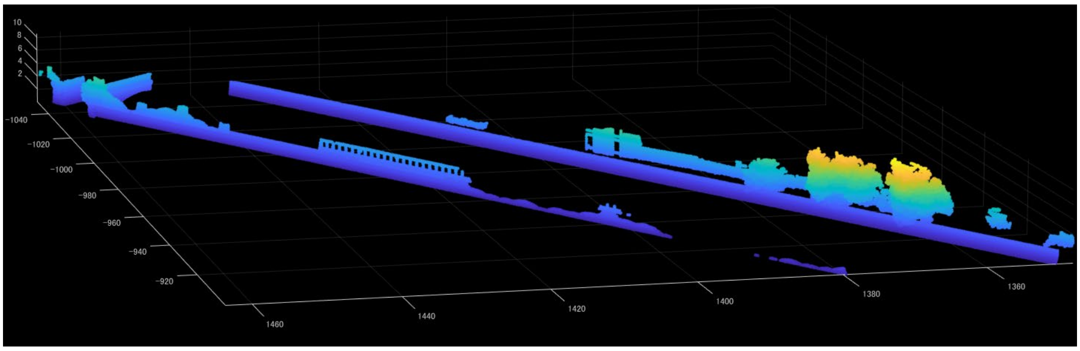

Figure 7: Point cloud and segmentation results around the Kanda River.

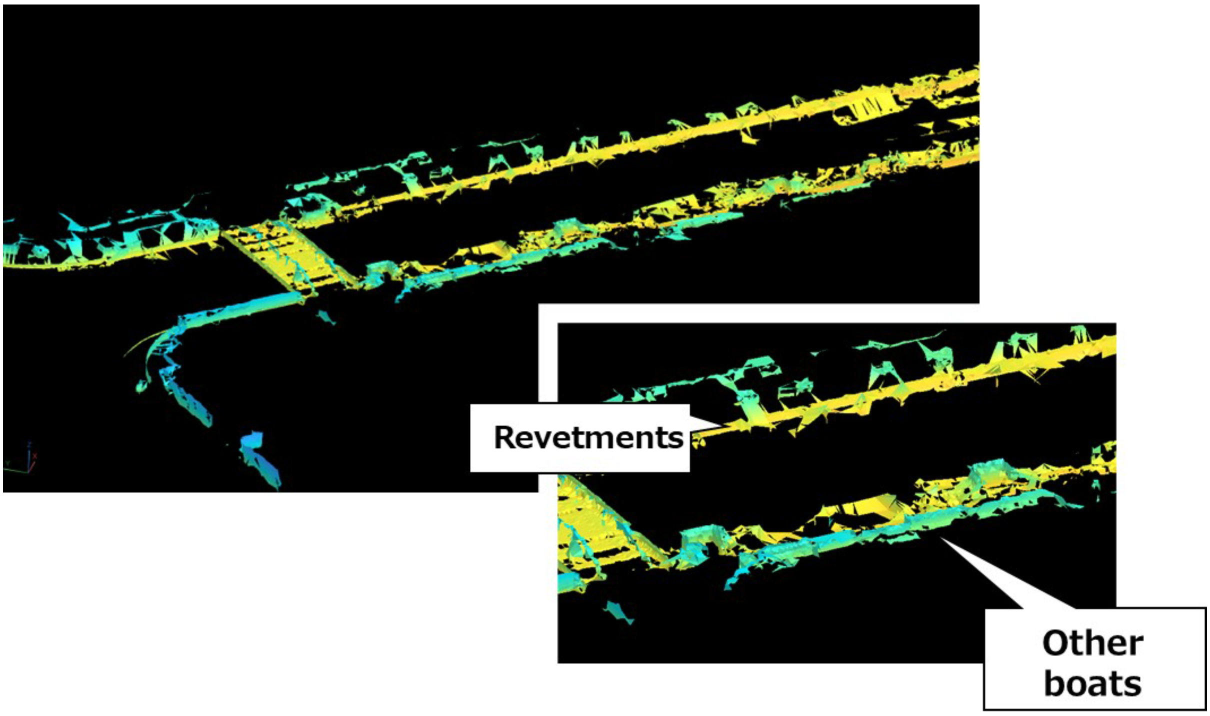

Figure 8: Result of polygon model generation.

b. Processing Time of Segmentation:

Our proposed segmentation was applied for water-borne MMS point clouds (2,098,080 point clouds/500 frames) and the hand-held LiDAR point clouds (1,000,000 points). The processing time results are shown in Table 1.

Table 1: Processing time.

Processing	Water-borne MMS Processing time (s)	Handheld LiDAR Processing time (s)
Environmental map construction(2D)	121.81	-
Environmental map construction(3D)	36.92	-
Preprocessing	-	10.53
Extraction of revetments	0.86	1.42
Extraction of piers	2.60	4.02
Extraction of slabs	1.87	2.44
Vector model generation	17.13	29.62
The number of point clouds (points)	2,098,080	1,000,000

d. Processing

The number of revetments, piers, and slabs of the Metropolitan Expressway extracted by segmentation are shown in Table 2. We classified the segmentation results into five evaluation items based on visual inspection using panorama images, as shown in Table 2.

Table 2: Extraction pattern classification in the same process.

Visual inspection results	Revetments	Piers	Slabs	Total
a) Correct matches: point clouds are represented as a segment	6	12	3	21
b) Separated segments: point clouds are represented as parallel segments	1	0	3	4
c) Merged segments: point clouds are represented as merged segments	1	1	4	6
d) No segment	0	0	4	4
e) No point clouds: no measured point clouds	The number of segments: 9			

Discussion

As a result of water-borne MMS point cloud acquisition and segmentation in an urban river, revetments, expressway slabs, and piers were extracted. However, as shown in Figure 2, not all surfaces could be measured and recognized because the geographic features were obscured by other geographic features depending on their relative positions to the intersections and navigation paths. The reasons for this could be that the experiment was conducted at low tide for safety reasons, the measurement height was low, and there were many occlusions in the river space because of the presence of various geological features and the presence of various geological objects. The low measurement height causes many occlusions because of various geological features in the river space. This low height may have prevented us from obtaining sufficient point clouds at a distance from the revetments and may have thus caused errors in the attribute estimation of the buildings around the revetments. On the other hand, the occlusion occurs because the point cloud was acquired by one-way navigation in this experiment. There are bridges in the river space that are shaped like glasses, and in some cases, only one side of the bridge can be observed when navigating these bridges. Multiple round-trip navigation may improve these problems by measuring from various viewpoints but may worsen measurement efficiency and processing time.

The accuracy of fixed LiDAR had a significant impact on segmentation accuracy because data transformation was performed using point clouds integrated by SLAM processing of the obtained data. Point cloud generation in urban river spaces is often monotonous or has a repetitive pattern of similar shapes and features. In the case of handheld LiDAR, the SLAM processing is not included; therefore, a speed-up can be expected. However, because it is a handheld measurement, the origin is not fixed, the local coordinate system cannot be fixed, and the relative positions of sensors such as other LiDAR and GNSS are not fixed. In this study, we are considering superposition with point clouds obtained by multibeam depth measurement, and other methods will be studied in the future. Therefore, if the origin is not fixed, the accuracy of the segmentation process may decrease in the future.

In the segmentation process, a lot of processing time is required to perform SLAM, which is mainly the process of integrating point clouds. In this method, the processing is performed by post-processing. Compared with general indoor measurements, the urban river has a long measurement range and a long measurement time because of the limited speed of boats, resulting in a very large amount of data. Therefore, the processing speed and efficiency must be improved to improve the real-time performance of processing and enable immediate reflection of cloud acquisition and segmentation.

Conclusion and Recommendation

In this study, we have investigated a semantic segmentation method for urban river map generation from water-borne MMS point clouds. We confirmed that approximately 40% of urban river features were estimated by conventional methods using water-borne MMS point clouds. The first issue was some in-river structures were not completely extracted because of the lack of feature points due to the limited measurement angle and position from the boat. Because the accuracy of attribute data estimation information depends on the results of point cloud acquisition and SLAM processing, several technical issues remained in object recognition and segmentation from point clouds. The second issue was vibration caused by water waves in laser scanning. False point clouds caused by bubbles captured by LiDAR on the water surface cause errors during segmentation. Such errors can interfere with other point clouds and superimposition with other sensors. Moreover, there are issues with processing time and accuracy in the preprocessing stages that may result in the wrong geographic feature being used as the answer during segmentation. To further improve accuracy, methods must reduce the number of the wrong geographic features.

In this study, we proposed a method for the segmentation of point cloud data acquired by a handheld SLAM-LiDAR and water-borne MMS on a river as a step in the mapping of urban rivers. We believe that the 3D model generated by the proposed method can contribute to the effective use of autonomous boats in urban rivers both in normal times and during disasters.

Acknowledgments

This work was supported by JSPS KAKENHI Grant Number JP 24K07704.

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.

Haala, N., Peter, M., Kremer, J., & Hunter, G. (2008). Mobile LiDAR mapping for 3D point cloud collection in urban areas—A performance test. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci*, *37*, 1119-1127.

Nikoohemat, S., Diakité, A. A., Zlatanova, S., & Vosselman, G. (2020). Indoor 3D reconstruction from point clouds for optimal routing in complex buildings to support disaster management. *Automation in construction*, *113*, 103109.

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. *The Japan Society for Precision Engineering 2013*, pp. 857-859.

Infrastructure, Transport and Tourism (MLIT), 2023. PLATEAU,

from https://www.mlit.go.jp/plateau/.

Shizuoka Prefecture, 2024. VIRTUAL SHIZUOKA Project by SHIZUOKA PREF.

from https://www.pref.shizuoka.jp/machizukuri/1049255/index.html

Kimura, N., Nakagawa, M., Ozeki, T., Kubo, N., Shimizu, E., (2023). Pose Graph Adjustment For Lidar-Slam with Boat Motion Characteristics In Urban Rivers. *The 44th 2023 Asian Conference on Remote Sensing (ACRS2023)*.

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.

Pref.shizuoka, 2024 , VIRTUAL SHIZUOKA,

from https://www.pref.shizuoka.jp/machizukuri/1049255/index.html

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.