CityGML-based Representation of Urban River Environments using Waterborne MMS Point Clouds with LOD-aware Modeling

Kantaro K.[1*], Tetsu Y.[1], Nobuaki K.[2], Etsuro S.[2] and, Masafumi N.[1]

[1] Shibaura Institute of Technology,3-7-5, Toyosu, Koto-ku, Tokyo 135-8548, Japan

[2] Tokyo University of Marine Science and Technology,2-1-6, Etchujima, Koto-ku, Tokyo 135-8548, Japan

ah20005@shibaura-it.ac.jp

Abstract: *CityGML consists of open standardized data in XML format that can be used for various purposes, such as urban planning and disaster simulations, in the form of a three-dimensional (3D) urban model. CityGML uses the concept of level of detail (LOD) to enable the integrated management of map data at different scales. However, LOD design and schema considerations for modeling landforms are insufficient in urban river environments. Additionally, traditional surveying methods such as aerial photogrammetry, mobile mapping systems (MMS), and ground-based laser surveying often encounter areas where observation is difficult. The current version of PLATEAU in CityGML has not sufficiently advanced the development of high-resolution 3D map data for river environments. To promote the effective use of urban rivers by describing them with CityGML, this study focuses on the 3D modeling of urban river environments. One consideration is autonomous boat navigation technology, which requires precise global navigation satellite system (GNSS)/non-GNSS seamless positioning technology, autonomous control technology, and 3D river maps. Therefore, this study aims to create 3D maps for autonomous boats in urban rivers. The proposed methodology consists of point cloud processing and 3D model generation based on LOD design. High-density point clouds were acquired using a waterborne MMS integrating GNSS and, simultaneous localization and mapping based on light detection and ranging (LiDAR), with handheld and roof-mounted LiDAR units. Synchronization of GNSS augmentation data and LiDAR time stamps enabled data integration and boat attitude estimation, producing an environmental map of the river space. Point cloud clustering was used to classify objects. The random sample consensus algorithm was also applied to extract planar structures, and Euclidean clustering was used for bridges, piers, and buildings. LOD0 to LOD2 models (flat plane, box-shaped, and detailed mesh were then designed. Missing data caused by occlusions were compensated using a principal component analysis-based planar approximation and circular interpolation. Finally, voxelization and mesh modeling methods such as the ball pivot algorithm were used to construct efficient and detailed 3D models. We also investigated methods for describing the generated 3D models using CityGML. These models can be applied as 3D maps for autonomous boats and may contribute to the maintenance and management of river infrastructure.*

Keywords: *Water-borne mobile mapping, CityGML, 3D modeling, LOD*

Introduction

PLATEAU is open data in CityGML format that can be used for various purposes, such as urban planning and disaster simulations [1]. CityGML is an open data model and international standard for three-dimensional (3D) city models developed by the Open Geospatial Consortium [2], an international organization that develops geospatial information standards. CityGML describes the conceptual structure of the features necessary for representing 3D city models in XML format. One notable feature of CityGML is the concept of level of detail (LOD), which enables the integrated management of map data at different scales and unified management of small- and large-scale map data. However, urban river spaces often have insufficient consideration of LOD design and schema in feature modeling. Additionally, traditional surveying methods such as aerial photogrammetry, mobile mapping system (MMS) surveying, and ground-based laser surveying often encounter areas where observation is difficult. The current version of PLATEAU has not sufficiently progressed in developing high-resolution 3D map data for river spaces. With a focus on the 3D modeling of urban river spaces, this study aims to promote the effective use of urban rivers by describing them using CityGML. In promoting the effective use of urban rivers, autonomous boat navigation is necessary to create new transportation services combining land and water, to ensure energy transportation during disasters, to establish temporary high-speed communication hubs on water, to automate infrastructure inspections around rivers, and to address operator shortages. To achieve this goal, in addition to the development of 3D river maps, high-precision positioning and autonomous control technologies are essential. Therefore, this study aimed to create 3D maps for autonomous boats in urban rivers by acquiring point clouds from a waterborne MMS, developing methodologies for assigning attributes to objects in urban river spaces, modeling the spaces at different LODs, and exploring techniques for describing the generated 3D models in CityGML. These models could be used as maps for autonomous boats and to support the maintenance and management of river infrastructure.

Literature Review

We previously developed a waterborne MMS that uses light detection and ranging (LiDAR) sensors mounted on a boat to acquire registered point clouds [3]. The waterborne MMS is based on simultaneous localization and mapping using LiDAR (LiDAR-SLAM) and consists of a centimeter-level augmentation service (CLAS) and laser scanning. Urban river spaces are surrounded by poor global navigation satellite system (GNSS) environments characterized by buildings and bridges, which makes acquiring 3D data difficult. Although point clouds can be acquired using a handheld LiDAR-SLAM system without GNSS positioning, global coordinate values are necessary for map management and boat navigation. Therefore, we developed a GNSS/non-GNSS seamless positioning methodology to improve the positioning performance of the waterborne MMS in terms of 3D data acquisition.

Existing research includes 3D urban models and CityGML extensions. Research on extending the CityGML standard to underground spaces [4] has demonstrated its applicability to various applications, including urban planning and underground real estate management. Furthermore, the application schema can handle data requirements related to legal boundaries in underground environments that cannot be processed by the standard schema. This clarifies and visualizes the relationship between legal and physical spaces in 3D space. The present study aimed to describe urban river environments using CityGML and promote the effective use of urban river spaces. Additionally, we are currently designing LODs for 3D models of the geographical features of urban river environments.

There are many methodologies for mapping attribute data to geographical features based on point clouds using deep learning. One example is SCPNeT, a neural network that performs semantic segmentation of point clouds [5]. This methodology can be applied to dynamic objects in dense scenes. Another example is the deep learning-based methodology for the semantic segmentation of 3D building models and subsequent conversion of the results into CityGML format by Rashidan et al. (2024). They trained their model using BuildingNet to identify architectural elements such as roofs, walls, windows, doors, and ground surfaces. As a result, high accuracy was achieved for large structural elements such as ground surfaces and roofs. However, challenges remain in identifying small, highly diverse elements such as doors. To demonstrate the potential of using deep learning to assign semantic information and generate CityGML automatically, which could be expected to contribute to the development of advanced automation techniques for urban spatial data modeling, we implemented the process of converting the obtained semantic labels to CityGML (LOD3) and verified their feasibility as standardized 3D urban models. In urban river spaces, preparing training data for deep learning is a technical challenge. Therefore, this study classified, geographical features using conventional point cloud processing methods based on a knowledge-based approach. We also aimed to generate LOD-specific models compliant with CityGML from MMS point clouds.

Methodology

The proposed methodology consists of two main parts: point cloud processing and 3D model generation (Figure 1). First, the acquired point clouds were classified by objects through point cloud processing. Next, voxelization was applied to the objects with assigned attributes, and 3D models are generated based on LOD design. A filling process is applied to the missing areas of the point clouds, making it possible to generate models for each LOD.

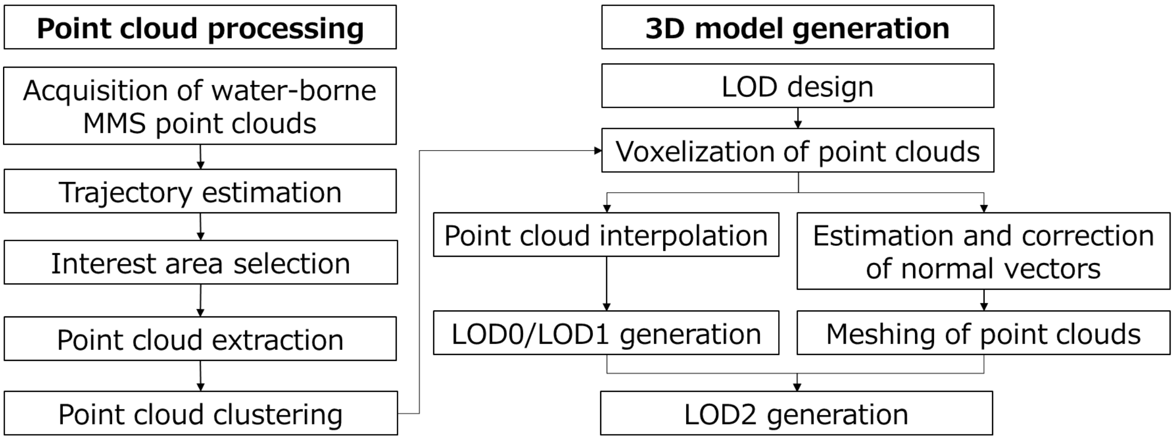


Figure 1: Proposed Methodology.

The proposed methodology is based on the use of dense point clouds acquired using a waterborne MMS [3] that integrates CLAS and LiDAR-SLAM or handheld LiDAR. The boat MMS used in this study, similar to conventional MMS, can acquire point clouds based on mobile measurement. Point cloud acquisition was performed using handheld LiDAR.

a. Point Cloud Clustering:

Point cloud clustering is used to classify landforms in urban river spaces. Although numerous studies have applied deep learning to assign attributes to landforms, creating learning data and performing measurements remain difficult. Thus, segmentation was performed based on point cloud processing. First, an area of interest is selected, and a local coordinate system is established. Next, to extract the target point clouds, a proximity search is performed using the distance threshold from the navigation point cloud to the landform. The Nihonbashi River is located under the Metropolitan Expressway and is surrounded by buildings and other structures; this results in poor LiDAR-SLAM accuracy and noise. Therefore, it is necessary to apply a feature classification methodology that does not depend on trajectory data. For this purpose, we introduced the random sample consensus (RANSAC) algorithm. RANSAC is a robust estimation methodology that identifies correct model parameters in noisy or outlier-containing data. In the present study, we extracted the deck using a planar model of the point clouds. We reduced the effects of terrain undulations and noise by sequentially applying RANSAC to slices within a certain range. Next, we performed clustering based on the Euclidean distance. Through these combined processes, we segmented different objects into individual clusters and grouped the point clouds by object. In this study, we extracted the Metropolitan Expressway, bridge piers, bridges, and buildings from the input point clouds.

b. LOD Design:

We designed an LOD for each location. Differences in LOD represent not only differences in appearance, but also the amount of information contained in the 3D model. The LOD design for buildings in 3D city models is divided into five levels, ranging from LOD0 to LOD4. LOD0 is a geometric projection of city objects onto a plane. LOD1 is a simple 3D model such as a box model. LOD2 includes shape information such as roofs. LOD3 is a more detailed model. LOD4 includes indoor information. These models can be used for various purposes such as simulations and urban planning. Referring to PLATEAU, each model is designed to be compatible with the CityGML format. In this study, three stages were set based on the acquired data, ranging from LOD0 to LOD2. LOD0 is a flat model generated by modeling the acquired point clouds to fit each object. LOD1 is a box-shaped model that extends LOD0 in the height direction. LOD2 is a mesh model that reproduces the detailed shape of the object and is generated by superimposing it on LOD1, as shown in Table 1.

Table 1: LOD Design for Each Object.

	LOD0	LOD1	LOD2
Revetments	Flat model	Polygon model	Mesh model
Bridge	Flat model	Polygon model	Mesh model
Building	Flat model	Polygon model	Mesh model
Plants	Flat model	—	Mesh model
Metropolitan Expressway	Flat model	Polygon model	Mesh model
Bridge pier	Flat model	Polygon model	—

Figure 2 shows the relationship between objects in urban river spaces. Extracted attributes can be classified based on their location (on land or underwater) and their potential to obstruct navigation. Furthermore, by establishing relative relationships between objects based on adjacency, objects can be managed according to their impact on navigation.

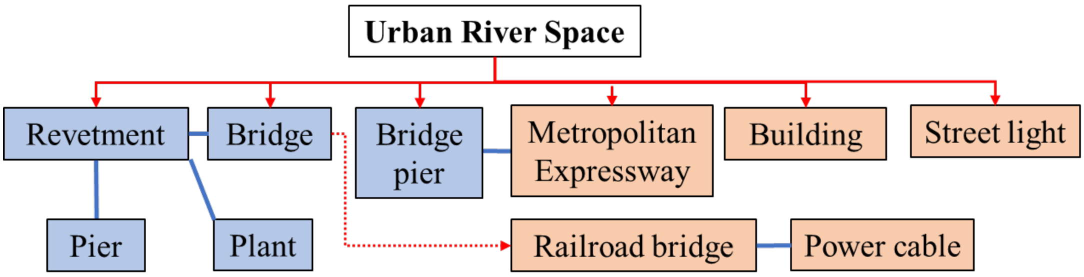


Figure 2: Relationships of River Features (Water-above Objects).

c. Point Cloud Completion:

When measuring urban river space, it is not possible to obtain point clouds of some land features when there are obstacles between the boat and land features, or when the boat cannot easily access certain areas. In these cases, geometric completion is performed based on the shapes of the target land features, and the point cloud data are corrected to generate LOD-specific models. Principal component analysis (PCA) is applied to the acquired point clouds of the deck plates, bridges, and buildings of the Metropolitan Expressway to approximate a rectangular shape. PCA extracts the main directions of the point cloud distribution and uses them to estimate the bounding rectangle or approximate plane. A grid-like point cloud is then generated within the obtained area. The piers are approximated as circular shapes with the center of gravity of the point clouds serving as the center and the average distance to each point serving as the radius. The interior is then uniformly interpolated with point clouds. We maintained the spatial consistency of the point clouds while interpolating missing areas and clarifying the shape based on the characteristics of the target structure, as shown in Figure 3. We compensated for the incompleteness of the acquired point clouds and enabled the description of a feature model corresponding to the CityGML LOD representation by selectively applying the interpolation methodology based on the shape characteristics of the target structure.

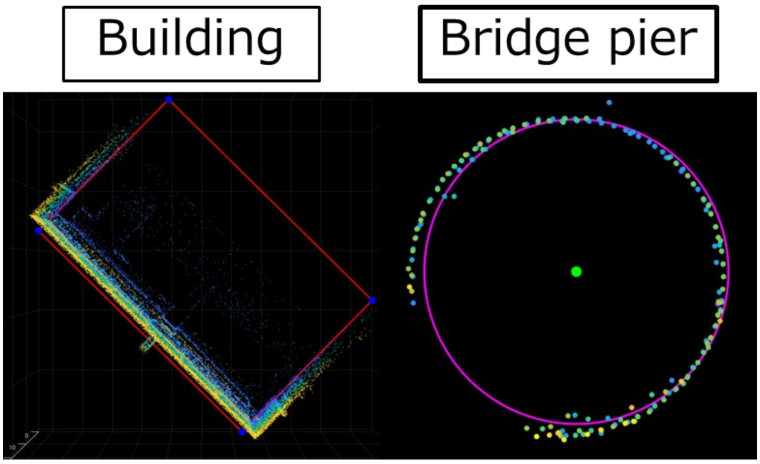


Figure 3: Example of Point Cloud Approximation.

d. LOD-specific model generation:

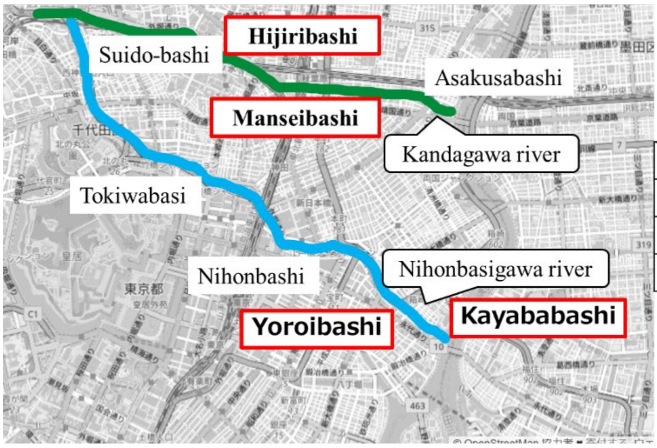
We apply 3D modeling based on LOD design to point clouds classified by terrain features. First, we performed voxelization processing on the target point cloud and replaced the point cloud with representative points in a 3D grid set at an arbitrary resolution. The original data were highly dense, which made generating a 3D model require excessive processing time therefore, so downsampling processing was applied. As the resolution of the voxels affects the identification and reproducibility of objects, it is important to set an appropriate threshold. If the resolution is too low, the boundaries of the objects become unclear and difficult to identify. On the other hand, if the resolution is too high, although the boundaries become clearer, the data size increases and the processing efficiency decreases. Therefore, the optimal voxel size must be selected according to the characteristics of the data and the purpose of the modeling.

Stepwise model construction corresponding to LOD is performed by applying geometric interpolation according to the shape of the target objects based on the acquired point clouds. This is then used as the LOD0 model to construct a basic plane model. Next, the plane model is layered along the Z-axis to reconstruct a 3D point cloud. Then, polygonization processing is applied to create the LOD1 box model. Furthermore, in LOD2, the normal vectors of the point cloud are estimated for mesh modeling. Errors in the normal vectors can cause some mesh surfaces to invert inconsistently, which prevents the generation of an accurate 3D model. Mesh data are considered geometric data consisting of points connected by edges and faces. While dense point clouds can express detailed shapes, the numbers of vertices and edges can increase, resulting in larger data volumes. One approach to mesh modeling is the ball pivot methodology. In this methodology, a model is generated by rolling a ball of radius r over point clouds to triangulate the set of points. First, three sample points are selected, and the ball is positioned to touch these points and form seed triangles. Then, the ball is rotated until it contacts another point while remaining in contact with two edges of the seed triangle. Each time the ball touches a new point, a new triangle forms between that edge and the point. This process continues until all adjacent point clouds are meshed into triangles and all accessible edges are covered. If any areas remain uncovered, another seed triangle is selected, and the same process is repeated. Superimposing the generated mesh model on a box-shaped model constructed at LOD1 allows for the reproduction of more detailed terrain shapes.

Experiment

We acquired dense point clouds using a waterborne MMS along the Nihonbashi River and Kanda River (Figure 4). The waterborne MMS consisted of a CLAS receiver and antenna, horizontal scanning LiDAR (VLP-32C; Velodyne) (Table 2), and oblique scanning LiDAR (VLP16; Velodyne) (Table 3). All of these components were mounted on the roof carrier of a battery-powered boat ("Raicho I") (Figure 5). Horizontal scanning LiDAR covers a wide horizontal area, while oblique scanning LiDAR covers the vertical surfaces of objects, such buildings, bridges, and revetments. We synchronized and integrated the CLAS and LiDAR data with GPS time and GNSS position data. The approximately 10-meter-long boat was suitable for narrow urban rivers. Its battery-powered propulsion system minimizes swaying and vibration, enabling stable measurements. Moreover, we used LiDAR-SLAM (Hovermap ST-X; Emesent) (Table 5.) to evaluate dense point clouds.

Most sections of the Nihonbashi River were located under the Metropolitan Expressway. By contrast, the Kanda River had a good GNSS environment, enabling stable reception of GNSS and CLAS signals. To verify the proposed methodology, we focused on two sections: the Manseibashi -Hijirbashi section, which has many types of objects, and the Kayababashi -Yoroibashi section, which runs under the Metropolitan Expressway.



	Date	Point clouds
Experiment1	November 6, 2022	41,364,456
Experiment2	October 13, 2023	20,786,717
Experiment3	October 28, 2024	—

Figure 4: Experiment Sections.

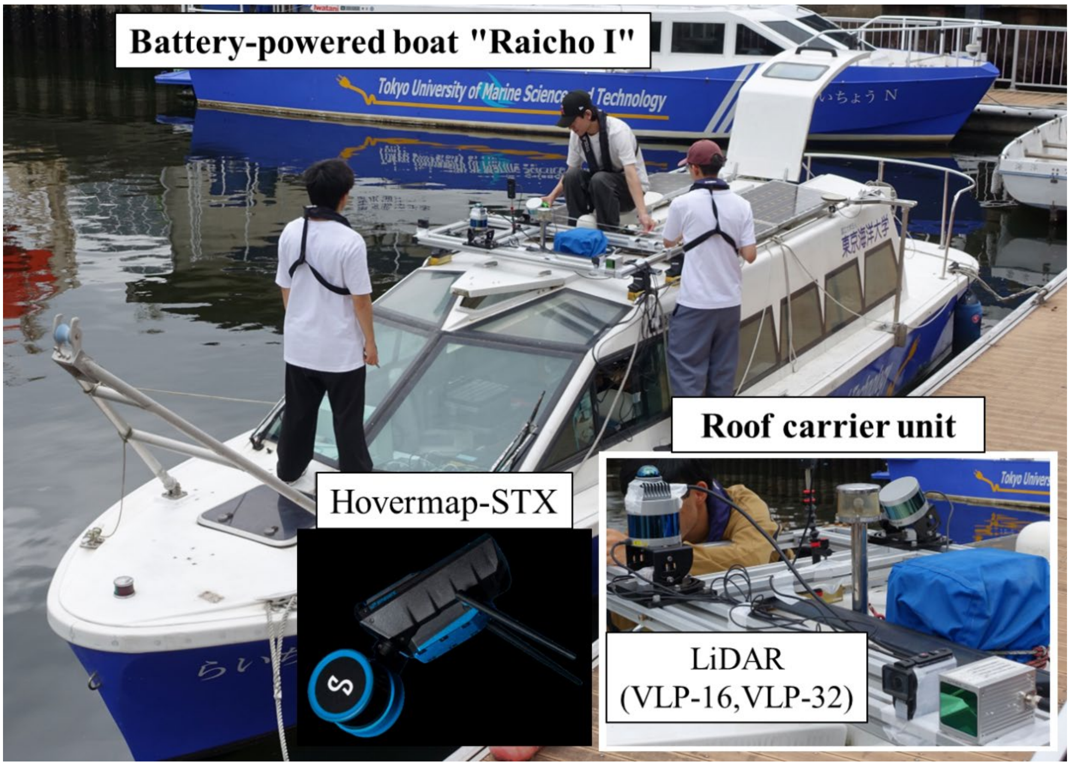


Figure 5: Waterborne MMS.

Table 2: Specifications of the LiDAR (VLP-16, VLP-32).

VLP-16		VLP-32	
Weight	830g	Weight	925g
Channels	16	Channels	32
Range of measurement	100m	Range of measurement	200m
Accuracy	Up to ±3cm (Typical)	Accuracy	Up to ±3cm (Typical)
Vertical Field of View	+15.0° to -15.0° （30°）	Vertical Field of View	40°(-25° to -15°)
Vertical angular resolution	2.0°	Vertical angular resolution	0.33°
Horizontal vision	360°	Horizontal vision	360°
Horizontal angular resolution	0.1° –0.4°	Horizontal angular resolution	0.1° –0.4°
Rotations	5 Hz – 20 Hz	Rotations	5 Hz – 20 Hz

Table 3: Specifications of the LiDAR (Hovermap STX).

Weight	1.57kg
Sensor channels	32
Measurement range	0.5m～300m
Mapping method	SLAM
Mapping accuracy	Outdoor environment ±15mm Indoor environment ±10mm Short range ±5mm
Viewing angle	360×290°
Data acquisition rate	640,000（points/second）
Sampling rate	5 Hz － 20 Hz

Results of the point cloud processing

Figure 6 shows the point clouds acquired using the waterborne MMS and LiDAR-SLAM. Although the point density differs depending on LiDAR performance, similar results were obtained when focusing on a two-dimensional map.

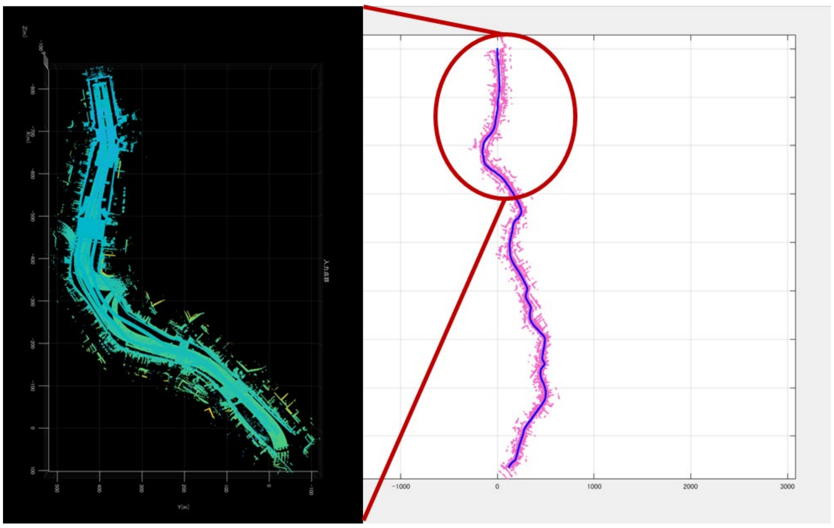


Figure 6: LiDAR-SLAM Results for the Nihonbashi River.

Figure 7 shows the results of feature classification and extraction using the proposed method for the Nihonbashi River. The color coding indicates the type of estimated features. Red indicates the capital bridge piers, blue indicates the bridge, and yellow indicates the capital bridge deck. Through this process, we were able to classify the capital bridge deck, piers, and bridge. Furthermore, the buildings mixed in with the extraction were identified as segmentation noise.

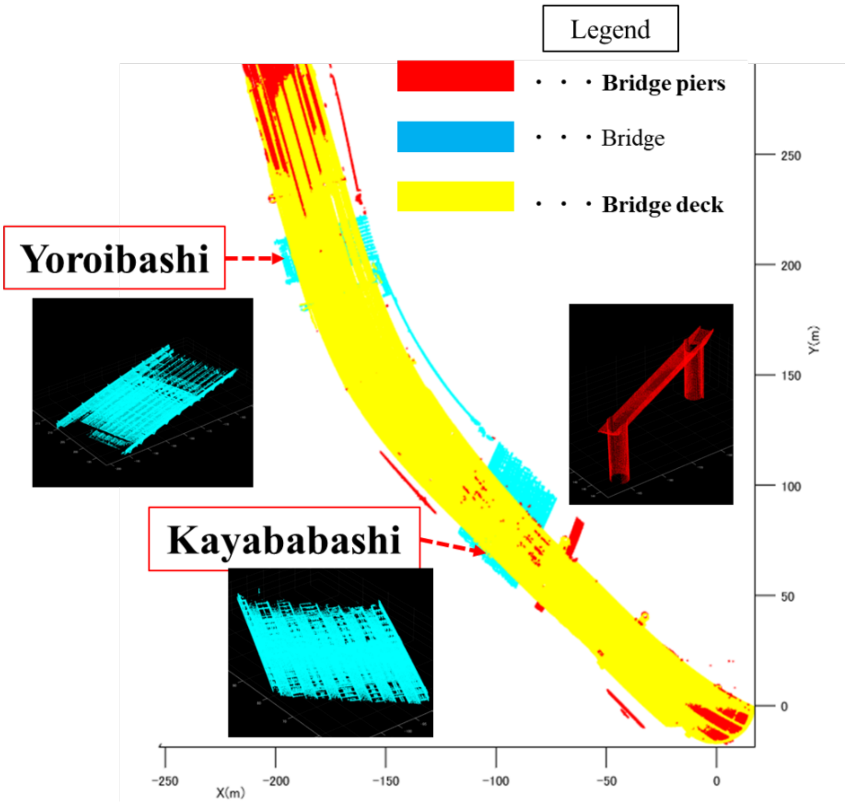


Figure 7: Processing Results of Point Clouds.

Results of the 3D model generation

The proposed methodology was applied to generate LOD-specific models of landforms in urban river spaces, as shown in Figure 8. The results for buildings, bridges, and bridge piers are shown below. For buildings, models were generated to fill in the gaps in the acquired point clouds.

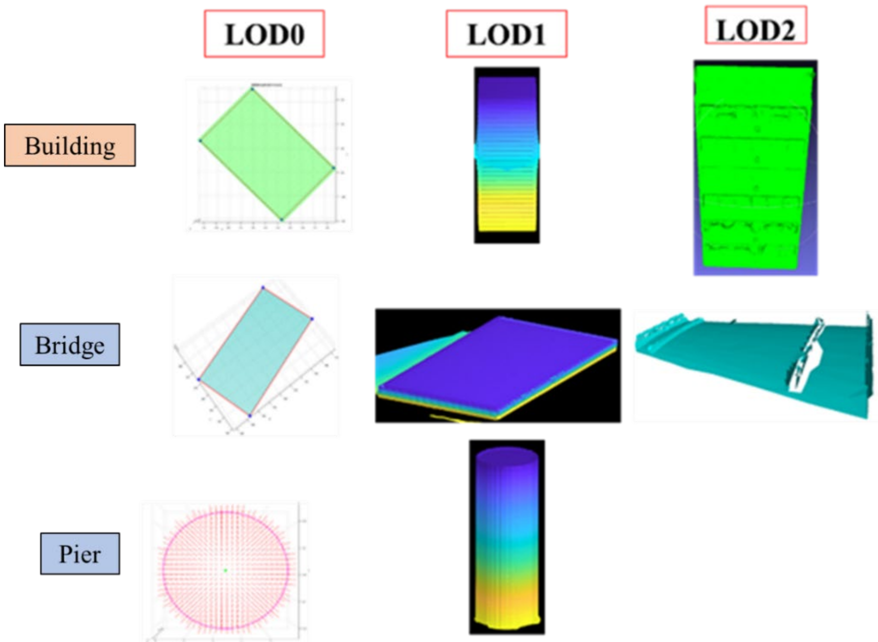


Figure 8: 3D Model Generation Results by LOD.

Discussion

We confirmed that 3D models can be generated in stages according to the LOD hierarchy. Using point cloud data acquired by the waterborne MMS, we performed object classification and 3D modeling based on the LOD design assuming a CityGML description in urban river spaces.

Because the Nihonbashi River runs under the Metropolitan Expressway, a feature extraction method independent of navigation data was necessary. The RANSAC algorithm was applied to detect planes for deck extraction. However, accurate extraction was difficult in some areas, as shown in Figure 9. The unique topographical undulations of urban river spaces caused local elevation differences along the Z-axis that obscured the distinction between the deck and the terrain. This could have led to incorrect plane fitting. Additionally, while many objects were composed of planes, many object shapes should have been treated as curved surfaces, which may have been a contributing factor.

Applying the processing to spatially continuous sections of the deck area is considered a way to achieve robust extraction despite the terrain undulations. In the future improving the algorithm by combining geometric features such as the distribution of normal vectors and point density, could enhance the accuracy of plane extraction.

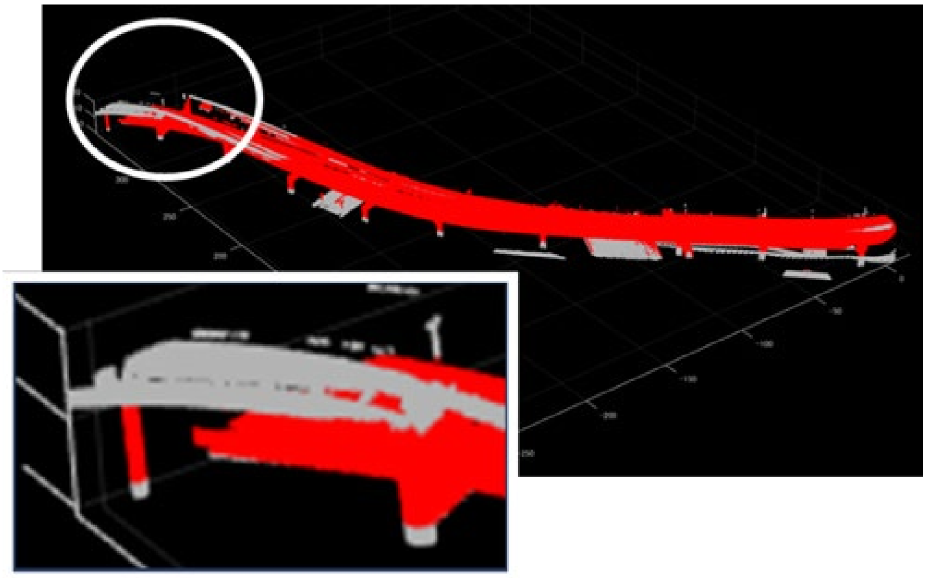


Figure 9: Example of a Failed Geo-Extraction.

Considerations for 3D model generation

A methodology for generating LOD0 models in urban river spaces was performed using point cloud completion with shape estimation based on PCA. Shape completion, including missing areas, was confirmed to be possible with a certain degree of accuracy for objects with simple geometric structures, such as bridges and buildings. However, for objects with complex structures and continuous shape changes, such as the deck of the Metropolitan Expressway, PCA-based region estimation produced simplified models that significantly deviated from the actual structure, as shown in Figure 10. Because PCA is applied uniformly to the entire target area, it cannot adequately capture nonlinearity or local shape changes. When the density and distribution of point clouds are uneven, the likelihood increases that the principal component axis will not accurately reflect the actual shape of the object. Therefore, rather than applying PCA uniformly to the entire target object, it is more effective to apply PCA to each local area, or to perform clustering according to the point cloud density.

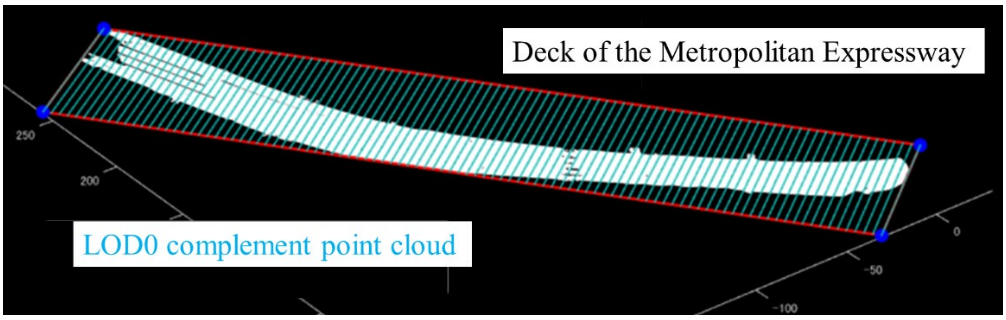


Figure 10. LOD0 of the Floor Slab of the Metropolitan Expressway.

Although geometric interpolation can effectively estimate areas missing as a result of occlusion, it can deviate from the actual measurement data. Therefore, interpolation is only a temporary measure, and optimizing the acquisition of point clouds is a more fundamental solution. Conducting round-trip measurements using a boat makes it possible to observe the same object from different perspectives and compensate for blind spots. This approach minimizes the defects caused by occlusion on a single route, resulting in high-precision point cloud acquisition independent of interpolation processing.

Considerations for LOD-based 3D model generation

In this study, we proposed a stepwise method for generating models from LOD0 to LOD2 based on CityGML descriptions of urban river spaces. According to the CityGML standard, LOD2 is constructed by adding components such as walls, roofs, and floors to LOD1. However, this study encountered an issue where attribute information could not be partially assigned because of the characteristics of point cloud acquisition using the waterborne MMS. Specifically, because the measurement depends on the boat's perspective, point clouds of the upper roofs cannot be acquired, resulting in insufficient roof modeling. One effective solution for this is to supplement the data by integrating it with data acquired from different perspectives, such as aerial photogrammetry or unmanned aerial vehicle LiDAR. Furthermore, regarding walls, they are currently simplified as a single plane. However, there is potential to identify openings by leveraging the reflectance intensity and incident angle characteristics in LiDAR measurements. The LiDAR-SLAM used in this study has a high point density, which enables the acquisition of detailed point clouds of structural elements. such as window frames and stairs. Accurate extraction and classification of these elements is essential for LOD3 generation. To improve the accuracy of object recognition, analysis of normal vectors and curvature in local areas can be performed. This enables the identification of flat walls, corners, and curved surfaces. It is also effective at detecting elements with local shape changes such as window frames and stairs. Furthermore, statistical and clustering methodologies have shown promise. Density-based clustering methodologies, including DBSCAN and Mean-Shift, can automatically extract spatial clusters such as building walls and bridge components. In addition, analyzing local changes in point density effectively identifies sparse areas such as windows and openings. These methodologies, which do not require deep learning, are considered robust recognition methodologies that directly utilize the geometric characteristics of point clouds.

On the other hand, for long bridges such as the Metropolitan Expressway, the setting of classification units in CityGML is problematic. Instead of simple segmentation based on routine physical characteristics, it is preferable to use segmentation based on geographical information and road structural divisions, such as interchanges, bridge pier locations, and river intersections. It is believed that more practical CityGML models can be generated by introducing semantic divisions based on geographical contexts using coordinate information and geographic information system data.

Future Prospects

Currently, LOD-based model generation is limited to local areas. However, there is a need to improve the algorithms so that they can be applied to entire urban rivers. In particular, in environments with continuous bridges and elevated structures, for example, it is essential to establish LOD update methods that consider spatial continuity.

When creating maps for boats, it is essential to consider underwater obstacles such as piers and rocks. It is also important to assess the risk of grounding properly. However, it is difficult to accurately identify and avoid these obstacles accurately based on visual observations from the boat, and LiDAR cannot acquire underwater point clouds. Therefore, a future challenge is to utilize multi-beam sonar, which is effective for depth surveys, to measure underwater environments in greater detail. Additionally, incorporating tidal level change data could enable the construction of more precise boat maps that dynamically reflect navigable areas. Some urban river spaces have sections where navigation is restricted because of tidal influences. Thus, integrating underwater and surface data is crucial for providing safe and efficient navigation support in these areas.

Conclusion

In this study, we assessed a modeling methodology for CityGML descriptions using waterborne MMS point clouds. The experimental results confirmed that our proposed methodology can generate LOD-specific 3D models of objects in urban river spaces. Future tasks include adding LOD3 objects and using bathymetry data to model underwater objects for boat navigation.

References

Ministry of Land, Infrastructure, Transport and Tourism：PLATEAU . Retrieved July 4, 2024, from https://www.mlit.go.jp/plateau

Standards-Open Geospatial Consortium Retrieved July 4, 2024, from https://www.ogc.org/standard/citygml

Naoto, K., Masafumi, N., Tomohiro, O., Nobuaki, K., Etsuro, S., (2022). Seamless Indoor-outdoor Positioning and Trajectory Interpolation using SLAM and PPP-RTK for River Mapping, The 43rd Asian Conference on Remote Sensing, 6 pages

Bahram S., Abbas R., Behnam A., Mohsen K., (2024). Managing underground legal boundaries in 3D - extending the CityGML standard Underground Space Volume 14, Pages 239-262

Zhaoyang X., Youquan L., Xin L., Xinge Z., Yuexin M., Yikang L., Yuenan H., Yu Q., (2023). SCPNet: Semantic Scene Completion on Point Cloud, Computer Vision and Pattern Recognition, arXiv:2303.06884v1

Hanis R., Ivin Amri M., Alias Abdul R., Volker C., Gurcan B., (2024). Semantic Segmentation of Building Models with Deep Learning in CityGML, The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences Volume XLVIII-4/W11