Relative Motion Detection of Boats based on LiDAR Scan Matching

Kazuki Ohira [1*], Naoto Kimura [1], Masafumi Nakagawa [1], Nobuaki Kubo [2], Etsuro Shimizu [2]

[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 135-8533, Japan

*ah20091@shibaura-it.ac.jp

Abstract: *This study investigates the development of autonomous boating technology, with a particular focus on improving automatic detection of surrounding boats and collision avoidance capabilities, which remain underdeveloped despite their importance. More than 40% of boating accidents in 2023 will be due to collisions, and the current reliance on crew members to monitor the boat's surroundings. The current reliance on manual observation by the crew to monitor the boat's surroundings highlights a significant gap. In this study, we propose a new approach to boat detection and tracking using LiDAR scan matching, a method that takes advantage of the real-time measurement capabilities and its application to multi-purpose mapping. Traditional methods often rely on image processing techniques such as optical flow, HOG features, and deep learning models such as Faster-RCNN. Image processing methods using deep learning rely heavily on pre-trained models and require training data from boats facing different directions, which remains a challenge. In contrast, the proposed LiDAR-based approach identifies moving boats by analyzing discrepancies in successive LiDAR scans, using the 3D distance between cluster centroids to distinguish stationary from moving objects. The detection process includes a filtering step to minimize false positives by evaluating the size of the identified clusters and excluding those that do not match the expected boat dimensions. For tracking, the system follows the movement of detected boats by continuously comparing centroids over different time frames. If a boat is lost, the system stops tracking and attempts to re-detect it in subsequent scans. In the experimental phase, a battery-powered boat equipped with LiDAR was used to navigate a designated route on the Kanda and Nihonbashi Rivers. The detection process occurred every second (i.e., every 10 frames), and the success rate varied from boat to boat, ranging from 48.65% to 89.66%. Despite the variability in detection rates, this study demonstrates the potential of LiDAR scan matching to improve autonomous navigation systems. This study will contribute to advanced maritime navigation with autonomous control to improve the safety and human operational workload.*

Keywords: *Scan matching, LiDAR, autonomous boats, object detection and tracking*

Introduction

In recent years, research and development of autonomous boats using ICT and sensor technologies have been active worldwide. Examples of related technological developments in Japan include automatic route-following control systems, automatic berthing systems using GNSS and LiDAR, and automatic avoidance navigation systems that operate day and night. However, boat accidents are still common. In 2008, 70% of all boating accidents involved small boats such as pleasure boats because many small boats rely on the operator's visual navigation without advanced navigational aids, compared to large vessels such as tankers and container ships. The operator of a small boat is required to comply with all regulations, from maneuvering to watching, which places a heavy burden on the operator while maneuvering the boat.

In addition, Tokyo is experiencing serious traffic congestion problems due to the concentration of its population. To alleviate the traffic problem, the Tokyo Metropolitan Government has begun to use rivers as commuter routes. The technical issues of river traffic include narrow waterways (the average width of rivers in densely populated areas is approximately 20 m) and the presence of many obstacles. Therefore, an object recognition function is required to avoid obstacles such as other boats, revetments, and piers. Conventional methods for collision avoidance between boats are position information sharing using GNSS position data and image processing with machine learning such as Faster-RCNN (deep learning model). Image processing with Faster-RCNN is a pre-trained object detection and recognition methodology. However, GNSS position sharing cannot be used in non-GNSS positioning environments, such as under bridges. Moreover, image processing with Faster-RCNN requires massive training data of boats facing in different directions to detect surrounding boats. Therefore, we propose a boat detection and tracking method using LiDAR mounted on an autonomous boat with a motion recognition method based on scan matching to achieve automatic avoidance of other boats. Furthermore, we have conducted experiments in an urban river using a LiDAR-mounted boat to evaluate our proposed method. The goal of this study is to enable automated detection of surrounding boats in urban river navigation using the proposed method.

Literature Review

There are various studies on the identification of small boats such as boat tracking using a panoramic camera (Kondo et al., 2020). In this study, a boat is identified from a panoramic image using Faster R-CNN, which is an efficient image processing method and a fast object detection algorithm based on deep learning. First, features are extracted from the input image using a 6-layer convolutional neural network. Next, a feature map is created with 2D data summarizing the obtained features. Then, the candidate object regions are output obstacles by RPN. The processes are repeated and trained to create a discriminator to detect boats from images. The training procedure for the discriminator using Faster R-CNN is as follows. First, other boats are extracted from an image to label the boat objects to be used as the training image. Next, the Faster R-CNN is trained to create a discriminator for detecting boats from images. The results of boat navigation experiments show that the Faster R-CNN can detect boats in various situations, however, the boat detection sometimes fails, depending on their orientation. Many other conventional boat detection methods mainly use cameras, including image processing methods such as optical flow, object detection with histograms of oriented gradients, and semantic segmentation using deep learning. However, the number of boat pixels in the captured images is not enough to improve the detection accuracy. The other issue is that the distance measurement to an object is not easy because of a single camera. Therefore, we aim to avoid the technical issues of small boat detection, such as "low resolution and orientation problems. In our research, we focus on LiDAR to estimate the positions, sizes, and relative speeds of other boats and obstacles.

Methodology

The proposed method consists of boat detection and boat tracking from a water-borne 3D measurement system based on scan matching processing, as shown in Figure 1.

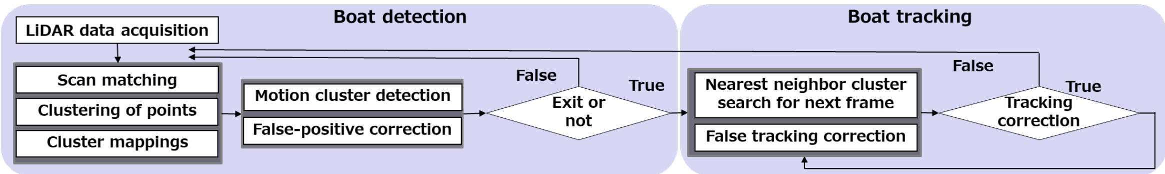

Figure 1: Proposed methodology.

a. Scan matching:

The boat detection shown in Figure 1 is based on scan matching between the current scan and the previous scan (the reference scan). Scan matching is the process of transforming (aligning) two different point clouds. The ICP algorithm is mainly used in scan matching, where the mapping of each scan point and the self-position estimation are alternately repeated between the current scan and a previous reference scan. In other words, the position posture with the smallest error between the corresponding points between the current scan and the reference scan is determined as the self-position. The procedure is as follows: first, the reference scan point with the smallest Euclidean distance from the current scan point is searched as a scan data mapping. Next, the average of the squares of the distances between the points is calculated according to the scan data matching. This is used as a cost function and the position and orientation that minimizes this function is found. In this optimization, the calculation is terminated if the difference between the cost function at the kth iteration and the cost function at the $k-1$st iteration is less than a threshold, the calculation is stopped. This is a nonlinear optimization problem and the simplest optimization method, the steepest descent method, is most often used. However, there are three challenges with the ICP algorithm. The first is that variations in point density reduce the accuracy of the positioning: because LiDAR measures the distance to an object by emitting lasers radially at equal angular intervals, the further away the object is, the lower the density at which the laser hits it. Such point density variations are generally undesirable. Scan matching has the weakness that the cost is calculated on the average of the scan points and consequently, thus, the influence of low-density regions is small. Second, scan point matching time consuming. In scan point matching, the search for the nearest neighbor of the current scan point must be performed, for all reference scan points. Therefore, in the scan matching discussed in this study, we apply an algorithm with some modifications to ICP. First, the spacing of the scan points is made uniform. By making the spacing between each scan point uniform, the density of the point clouds can be made uniform. This allows even low-density scan point clouds can be used for high-precision positioning over a wide range of geometry. Next, when searching for the nearest neighbor of the current scan point in the scan point mapping, the Euclidean distance is calculated only for reference scan points that lie within a circle of arbitrary size centered on the current scan point. This is more efficient because it does not require a linear search of all reference scan points. Furthermore, instead of the average of the squares of the Euclidean distances between the corresponding points, the average of the squares of the vertical distances between the corresponding points is set as the cost function. This is because in an environment with many planes, the matching accuracy tends to be better when the length of the perpendicular line (vertical distance) down to the tangent line of the reference scan point is used as the cost function.

The cost function used for scan matching is shown in Equation 1. This function uses a rotation matrix R and a translation vector t and is evaluated to minimize the error between the current scan point coordinates p_i and the corresponding reference scan point coordinates q_i. The vertical distance between the new scan point coordinates and the reference scan point coordinates is calculated by taking the inner product of the difference and the normal vector n_i of the reference scan point. Finally, the sum of the squares of the vertical distances between each corresponding point is divided by the number of points N to obtain the mean square value. In conventional scan matching, the distance between the coordinates of the new scan point and the coordinates of the reference scan point is evaluated as the point-to-point distance. In the conventional method, the distance between corresponding points is evaluated to optimize the point cloud alignment parameters, as shown in Figure 2a. At points q1 and q2, the alignment parameters may not be estimated due to opposite forces acting along the line between the new scan point and the reference scan point. In contrast, Figure 2b shows that point-to-point matching can be used to estimate the point cloud alignment parameters even if the corresponding points cannot be matched correctly.

$$G(R,t) = \frac{1}{N}\sum_{i=1}^{N}\|n_i \cdot (Rp_i + t - q_i)\|^2 \quad \text{(Equation 1)}$$

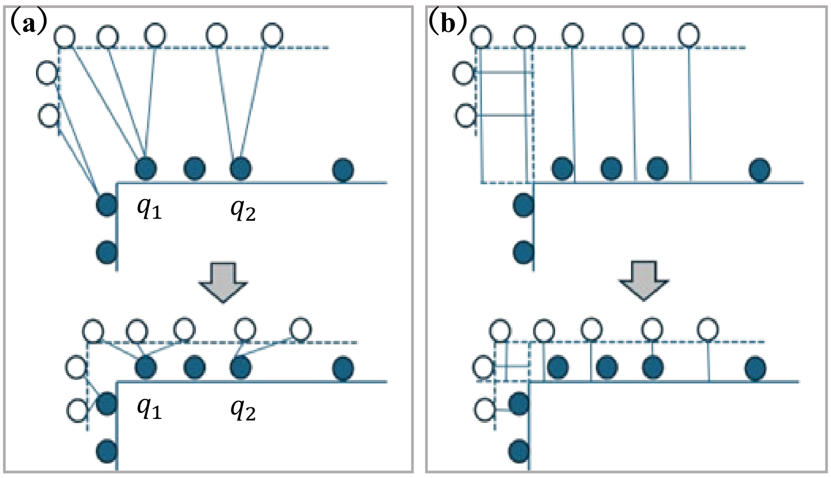

Figure 2: Scan matching (left: conventional method using distance between points, right: point-to-line matching).

In the proposed method, the optimization of the cost function is performed by combining the steepest descent method and the linear search method (Figure 3). Optimizing the cost function means searching for a transformation matrix that minimizes the cost function. Conventional methods use the steepest descent method as the optimization method, which is easy to implement but has the problem of slow convergence. Although many optimization methods are superior to the steepest descent method, the proposed method uses the steepest descent method, which is easy to implement, and uses linear search to speed up the process. First, numerical differentiation of the cost function in the initial transformation matrix is used to determine the direction of the gradient toward the minimum of the cost function. Next, the step size of how much to move the transformation matrix in the obtained gradient direction is determined by increasing the step size if the Armijo condition, which requires the cost function value to decrease sufficiently on a straight line in the search direction, is satisfied, and decrease, the step size while searching for the minimum if it is not.

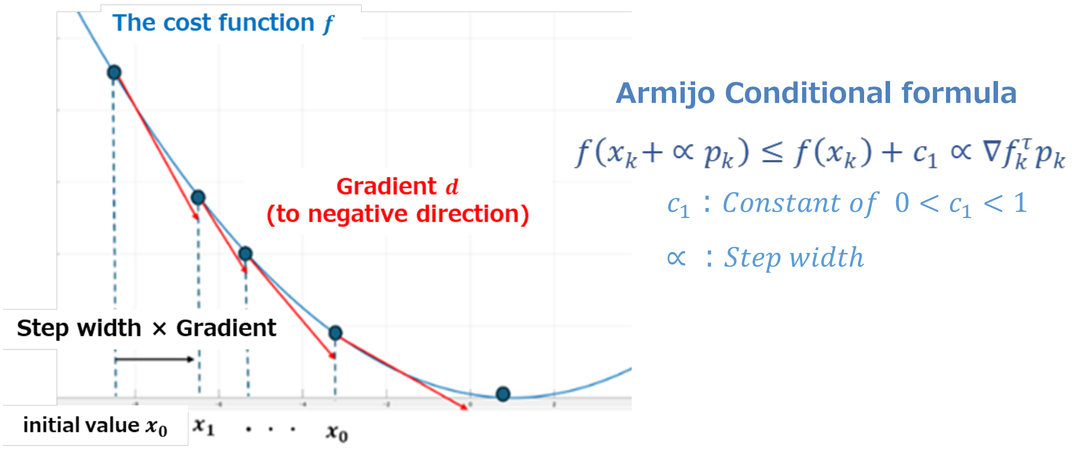

Figure 3: Steepest descent method.

b. Boat detection and tracking:

Although scan matching is used in point cloud alignment, in this study, scan matching is used to detect moving boats. In scan matching, the two point clouds are perfectly matched only for the stationary object being scanned. By contrast, point clouds of moving objects are never matched, and the matching result is considered a scan matching error. We used these errors to detect moving boats from streaming point clouds. In this method, the 3D distance between the centers of the corresponding clusters among the point clouds after scan matching is evaluated to classify into the background and moving objects such as boats. Figure 4 shows an example of moving cluster detection. The left example in Figure 4 shows that the selected point clouds are static objects due to the short 3D distances between the corresponding cluster centers. By contrast, the right example in Figure 4 shows that the selected point clouds are moving objects due to the long 3D distances between the corresponding cluster centers. However, in LiDAR point cloud acquisition, when the cluster centroids of bridges and revetment areas are hidden between frames due to the occlusion, the revetments and other objects are often detected as a moving boat as shown in Figure 5. In the proposed method, a false positive correction is performed as a countermeasure against this problem. The false positive correction uses an algorithm with a threshold on the size of clusters in motion detection to reject large clusters as errors.

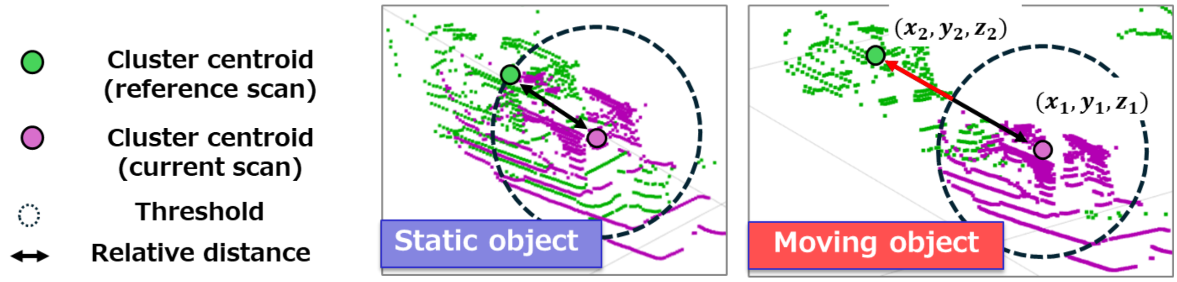

Figure 4: Match evaluation of point cloud clusters.

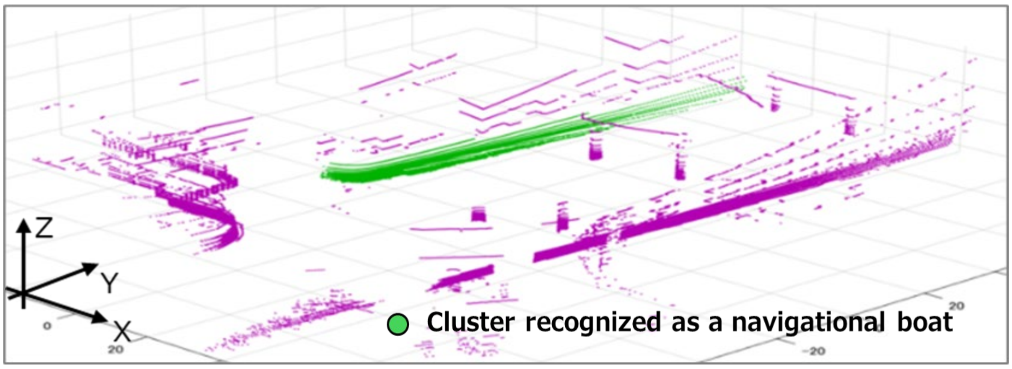

Figure 5: Example of missed boats.

After boat detection, a boat tracking task is performed. In the proposed method, the moving object is tracked by searching for the nearest centroids in the next scene, based on the measurement scenario consisting of slow moving boats using a LiDAR with 10Hz sampling rate. In the tracking task, in LiDAR point cloud acquisition, when the cluster centroid of a bridge or revetment is hidden between frames due to occlusion in the LiDAR point cloud acquisition, piers and other objects are often tracked as moving vessels, as shown in Figure 6. In this case, Figure 6 shows that the distance between the center of gravity of clusters tracked as moving objects between the current and reference scans is extremely large compared to the case of correct tracking. This is used to compensate for false tracking by aborting tracking when the distance between the centroids of the clusters tracked as moving objects between the current and reference scans exceeds an arbitrary threshold value. If tracking is aborted, the system will search for the boat again using scan-matching motion detection and will track the boat again if it is found.

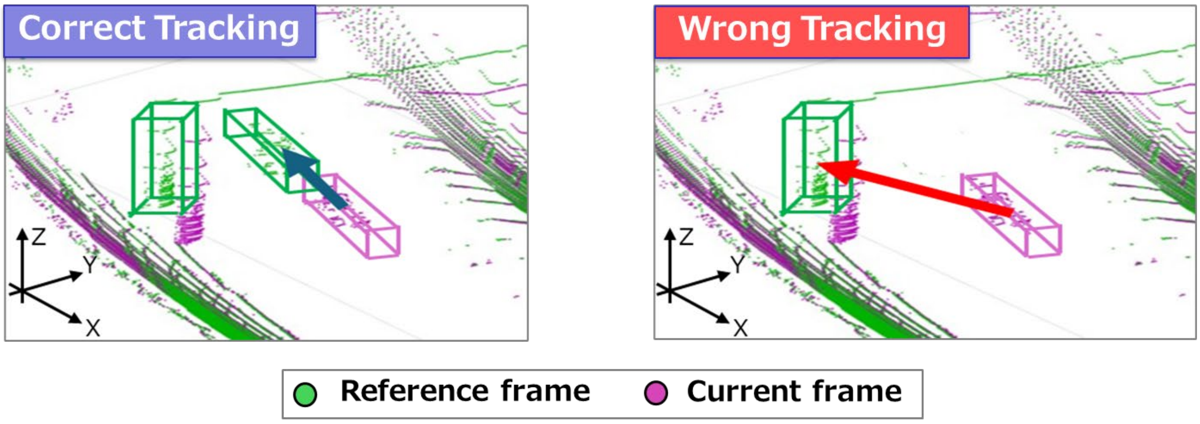

Figure 6: Tracking result of boats (left: successful case, right: failed case).

Experiments

We conducted experiments using a LiDAR mounted on a battery-powered boat “Raicho I” as a water-borne mobile mapping system (MMS) to acquire point clouds of landmarks and other boats (Figure 7).

Figure 7: Overall view of the measurement system for the experiment.

The experiment was conducted along the Kanda River and Nihonbashi River. We also set, the route from Asakusa Bridge to Suidobashi, Kanda Bridge, and Eitaibashi Bridge, as shown in Figure 8. The Kanda River consisted of an open sky environment. Therefore, LiDAR acquired point clouds of buildings and revetments for stable scan matching along the river with GNSS data. The section of the Sumida River from Etchujima was excluded because the river was too wide to acquire point clouds with LiDAR used in our experiments. The Nihonbashi River from the Minato Ohashi Bridge to the Koishikawa Bridge was covered by the Metropolitan Expressway. The Metropolitan Expressway had complex shapes due to the many ramps and interchanges of the expressway along the river. Although the river was wide, the Nihonbashi River had suitable sections for measuring oncoming boats due to many road bridges. We observed six other boats (from Boat 1 to Boat 6) in the experimental sections, as shown in Figure 10. Boat 1 was observed near the Suimon Terrace Liaison Boat 2 was observed near the Eitai Bridge, Boat 3 was observed near the Sumida River Bridge, Boat 4 was observed near Kiyosubashi Bridge, Boat 5 was observed near Ryogoku Bridge, and Boat 6 was observed near Sagara Bridge.

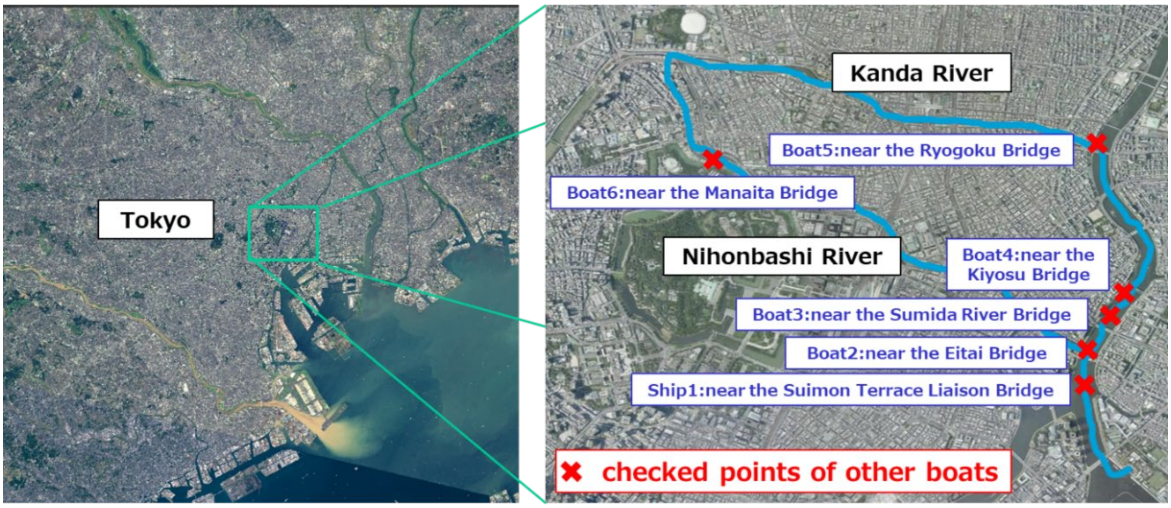

Figure 8: Location of the experiment site.

a. Platform

In this study, we used a battery-powered boat "Raicho I" as a platform for a water-borne mobile mapping system (MMS). The size of the boat was small (10 m in length) to safely conduct experiments in narrow rivers. Moreover, the battery-powered propulsion was suitable for stable point cloud acquisition due to low rocking and vibration.

Table 1: Specifications of "Raicho I".

Length	Approx. 10m
Width	Approx. 2.3m
Height of underwater	Approx. 1.2m
Motor power	25kW
Gross ton	2.5ton

Maximum speed	10kn
Boat type	Battery-powered propulsion, quick-charge capability
The number of passengers	Crews: 2, Passengers: 10

b. LiDAR

We used the VLP-32C (Velodyne), to acquire streaming point clouds because of its wide and long measurement range suitable for object detection and 3D mapping in river environments.

Table 2: Specifications of the VLP-32C.

Size	Diameter 103mm, height 86.9mm
Weight	925g
The Number of channels	32
Measurement range	200m
Distance measurement accuracy	Up to±3cm (Typical)
Vertical field of view	40°(-25°to ±15°)
Vertical angle resolution	0.33°
Horizontal field of view	360°
Horizontal angle resolution	0.1°-0.4°
Scanning speed	5Hz-20Hz

c. Omni-directional camera

An omni-directional camera (Ladybug 5+, FLIR) was used to capture panoramic images of the environments. The resolution of the image composition was 2 mm/10 m.

Table 3: Specifications of the Ladybug 5+.

Size	Diameter 197mm, height 160mm
Weight	3.0kg
Resolution	2048×2464
Frame rate	30
Sensor	Sony IMX264， CMOS， 2/3”
Shutter	Global shutter
Dynamic range	70dB

Results and Discussion

The locations of the six boats identified in the experimental section are shown in Figure 8.

The measured views of the boats are also shown in Figure 9.

Figure 9: A scene of boat measurement.

Figure 10 shows an example of successful boat detection. The green color indicates the detected boats, and the pink color indicates background objects, such as revetments and bridges. In this example, only the point cloud of the boat cluster is correctly identified as “boat”. Figure 11 shows the tracking results. The red bounding box indicates the detected boat, and the red line indicates the estimated trajectory of the boat.

Figure 10: Successful example of boat detection.

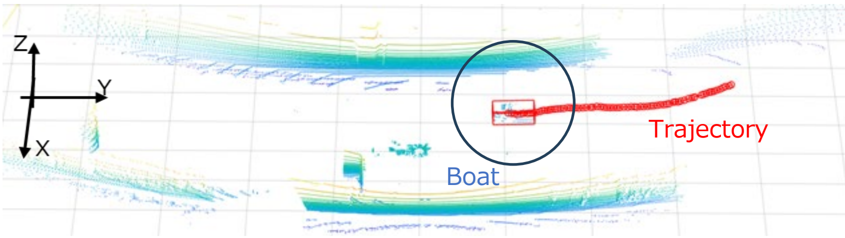

Figure 11: Boat tracking result.

In this study, boat detection using point clouds was performed on objects within a 100-meter radius of around the LiDAR. The detection radius was determined based on the acquisition of sufficient point clouds for detection and avoidance maneuvers. Although the selected LiDAR is capable of acquiring point clouds at 10Hz, object detection in this study was, performed at 1 Hz (sampled every 10 frames) to balance object detection performance and real-time processing. The detection rate was calculated using Equation 2. The detection rates for six moving boats are shown in Table 4. The best result was 89.66%, and the worst result was 48.65%.

$$Detection\ rate = \frac{The\ number\ of\ frames\ in\ which\ navigational\ boats\ were\ correctly\ detected}{The\ number\ of\ frames\ showing\ boats\ underway}$$

(Equation 2)

Table4: Detection rate of boats.

	Boat 1	Boat 2	Boat 3	Boat 4	Boat 5	Boat 6
Detection Rate[%]	53.84	77.27	89.66	50.00	48.65	66.67
The number of Frames	26	22	29	20	37	48

There are many cases of undetected boats and false positives were present in all scenes. Figure 12 and Figure 13 show examples of object detection. Figure 12 shows that a moving boat was not detected from the point clouds. The moving boat was detected as a background object because the point cloud cluster of the moving boat was too large to fit into the bounding box used for false detection correction. In this case, point clouds of water surfaces adjacent to the boat were merged into point clouds of the moving boat. Therefore, the merged point clouds were filtered as background data. To remove water surface point clouds of water surfaces, we applied a height filter. However, the point clouds of boats up to 50 m away from the boat were removed by the height filtering. Therefore, it is necessary to develop a method to remove the water surface point clouds with horizontal rectification of LiDAR. In addition to this example, there were also detection failures due to an insufficient number of points in the moving boat cluster also occurred. Since the identification of oncoming boats in our method was determined by the corresponding clusters between the current scan and the reference scan after scan matching, point cloud clustering was required before the object identification processing. In Figure 13, the object identification failed because the number of moving boat point clouds was too small for point cloud clustering. Based on LiDAR scanning, the point cloud density decreases at distant points. Therefore, the performance of object detection depends on the LiDAR measurement distances to identify boats. Figure 13 also shows that part of the revetment was detected as a part of the boat. One of the reasons for this is that there is a problem in the evaluation method for the degree of conformity of the corresponding clusters. In our method, the corresponding clusters were evaluated by the movement value of the cluster center of gravity. However, the center of gravity of the cluster does not correspond to the center of the object if not all the surfaces are measured. Therefore, it is necessary to improve the cluster matching evaluation method by using the average distance between the centroids of the corresponding clusters.

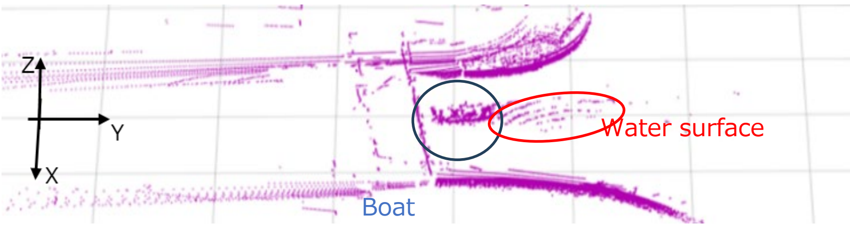

Figure 12: Failed example 1 (boat was not detected).

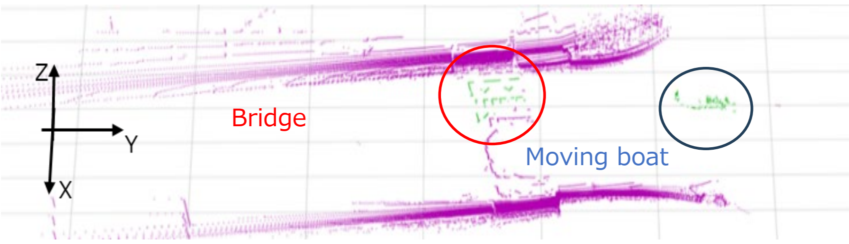

Figure 13: Failed example 2 (bridge was detected as a boat).

An example of a scan matching failure is shown in Figure 14. The green dots indicate the current scan data, and the pink dots indicate the reference scan data. We confirmed that the scan matching failures occurred along a curved trajectory. The direction of the boat motion was determined by the motion estimated from the subtraction between the current and previous scan. Thus, the optimization algorithm was not run to estimate the rotation and translation matrices of the relative boat motion.

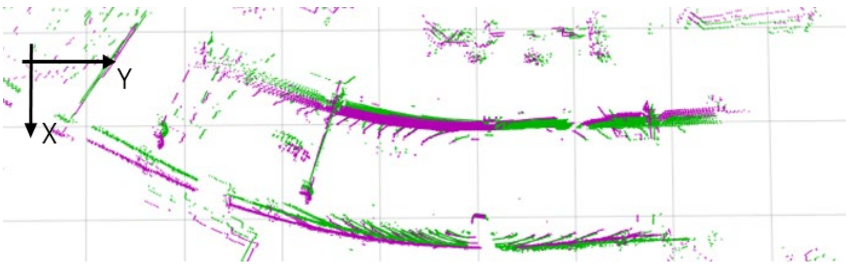

Figure 14: Failed example 3 (curved trajectory).

Conclusion

In this study, a moving object recognition method based on scan matching using LiDAR was applied to the detection and tracking of boats to achieve an automated avoidance function. We evaluated our methodology and the average detection rate of boats in the experimental area was 64.29%. Moreover, we confirmed that scan-matching failures frequently occurred when the boat mounted on the LiDAR turned. Through our experiments, we confirmed that water surface points remained in the object detection processing. Thus, we focus on the point cloud filtering before the moving object detection. We also confirmed that technical issues remained in the scan matching. Improved cost functions and optimization methodologies are also needed to precisely match clusters. Therefore, we will propose a method to improve boat detection and scan matching, as our future works.

Reference:

M. Shinbo, M. Hirasawa, and M. Oshima, (2005). Extraction of Navigational Boats from Time-Series Images of Navigational Environment Acquired by Multi-sensors. Transactions of the Nautical Institute of Japan, No.114, pp.65-72.

Lihui Zhou, Masatoshi Shinbo, and Masatake Oshima, (2007). Real-Time Boat Extraction Method from Time-Series Images of Navigational Environment. Transactions of the Nautical Institute of Japan, No.117, pp.175-182.

Yohei Matsumoto, (2013). Ship Image Recognition Using HOG Features. Transactions of the Nautical Institute of Japan, No.129, pp.105-112.

M. Kondo, T. Miyashita, R. Imai, M. Inaishi, M. Furuya, and M. Nakagawa, (2020). A Study on Ship Tracking and Course Prediction Using Panorama Camera -For River Traffic-. Transactions of the Japan Society of Navigation Engineers, No.143, pp.31-41.