Wide-area Accident Risk Estimation System using Dynamic Coordination of Multiple LiDARs

Kenji Murase
Ritsumeikan University
Graduate School of Information Science and Engineering
Osaka, Japan
is0582xi@ed.ritsumei.ac.jp

Hideaki MIYAJI   Hiroshi YAMAMOTO
Ritsumeikan University
College of Information Science and Engineering
Osaka, Japan

***Abstract*—In a society facing rapid population decline and aging, the use of autonomous robots is being considered, but there remains a risk of unpredictable accidents such as contact accidents with pedestrians. Especially, in indoor corridor intersections, the robots cannot easily perceive pedestrians and other robots passing through, which can result in contact accidents with them. To address this problem, systems are needed that identify conditions of blind spots and provide early warnings of potential accidents. Existing studies propose installing a LiDAR and camera near outdoor intersections for preventing accidents related to vehicles. However, since these sensors are fixed and blind spots remain, positional relationships between pedestrians and vehicles cannot always be observed. Therefore, this study proposes a new observation system in which multiple three-dimensional LiDARs capable of measuring distance, position, and shape are mounted on facilities, pedestrians, or robots. Additionally, the system coordinates with autonomous mobile robots equipped with similar sensors to enable detailed observation of the internal state of intersections. By sharing and integrating the observation results among the LiDARs in real time, the system captures detailed conditions of the intersections without blind spots. In the proposed system, the coordinate systems of point-cloud data observed from all LiDARs are unified based on relative positions estimated from IMU data, then merged into a unified map. And then, the clusters of points corresponding to moving objects such as humans and other robots are extracted from the constructed point-cloud map using clustering algorithms. Finally, machine learning models are applied to the cluster of moving objects to identify type of pedestrians such as carrying luggage, providing information for preventing the accidents to surrounding pedestrians.**

***Index Terms*—Digital Twin, IoT, 3D LiDAR, 3D point cloud analysis, human detection, Robot**

I. Introduction

In a society experiencing rapid population aging and declining birth rates, the use of robots is being considered, but there exists a risk of collisions with pedestrians. Particularly at intersections within building corridors, not only predestrians but also robots cannot perceive the state of pathways intersecting perpendicularly to their moving directions, potentially leading to contact accidents with people or other robots.

Regarding the contact accidents, the survey by an insurance company in Japan reveals that the ratio of accidents occurred when the vehicles turning right at intersections is 39% of the total [1]. The primary cause of the high rate of the contact accidents during right turns is the existence of the blind spots due to surrounding obstacles, preventing drivers from detecting pedestrians on the roadside in advance. Similar risks exist at indoor intersections. To prevent this, systems are needed at intersections that can observe situations in blind spots and provide early warnings of potential accidents.

Existing researches propose systems that install LiDAR (a type of laser radar capable of acquiring distance, position, and shape data of objects) and cameras at outdoor intersections to monitor the movement of pedestrians and vehicles near the intersection [2]. However, since LiDAR and cameras are installed in fixed positions, it is difficult to observe the conditions blind spots within the intersection, potentially preventing the provision of information regarding the relative positions of pedestrians and vehicles.

Therefore, in the proposed system, the coordinate systems of point-cloud data observed from all LiDARs are unified based on relative positions estimated from IMU data, then merged into a unified map. And then, the clusters of points corresponding to moving objects such as humans and other robots are extracted from the constructed point-cloud map using clustering algorithms. Finally, machine learning models are applied to the cluster of moving objects to identify type of pedestrians such as carrying luggage, providing information for preventing the accidents to surrounding pedestrians.

II. Related Works and Objectives of Our Study

A. Research on Human Detection Using LiDAR and Infrared Cameras

Saito et al.'s study proposes a system that 3D LiDAR and infrared cameras are installed near traffic signals at intersections to detect moving objects such as people and vehicles, as well as obstacles [2]. However, the 3D LiDAR used in this system has a narrow field of view of 30 degrees, making it difficult to capture spatial information over a wide area. Furthermore, since the 3D LiDAR and infrared camera are installed in fixed positions such as on the roadside, if large trucks or similar vehicles are present at the intersection, the observable range becomes limited, creating difficulties in observing conditions in blind spots.

B. Research on Human Detection Using LiDAR and Cameras

Yingwei Li et al. propose a method to improve object detection accuracy using LiDAR and RGB cameras [3]. Specifically, the proposed method enhances object detection accuracy through machine learning by adding color information from the RGB camera to the point cloud data collected by the LiDAR. However, this approach faces the challenge of unstable object detection in areas with fluctuating ambient brightness.

C. Research on a Three-Dimensional Behavior Observation Network System Utilizing Multiple LiDARs

In our previous study, Mito et al. propose a system that observes the positions and movement trajectories of people across a wide area of indoor environments by installing multiple LiDARs [4]. Specifically, point cloud data from multiple LiDARs is integrated to form a single point cloud map, from which points corresponding to dynamic objects (people) are extracted. Furthermore, by observing the centroid coordinates of the cluster of points corresponding to each person, the system tracks their movement trajectories. However, since multiple LiDARs are installed at elevated positions, numerous blind spots (such as shadows cast by obstacles) exist in rooms with many obstacles (desks, bookshelves, etc.). Consequently, it is difficult to acquire point cloud data corresponding to people in all locations within the room.

D. Objectives of Our Research

Existing research faces the problem that observation devices such as LiDAR are installed in fixed positions such as on the roadside, resulting in existence of blind spots where observation results cannot be obtained. Furthermore, processing the large amounts of point cloud data generated by multiple LiDARs results in high computational loads and low real-time performance.

Therefore, this study proposes a new observation system that integrates point cloud data acquired in real time from multiple LiDARs fixed within a facility and mounted on autonomous robots moving within the facility. The proposed system enables real-time understanding of the state around indoor intersections and provides information to avoid contact accidents between pedestrians and robots. Specifically, point cloud data from multiple 3D LiDARs is integrated based on relative positions estimated via IMU and point cloud registration algorithms to construct a unified point cloud map of the observation environment. The constructed point cloud map data is then analyzed to extract clusters of points corresponding to dynamic objects like people and robots. The extracted clusters are then input into a machine learning model to identify which category they correspond to, such as pedestrians. Furthermore, based on the identification results, the system estimates the risk level of the accidents occurring near the intersection, and notifies nearby pedestrians.

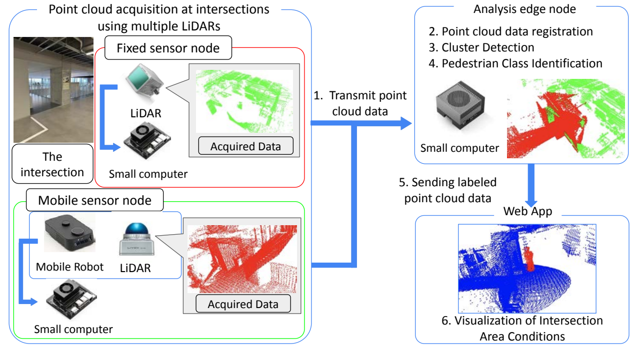

Fig. 1: Overview of the proposed system.

TABLE I: Specifications of 3D LiDAR (LIVOX AVIA)

Specification	Value
Power Supply Voltage	10–15V
Weight	498g
Maximum Range	450m
Field of View	70.4° (Horizontal) × 77.2° (Vertical)
Point Cloud Data Rate	240,000 points/sec
Range Accuracy	2cm

III. PROPOSED INDOOR INTERSECTION OBSERVATION SYSTEM

A. Overview of the Proposed System

Figure 1 shows the overall architecture of the system proposed in this study. Fixed sensor nodes and mobile sensor nodes acquire point cloud data from their respective mounted LiDARs. Each edge node performs preprocessing such as downsampling on the point cloud data and transmits the processed data to the analysis edge node. The analysis edge node constructs the received point cloud data into a unified point cloud map and performs clustering processing and pedestrian type identification from the point cloud using machine learning models.

B. Configuration of Fixed Sensor Nodes

The fixed sensor node consists of a single board computer (Nvidia Jetson Orin Nano) and a 3D LiDAR (Livox Avia). The sensor node is installed on roadside units at intersections to capture point cloud data of specific areas (such as corridors). As the 3D LiDAR, the Livox Avia is selected because, it can capture high-density point cloud data within its observation range, as shown in Table I.

C. Configuration of Mobile Sensor Nodes

The mobile sensor node consists of a single board computer (Jetson Orin Nano) and a 3D LiDAR (Livox MID-360). The sensor node is installed on robots moving near intersections. As shown in Table II, the MID-360 is suitable for acquiring surrounding point cloud data to create point cloud maps and for performing self-localization using LIO (LiDAR-Inertial Odometry) algorithm because the LiDAR enables omnidirectional observation.

TABLE II: Specifications of 3D LiDAR (LIVOX MID-360)

Specification	Value
Power Supply Voltage	9–27V
Weight	265g
Maximum Range	40m
Field of View	360° × -7°∼52°
Point Cloud Data Rate	200,000 points/sec
Range Accuracy	2cm

D. *Configuration of Edge Nodes for Analysis*

As the edge node for analysis, a Jetson AGX Orin is selected for the high-performance point-cloud analysis and processing of machine learning. The edge node can handle high-performance processing such as analyzing large-scale data like point cloud data and executing machine learning models. For example, this includes point cloud alignment processing and background subtraction processing.

E. *Workflow for Analyzing Point Cloud Data*

Fixed sensor nodes and mobile sensor nodes acquire point cloud data from their LiDARs at the same cycle (e.g., 3Hz, 5Hz). The sensor nodes also analyze the point cloud data acquired from the LiDAR to construct a point cloud map of the surrounding area. Additionally, the mobile sensor node simultaneously performs self-localization within that point cloud map. Subsequently, each sensor node transmits the generated point cloud map to the edge node of analysis. The edge node of analysis integrates the received point cloud maps to construct an overall point cloud around the intersection and performs processing such as extracting points corresponding with each moving objects on the point cloud map. Details of the processing are described in the Section IV.

IV. POINT CLOUD DATA ANALYSIS METHODS

A. *Method for Generating Point Cloud Maps Around Mobile Sensor Nodes*

This section describes a method for integrating point cloud data acquired by mobile sensor nodes to generate a point cloud map of their surroundings. In this study, the mobile sensor nodes constantly estimate their relative position from the starting point and use an algorithm called Fast-LIO2 to integrate point cloud data based on the estimated position [5]. This algorithm continuously estimates the relative position using three-dimensional acceleration and angular velocity data acquired at 3-millisecond intervals from the IMU embedded on the LiDAR. Furthermore, the continuously obtained point cloud datasets are aligned based on this relative position and incorporates a mechanism to gradually construct a point cloud map by correcting misalignments between consecutively acquired point cloud data sets using EKF (Extended Kalman Filter) [6].

B. *Method for Combining Point Cloud Data Between Sensor Nodes*

This section describes the method for integrating point cloud data obtained from fixed-installation sensor nodes and mobile sensor nodes. The point cloud map composed by the sensor nodes is merged with a global point cloud map constructed by the edge node. Here, the point cloud map obtained from each sensor node is designated as the source point cloud, and the pre-constructed global point cloud map is designated as the target point cloud.

In the proposed method, the coordinate system of the source point cloud is transformed into that of the target point cloud. First, for both the source point cloud and the target point cloud, the FPFH (Fast Point Feature Histograms) feature vector, which represents the spatial relationships between points, is calculated. The FPFH is a 33-dimensional vector capturing the spatial relationships between each point in the point cloud and its neighboring points. This feature vector is used to evaluate the similarity of the 3D structure between point clouds [7]. Based on these FPFH features, a rotation matrix is estimated using the RANSAC (Random Sample Consensus) algorithm to align the coordinate system of the source point cloud with that of the target point cloud [8]. The RANSAC algorithm is designed to derive a rotation matrix that fits representative shapes (e.g., planes) across different datasets. As shown in Fig. 2, the algorithm randomly selects pairs of points from both point clouds and calculates a transformation matrix (rotation + translation) so that bot point-cloud datasets overlap as much as possible. Next, this transformation is applied to all points in the source point cloud, and the number of points that end up in close proximity (the number of inliers) is evaluated. By repeatedly performing the process the algorithm becomes less susceptible to outliers and noise, enabling it to achieves a roughly correct alignment by identifying a transformation matrix so as to maximize the number of inliers.

Subsequently, precise alignment is performed using the ICP (Iterative Closest Point) algorithm. This algorithm rotates the coordinate systems to match the closest points between the source and target point clouds by minimizing the distance between the a small number of points randomly extracted from both point-cloud dataset as shown in Fig. 4. Through these processes, the rotation matrix is obtained for each source point cloud. After applying the rotation matrix to each source point cloud, they are integrated to generate a unified point cloud map on the global coordinate system of the target point cloud.

C. *Method for Extracting Clusters Corresponding to Pedestrians*

This section describes a method for extracting cluster of point corresponding to pedestrians from point cloud data. First, a point cloud map generated beforehand when no pedestrians are present is treated as the background point cloud. Next, for a newly generated point cloud map, background difference processing is applied based on the background point cloud to extract points corresponding to moving objects, including pedestrians and robots. The background subtraction processing

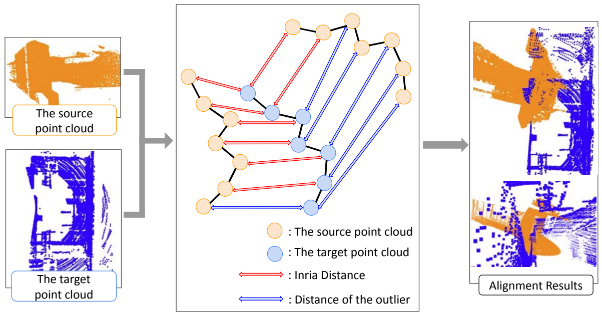

Fig. 2: Overview of Point Cloud Data Registration Using the RANSAC Algorithm

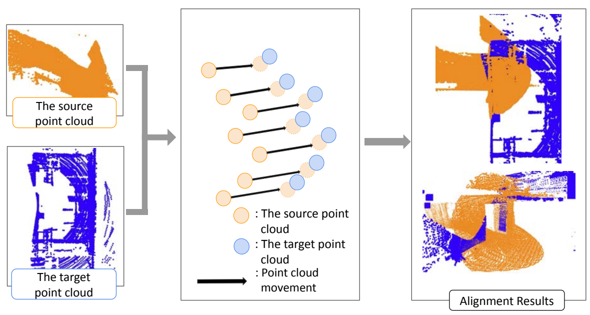

Fig. 3: Overview of Point Cloud Data Registration Using the ICP Algorithm

is a technique that extracts points present only in the newly generated point cloud by taking the difference of the nearest points between the background and the newly generated point cloud maps.

Subsequently, the DBSCAN (Density-Based Spatial Clustering of Applications with Noise) algorithm is applied to the points corresponding to the moving objects to build clusters based on point density [9]. As shown in Fig. 4, a core point is randomly selected from the point cloud. If a number of points greater than the threashold (min_points) are contained within a distance eps from the core point, the set of points are treated as a same cluster. Finally, the maximum value along the Z-axis (height) is calculated for the points within each cluster. If this value exceeds a predefined threshold (100 cm), the cluster is judged to correspond to a pedestrian.

D. Classification of Cluster Types Using Machine Learning Model

The proposed system identifies the pedestrian type of clusters extracted from point cloud maps by applying a machine learning model to them. The model identifies the clusters into three types (e.g., pedestrian, cargo carrier, pulling a suitcase) in order to assess the potential for the contact accident at intersections and the severity of such collisions. For example, pedestrians with obstructed forward vision may collide due to inattention.

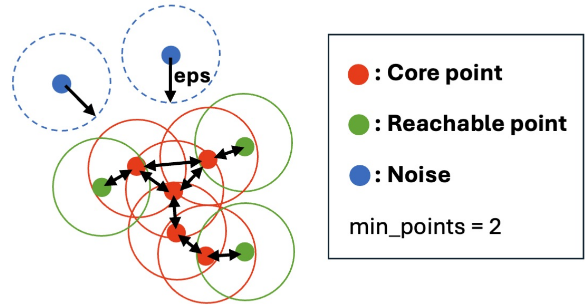

Fig. 4: Overview of Point Cloud Clustering Using the DBSCAN Algorithm

First, the cluster of points is vertically divided into two parts based on the average height along the Z-axis. For each part of the cluster, sizes in the x, y, and z axis directions and the number of points are calculated. Additionally, to obtain geometric features of the cluster point cloud, the Global Radius-based Surface Distribution (GRSD) is derived. The GRSD is a feature vector that represents the shape of the point cloud formed by each point as a feature. As shown in Fig. 5, it calculates the the GRSD is calculated by identifying geometric structures such as “plane”, “edge” and “cylinder” at various locations within the point cloud cluster, generating 21-dimensional features [11]. As a result, the proposed method derives the number of point clouds, the height along the Z-axis, and the geometric information derived from GRSD for both the vertically split point cloud clusters and the entire point cloud.

Furthermore, the proposed method adopts LightGBM as the machine learning model for classifying pedestrian types [10]. The lightGBM is an open-source machine learning framework based on Gradient Boosting Decision Trees (GBDT) developed by Microsoft capable of efficiently handling large datasets and high-dimensional features. By inputting the features derived from each cluster to the model, the system identifies the type of the cluster.

V. Performance Evaluation of the Proposed Method

A. Experimental Objectives

It is necessary to verify whether the point cloud data obtained from the mobile edge node and the pre-prepared point cloud map are accurately aligned during the registration process. Additionally, it is necessary to verify the accuracy rate of the machine learning model in correctly identifying pedestrian types from the point cloud data. Finally, it is necessary to verify whether the system’s sequence of processing steps can be executed in real time.

B. Performance Evaluation Settings

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \tag{1}$$

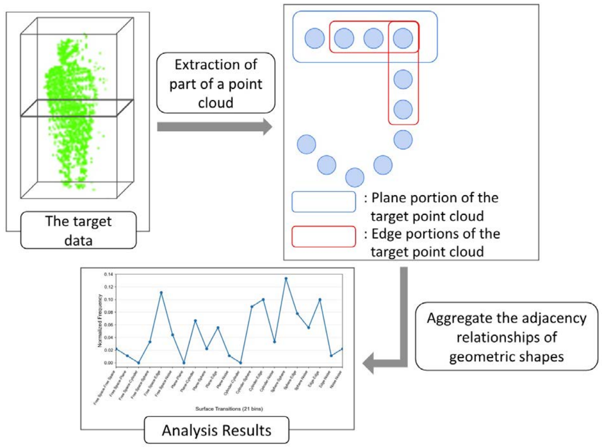

Fig. 5: Overview of Feature Extraction Using the GRSD Algorithm

In this experiment, the point cloud map obtained from the mobile sensor node is used as the source point cloud, and the point cloud constructed beforehand using the MID-360 LiDAR serves as the target point cloud. As explained in Section IV-A, the point cloud map of the mobile sensor node is generated by Fast-LIO2, from point cloud data observed over a 30-second period. The experimental environment is prepared in a Ritsumeikan University building and is assuming as an intersection, as shown in Fig. 6. One fixed sensor node and one mobile sensor node are deployed on the intersection. After performing alignment using the method described in Section IV-A, such as wall corners and floor heights are selected, as shown in Eq. 1, and multiple points are randomly acquired from these areas. The alignment accuracy is evaluated by calculating the Root Mean Squared Error (RMSE) between corresponding points. In Eq. 1, y_i is the point coordinate from the point cloud map obtained by Fast-LIO2, and $\hat{y}_i$ is the point coordinate from the pre-acquired point cloud map. Furthermore, the alignment success rate is evaluated by visually confirming whether randomly selected corresponding points are correctly aligned, assessing the percentage of successful completions out of 20 alignment processes. Additionally, for evaluating the accuracy of the machine learning model in identifying pedestrian types, clusters corresponding to each class people carrying items in front, and people pulling suitcases, extracted from the point cloud map, are prepared. For training the machine learning model, 240 datasets corresponding to each type are utilized, and 60 datasets are utilized for performance evaluation. An example of the prepared clusters is shown in Fig. 7. As performance metrics of the proposed method, the accuracy rate, which represents the percentage of subjects correctly identified is used.

C. Evaluation of Point Cloud Map Alignment

This section presents the evaluation results for RMSE and success rate in the point cloud registration method. As shown in Tab. III, the root mean square error (RMSE) in the proposed

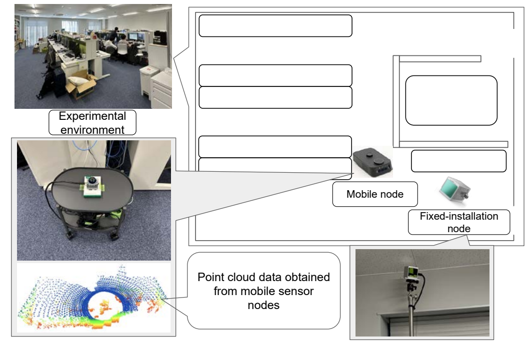

Fig. 6: Experimental Environment and Sensor Node Configuration

TABLE III: Results of Point Cloud Map Alignment Accuracy

Point Cloud Data Type	RMSE
With Point Cloud Map	5.48cm
Real-time Point Cloud without Map	27.23cm

point cloud registration is 5.48 cm. On the other hand, this table also presents the results when raw point-cloud data is used without constructing the map, and the RMSE in that case is 27.23 cm.

Furthermore, the alignment success rate was 80% when using the proposed method, compared to 35% when using the raw point cloud data. These results demonstrate that constructing point cloud maps with three-dimensional structural characteristics facilitates the identification of key feature points, leading to reduced alignment errors and improved success rates.

D. Evaluation of pedestrian Identification Algorithms

This section evaluates the accuracy of human type identification using LightGBM for clusters of moving objects extracted from point cloud maps. Figure 8 shows the accuracy rate of pedestrian type identification in cases of acquisition frequency of 3Hz and 5Hz. In all cases of the frequency, the identification accuracy is close to 90%, indicating high-precision inference capability. Future research will aim to

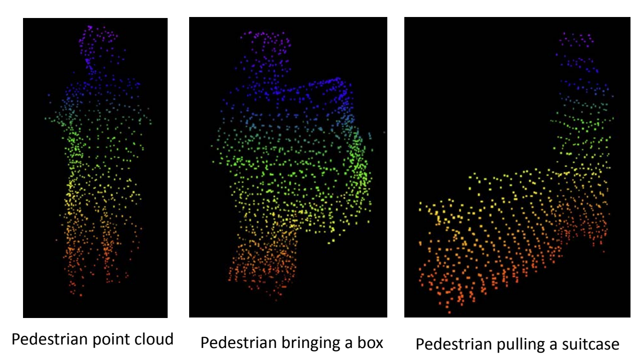

Fig. 7: Example of point cloud data for pedestrians.

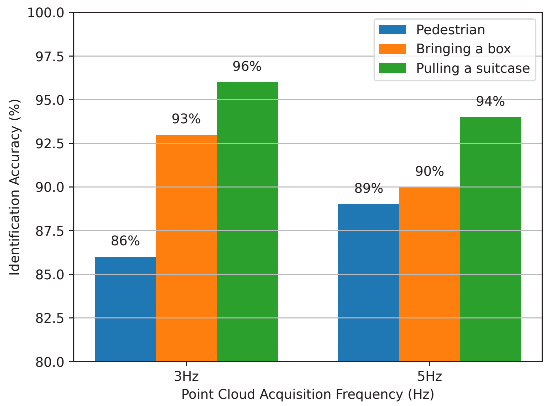

Fig. 8: Evaluation of the accuracy of human type identification.

TABLE IV: Results of Processing Time for Each Operation

Processing Content	Processing Time (s)
Point Cloud Map Construction	0.01
RANSAC	19.0
ICP	0.007
Voxelization	0.008
Background Subtraction	0.108
Cluster Extraction	0.015
Feature Extraction	0.03
Human Type Classification	0.00001

improve pedestrian type identification accuracy by collecting more data and refining feature extraction methods from point cloud data.

E. Processing Time Evaluation

This section evaluates the processing time required for constructing point cloud maps using Fast-LIO2, the processing time required for each operation on the point cloud maps, and the processing time for cluster identification using LightGBM. The Jetson AGX Orin, which constitutes the edge node, is used for the evaluation.

The average processing time for each task is shown in Tab. IV. This figure shows that the RANSAC process takes longer time of about 19 seconds from others. However, this process only needs to be performed during the initial alignment, and the subsequent alignments only require the registration by ICP. Excluding the initial RANSAC processing, the processing time falls within the 300ms processing interval, ensuring real-time capability.

VI. Conclusion

This study proposed a new observation system that detects moving objects near intersections with blind spots and identifies pedestrian types by integrating and analyzing point cloud data acquired from multiple LiDARs installed at indoor intersections and on robots traversing those intersections. Furthermore, to demonstrate the effectiveness of the proposed system, a proof-of-concept experiment was conducted in a simulated indoor intersection space. As a results, it was confirmed that the system can accurately construct point cloud maps, enable identification of pedestrian types extracted from these maps, and execute these processes in real time. In the future study, we aim to further improve real-time performance of the point-cloud analysis by proposing algorithms to optimize RANSAC and ICP processing. Additionally, the system will be enhanced to identify a wider variety of pedestrian types and improve the identification accuracy. As a method, we will install additional fixed-mount LiDAR units to determine how much they improve pedestrian type identification.

Acknowledgement

This work is supported by Japan Society for the Promotion of Science (JSPS) KAKENHI Grant Number JP24K20774, JP24K02916, and JP23K28078.

References

[1] Sompo Japan Insurance Inc., “Dangers Hidden in Right Turns at Intersections”, February 2024, https://image.sompo-rc.co.jp/reports/Monthly2402.pdf (in Japanese).

[2] Mai Saito, Chotaro Yamamoto, Hiroto One, Takumi Masamoto, Ang Li, Emil Dobber, Houcheng Li, Toshiro Ito, Hiroyuki Nomura, Kazuhiro Oda, “Infrastructure Alert System Based on Intersection Monitoring Using Measurement Sensor and Infrared Camera,” Journal of Industrial Information Integration, November 2022.

[3] Haibin Liu, Chao Wu, Huanjie Wang, “Real time object detection using LiDAR and camera fusion for autonomous driving,” scientific reports, May 2023.

[4] Syosei Mimoto, Hideaki Miyaji, Hiroshi Yamamoto, “Three-Dimensional Behavior Observation Network System Using Multiple LiDARs for Indoor Digital-Twin”, IEEE International Conference on Consumer Electronics, January 2025.

[5] Wei Xu, Yixi Cai, Dongjiao He, Jiarong Lin, Fu Zhang, “FAST-LIO2: Fast Direct LiDAR-inertial Odometry,” IEEE TRANSACTIONS ON ROBOTICS, January 2022.

[6] S. Rusinkiewicz, M. Levoy, “Efficient variants of the ICP algorithm, ” Proceedings Third International Conference on 3-D Digital Imaging and Modeling, pp. 145, 2001.

[7] Radu Bogdan Rusu, Nico Blodow, Michael Beetz, “Fast Point Feature Histograms (FPFH) for 3D registration, ” IEEE International Conference on Robotics and Automation, May 2009.

[8] Martin A. Fischler, Robert C. Bolles, “Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography, ” Readings in Computer Vision, pp. 726-740, 1987.

[9] Erich Schubert, Jorg Sander, Martin Ester, Hans Peter Kriegel, Xiaowei Xu, “DBSCAN Revisited, Revisited: Why and How You Should (Still) Use DBSCAN, ” ACM Transactions on Database Systems, pp. 1-21, July 2017.

[10] Guolin Ke, Qi Meng, Thomas Finley, Taifeng Wang, Wei Chen, Weidong Ma, Qiwei Ye, Tie-Yan Liu, “LightGBM: A Highly Efficient Gradient Boosting Decision Tree, ” Conference on Neural Information Processing Systems, 2017.

[11] João Paulo Silva do Monte Lima, Veronica Teichrieb, “An Efficient Global Point Cloud Descriptor for Object Recognition and Pose Estimation, ” SIBGRAPI Conference on Graphics, October 2016.