PAIDS: toward pedestrian high-precision position and attribute information detection

Zhengshu Zhou*

Graduate School of Informatics,
Nagoya University,
Nagoya, Japan
Email: shu@ertl.jp
*Corresponding author

Saya Kitamura

School of Engineering,
Nagoya University,
Nagoya, Japan
Email: kitamura@ertl.jp

Yousuke Watanabe and Shunya Yamada

Institute of Innovation for Future Society,
Nagoya University,
Nagoya, Japan
Email: watanabe@ertl.jp
Email: s_yamada@ertl.jp

Hiroaki Takada

Graduate School of Informatics,
Nagoya University,
Nagoya, Japan
Email: hiro@ertl.jp

Abstract: Pedestrian detection sensors in road infrastructure and smartphone's built-in sensors have been used to detect and track pedestrians for road safety. Nevertheless, although pedestrian detection sensors in road infrastructure can detect pedestrians' high-precision position, they cannot acquire the accurate attribute information of pedestrians. On the other hand, smartphone sensors can send location information, user identifier, and the attribute information of a user, but it has a significant margin of error in GPS data. The defects of LiDAR and smartphone render acquiring a pedestrian's high-precision location and attribute information simultaneously impossible. Currently, few studies on the simultaneous acquisition of pedestrian high-precision position and attribute information have been conducted. In this paper, the authors propose a pedestrian position and attribute information detecting system to extract both pedestrian high-precision position and attribute information in real-time based on LiDAR and smartphone sensor fusion. Moreover, an experiment is carried out to evaluate the system.

Keywords: pedestrian positioning; pedestrian information; smartphone sensor; LiDAR; sensor fusion.

Reference to this paper should be made as follows: Zhou, Z., Kitamura, S., Watanabe, Y., Yamada, S. and Takada, H. (2021) 'PAIDS: toward pedestrian high-precision position and attribute information detection', *Int. J. Mechatronics and Automation*, Vol. 8, No. 4, pp.187–199.

Biographical notes: Zhengshu Zhou received his Master's in Informatics from the Nagoya University, Japan, where he is currently pursuing his PhD at Graduate School of Informatics. Previously, he was a software engineer at a Japanese software industry. His research interests mainly include software engineering, autonomous driving and system dependability.

Saya Kitamura received her BE in Electrical and Electronic Engineering and Information Engineering from the Nagoya University, Nagoya, Japan, in 2020. She is currently an engineer with the Subaru Corporation. Her research interests mainly include dynamic map, sensor fusion and autonomous driving.

Yousuke Watanabe received his ME and DrE degrees from the University of Tsukuba, in 2003 and 2006. In 2014, he joined the Institute of Innovation for Future Society, Nagoya University, as a Designated Associate Professor. His research interests include data stream processing and information integration. He is a member of the Database Society of Japan, IEICE and ACM.

Shunya Yamada received his PhD in Informatics from the Nagoya University, Aichi, Japan, in 2020. Currently, he is a researcher at the Institute of Innovation for Future Society, Nagoya University. His research interests include sensing for intelligent space. He is a member of RSJ, JSME and IEICE.

Hiroaki Takada received his PhD in Information Science from the University of Tokyo, Tokyo, Japan, in 1996. He is a Professor at the Institute of Innovation for Future Society, Nagoya University, Nagoya, Japan, where he is also a Professor and Executive Director of Center for Embedded Computing Systems (NCES), Graduate School of Informatics. He was a Research Associate at the University of Tokyo from 1989 to 1997, and was a Lecturer and then Associate Professor at Toyohashi University of Technology, Toyohashi, Japan, from 1997 to 2003. His research interests include real-time operating systems, real-time scheduling theory and embedded system design. He is a Fellow of the IPSJ and JSSST, and member of ACM, IEICE and JSAE.

This paper is a revised and expanded version of a paper entitled 'Extraction of pedestrian position and attribute information based on the integration of LiDAR and smartphone sensors' presented at 2021 IEEE International Conference on Mechatronics and Automation (ICMA), Takamatsu, Japan, 8–11 August 2021.

1 Introduction

As everyone knows, driving automation system (DAS) is a typical safety-critical system. Accurate and timely acquisition of road information is the key to ensure road safety. As obtaining the location of pedestrians around autonomous vehicle (AV) is vital for collision avoidance, pedestrian high-precision position is an important part of road information. Meanwhile, pedestrian attribute information (e.g., age, language, health status, etc.) is essential for DAS to predict hazard and control AVs to accomplish their task safely according to the surroundings.

Hazard prediction can be enhanced using the attribute information of a pedestrian. If DAS is implemented with a pedestrian high-precision position and attribute information detecting function, safer autonomous driving can be realised. For instance, when pedestrians are walking on a sidewalk near an autonomous driving vehicle, children, older people, and disabled people should be paid more attention to than regular pedestrians. Another example is detecting a pedestrian who wants to pass in front of the autonomous driving vehicle. If the attribute information implies that the pedestrian is an old person, as the old walk slower than the young, DAS should turn off the engine to avoid idling when stopped to save energy.

Usually, communication between vehicle and pedestrian is based on driver's facial expressions and gestures. Pedestrian information can also be effectively used for communication between pedestrians, cyclists, and vehicles. For example, when a vehicle gives way to another vehicle, or a vehicle gives way to pedestrians or cyclists who want to cross the road, usually the driver shows by expression and gesture that the car is stopped and that it is okay to proceed. However, since there is no driver in the automatic driving vehicle, DAS has to use the car window as a display or other means to convey the message. However, when the target of the message conveyed is a pedestrian with visual impairment, this may be ineffectual or even dangerous. Therefore, if the DAS can acquire detailed information about the pedestrian, other means to convey the message can be considered. Besides that, when pedestrian near AV is a foreigner, he or she may not understand the usual voice instructions given by the AV. If the pedestrian's idiomatic language information can be obtained, DAS will give voice instructions in that language. In this way, the pedestrian's attribute information can be used to predict the danger or achieve fine control according to the situation. Furthermore, high-precision location information and attribute information of pedestrians can be used for the integration of DAS and smartphone applications. The location and attribute information of pedestrians shall preferably be timely and accurate enough, to provide quick and proper mobility service.

Although obtaining pedestrian high-precision position and attribute information simultaneously can be used to ensure road safety and provide rapid and proper mobility service, at present no research paper has been found on the topic. To fill this gap, the paper proposes a method to extract both pedestrian high-precision position and attribute information in real-time based on LiDAR and smartphone sensor fusion. It should be pointed out that the paper is a revised and expanded version of a conference paper (Zhou et al., 2021).

The paper is organised as follows. Section 1 introduces the background and motivation of our research in this paper. Section 2 describes related work. Section 3 proposes a system to acquire high-precision position and attribute information of pedestrians. The data processing and pedestrian matching method of the proposed system are introduced in Section 4 and Section 5, respectively. Section 6 presents the experimental evaluation of the proposed system. Finally, Section 7 summarises the paper by discussing the effectiveness and limitation of the proposed system and directing future work.

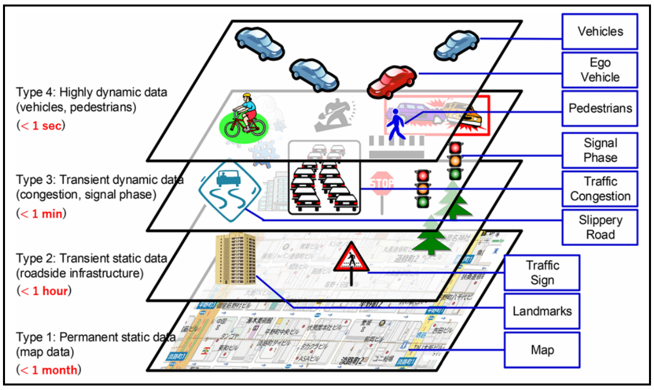
2 Preliminaries

This section explains the research related to pedestrian detection and pedestrian attribute information extraction.

2.1 Dynamic map

Our study in the paper is a part of an ongoing project of Nagoya University called Dynamic Map (DM2.0) platform (Shimada et al., 2015) which aims at integrating and sharing traffic-related information for autonomous driving and traffic situation analysis. As shown in Figure 1, DM2.0 platform has a four-layer architecture: highly dynamic data, transient dynamic data, transient static data, and permanent static data.

Figure 1 Four-layer structure of the dynamic map platform (see online version for colours)



As one of the contents of highly dynamic data, pedestrian real-time location and attribute information are undoubtedly of great significance to autonomous driving, road safety, and traffic condition analysis. Nevertheless, to the best of authors' knowledge, there are no existing methods or guidelines for simultaneous detection of pedestrian high-precision location and attribute information. We therefore conducted a literature review to capture the latest research progress (see Section 2.2).

2.2 Related work

2.2.1 Pedestrian detection and tracking

Approaches to detect and track pedestrians have attracted the interest of many scholars. Premebida et al. (2007) proposed an architecture to detect, track, and classify entities by using in-vehicle LiDAR and monocular vision. Subsequently, they presented a multi-sensor-based pedestrian detection method in an urban scenario using exclusive LiDAR-based features (Premebida et al., 2009). Cho et al. (2014) developed a moving object detection and tracking system to detect pedestrians, cyclists, and vehicles based on radar, LiDAR, and vision sensors. Kang and Han (2015) proposed a smartphone-based pedestrian dead reckoning approach to track pedestrians using data from inertial sensors embedded in smartphones. Zhou et al. (2015) proposed an activity-sequence-based indoor pedestrian localisation approach using smartphones. Kwon et al. (2016) presented a LiDAR and radar sensor fusion scheme for detecting partially hidden pedestrians. Lahmyed and Ansari (2016) designed a LiDAR and vision-based pedestrian detection system. Jin et al. (2016) presented a method to reduce erroneous pedestrian detection by moving vehicles. Lin and Lin (2016) proposed a fusing approach to improve the reliability of pedestrian detection using a 3D sensor and a camera. Chavez-Garcia and Aycard (2016) presented a fusion approach to detect and track moving objects using radar, LiDAR, and camera sensors. Wang et al. (2017) proposed a pedestrian detection and counting method to detect pedestrians using machine learning. Ghosh et al. (2017) presented an approach to detect pedestrians when they are close together or occluding one another using deep convolutional neural network. Kwon et al. (2017) proposed a LiDAR-radar sensor fusion scheme to detect pedestrians. Shin et al. (2017) presented a track management method to solve the discontinuous tracking problem caused by occlusions. Matti et al. (2017) developed a pedestrian detector for AV that exploits LiDAR data and visual information. Wu et al. (2017) proposed a pedestrian detection approach based on LiDAR and camera sensor fusion. Navarro et al. (2017) presented a machine learning method to detect pedestrian for autonomous vehicles using high-definition 3D range data. Zhang et al. (2018) proposed a multi-class pedestrian detection network for pedestrian detection in a distorted field of vision. Zeng et al. (2018) proposed a smartphone fusion location method optimised by indoor/outdoor pedestrian detection. Nauth et al. (2019) designed a smart pedestrian detection method using ultrasonic signals analysis. Han et al. (2020) proposed a deep small-scale sense network to detect small-scale pedestrians who are relatively far from cameras in practical applications.

2.2.2 Pedestrian feature recognition

There are also quite a few published studies on extraction and analysis of pedestrian features. Hariyono et al. (2014) proposed a method to detect pedestrians from a single camera mounted on the vehicle then classify the location of the pedestrian for the driver assistance system. Li et al. (2018) developed a smartphone-based system that detects the walking behaviour of pedestrians by leveraging the sensors and front camera on smartphones, improving the safety of pedestrians staring at smartphone screens. Tung and Shin (2018) proposed a system to detect distracted walkers by using a smartphone's built-in sensors and applications. Wu et al. (2020) proposed a tracking enhanced detection method to recognise people using their smartphones while walking. Junejo and Ahmed (2020) presented a multi-branch convolutional neural network that uses depthwise separable convolution layers to conduct pedestrian attribute recognition. Wong et al. (2021)

proposed a methodology of pedestrian tracking and attribute recognition based on computer vision and deep learning, facilitating the analysis of pedestrian walking behaviour. The studies tend to extract the state of pedestrians rather than their specific attributes. Additionally, most of the methods to obtain pedestrian attribute information have used techniques such as video recognition and deep learning. Although these techniques can extract pedestrian attributes, they are not always accurate and reliable. For example, computer vision can distinguish ethnic differences of pedestrians, but it is unable to tell pedestrian's nationality and preferred language.

2.2.3 Pedestrian safety

Pedestrians are vulnerable road users who need proactive protection in a hazardous environment. Yoshida et al. (2015) designed a system to predict and prevent bicycle-pedestrian collision by using smartphone built-in GPS. Jang and Lee (2017) proposed an analysis system to assess pedestrian–vehicle interaction risk levels using drone videos. Pottier et al. (2017) designed a pedestrian detection strategy to anticipate vehicle–pedestrian collision in urban areas by using capacitive probe. Lv et al. (2019) presented a method to generate high-resolution traffic trajectories from roadside-deployed LiDAR to predict vehicle–pedestrian conflicts. Zhao et al. (2019) presented a modified naive Bayes approach to conduct probabilistic prediction of pedestrian crossing intention using roadside LiDAR data. Ulak et al. (2021) proposed a pedestrian safety index for public transportation bus stops, to identify high-risk locations in a proactive manner. Zhu et al. (2021) proposed an agent-based framework for evaluating pedestrian safety at unsignalised crosswalks. Gruden et al. (2021) analysed the effects of digital distraction (e.g., smartphone) on pedestrians as they approach unsignalised intersections located on roundabout entrances and exits.

In summary, although there have been many studies on pedestrian detection and tracking, pedestrian feature recognition, and pedestrian protection, none of them can extract the real-time high-precision position and detailed attribute information of pedestrians at the same time. To put forward such a method to narrow the gap is the motivation of our research in the paper.

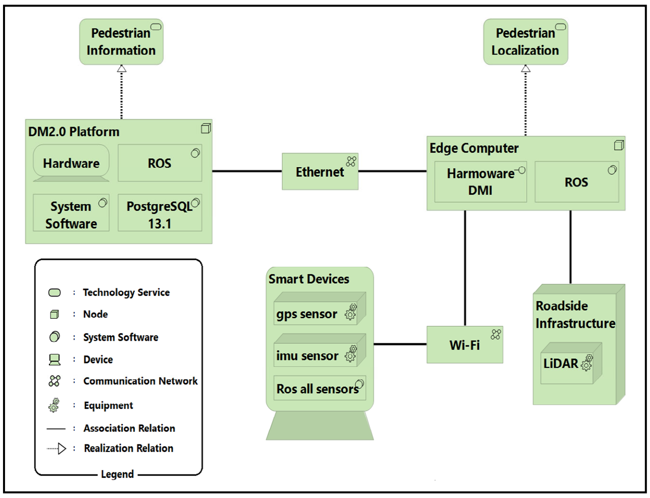
3 PAIDS: the pedestrian position and attribute information detecting system

The section introduces the architecture and processes of our proposal: the pedestrian position and attribute information detecting system (PAIDS).

3.1 System architecture

Figure 2 presents the system architecture of PAIDS. The system depends on the DM2.0 platform. The smartphones used in PAIDS are Google Pixel 4XL, the operating system is Android 10, and the application used to connect Android operating system and the robot operating system (ROS) is ROS-all-sensors. The LiDAR used is Velodyne VLP-16. The operating system used in the edge computer is Ubuntu 16.04 LTS. The ROS version is Kinetic Kame that corresponds to Ubuntu 16.04. LiDAR is connected to the edge computer through Ethernet. Smart devices and the edge computer, and the edge computer and DM2.0 platform, are connected by wireless networks.

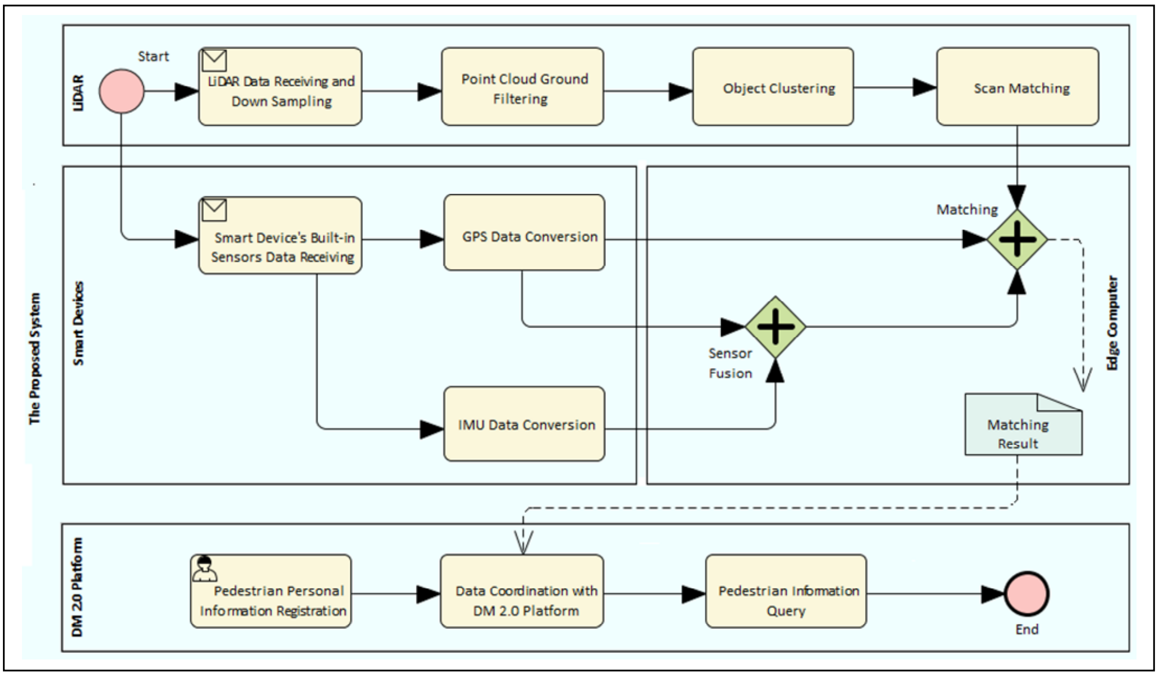
Figure 2 System architecture of PAIDS (see online version for colours)



3.2 System processes

The system process flow of PAIDS is shown in Figure 3. LiDAR data processing and smartphone data processing are performed simultaneously. Point cloud data is received from the LiDAR and then down-sampling is implemented. Ground point cloud is removed to extract non-ground point cloud for accurate object clustering. We use several human feature filters to extract pedestrians from the detected LiDAR clusters. After pedestrian detection, normal distributions transform (NDT) scan matching converts LiDAR coordinates to the Japanese plane rectangular coordinate (a world coordinate used in Japan). At the same time, GPS data is received from the pedestrian's smartphone sensor. We convert the raw GPS data from longitude and latitude values to the same coordinate system of converted LiDAR point cloud data. Next, the converted GPS position of the smartphone is matched with the pedestrian's position detected by LiDAR to calculate the probability that the detected pedestrian is the smartphone user, and the accuracy of matching is improved by GPS and IMU sensor fusion. The result of the pedestrian position matching is sent to the DM2.0 platform, to be linked with the personal attribute information registered by the system user and published to AVs for road safety.

Figure 3 System processes of PAIDS (see online version for colours)



4 LiDAR data and smartphone sensor data processing

The section describes the detailed implementation method of PAIDS based on ROS.

4.1 LiDAR data processing

We implement three ROS nodes to process LiDAR point cloud data. The ROS nodes are used to extract non-ground point cloud for accurate object clustering, to conduct NDT scan matching to transform the point cloud data from the LiDAR coordinate system into the global coordinate system, and to cluster and extract objects from the point cloud, respectively.

We adopt ray ground filter (Himmelsbach et al., 2010), a point cloud segmentation algorithm to remove ground points. The brief process of the algorithm is as follows:

1 Subscription and pre-processing of point cloud data: Gain access to original point cloud data by subscribing to the relevant ROS topic, and down sample the original point cloud.

2 Clipping and filtering point cloud after down-sampling: Set the upper threshold of height for point cloud clipping according to the height of the LiDAR and the estimated maximum height of pedestrians.

3 Segmentation of ground and non-ground point cloud: Sort the adjacent points on the same LiDAR ray by radius value (distance from point to LiDAR) and then judge whether the slope of the adjacent two points is greater than the preset slope threshold, to determine whether the points are ground points or not.

We transform the LiDAR point cloud data from the LiDAR coordinate system into the global coordinate system using high-precision 3D map data and NDT scan matching. Although the Point Cloud Library (PCL) includes the function of NDT matching, we implement the ndt_omp, a high-speed matching method using multithreading and novel search algorithms, to achieve higher real-time performance.

Thereafter, we use EuclideanClusterExtraction function of PCL for Euclidean clustering. Since the maximum sensing distance of the Velodyne VLP-16 LiDAR is about 100 metres, we remove the point cloud data with a distance of 100 metres or more from the LiDAR for better processing performance. In addition, distance sensors such as 3D-LiDAR have the characteristic that the density of the point cloud becomes sparser in areas farther from the sensor; therefore, different clustering distance thresholds are used in regions with different distances from the LiDAR to achieve better object detection accuracy. Next, we extract pedestrians from the detected object cloud by using human feature filters, which will be introduced in detail in Section 5. We send the detect pedestrians' LiDAR data to the DM2.0 platform via a dedicated communication API. The LiDAR data format is jsk_recognition_msgs/BoundingBoxArray.msg.

4.2 Smartphone data processing

To receive GPS data from multiple smartphones, we implement a script program to monitor all ROS topics, and then use a regular expression to extract ROS topics containing {device_id}/android/fix to subscribe to the GPS position message of each smartphone by specifying {device_id} as a parameter in the callback function of the ROS node. Since the acquired GPS position data is in the format of longitude and latitude, it is converted to the Japanese plane rectangular coordinate so that it can be matched with LiDAR data.

To calculate the probability that the detected cluster is a smartphone user of our proposed system, we use a bivariate normal distribution of GPS data as shown in Table 1. μ_1 is the expected value of the x-coordinate (plane rectangular coordinate), μ_2 is the expected value of the y-coordinate (plane rectangular coordinate), σ_1 is the variance of the x-coordinate, σ_2 is the variance of the y-coordinate, and ρxy is the correlation coefficient of the x and y coordinates. We use a custom ROS message (see Table 1) and the dedicated communication API to send pedestrians' GPS data and LiDAR data to the DM2.0 platform.

Table 1 Definition of ROS custom message for GPS data

Device ID	*Header*	*Frame ID*	μ_1	μ_2
String	Long	String	Double	Double
Device ID	σ_1		σ_2	ρxy
String	Double		Double	Double

5 Pedestrian detection and identification

As shown in Figure 4, among the LiDAR clusters, in addition to pedestrian objects, there are also a lot of garbage data which interferes with pedestrian detection and matching. Therefore, it is necessary to screen LiDAR clusters to extract pedestrians. In the LiDAR data processing of the previous section, to extract pedestrians from LiDAR point cloud data we calculated and saved the features of the detected objects while clustering. This section describes the human feature filters we designed to extract pedestrians from the detected objects, and the matching method to link pedestrians' high-precision position with their personal attribute information.

Filters 1~4 are designed for the number of points in the LiDAR point cloud cluster, the size of the detected cluster, the standard deviation value of the points' coordinates of the cluster, and the shape of detected cluster, respectively. The filters' thresholds are determined while tuning to reach a relatively good filtering effect. Table 2 presents the details of the filters. And the quantitative evaluation results of the filtering effect are presented in the next section.

After pedestrian extraction by cluster filters, matching is conducted using the converted LiDAR data and GPS data (see Figure 5). Pedestrian matching is employed to collate and link the position of an object detected by LiDAR with the GPS position information transmitted by the pedestrian's smart device. Next, the probability that the detected object is a pedestrian with the received GPS location is calculated and published as a ROS topic. The calculation method of the probability that a GPS source pedestrian exists in a certain range is integrated into the probability density function in the range.

Figure 4 Pedestrian detection with and without object filters (see online version for colours)

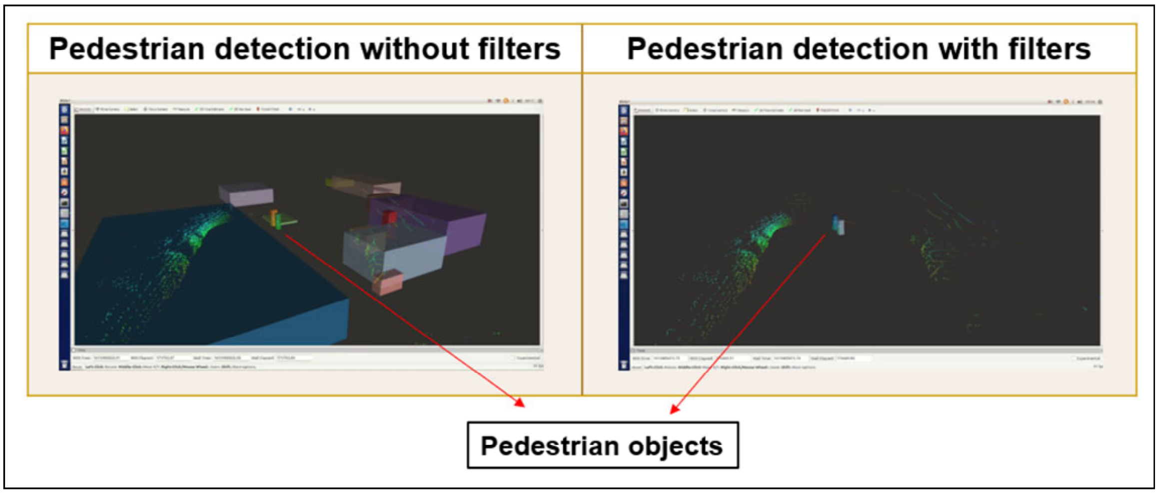


Figure 5 Matching of LiDAR data and smartphone data (see online version for colours)

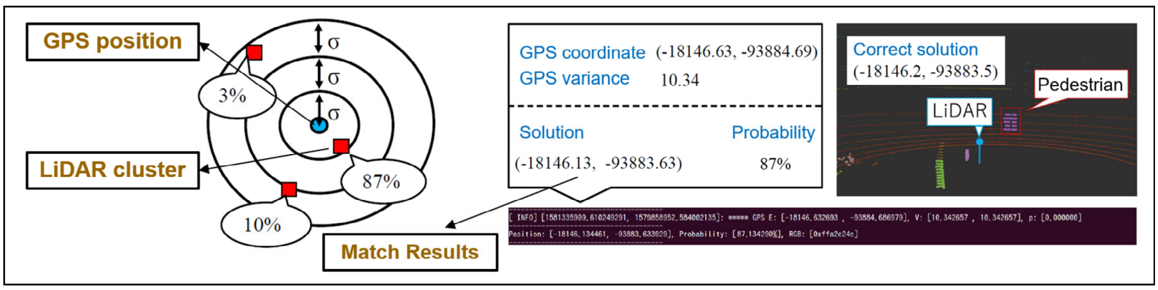


Table 2 Human feature filters to extract pedestrians

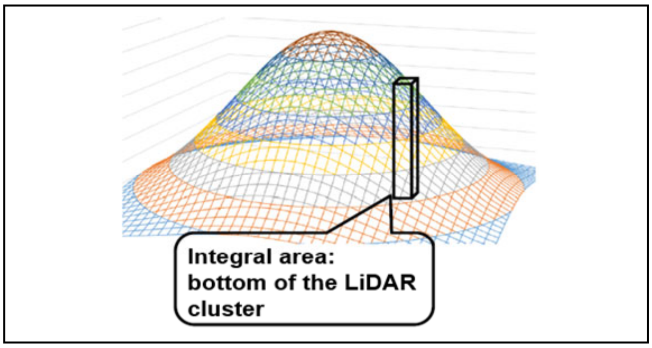
Filter	*Feature*	*Obstacles to be filtered*
1	Number of points in LiDAR data	Buildings, large vehicles
2	Size	Buildings, vehicles
3	Standard deviation of LiDAR data	Buildings, vehicles, motorcycles, bikes
4	Body type (proportional between body and head size)	Cylinders (e.g., roadside trees and wire poles)

We considered integrating probability density function that represents the existence probability of a pedestrian and employed the two-dimensional normal distribution. Normal distribution is used because the random variables of the coordinates where GPS actually existed follow the normal distribution of the coordinate and variance values transmitted by the GPS. The reason why we use two-dimensional instead of three-dimensional is that the vertical sensing accuracy of smartphone GPS is quite low. The z-coordinate (altitude) is therefore discarded and only the horizontal planar *x*-coordinate and *y*-coordinate positions are used. Furthermore, the integral is approximated by calculating the volume of a cuboid whose bottom surface is a 1 m × 1 m square and height is the value of the probability density function corresponding to the average x, y coordinate value.

As shown in Figure 6, by integrating the probability density function representing the two-dimensional normal distribution in a certain range on the *x*-*y* plane, the probability that a pedestrian exists in that area can be obtained. The probability density function of the two-dimensional normal distribution uses the *x* and *y* coordinates obtained from the GPS data as the average value (μ_1, μ_2), variance σ1 of the x-coordinate, variance σ_2 of the y-coordinate, and the correlation coefficient ρ (see Table 1). The formula is as follows:

$$f(x_1, x_2) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \times \exp\left\{-\frac{1}{2(1-\rho^2)}\left(\frac{(x_1-\mu_1)^2}{\sigma_1^2} - 2\rho\frac{(x_1-\mu_1)(x_2-\mu_2)}{\sigma_1\sigma_2} + \frac{(x_2-\mu_2)}{\sigma_2^2}\right)\right\} \quad (1)$$

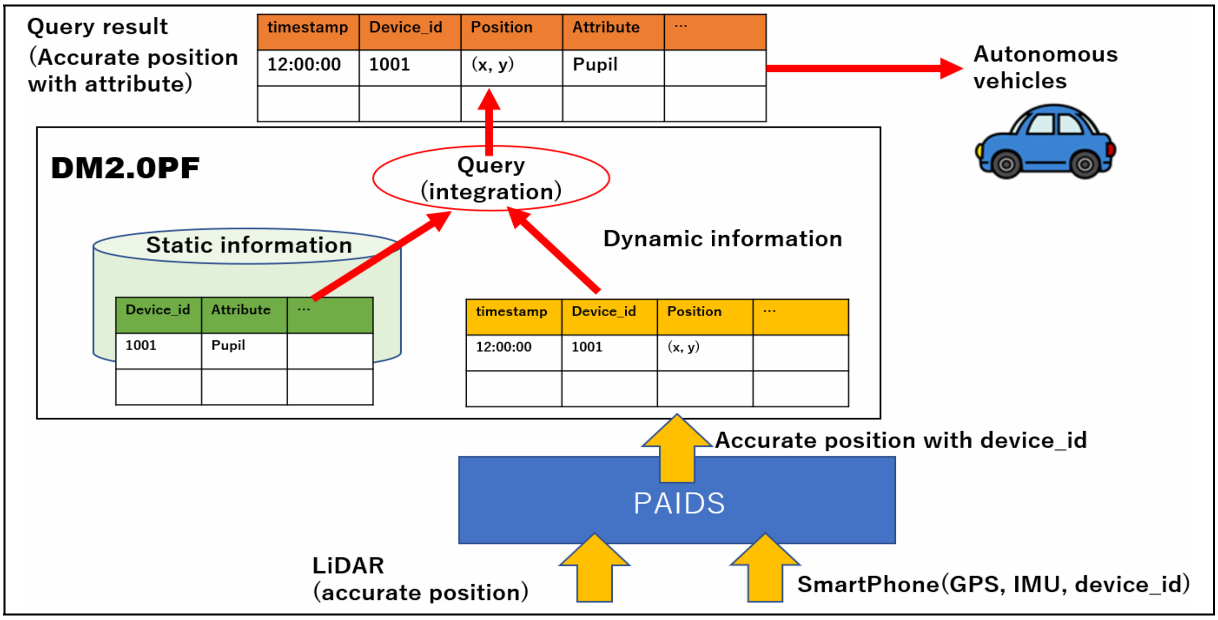
Figure 6 Integral region of the probability density function (see online version for colours)



To calculate the approximate value of the probability that a LiDAR cluster is a pedestrian of a GPS source, let the number of clusters within 3σ from the LiDAR position be n; then, the approximate value of the probability that the cluster i is a pedestrian of a GPS source can be obtained as follows:

$$\text{Probability }(i) = \frac{f_i(x, y)}{\sum_{i=1}^{n} f_i(x, y)} \quad (2)$$

Figure 7 Integration of pedestrian high-precision position and attribute information



Nevertheless, single person identification is not able to be realised when there are multiple pedestrians in a small area at the same time. For example, when five pedestrians gather together within the detection area of the system, we can only state that the probability of each pedestrian's existence is about 20%, but we cannot determine which person is the one we are interested in. Therefore, we use Kalman filter to conduct IMU and GPS sensor fusion to fix the pedestrian's velocity and course to obtain the pedestrian's posture for single person recognition (Sola, 2017). We then compare the fixed velocity and course with those of the clusters detected by LiDAR. The GPS source and LiDAR cluster that have the nearest posture are cross-matched to determine each pedestrian. The result of pedestrian position matching will be sent to the DM2.0 platform, to be related to the pedestrian attribute information in database and published to AVs (see Figure 7).

6 Experiments and results

We conducted three groups of experiments on a section of sidewalk of our university campus to evaluate the proposed system on cluster filtering, pedestrian existence probability calculation and single person matching, respectively. Figure 8 shows the experimental site, experiment devices and one of the experiment scenes.

Five groups of experiments were carried out to evaluate the filtration of using the four human feature filters separately and simultaneously (see Figure 9). We cumulatively counted the total number of pedestrians and obstacles every three seconds, and then investigated the number of detected pedestrians and obstacles to evaluate the filtering effect. Table 3 shows the experimental results. The results indicate that the four filters can filter non-pedestrian objects, and the filtering effect order is: all filters > filter 2 > filter 4 > filter 1 > filter 3. However, application of the filters will also lead to an increase in the miss detection rate. In particular, when the four filters are applied simultaneously, the miss detection rate is about 15.8%.

To verify the pedestrian matching function, we collected four datasets 1–4 that evaluate the detection of multiple pedestrians in close proximity, single pedestrian, single user of the proposed system/non-user, and multiple users of the proposed system/non-users, respectively. We counted the cumulative number of pedestrians and the number of correct pedestrian matches. Figure 10 presents the pedestrian matching accuracy. From the results, it is obvious that the matching effect is the best when there is only a single pedestrian. When there are multiple pedestrians, matching accuracy will decrease. We consider that the reason is pedestrians' body occlusion. Furthermore, when there are non-system users, the accuracy will also decrease. Therefore, it is necessary to popularise the system to attract more users for better matching accuracy.

Figure 8 Experiment devices and experiment scene (see online version for colours)

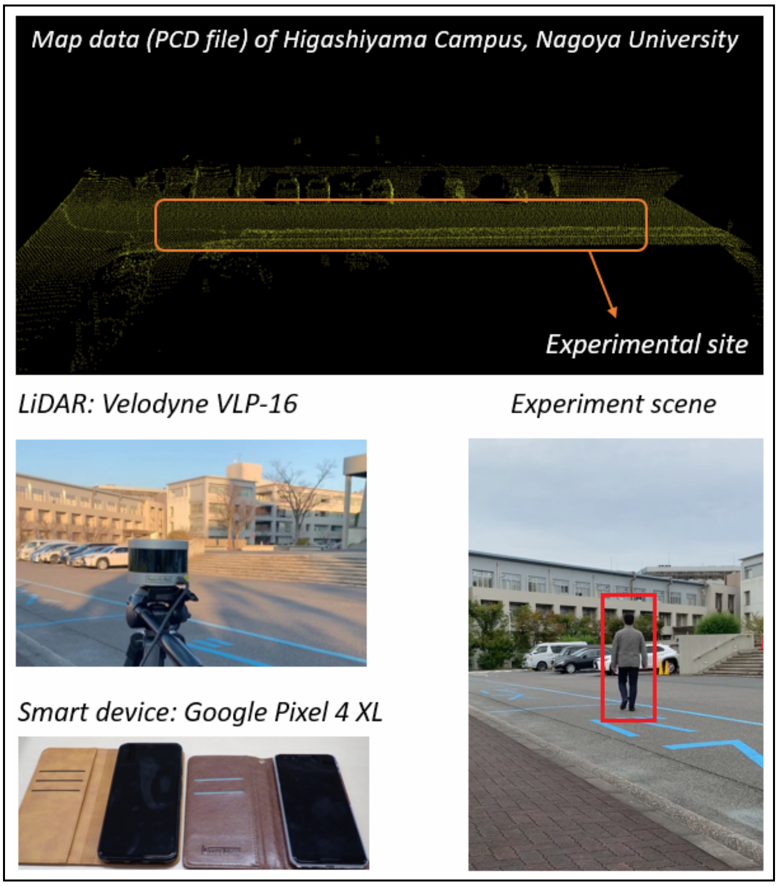


Figure 9 Experimental results of LiDAR cluster filtering (see online version for colours)

No filters	*Filter 1, 2, 3, 4 and all*

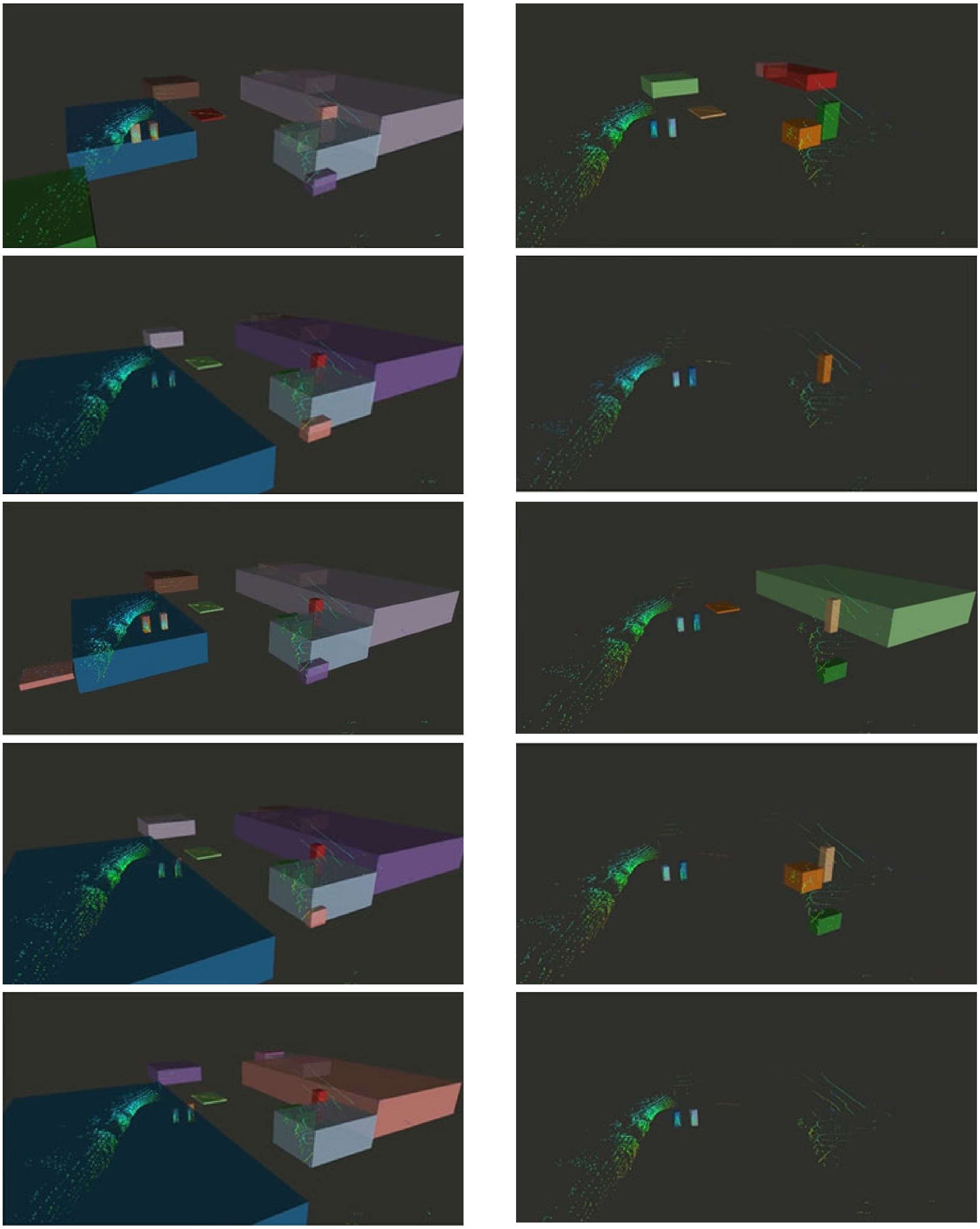


Table 3 Evaluation of cluster filtration for pedestrian detection

Filter ID	*1*	*2*	*3*	*4*	*All filters*
No. of pedestrians	24	23	22	23	19
No. of detected pedestrians	21	22	21	22	16
No. of detected obstacles	75	7	73	43	0
Correct detection rate	87.5%	95.7%	95.5%	95.7%	84.2%
Missed detection rate	12.5%	4.3%	4.5%	4.3%	15.8%
False detection rate	312.5%	30.4%	331.8%	187.0%	0.0%

Figure 10 Pedestrian matching accuracy (see online version for colours)

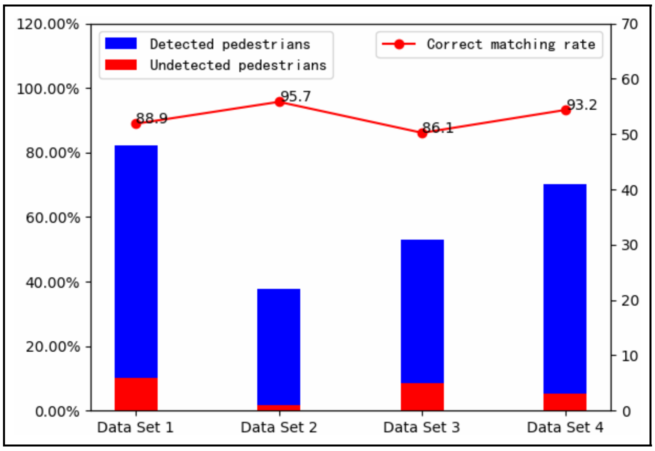
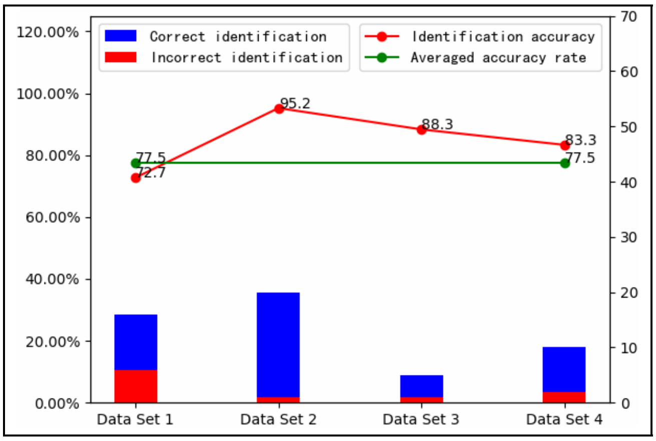


Figure 11 Single person identification accuracy (see online version for colours)



Note: Dataset 2 is collected for single pedestrian detection, so experimental results of this dataset are not included in the average accuracy rate.

The results of single person identification are presented in Figure 11. The results show that when pedestrians are close to each other (dataset 1), it is difficult to identify them. This is because the method to identify a single pedestrian is to compare and match their speed and course; when they walk side by side, the speed and direction tend to be the same, which makes it difficult to distinguish the pedestrians from each other. In this case, more personal feature information (e.g., height, body size) is needed to identify pedestrians.

7 Conclusions and future work

7.1 Conclusions

The paper proposed and evaluated a system to extract pedestrian high-precision position and attribute information using LiDAR and built-in sensors of smart devices. First, we used ROS to subscribe to and obtain raw LiDAR point cloud data, GPS, and IMU data. Subsequently, we designed several ROS nodes to realise the coordinate transformation and clustering of LiDAR and GPS data. We also developed a node that matches the pre-processed LiDAR and GPS data and outputs the coordinates and probability of a pedestrian's correct identification and location. Moreover, to determine the identity of a single pedestrian, we used GPS&IMU sensor fusion to capture the speed and course of the pedestrian to improve the accuracy of matching. Finally, we conducted an experiment to confirm whether the proposed system can detect pedestrians and analysed the experimental data. From the results, matching is proved to be possible. Furthermore, because matching accuracy may decline due to the existence of obstacles other than pedestrians, we designed four human feature filters to exclude non-pedestrian clusters from the matched objects, and the matching accuracy increased. We summarise the characteristics of the system as follows:

- PAIDS calculates the probability that a LiDAR cluster is a GPS source by integrating probability density function
- human feature filters are designed and used in PAIDS to extract pedestrians from LiDAR clusters
- sensor fusion is implemented for accurate position detection.

Since the result of literature survey shows that there are no existing approaches to extract pedestrian position and attribute information at the same time, we cannot compare our system with similar methods. However, the experimental results indicate that our system can be put into operation.

7.2 Limitations and future work

Although matching accuracy improved by applying filters related to cluster size, shape, and discrete condition, the miss detection rate also increased as a consequence of applying the filters. This implies that the threshold values are strictly defined, and sometimes pedestrians are incorrectly excluded. To improve the correct detection rate and reduce the false detection rate of the filters, further tuning and refining the threshold value of the filter are necessary. Obtaining the attribute information of the target pedestrian from the DM2.0 platform beforehand and using it in the filters may also be a good idea to identify specific pedestrians.

In addition, when the distance between the LiDAR and pedestrians is too close or too far, when pedestrians are close to obstacles, when pedestrians are in a blind-spot of LiDAR, or when pedestrians gather together, missing detection occurs. Furthermore, in situations with a large number of pedestrians, or if the GPS error is large, the proposed system may not be able to match pedestrians correctly. Because the traffic conditions of the university are relatively simple, conducting experiments under more complex traffic conditions is necessary. Further experiments to evaluate the proposed system and improve matching accuracy are the future directions of our research.

At present, pedestrians are using the ROS-all-sensors application (available in android application market) to communicate with ROS. The application can only send device ID and sensor data. System administrators have to register user's device ID and attribute information into the database of DM2.0 platform via a dedicated user management system. Therefore, developing applications that can directly send sensor information and detailed user information is also one of our tasks in the future.

Acknowledgements

This work was partially supported by JST, COI, JPMJCE1317 and JST, OPERA, JPMJOP1612.

References

Chavez-Garcia, R.O. and Aycard, O. (2016) 'Multiple sensor fusion and classification for moving object detection and tracking', *IEEE Transactions on Intelligent Transportation Systems*, Vol. 17, No. 2, pp.525–534, DOI: 10.1109/TITS.2015.2479925.

Cho, H., Seo, Y., Kumar, B.V.K.V. and Rajkumar, R.R. (2014) 'A multi-sensor fusion system for moving object detection and tracking in urban driving environments', *2014 IEEE International Conference on Robotics and Automation (ICRA)*.

Geospatial Information Authority of Japan (2013) *Converting Latitude and Longitude to Obtain Japanese Plane Rectangular Coordinate* [online] https://vldb.gsi.go.jp/sokuchi/surveycalc/main.html.

Ghosh, S., Amon, P., Hutter, A. and Kaup, A. (2017) 'Detecting closely spaced and occluded pedestrians using specialized deep models for counting', Paper presented at the *2017 IEEE Visual Communications and Image Processing (VCIP)*.

Gruden, C., Otković, I.I. and Šraml, M. (2021) 'Pedestrian safety at roundabouts: their crossing and glance behavior in the interaction with vehicular traffic', *Accident Analysis & Prevention*, Vol. 159, No. 1, p.106290.

Han, B., Wang, Y., Yang, Z. and Gao, X. (2020) 'Small-scale pedestrian detection based on deep neural network', *IEEE Transactions on Intelligent Transportation Systems*, Vol. 21, No. 7, pp.3046–3055, DOI: 10.1109/TITS.2019.2923752.

Hariyono, J., Hoang, V., and Jo, K. (2014) 'Location classification of detected pedestrian', Paper presented at the *2014 14th International Conference on Control, Automation and Systems (ICCAS 2014)*.

Himmelsbach, M., Hundelshausen, F.V. and Wuensche, H. (2010) 'Fast segmentation of 3D point clouds for ground vehicles'. Paper presented at the *2010 IEEE Intelligent Vehicles Symposium*.

Jang, J.A. and Lee, H.M. (2017) 'An analysis system of pedestrian-vehicle interaction risk level using drone videos', Paper presented at the *2017 International Conference on Information and Communication Technology Convergence (ICTC)*.

Jin, C., Cui, X., Woo, T. and Kim, H. (2016) 'Method for pedestrian detection using ground plane constraint based on vision sensor'. Paper presented at the 2016 *International Conference on Electronics, Information, and Communications (ICEIC)*.

Junejo, I.N. and Ahmed, N. (2020) 'A multi-branch separable convolution neural network for pedestrian attribute recognition', *Heliyon*, Vol. 6, No. 3, p.e03563.

Kang, W. and Han, Y. (2015) 'SmartPDR: smartphone-based pedestrian dead reckoning for indoor localization', *IEEE Sensors Journal*, Vol. 15, No. 5, pp.2906–2916, DOI: 10.1109/JSEN.2014.2382568.

Kwon, S.K., Hyun, E., Lee, J., Lee, J., and Son, S.H. (2016) 'A low-complexity scheme for partially occluded pedestrian detection using LIDAR-RADAR sensor fusion', Paper presented at the *2016 IEEE 22nd International Conference on Embedded and Real-Time Computing Systems and Applications (RTCSA)*.

Kwon, S.K., Son, S.H., Hyun, E., Lee, J. and Lee, J. (2017). 'Radar-Lidar sensor fusion scheme using occluded depth generation for pedestrian detection', Paper presented at the *2017 International Conference on Computational Science and Computational Intelligence (CSCI)*.

Lahmyed, R. and Ansari, M.E. (2016) 'Multisensors-based pedestrian detection system', Paper presented at the *2016 IEEE/ACS 13th International Conference of Computer Systems and Applications (AICCSA)*.

Li, Y., Xue, F., Fan, X., Qu, Z. and Zhou, G. (2018) 'Pedestrian walking safety system based on smartphone built-in sensors'. *IET Communications*, Vol. 12, DOI: 10.1049/iet-com.2017.0502.

Lin, B. and Lin, C. (2016) 'Pedestrian detection by fusing 3D points and color images', Paper presented at the *2016 IEEE/ACIS 15th International Conference on Computer and Information Science (ICIS)*.

Lv, B., Sun, R., Zhang, H., Xu, H. and Yue, R. (2019) 'Automatic vehicle-pedestrian conflict identification with trajectories of road users extracted from roadside LiDAR sensors using a rule-based method', *IEEE Access*, Vol. 7, pp.161594–161606, DOI: 10.1109/ACCESS.2019.2951763.

Matti, D., Ekenel, H.K. and Thiran, J. (2017) 'Combining LiDAR space clustering and convolutional neural networks for pedestrian detection', Paper presented at the *2017 14th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS)*.

Nauth, P.M., Pech, A.H., and Michalik, R. (2019) 'Research on a new smart pedestrian detection sensor for vehicles', Paper presented at the *2019 IEEE Sensors Applications Symposium (SAS)*.

Navarro, P.J., Fernández, C., Borraz, R. and Alonso, D. (2017) 'A machine learning approach to pedestrian detection for autonomous vehicles using high-definition 3D range data', *Sensors*, Vol. 17, No. 1, DOI: 10.3390/s17010018.

Pottier, B., Rasolofondraibe, L. and Kerroumi, S. (2017) 'Pedestrian detection strategy in urban area: capacitance probes and pedestrians' signature', *IEEE Sensors Journal*, Vol. 17, No. 17, pp.5663–5668, DOI: 10.1109/JSEN.2017.2718734.

Premebida, C., Ludwig, O. and Nunes, U. (2009) 'Exploiting lidar-based features on pedestrian detection in urban scenarios', *12th International IEEE Conference on Intelligent Transportation Systems*.

Premebida, C., Monteiro, G., Nunes, U. and Peixoto, P. (2007) 'A lidar and vision-based approach for pedestrian and vehicle detection and tracking', Paper presented at the *2007 IEEE Intelligent Transportation Systems Conference*.

Shimada, H., Yamaguchi, A., Takada, H. and Sato, K. (2015) ''Implementation and evaluation of local dynamic map in safety driving systems', *Journal of Transportation Technologies*, Vol. 5, pp.102–112, DOI: 10.4236/jtts.2015.52010.

Shin, S., Ahn, D. and Lee, H. (2017) 'Occlusion handling and track management method of high-level sensor fusion for robust pedestrian tracking', *Paper presented at the 2017 IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI)*.

Sola, J. (2017) *Quaternion Kinematics for the Error-State Kalman Filter*, arXiv preprint arXiv:1711.02508.

Tung, Y. and Shin, K.G. (2018) 'Use of phone sensors to enhance distracted pedestrians safety', *IEEE Transactions on Mobile Computing*, Vol. 17, No. 6, pp.1469–1482, DOI: 10.1109/TMC.2017.2764909.

Ulak, M.B., Kocatepe, A., Yazici, A., Ozguven, E.E. and Kumar, A. (2021) 'A stop safety index to address pedestrian safety around bus stops', *Safety Science*, Vol. 133, p.105017.

Wang, Y., Guo, S. and Huang, H. (2017) 'The pedestrian detecting and counting system based on automatic method of CCD calibration'. Paper presented at the *2017 9th International Conference on Advanced Infocomm Technology (ICAIT)*.

Wong, P.K-Y., Luo, H., Wang, M., Leung, P.H. and Cheng, J.C.P. (2021) 'Recognition of pedestrian trajectories and attributes with computer vision and deep learning techniques', *Advanced Engineering Informatics*, Vol. 49, p.101356.

Wu, J., Tamura, Y., Wang, Y., Woo, H., Moro, A., Yamashita, A. and Asama, H. (2020) 'Smartphone zombie detection from LiDAR point cloud for mobile robot safety', *IEEE Robotics and Automation Letters*, Vol. 5, No. 2, pp.2256–2263, DOI: 10.1109/LRA.2020.2970570.

Wu, T., Tsai, C., and Guo, J. (2017) 'LiDAR/camera sensor fusion technology for pedestrian detection', Paper presented at the *2017 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC)*.

Yoshida, H., Hoshina, A., Nakano, M. and Sugaya, M. (2015) 'Collision detection for bicycle and pedestrian exchange GPS location in smartphone', Paper presented at the *Adjunct Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2015 ACM International Symposium on Wearable Computers*, Osaka, Japan.

Zeng, Q., Wang, J., Meng, Q., Zhang, X. and Zeng, S. (2018) 'Seamless pedestrian navigation methodology optimized for indoor/outdoor detection', *IEEE Sensors Journal*, Vol. 18, No. 1, pp.363–374, DOI: 10.1109/JSEN.2017.2764509.

Zhang, J., Xiao, J., Zhou, C. and Peng, C. (2018) 'A multi-class pedestrian detection network for distorted pedestrians', Paper presented at the *2018 13th IEEE Conference on Industrial Electronics and Applications (ICIEA)*.

Zhao, J., Li, Y., Xu, H. and Liu, H. (2019) 'Probabilistic prediction of pedestrian crossing intention using roadside LiDAR data', *IEEE Access*, Vol. 7, pp.93781–93790, DOI: 10.1109/ACCESS.2019.2927889.

Zhou, B., Li, Q., Mao, Q., Tu, W. and Zhang, X. (2015) 'Activity sequence-based indoor pedestrian localization using smartphones'. *IEEE Transactions on Human-Machine Systems*, Vol. 45, No. 5, pp.562–574, DOI: 10.1109/THMS.2014.2368092.

Zhou, Z., Kitamura, S., Watanabe, Y., Yamada, S. and Takada, H. (2021) 'Extraction of pedestrian position and attribute information based on the integration of LiDAR and smartphone sensors', Paper presented at the *2021 IEEE International Conference on Mechatronics and Automation (ICMA)*.

Zhu, H., Almukdad, A., Iryo-Asano, M., Alhajyaseen, W.K.M., Nakamura, H. and Zhang, X. (2021) 'A novel agent-based framework for evaluating pedestrian safety at unsignalized mid-block crosswalks', *Accident Analysis & Prevention*, Vol. 159, p.106288 [online] https://doi.org/10.1016/j.aap.2021.106288.

Appendix

Geospatial Information Authority of Japan (2013) provides the method of computation to convert longitude and latitude to Japanese plane rectangular coordinate.

Input φ: x-coordinate (longitude), λ: y-coordinate (latitude).

Output x – x-coordinate (plane rectangular coordinate), y – y-coordinate (plane rectangular coordinate).

$$x = \overline{A}\left(\xi' + \sum_{j=1}^{5} \alpha_j \sin 2j\xi' \cosh 2j\eta'\right) - \overline{S}_{\varphi 0}$$

$$y = \overline{A}\left(\eta' + \sum_{j=1}^{5} \alpha \cos 2j\xi' \sinh 2j\eta'\right)$$

where

$$n = \frac{1}{2F - 1}$$

$$t = \sinh\left(\tanh^{-1} \sin\varphi - \frac{2\sqrt{n}}{1+n} \tanh^{-1}\left[\frac{2\sqrt{n}}{1+n} \sin\varphi\right]\right)$$

$$\overline{t} = \sqrt{1+t^2}$$

$$\xi' = \tan^{-1}\left(\frac{t}{\lambda_c}\right), \eta' = \tanh^{-1}\left(\frac{\lambda_s}{\overline{t}}\right)$$

$$\sigma = 1 + \sum_{j=1}^{5} 2j\alpha_j \cos 2j\xi' \cosh 2j\eta',$$

$$\tau = \sum_{j=1}^{5} 2j\alpha_j \cos 2j\xi' \cosh 2j\eta'$$

$$\alpha_1 = \frac{1}{2}n - \frac{2}{3}n^2 + \frac{5}{16}n^3 + \frac{41}{180}n^4 - \frac{127}{288}n^5$$

$$\alpha_2 = \frac{13}{48}n^2 - \frac{3}{5}n^3 + \frac{557}{1{,}440}n^4 + \frac{281}{630}n^5$$

$$\alpha_3 = \frac{61}{240}n^3 - \frac{103}{140}n^4 + \frac{15{,}061}{26{,}880}n^5$$

$$\alpha_4 = \frac{49{,}561}{161{,}280}n^4 - \frac{179}{168}n^5, \quad \alpha_5 = \frac{34{,}729}{80{,}640}n^5$$

$$\bar{S}_{\varphi_0} = \frac{m_0 a}{1+n}\left(A_0 \frac{\varphi_0}{\rho''} + \sum_{j=1}^{5} A_j \sin 2j\varphi_0 \right), \quad \bar{A} = \frac{m_0 a}{1+n} A_0$$

$$A_0 = 1 + \frac{n^2}{4} + \frac{n^4}{64}, \quad A_1 = -\frac{3}{2}\left(n - \frac{n^3}{8} - \frac{n^5}{64} \right)$$

$$A_2 = \frac{15}{16}\left(n^2 - \frac{n^4}{4} \right)$$

$$A_3 = -\frac{35}{48}\left(n^3 - \frac{5}{16}n^5 \right), \quad A_4 = \frac{315}{512}n^4, \quad A_5 = \frac{693}{1{,}280}n^5$$

φ_0, λ_0 are latitude value and longitude value of the origin of plane rectangular coordinate, respectively. F is the reciprocal flattening, and a is the semimajor axis length of the earth. m_0 is the scale factor on the x axis of the plane rectangular coordinate.