Human Interaction and Emerging Technologies (IHIET 2024), Vol. 157, 2024, 528–534
https://doi.org/10.54941/ahfe1005512

Human Detection Method by 3D-LiDAR With Low Calculation Costs

Haruki Mochizuki[1] and Ryozo Kiyoharara[2]

[1]Graduate School of Kanagawa Institute of Technology, Atsugi, Kanagawa, Japan
[2]Kanagawa Institute of Technology, Atsugi, Kanagawa, Japan

ABSTRACT

Robot technologies are developing and it is becoming possible to perform not only simple tasks but also tasks that involve interaction with humans. In particular, robots are expected to be introduced in patrol security, which requires a large number of personnel, as much of the work involves confirming that there are no problems. In Japan, for example, robots that serve food at restaurants move around in the flow of people. In terms of security, if it cannot be confirmed by robots that there is no problem, a security guard is required to go and check. In order to reduce such cases, improving the accuracy of confirmation is an absolute requirement. In addition, it is necessary to reduce the cost of introducing and operating the robot as much as possible. Technology for recognizing people using cameras mounted on robots and 3D-LiDAR (Light Detection And Raging) has already been established and is in use. However, no technology has been established for use in poorly lit areas or by low-cost 2D-LiDAR recognition. In this paper, we propose a new method for reducing calculation cost and evaluate our method.

Keywords: Human detection, 3D-LiDAR, Security robots

INTRODUCTION

In many developed countries, aging populations and slow immigration are predicted to result in labor shortages. Labor shortages are already a problem in Japan. In the security industry, robots are being introduced for simple patrol tasks. Patrol task is primarily crime deterrent. If suspicious humans are found, its information have to be reported. Robot carts that can freely move outdoors at night are equipped with 3D-LiDAR (Light Detection And Raging) in many cases. Moreover, they move while estimating its own position using SLAM (Simultaneous Localization and Mapping) technologies.

SLAM with 3D-LiDAR requires many calculation costs. Therefore, human detection processing should have as little computational cost as possible. Therefore, we propose a human recognition method that uses only part of the pointcloud data obtained from the on-board 3D-LiDAR as data obtained from multiple 2D-LiDARs. If 3D-LiDAR collects pointcloud data from 16 different heights, using only one can reduce the number of pointcloud data to 1/16. Furthermore, the computational cost of machine learning evaluation is also reduced. Therefore, we evaluate the validity of the proposed method by evaluating the person recognition rate using multiple 2D-LiDARs. Moreover, by evaluating the learning required for the human orientation and distance from LiDAR, we evaluate whether it is possible to recognize people in any orientation or distance.

LIDAR

LiDAR emits laser light and measures distance based on the phase difference of the reflected light. By continuing to rotate the irradiation, pointcloud data can be collected. LiDAR is installed on the Robot Cart as shown in Figure 1. Robot Cart estimate the location by SLAM with Pointcloud data to move.

Figure 1: Robot cart with 3D-LiDAR as a security guard.

RELATED STUDIES

Many studies were conducted regarding human recognition. Human recognition using image processing is one of the most commonly used methods. However, Robot as a security guard moves in midnight. Therefore, there are not enough light. There are some studies on human recognition with 3D-LiDAR. Pedestrian recognition using 3D-LiDAR for autonomous vehicles (Nagai, 2022) uses a support vector machine (SVM) for human recognition. Other approaches use 3D LiDARs as multiple 2D- LiDARs (Chung, 2012). This study gathered pointcloud data from many different heights and achieved a classification rate of up to 96%. An example of using reduced pointcloud data is dividing 3D pointcloud data in the target into parts, that is, the trunk and legs, to achieve robust in situ pedestrian recognition (Asad, 2020).

All of these studies use 3D pointcloud data, which requires a lot of CPU resources. 3D-LiDAR can obtain a pointcloud of the entire person. Therefore, it can identify the person regardless of the shape of the clothes or other factors. However, there is no need to have multiple person shapes that can be obtained from similar height locations. Therefore, in this study, we propose a person recognition method that uses multiple 2D-LiDARs to reduce computational costs.

In our previous study (Nagai, 2022), we proposed a person detection method based on low-resolution 2D-LiDAR and a PointNet-based autoencoder (Qi, 2017). We used 2D-LiDAR to collect data from both the front and side from a height of 90 cm, which is considered the typical waist height of a human. The proposed method successfully detected people within 5 m of the LiDAR.

In a previous study (Watanabe, 2023), we proposed the human recognition method by multiple 2D-LiDARs. However, both methods require to recognize from exact front. Moreover in a previous study (Mochizuki, 2024), we showed possibility of our method. In this paper, we propose a method for recognizing people regardless of the sensing angle and distance.

PROPOSED METHOD

Based on the case at Kanagawa Institute of Technology, human to be detected by patrol security is an elderly human who enters the campus without malicious intent. By focusing on human between 135 cm and 175 cm in Japan, which is the average height of elderly people by age, it is possible to detect most of the elderly human who intrude without malicious intent. Human characteristics are broadly classified into head, torso, arms, and legs. The head has a smaller proportion on the body than the torso, arms, and legs. Moreover the height used to obtain pointcloud data about the head tends to change depending on the height. Therefore, we obtain pointcloud data about human torso, arms, and legs. The pointcloud data is acquired by setting the range in which the torso and arms can be acquired in the height range of 135 cm to 175 cm, with the torso and arms being 80 cm to 110 cm in height, and the legs being 20 cm to 55 cm in height.

In this paper, limited number of lasers irradiated by 3D-LiDAR can reduce the amount of pointcloud data to obtain for reducing calculation cost. The laser is selected by distance from obstacles and installed height. Therefore, the distance should be less than approximately 25 m to select more than three rays which are for arms, torso and legs. Moreover, if the distance is more than approximately 37 m, it is impossible to get pointcloud data for legs. Therefore, we assume that the robot only recognizes people within a distance of 25m (see Table 1).

Table 1. Distance and height relations for 3D LiDAR(height : cm).

Laser No.	2 m	5 m	15 m	25 m	37 m
7	89.5	126.3	249.1	378.1	525.4
5	82.5	108.7	196.2	288.1	393.0
3	75.4	91.2	143.6	198.6	261.5
1	68.4	73.7	91.1	109.5	130.4
–1	61.5	56.27	38.8	20.4	0.00
–3	54.5	38.8	0.0	0.0	0.0
–5	47.5	21.2	0.0	0.0	0.0

When a human is detected, the security robot notifies the security guard. Figure 2 shows the recognition process used in patrol security to notify security guards when a person is detected.

First, a robot equipped with LiDAR and a computer patrols and collects pointcloud data of the surrounding area. Background subtraction processing is performed on the computer equipped with the acquired pointcloud data.

Next, the extracted pointcloud data of obstacles such as people and objects are sent to the server through the wireless network, and the server side uses PointNet to identify human or obstacle. If the target is recognized as a human, a security guard is notified, the security guard heads to the robot's location, and the robot tracks the human.

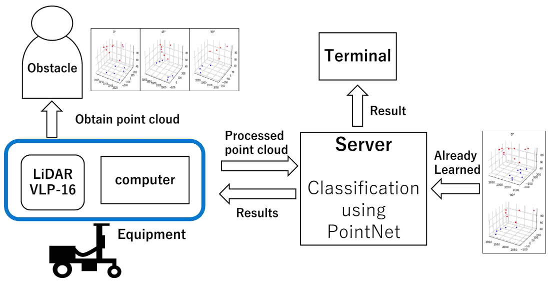

Figure 2: Notification flow for security guard when human is detected.

On the server, in order to perform human recognition using the acquired pointcloud data, they are classified using machine learning technologies. To build a system for classification, use "Classify pointcloud using PointNet," which is published in the MATLAB library. This is a method that uses PointNet to compress and restore pointcloud data and extract only the features for classification.

Classification using PointNet is supervised learning, so it is necessary to learn pointcloud data to be classified in advance. Therefore, it is learned pointcloud data about humans and trees, and classify the acquired pointcloud data into humans and trees. Since the number and shape of pointcloud data differ depending on the size of the target object, learning is performed by acquiring point cloud information of various types of humans and trees. Furthermore, the number and shape of pointcloud data that can be acquired differs depending on the orientation of the person. When humans learn one direction, it is difficult to recognize directions that have not been learned.

In the case of front facing, pointcloud data of two human legs are obtained. Pointcloud data for the two legs are obtained in the same way even in orientations other than landscape. Therefore, by learning the front facing direction, it is possible to recognize human with a certain degree of accuracy in situations other than those facing sideways. Moreover, since the learned orientation can be determined in the case of landscape orientation, it can be determined by learning the landscape orientation.

Therefore, we think that by learning two directions, frontal and sideways, the range of human orientations that can be determined will be increased, making it possible to identify human even in a variety of orientations.

EVALUATION

Table 2 shows the specification of the LiDAR for our evaluation.

Table 2. Specification of the LiDAR for our evaluation.

Name	VLP-16 (by Velodyne)
Range	100m
View Angle (Vertical)	30°(±15°)
View Angle (Horizontal)	360°
Rotational Speed	5Hz-20Hz

First, the number of pointcloud data have been compared between 16 lasers and 2 lasers for standing human at various distance. Moreover, we also have been compared the number of pointcloud data acquired when using two 3D-LiDAR lasers and when using two 2D-LiDAR for standing human at various distance. We have been able to confirm to reduce the amount of pointcloud data.

Next, accuracy of human recognition is most important. Then, we have confirmed the accuracy of human recognition by two lasers of 3D-LiDAR. The distance range of pointcloud data to be acquired has been set from 2 m to 5 m, and training data and test data have been acquired separately. In the training data, pointcloud data for humans and trees have been used to prepare pointcloud data for each class classification. For the human pointcloud data in the training data, we have prepared 20 pieces of tree pointcloud data, 10 pieces each of frontal and sideways pointcloud data as the two orientations to be learned. Moreover, pointcloud data of humans and trees acquired from a distance of 3 m from LiDAR has been used as training data. The number of epoch has been 150 for each. As test data, we have prepared16 pieces of pointcloud data of humans facing forward (0°), diagonally (45°), sideways (90°), and trees, acquired at distances of 1 m each from 2 m to 5 m. Table 3 shows the F value of the result of learning the training data in advance and classifying it.

Table 3. F Value by limited 3D LiDAR for our evaluation.

	0°	45°	90°
2m	1.00	1.00	1.00
3m	1.00	1.00	1.00
4m	1.00	1.00	1.00
5m	1.00	1.00	1.00

From this result, sufficient accuracy has been able to be obtained when acquired using 2 lasers in 3D-LiDAR.

We have confirmed whether 3D-LiDAR has been able to measure human torso, arms, and legs, and has been able to recognize human in various orientations at a maximum measurement distance of 25 m. The distance range of the pointcloud data to be acquired has been set from 5 m to

25 m, and training data and test data have been acquired separately. For training data, we have prepared pointcloud data for humans and trees for class classification. For the training data, we have prepared 10 pieces of human pointcloud data for each of the two orientations to be learned, frontal and sideways, and 20 pieces of tree point cloud data. Moreover, pointcloud data of humans and trees acquired from a distance of 3 m from LiDAR has been used as training data. The number of epoch have been 150. For the test data, we have prepared16 points each of pointcloud data of a human facing forward, diagonally, sideways, and a tree, taken at 5 m intervals from a distance of 5 m to 25 m. Table 4 shows the F value of the result of learning the training data in advance and classifying them.

Table 4. F Value by distance for our evaluation.

	0°	45°	90°
5m	1.00	1.00	1.00
10m	0.91	0.91	0.00
15m	1.00	0.67	0.75
20m	0.91	0.67	0.91

According to the recognition results, for diagonal orientations, the accuracy decreased to 0.67 over 15 m. Moreover, although it has been possible to distinguish between humans and trees at other distances when facing sideways, it has not been possible to distinguish between humans and trees at 10 m. We think that the cause has been the low rotation speed and the small number of acquired pointcloud data, which led to a decrease in accuracy. In other words, the pointcloud data used for training data has been too little and it has not been possible to distinguish between landscape orientation and tree features. Moreover, because the number of test data has been too little, it is thought that the shape shown by the pointcloud data are biased. Therefore, it necessary to find an appropriate amount of data.

CONCLUSION

In order to relieve the burden of patrolling security at universities due to the labor shortage problem in the security industry, we have verified the human recognition rate in various orientations using a security robot equipped with 3D-LiDAR. We have proposed a method that have been able to reduce the calculation cost and recognize human with high accuracy by acquiring fewer pointcloud data than the original number for person recognition using 3D-LiDAR.We have showed that by selecting two lasers, it was possible to sufficiently recognize human even in situations where the amount of pointcloud data is little. However, the accuracy have not been stable when the distance was large, so it has been necessary to acquire more point cloud data and perform classification.

In the future work, the influence of clothing and human posture on recognition accuracy should be verified. In particular, jackets and skirts can worsen recognition accuracy.

REFERENCES

Asad, L. Ehsan, N. Luis F. M-M. (2020) “Development and evaluation of a real-time pedestrian counting system for high-volume conditions based on 2D LiDAR,” Transportation Research Part C: Emerging Technologies, Vol. 114, pp. 20–35.

Chung, W. Kim, H. Yoo, Y. Moon, C. –B. and Park, J. (2012) “The Detection and Following of Human Legs Through Inductive Approaches for a Mobile Robot With a Single Laser Range Finder,” IEEE Trans. on Industrial Electronics, vol. 59, no. 8, pp. 3156–3166.

Mochizuki, H. Watanabe, T. Horii, M. and Kiyohara R. (2024),“Stationary Human Detection Method by 2D-LiDARs,” IEEE International Conference on Information Networking (ICOIN).

Nagai, Y. Sawano, Y. Terashima, Y. Suzuki, T. and Kiyohara, R. (2022). “A Method for Detecting Human by 2D-LiDAR,” IEEE International Conference on Consumer Electronics (ICCE).

Qi, R. C. Su, H. Mo, K. Guibas, J. L. (2017). “PointNet: Deep Learning on Point Sets for 3D Classification and Segmentation,” IEEE Conference on Computer Vision and Pattern Recognition (CVPR).

Watanabe, T. Terashima Y. and Kiyohara R. (2023), “Standing Human Detection Method Using 2D-LiDARs,” IEEE 47th Annual Computers, Software, and Applications Conference (COMPSAC), pp. 1260–1265.