Time-series LIDAR Data Superimposition for Autonomous Driving

Shuhei AYUKAWA, Naruaki TOKUDOME, Shuichi ENOKIDA, and Takeshi NISHIDA

***Abstract*— LIDAR (light detection and ranging) sensors are important for autonomous driving because it is possible to obtain wide-ranging distance information via three-dimensional laser irradiation. However, because this sensor irradiates lasers radially, the quantity of measurement data obtained from distant objects decreases. Therefore, when an object recognition algorithm is applied to the data, the recognition ratio of distant objects declines. The performance of present LIDAR is not adequate to obtain the data for autonomous driving for car speeds of 60 (km/h) or greater. Thus, we propose two methods of superimposing time-series LIDAR data. In the first method, the posture transition of LIDAR during a sampling time is estimated and used for superimposition of several measuring frames. In the second method, target object extraction results in each measurement are used to merge several frames. We describe these methods in detail and show several experimental results.**

I. Introduction

The next-generation driving support system requires two functions: a function for predicting the risk associated with pedestrians or cyclists and an environment recognition function for the avoidance behavior. LIDAR (light detection and ranging) is widely used in environment recognition for autonomous driving. This sensor is capable of measuring the object distance and the reflection intensity of the measurement object. It is also capable of scanning a greater distance and wider range than other sensors. In addition, the performance degradation on the natural environment change is small. Several studies have been conducted on person recognition using LIDAR point cloud data [1, 2]. However, the problem of the reduction in recognition precision with increase in distance to the measured object persists.

In this research, we present a method that superimposes time-series LIDAR data, and we tried to improve the performance of various recognition algorithms using this method. The following describes the results of these validation experiments.

II. LIDAR

In this research, a LIDAR sensor, HDL-64E (Velodyne), was used. The overview is shown in Fig. 2. The HDL-64E can measure all distances in the horizontal plane. The vertical measurement region is from +2 (°) to -24.8 (°), and it is capable of laser irradiation rates of 1.33 million per second. Then, point sets of three-dimensional (3D) Euclidean space (point clouds) are acquired by the measurement of the LIDAR. Since this sensor irradiates lasers radially, as distance increases, the number of measurement data points obtained at a distant object decreases. Therefore when an object recognition algorithm is applied to the data, the recognition ratio of distant objects declines.

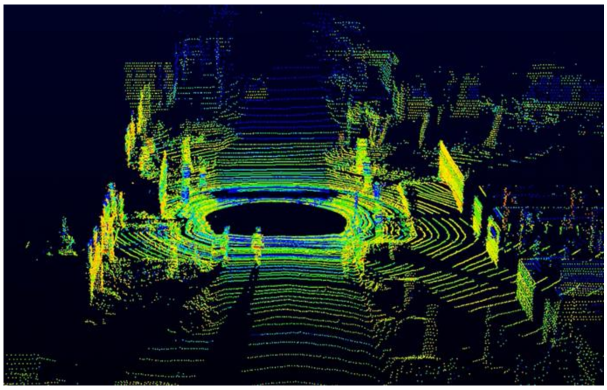


Fig. 1. Example of LIDAR measurment of a road environment.

III. Superimposing

When the distance to the target is far, the density of laser spots becomes sparse (Fig. 3). For example, in the case of the person who stands at distance of 50 (m) from the LIDAR, the number of irradiated laser spots is only ten, as shown in Fig. 4. For pedestrian recognition by LIDAR measurement, more than 30 laser spots are required. This number is derived from the results of the recognition experiment and based on a pedestrian recognition method from research literature [1, 2].

In this paper, in order to increase the measurement and the recognition performance over long distance (about 50 (m)), we propose a superimposition method of multiple time-sequence measurement data. Strategies of the superimposition method can be divided broadly into two categories. The first one is a method of superimposing all collected data in multiple frames, while the second is a superimposing cut-out of only the measurement point cloud of distant targets. The former has high accuracy and high computational cost owing to the large

Fig. 2. Overview of Velodyne HDL-64E and a vehicle equipped with the LIDAR.

S. Ayukawa, N. Tokudome, and T. Nishida are with the Kyushu Institute of Technology, Kitakyushu, Fukuoka, Japan (corresponding author e-mail: nishida@cntl.kyutech.ac.jp).

S. Enokida is with the Kyushu Institute of Technology, Iizuka, Fukuoka, Japan.

number of data points being processed. On the other hand, the latter executes the convolution of only target objects, its accuracy is lower than the former, and it is possible to lower the number of calculations. This method can even account for pedestrians jumping out in front of a vehicle and other rapid changes in traffic conditions. Furthermore, the first method results in finding a rigid-coordinate transformation of the corresponding unknown 3D point cloud, while the second method has a problem detecting and tracking target objects. In this research, we constructed and evaluated an algorithm using the first method. In this method, target recognition and extraction is executed in each frame, and estimated movement is executed by the extracted results collated between the current and latest frames. Moreover, the superimposing of multiple frames is executed based on the estimated amount of movement and the results are forwarded to the target recognition function.

A. Estimation of the amount of the movement

The measurement of LIDAR fluctuates in each frame because the positional relationship between the LIDAR mounted on the car and environment continuously changes. Thus, an optimization matching algorithm for measured point clouds is required. In this research, we tried to estimate the amount of movement by using the iterative closest point (ICP) algorithm [3] and the normal distribution transform (NDT) algorithm [4]. The ICP algorithm is an optimization method of the rotation and translation transformation between the point clouds by an iterative procedure for reducing a cost function. The NDT algorithm executes probabilistic matching for measurement point clouds in multiple frames, including the rotation and translation matrix, which means that the amount of movement is estimated.

Examples of application results of these algorithms are shown in Fig. 6. In these experiments, we executed an offline matching procedure by using the frames Fig. 6 (a) and (b) measured by the HDL-64E. The car speed was 20 (km/h) as measured by speedometer, the sampling frequency of LIDAR was 10 (Hz), and the mounting position of the LIDAR was 1.7 (m) height. In this condition, the car travels 0.556 (m) during one sampling interval, and this value was used as the initial transition of the ICP and the NDT algorithms. The parameters

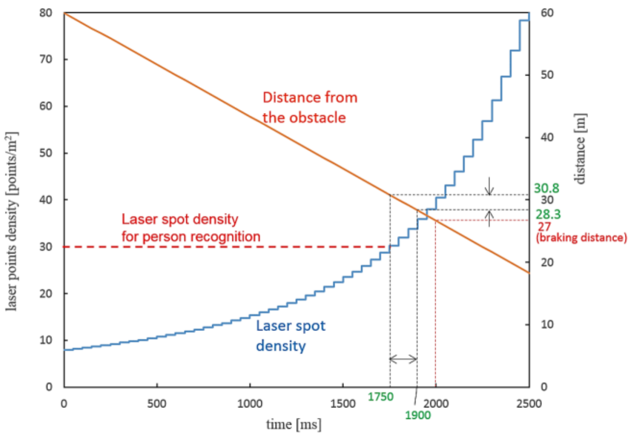


Fig. 3. Relationship of laser spot density and distance.

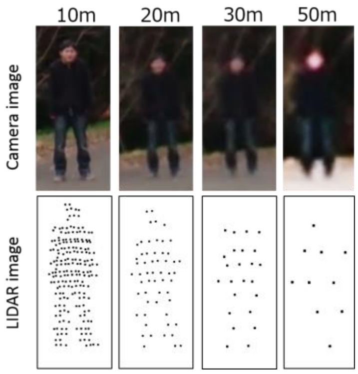


Fig. 4. Example of LIDAR measurment of a pedestrian.

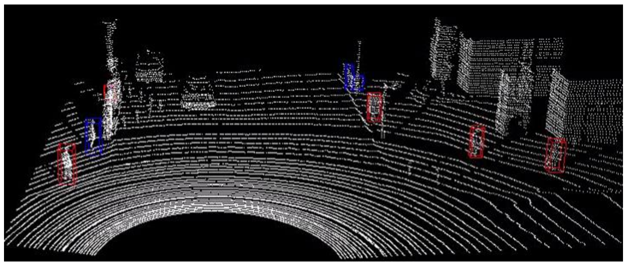


Fig. 5. Example of pedestrian recognitin by LIDAR measurment.

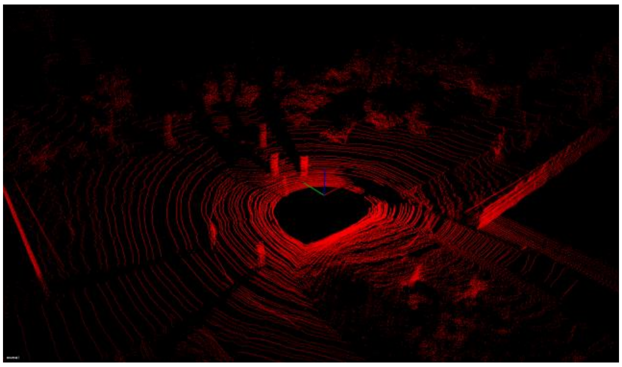
of the ICP algorithm were set as follows considering the balance of accuracy and processing speed: the maximum number of iterations was 10, the transformation epsilon was 0.001, the fitness epsilon was 0.00001, and the maximum correspondence distance was 1.0 (m). Moreover, the parameters of the NDT algorithm were set as follows: the grid resolution was 2.0 (m), the maximum number of iterations was 3, the fitness epsilon was 0.2, and the maximum step size was 1.0 (m). The cost function is defined as follows:

$$f(\boldsymbol{R}_k, \boldsymbol{t}_k) = \|(\boldsymbol{R}_k \boldsymbol{X}_k + \boldsymbol{t}_k) - \boldsymbol{X}_{k-1}\|_F / N$$

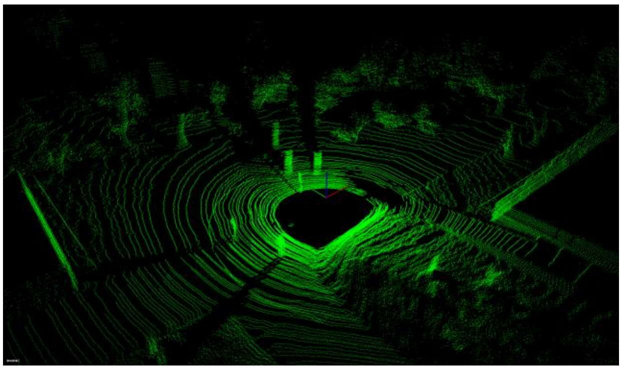
where $\boldsymbol{X}_k \triangleq [\boldsymbol{x}_k^1 \quad \cdots \quad \boldsymbol{x}_k^N]^T \in \mathbb{R}^{3\times N}$ is a matrix configured 3D point $\boldsymbol{x}_k^n$ at discrete time k, $\|\cdot\|_F$ is Frobenius norm, N is the number of measurement points, and $\boldsymbol{R}_k$ and $\boldsymbol{t}_k$ are the estimated rotation matrix and translation vector, respectively. A smaller cost corresponds to improved matching. The matching accuracy and processing time of these experiments are shown in Table I. It is found from these results that the ICP algorithm was better in the accuracy and the processing time than the NDT algorithm. However, these experimental results were obtained in an almost invariable environment. If the measured objects were moving, there is a possibility that the performance of the NDT algorithm would be better.

TABLE I. EXPERIMENTAL RESULT

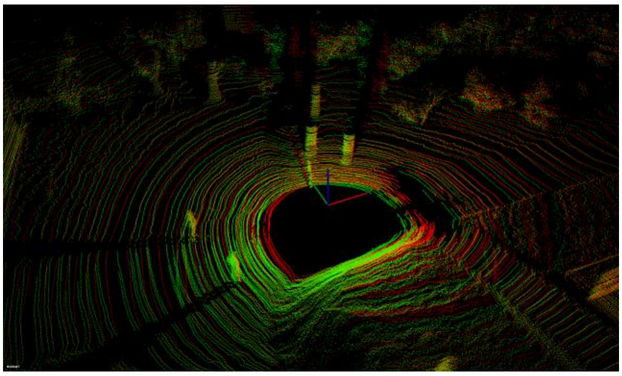
Algorithm	**Performance**	
	Cost value	***Processing time* [s]**
ICP	0.101	0.846
NDT	0.306	14.194



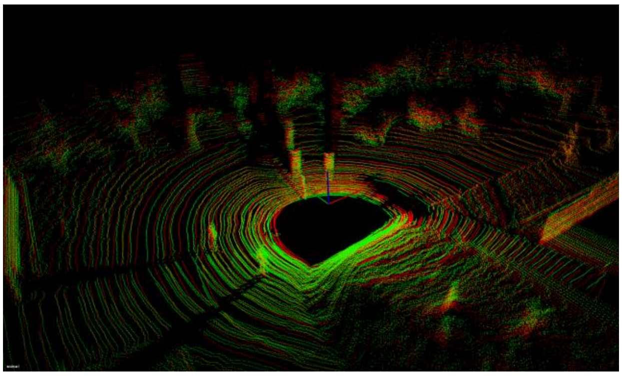
(a) 1st LIDAR image



(b) 2nd LIDAR image



(c) Superimposition by ICP



(d) Superimposition by NDT

Fig. 6 Example of superimposing of LIDAR measurment.

B. Extraction of person point cloud clusters

The processing flow of extraction and superimposition of pedestrian point cloud proposed in this paper is shown in Fig. 7. In this method, the movement estimation and pedestrian detection are executed in parallel in the point cloud obtained by LIDAR. In the movement estimation procedure, the pedestrian detection results of the last and current frame are integrated and the difference calculated. In the pedestrian detection procedure, the point cloud clustering is executed for their categorization. First, the measurement points on the road are removed. Next, multiple point clouds are obtained in isolation and the clustering process executed for them. Additionally, the voxel fitting process is executed. The values of width, height, and depth of the voxel are denoted as w (m), l (m), h (m) respectively. A voxel that satisfies the following conditions is determined as a pedestrian: $w \leq 1.0$ (m), $l \leq 1.0$ (m), and 0.45 (m) $\leq h \leq 2.0$ (m).

C. Identity determination and superimposing of point clouds

From the pedestrian detection results of consecutive k and $k+1$ frames, the voxel is determined to be the same person. The position of the center of gravity and the size of each side of the voxel were used for the evaluation of their identity. By the fluctuation of the measurement of LIDAR, the number of laser spots on the same pedestrian changes for each frame. Therefore, the size of voxels in each frames changes. Thus, allowable errors are set to compare voxels as follows:

$$h_{i,k}(1-\alpha) \leq h_{i,k+1} \leq h_{i,k}(1+\alpha)$$

$$w_{i,k}(1-\beta) \leq w_{i,k+1} \leq w_{i,k}(1+\beta)$$

where $h_{i,k}$ and $w_{i,k}$ represent the height and width of a voxel at time k, $\alpha = 0.27$, $\beta = 0.15$, and i is index of the voxel. By assuming the distance of pedestrian movement per sample, the distance is incorporated in these parameters. Voxels determined to represent the same pedestrian are superimposed by the coordinate transformation to align their centroid. The number of laser spots irradiated to the pedestrian can be increased by these processes. Next, the determination of the identity of the voxels in consecutive frames is carried out by the following condition,

$$\|\boldsymbol{x}_{k-1}^{c} - \boldsymbol{x}_{k}^{c} - \boldsymbol{l}_k\| \leq r$$

where $\boldsymbol{l}_k \triangleq [l_{x,k} \quad l_{y,k}]^T$ is the estimated vehicle movement, $\boldsymbol{x}_k^c \triangleq [x_k^c \quad y_k^c]^T$ is center of gravity coordinates of a voxel at k, and r is the range of movement of the pedestrian and the measurement error. Voxels nearest this expression when satisfied are regarded as identical.

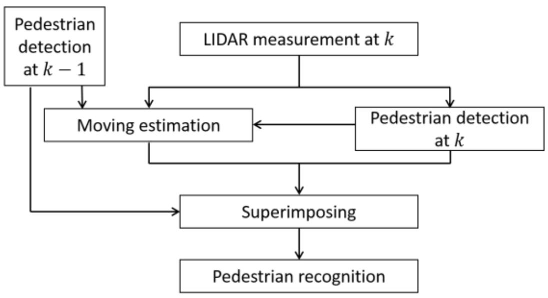


Fig. 7 Superimposing flow chart.

IV. Experiment

Several experiments were conducted by using measurement data of the LIDAR. In the experiment, the car had a LIDAR mounted on the roof and was traveling at a constant speed (20 (km/h)) as depicted in Fig. 8. The result of the superimposition is shown in Fig. 9, and the integrated point clouds are shown in Fig. 10.

In the case of a moving pedestrian, three frames were superimposed by the proposed method, and the number of data points increased by 2.4 times. For the same experiments conducted with a pedestrian walking 40 (m) to 60 (m) away from the LIDAR, the average number of data points was 9.94 in a voxel at single frame, and increased to 29.3 by superimposing three frames.

In the case of a stationary pedestrian, three frames were superimposed by the proposed method and the number of data points increased by 3.2 times. For the same experiments conducted with a stationary pedestrian 40 (m) to 60 (m) away from the LIDAR, the average number of data points was 10.9 in a voxel at single frame, and increased to 34.8 by superimposing three frames.

V. Conclusion

In this research, we proposed a superimposing method of multiple LIDAR measurements for increasing of the accuracy of pedestrian recognition. In this method, the clustered point clouds are extracted from LIDAR measurement data and voxel fitting is executed. Based on the voxel size, the pedestrian voxels are extracted and collated with identical ones in the previous frame. The point clouds included in the collated voxels are superimposed, and the movement estimation is executed using the collated result. It was confirmed by the experimental results that superimposing pedestrian point clouds in three frames can be stably executed. Moreover, the number of laser spots from the HDL-64E to a pedestrian 50 (m) away could be increased from 16 to 39 points.

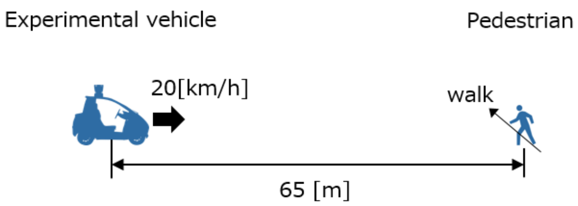


Fig. 8 Outline of the exprimental conditions.

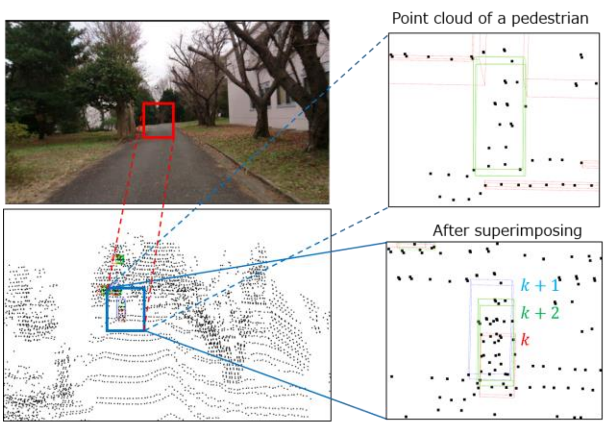
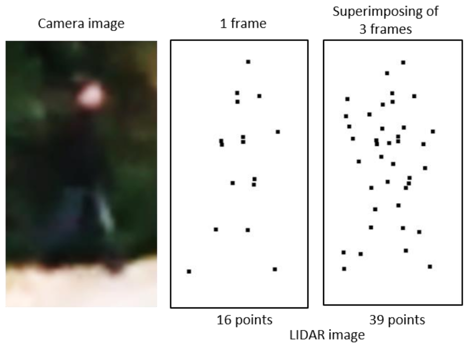
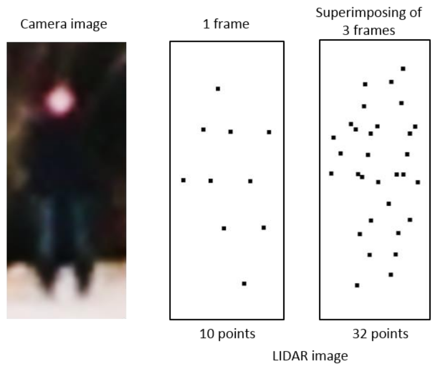


Fig. 9 Superimposing result for pedestrian detection.



(a) Walking pedestrian



(b) Standing pedestrian

Fig. 10 Superimposing result.

This research is part of the results of a project commissioned by the Ministry of Economy, Trade and Industry, “2015: The Next-generation Advanced Driver Assistance Systems Research and Development Demonstration.”

References

[1] H. Yokoyama, H. Date, S. Kanai, H. Takeda, “Detection and Classification of Pole-like Objects from Mobile Laser Scanning Data of Urban Environments”, Int. Jour. of CAD/CAM, Vol. 13, No. 2, pp. 31-40, 2013.

[2] K. Kidono, A. Watanabe, T. Naito, J. Miura, “Pedestrian Recognition Using High-definition LIDAR”, JRSJ, Vol. 29, No. 10, 2011.

[3] P. J. Besel, N. D. McKay, “A Method for Registration of 3-D Shapes”, IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. 2, No. 14, pp. 239–256, 1992.

[4] P. Biber, W. StraBer, “The Normal Distributions Transform: A New Approach to Laser Scan Matching”, Proc. of IROS, pp. 2743-2748, 2003.