Object recognition by a laser scanner using multimodal information

Shuji Oishi, Naoaki Kondo
Graduate School of Information Science
and Electrical Engineering
Kyushu University, Fukuoka, Japan

Yumi Iwashita, Ryo Kurazume
Graduate Faculty of Information Science
and Electrical Engineering
Kyushu University, Fukuoka, Japan

Abstract—This paper presents a new object recognition technique using multimodal information by a laser scanner. A ToF laser scanner captures a three dimensional (3D) shape of a target object precisely by measuring the round-trip time of a laser pulse. In addition, a laser scanner also collects the strength of the reflected laser (reflectivity) as side-product of range measurement. Focusing on the fact that the reflectivity contains rich appearance information of the target object, the proposed technique utilizes not only range but also reflectance information to improve the recognition performance. To evaluate the validity of the proposed technique, we applied it for the classification of moving objects on a road as a traffic census application.

I. INTRODUCTION

Laser scanners have been widely used for a variety of applications such as 3D digital modeling, obstacle detection for mobile robots, and people tracking. A laser scanner measures a range data from the sensor to a target sequentially by a round-trip time of a laser pulse, and captures the 3D shape of the target object. Many object recognition/classification techniques based on range information have been proposed so far [1][2][3][4].

However, it is quite difficult to distinguish objects belonging to different classes but with similar 3D shapes such as a remote control and a cell phone using range data only. To improve the recognition performance, appearance information taken by a camera is quite useful. Mohottala et al.[5] developed a recognition system that utilizes a laser scanner and a camera. This system offered excellent results by acquiring textured 3D information of an object. However, since range and texture images are obtained by individual sensors, a calibration process between these sensors is necessary.

On the other hand, laser reflectance, which is the power of laser pulse reflected on a surface of a target object, is available for most laser scanners. A reflectance image that depicts reflectance values as a gray-scale image contains appearance information of the target, and range and reflectance images are fundamentally aligned since both images are obtained by a laser scanner simultaneously. Focusing on the reflectivity, we propose an object recognition technique by a laser scanner using multimodal information: structure from range image and appearance from reflectance image. We apply the proposed technique to a traffic census application, and show that proposed method improve the recognition accuracy.

The rest of this paper is organized as follows. In section 2, we propose a new object recognition technique utilizing
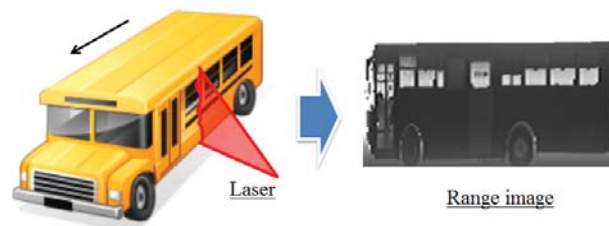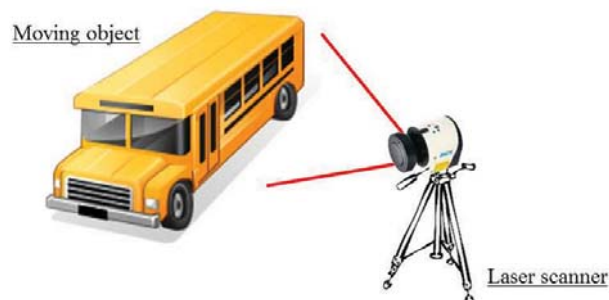

Fig. 1. Traffic census with a laser scanner

multimodal information. In section 3, experimental results for a traffic census application are shown, and the effectiveness of the proposed technique is verified.
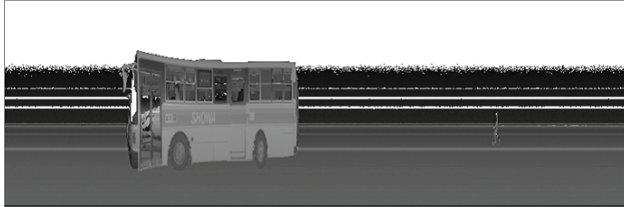
II. PROPOSED METHOD

In this section, we describe the procedure of the proposed method, especially about the reflectance image, the segmentation technique by background subtraction based on a range data, and the feature extraction for multimodal information. Note that we assume that range and reflectance images of moving objects are obtained from a 2D slit-type laser scanner placed on a roadside (Figure 1). In this system, the 2D slit laser beam is scanned along the vertical direction and line images of the range and reflectance of an object moving in front of the laser scanner are acquired sequentially.

A. Reflectance image

ToF optical range sensors acquire range data by measuring a round-trip time of a laser pulse reflected on an object surface.
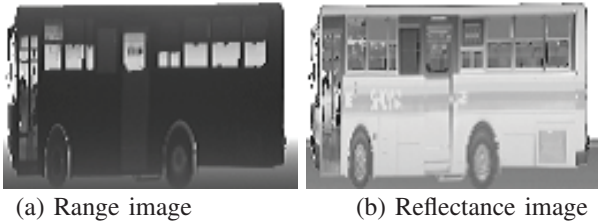
(a) Range image

(b) Reflectance image

Fig. 2. Range and reflectance images

(a) Range image   (b) Reflectance image

Fig. 3. Extracted range and reflectance images

(a) Variances of range values in each region ($k = 2$)

(b) Height of a foreground object

Fig. 4. Features of a range image

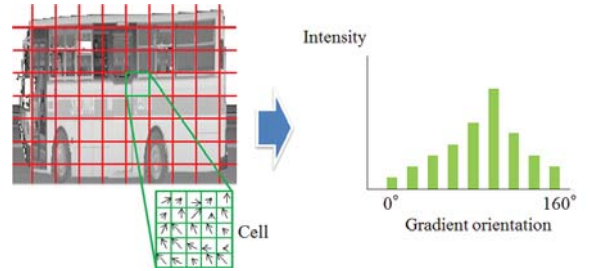Fig. 5. Histogram of oriented gradients (HOG) feature for a reflectance image

Figure 2(a) shows an example of a range image acquired by a laser scanner (Fig.1). On the other hand, most of optical range sensors record the strength of the reflected laser pulse (reflectivity).

Reflectance value indicates an intensity on the surfaces points of the targets under a single-frequency light source. Based on the dichromatic reflection model, the reflectance value consists of diffuse and specular reflection. Assuming the specular component can be ignored and Lambertian reflection is used as a model for diffuse reflection, a reflectance value is expressed as follows[6]:

$$I = k_d \frac{I_q}{r^2} cos\alpha \qquad (1)$$

where k_d is the diffuse reflection coefficient, I_q is the power of the light source, r is the distance of the light source toward the target, and α is the incident angle against the surface normal. Figure 2(b) shows an example of a reflectance image that depicts reflectance values as a gray-scale image. As mentioned above, the reflectance value is determined uniquely for each pixel in the range image. In other word, the range image and the reflectance image are fundamentally aligned.

B. Segmentation by background subtraction of range data

To extract features of a target object, we first cut out an object region from range and reflectance images. To do so, we utilize a range image since the laser scanner is fixed

on a roadside and moving targets can be extracted easily as foreground regions by background subtraction technique for range images. In addition, the appearance information is available by referring the same region in a corresponding reflectance image (Fig.3). Figure 7 shows example range and reflectance images of moving objects.

C. Feature extraction

We extract two kinds of features from segmented range and reflectance images respectively. In order to describe the structure of a target object, we first divide the range image into $k \times k$ areas, and then extract a variance of range values for each area. A height of the target object in the range image is also added to the feature vector. Consequently, we extract $k^2 + 1$ features from a range image as shown in Figure 4. For appearance information, we calculate Histogram of Oriented Gradients (HOG) features[7] from the reflectance image. HOG features describe local object appearances robustly by extracting the histogram of gradient orientation and the intensity (Fig.5).

D. Proposed method

Consequently, the proposed recognition technique is summarized as follows. We use Support Vector Machine (SVM) to identify classes of moving objects.
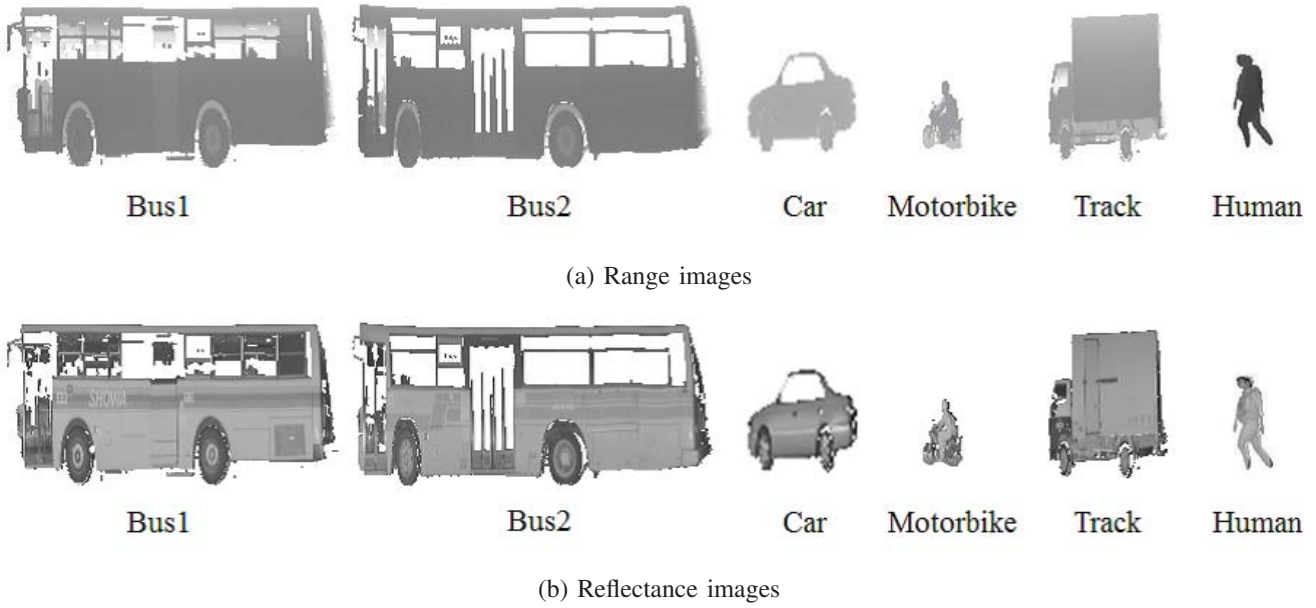
(a) Range images

(b) Reflectance images

Fig. 7. Categories of moving objects

TABLE II. RECOGNITION RESULTS

| | Range and Reflectance images (Proposed) | | | | | | Range image | | | | | |
	Bus1	Bus2	Car	Motorbike	Track	Human	Bus1	Bus2	Car	Motorbike	Track	Human
Images	24	13	72	43	16	50	24	13	72	43	16	50
Detected	24	13	72	43	16	50	23	17	69	40	19	50
Correct	22	11	72	43	16	50	18	10	68	37	16	45
Precision	92%	85%	100%	100%	100%	100%	78%	59%	99%	93%	84%	90%
Recall	92%	85%	100%	100%	100%	100%	75%	77%	94%	86%	100%	90%

TABLE I. FEATURES OF LMS151

Field of view	270 [°]
Frequency	50 [Hz]
Resolution	0.5 [°]
Operating range	0.5 - 50 [m]
Systematic error	± 30 [mm]
Statistical error	± 12 [mm]
Light source	Infrared (905 [nm])

1) Acquire continuous range and reflectance image through a laser scanner.
2) Segment a region of a object by background subtraction based on a range image.
3) Extract structural features from the range image and appearance features from the reflectance image.
4) Recognize a target object using SVM.

III. EXPERIMENTS

In this section, we introduce the experimental results of the proposed technique for the traffic census application. Figure

Fig. 6. Experimental setup

6 shows experimental setup. We use a laser scanner (SICK, LMS151) that has a field of view of 270 ° and a scanning frequency of 50 Hz with a resolution of 0.5 ° (Table I). Sequential range and reflectance image are obtained simultaneously. In the experiments, range and reflectance images are normalized from 0 to 255. The HOG feature consists of 7200 dimensions, and the parameter is set as $k = 4$. In this experiment, moving objects

Fig. 8. Recognition results after applying PCA to HOG features

TABLE III. RECOGNITION RESULTS USING 50 DIMENSIONS OF HOG FEATURES

| | Range and Reflectance images (Proposed) | | | | | | |
	Bus1	Bus2	Car	Motorbike	Track	Human	Total
Images	24	13	72	43	16	50	218
Detected	27	12	70	43	16	50	
Correct	22	10	70	43	16	50	211
Precision	81%	83%	100%	100%	100%	100%	
Recall	92%	77%	97%	100%	100%	100%	97%

consist of 6 classes, that is, Bus1, Bus2, Car, Motorbike, Track, and Human. We consider Bus1 and Bus2 which have slightly different shapes as different classes since these two buses belong to different companies.

Table II shows the classification results of the moving objects with and without the reflectance image, respectively. As shown in this table, the recognition accuracy is improved by using multimodal information. Especially, the classification accuracy for Bus1 and Bus2 is enhanced using the laser reflectivity. Next, we applied Principal Component Analysis (PCA) for HOG features to reduce the dimension. The results using PCA are shown in Figure 8 and Table III. From these results, we verified that the proposed method works well although the dimension of HOG features is reduced.

IV. CONCLUSION

This paper proposed a new object recognition method by a laser scanner utilizing multimodal information. Using a reflectance image which is obtained as a side-product of range scanning, the proposed technique improves the recognition accuracy even if targets have similar 3D shapes. We conducted experiments for the traffic census application, and verified that the proposed technique enable to classify six kinds of moving objects with higher accuracy than the case using range information only. In the future, we will discuss about the optimum parameters for the proposed technique, and apply the proposed technique for a variety of moving objects.

REFERENCES

[1] A. Johnson and M. Hebert, "Using spin images for efficient object recognition in cluttered 3d scenes," in *IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE*, 1999, vol. 21.

[2] R. Triebel, K. Kersting, and W. Burgard, "Robust 3d scan point classification using associative markov networks," in *IEEE International Conference on Robotics and Automation*, 2006.

[3] B. Steder, R. B. Rusu, K. Konolige, and W. Burgard, "Point feature extraction on 3d range scans taking into account object boundaries," in *IEEE International Conference on Robotics and Automation*, 2011.

[4] Volker Steinhage Jens Behley and Armin B. Cremers, "Performance of histogram descriptors for the classification of 3d laser range data in urban environments," in *IEEE International Conference on Robotics and Automation*, 2012.

[5] Masataka Kagesawa Shirmila Mohottala, Shintaro Ono and Katsushi Ikeuchi, "Fusion of a camera and a laser range sensor for vehicle recognition," in *Computer Vision and Pattern Recognition Workshops*, 2009.

[6] M.Bohme, M.Haker, T.Martinetz, and E.Barth, "Shading constraint improves accuracy of time-of-flight measurements," *Computer Vision and Image Understanding*, vol. 114, pp. 1329–1335, 2010.

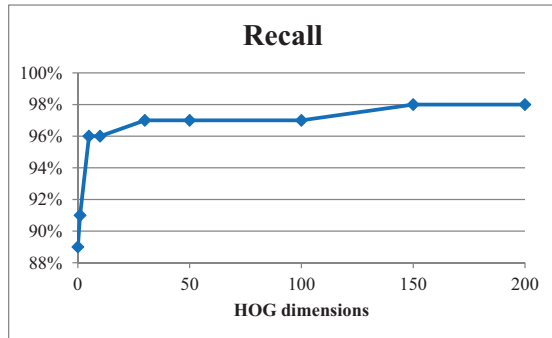[7] Navneet Dalal and Bill Triggs, "Histograms of oriented gradients for human detection," in *Computer Vision and Pattern Recognition*, 2005, vol. Vol.1.