Special Feature: Active Safety

Research Report

Road Environment Recognition for Advanced Driver Assistance Systems

Takashi Naito, Kunihiro Goto and Kiyosumi Kidono

Report received on Feb. 20, 2012

ABSTRACT Due to the active economy and growing population of developing countries, worldwide auto sales are expected to increase significantly for the next several years. With increasing automobiles, there are growing fears of increasing traffic accidents. In order to reduce traffic accidents, various advanced driver assistant systems (ADAS) have been becoming popular in recent years and they have contributed to reducing damages from traffic accidents as well as preventing the accidents. However, more intelligent functions for road environment recognition are required as future ADAS. For example, prediction of pedestrian behaviors can be used to estimate the probability of a collision precisely. In this paper, our research achievements for recognizing the road environment utilizing camera and LIDAR are introduced. Some experimental results show that our developed algorithms are competent as key roles of future ADAS.

KEYWORDS ADAS, Image Processing, Pattern Recognition, Laser Radar, Sensor Fusion

1. Introduction

Recently, the annual auto sales in Japan have remained unchanged, leveling at approximately 4,000,000 to 5,000,000 automobiles sold per year, due to economic depression, low birthrate and an aging population. However, from a global viewpoint, due to the active economy and growing population of developing countries, worldwide auto sales are expected to increase significantly for the next several years.

With increasing number of automobiles, there are growing fears of increasing traffic accidents, especially in developing countries. For example, in China, the annual number of fatalities resulting from road traffic accidents appeared to be approximately 90,000 in 2006.(1) As a result, safety technologies for automobiles have been becoming more important.

Safety technologies are divided into two categories, namely, passive and active safety technologies. The former alleviates damages from an accident as much as possible, such as improvement of the stiffness of vehicle body and air-bag system. Meanwhile, the latter prevents the occurrence of traffic accidents.

For example, the anti-lock breaking system (ABS) and electronic stability control (ESC) system help drivers to control their automobiles stably even in emergency situations. In particular, systems that support drivers by recognizing the road environment using sensors installed on the automobiles are becoming popular recently. The lane departure warning (LDW) system(2) gives a warning to drivers when the vehicle is about to deviate from a traffic lane. In addition, the pre-crash safety (PCS) system(3) reduces damages from a frontal collision using sensors such as a millimeter wave radar and a stereo camera. The above-mentioned active safety systems are also known as advanced driver assistant systems (ADAS).

In recent years, various kinds of ADAS have been introduced into the commercial market and those systems have been confirmed to reduce the damage of traffic accidents as well as prevent them. However, in order to eliminate traffic accidents (known as the "Zeronize" concept (4)), which is the ultimate goal for us, more intelligent functions for road environment recognition are required as future ADAS. For example, prediction of pedestrian behaviors can be used to estimate the probability of a collision precisely.

This paper introduces our research achievements to realize future ADAS, particularly on three developed recognition algorithms utilizing a camera and a light detection and ranging sensor (LIDAR).

In section 2, pedestrian detection and direction estimation by camera images are introduced to predict pedestrian behaviors. In section 3, algorithms for recognizing road environment by high-definition LIDAR are proposed, particularly to recognize moving objects and static objects including curbs, guard rails and lane marks. Furthermore, in section 4, a novel pedestrian recognition algorithm based on several information acquired from both range data and reflection intensity is explained, focusing on pedestrian recognition. Finally, in section 5, we will summarize our present study and discuss our future studies.

2. Pedestrian Detection and Direction Estimation

In the development of ADAS, protecting pedestrians is an important issue to reduce traffic accident fatalities. A wide variety of pedestrian detection methods by camera images have been proposed.[5] However, these methods do not estimate the direction in which a pedestrian is moving or facing. It is possible that more assistance can be offered when the imminent behavior can be inferred by the direction of the pedestrian.

We have proposed a pedestrian detection method and a direction estimation method by a cascade detector using a feature interaction descriptor (FIND),[6] which can capture the high-level properties of an object's appearance in comparison with common image features such as the histogram of oriented gradients (HOG).[7]

FIND improves the pedestrian classification performance, but it needs much more computational cost than HOG due to the computation of the feature interaction of adjacent histogram elements. Therefore, for computational efficiency, a three-stage cascade detector, shown in **Fig. 1** has been proposed.[8] Each stage uses a different image feature taking the classification performance and computational cost into consideration. The first stage of the cascade detector applies "sparse" HOG without block overlapping for fewer dimensions. The second stage of the cascade detector applies HOG with block overlapping. In this stage, the number of pedestrian candidates is narrowed down by HOG, because its classification performance is better than that of "sparse" HOG. The purpose of the first and second stages is to efficiently eliminate obvious non-pedestrians from the vast number of candidates in a short amount of time. In the final stage, FIND is applied to verify the small number of remaining pedestrian candidates. In the training phase, the classifier of each stage is trained by a linear support vector machine (SVM).[9]

In order to estimate the directions of pedestrians at different distances, multi-classifiers specialized in pedestrian directions and distances are employed. **Figures 2**(a) and (b) show examples of pedestrians in different directions and distances. There are notable differences in the pedestrians' appearance based on their direction and distance. For the cascade detector shown in Fig. 1, distance-related classifiers are applied to each stage. Meanwhile, the four direction-related classifiers (right, left, front, back) are applied only in the last stage. The direction is determined from the classifier whose score is maximal.

We evaluate our proposed method using a large video sequence data set with a resolution of 640 × 480 pixels acquired in the daytime and with several road conditions. The data set consists of 8,166 pedestrians.

Figures 3(a) and (b) show the experimental results for pedestrian detection and direction estimation by the proposed method. **Figure 4** shows the quantitative performance of pedestrian detection in comparison to the method without multi-classifiers. It is confirmed that a combination of the distance-related and direction-related classifiers further improves the detection performance. **Figure 5** shows the direction estimation results. It can be seen that the direction of over 85% of the pedestrians was estimated correctly. These experimental results show that superior detection and direction estimation performances are obtained by the proposed multi-classifier method.

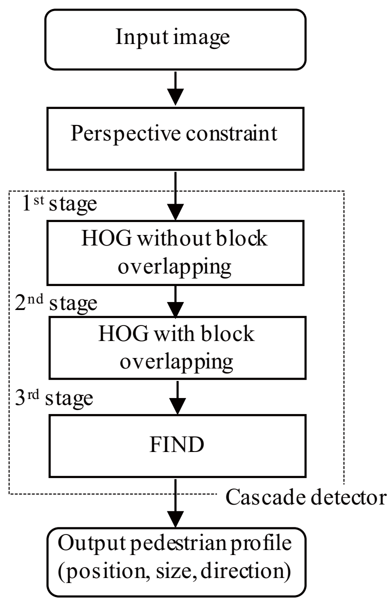


Fig. 1 Flow chart of the proposed cascade detector.

3. Road Environment Recognition by LIDAR

A horizontally scanning laser scanner, known as a LIDAR system, is often used for recognition of the road environment. For example, in the DARPA urban challenge,[10] most vehicles that participated in the competition had LIDAR for autonomous driving. LIDAR has excellent advantages of high spatial

Fig. 2 Examples of pedestrian profiles with respect to direction (a) and distance (b).

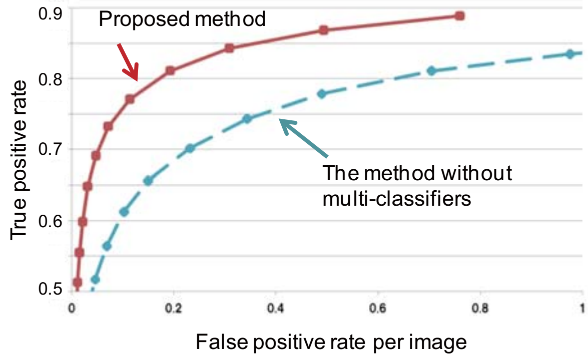


Fig. 4 Quantitative performance for pedestrian detection by the proposed method compared to that by the method without multi-classifiers.

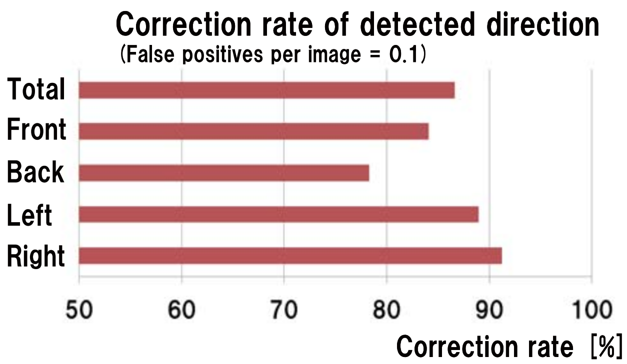


Fig. 5 Results of the direction estimation for the detected pedestrians.

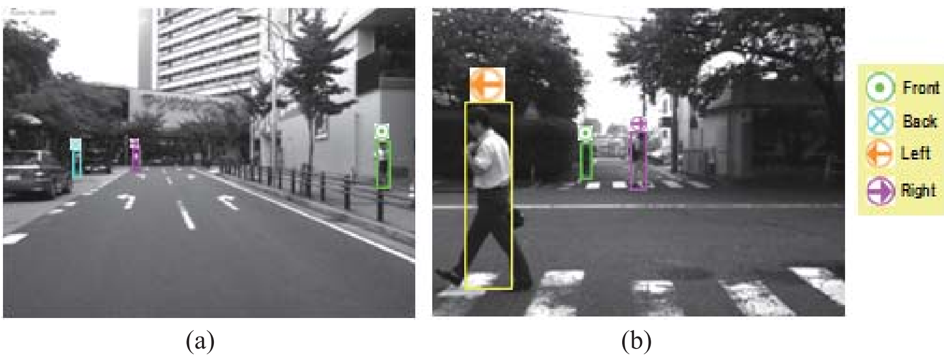


Fig. 3 Results of the pedestrian detection. Rectangles show the detected pedestrian and symbols on the top of the rectangles show the estimated direction by our proposed method.

resolution and high range accuracy compared with other sensors such as a millimeter wave radar and a camera, respectively.

In this section, we propose a road environment recognition system by high-definition LIDAR (Velodyne), especially for recognizing static and moving objects comprehensively in urban areas.(11,12)

The specifications of Velodyne are listed in **Table 1**. It has 64 scanning lines at approximately 0.4 degree intervals aligned in the vertical plane and it can obtain dense 3D range data by horizontal scans of 360 degrees. The sensor is mounted on the roof carrier of the experimental vehicles, whose height is approximately 2 m. **Figure 6** shows a sample of the measured 3D range data that is composed of points with 3D position and reflection intensities.

The processing flow of our proposed method is shown in **Fig. 7**. First, the acquired 3D point cloud as shown in Fig. 6 is divided into two classes, ground data and obstacles, using an occupancy grid map.(13) All the 3D points are projected onto a 2D occupancy grid, which is parallel to the ground plane. The difference between the maximum height and the minimum height in each cell is investigated. If the difference is larger than the threshold, the points in the cell are segmented as obstacles. However, if the difference is smaller, the points are assigned as ground data including the ground plane and curbs. Next, the motion of the vehicle is estimated. Based on the hypothesis that the cells segmented as obstacles contain many static objects such as trees, poles, guardrails, buildings, and

Table 1 Specifications of LIDAR.

Item	Specification
Scanning rate	10 scans/s
Horizontal field of view	360°
Horizontal angular resolution	0.23°
Vertical field of view	26.8°
Vertical angular resolution	0.4° (64 lines)
Detection range	40 m for pavements 120 m for cars
Range accuracy	0.02 m
Wavelength of laser beam	905 nm

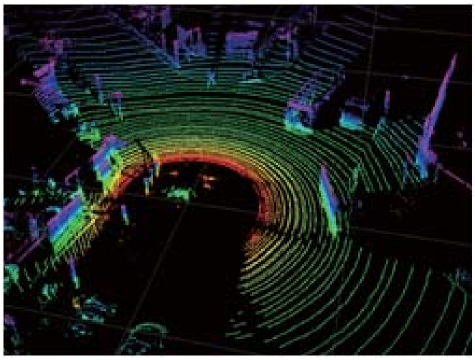


Fig. 6 Example of 3D range data around the vehicle installed with high-definition LIDAR.

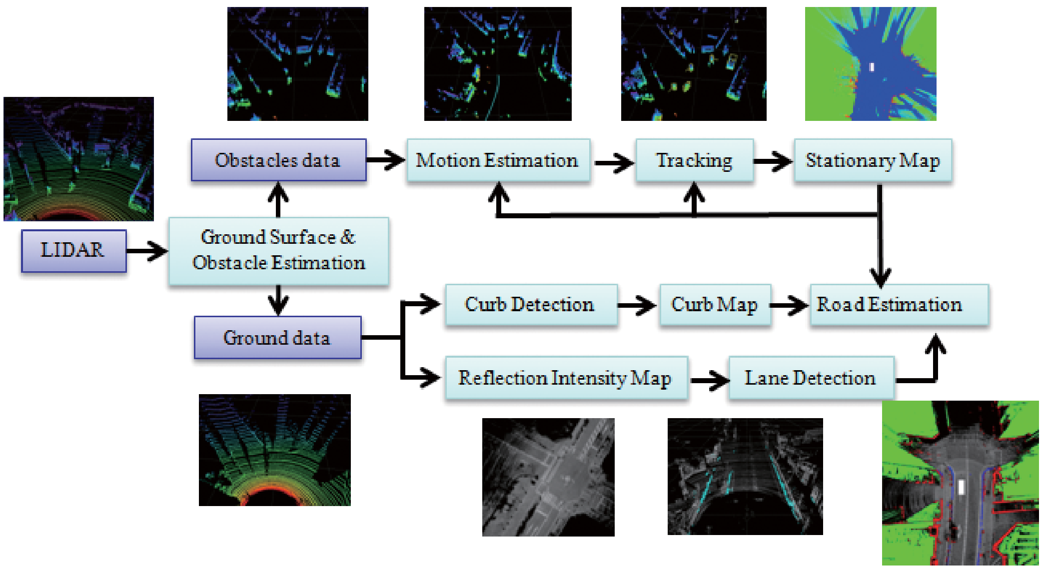


Fig. 7 Processing flow of the road environment recognition system with high-definition LIDAR.

stopped cars, the ICP algorithm[14] is applied to estimate the motion parameters between two adjacent obstacle data. After estimating the motion of the vehicle (sensor) by subtraction, the moving objects can be tracked from the obstacle data and the stationary map can be recovered simultaneously.

Meanwhile, the curbs are detected from ground data whose height is between 10 and 25 cm. Moreover, by accumulating reflection intensities over time and by applying the conventional image processing method, the reflection intensity map is generated as shown in Fig. 7 and the lane can be detected, respectively. Finally, the road parameters are estimated by combining the curb and lane information.

Figure 8 shows an example of the results based on road environment recognition by our proposed method. Figures 8(a) and (b) show the 3D range data depicted from a driver viewpoint and the processing result, respectively. It is confirmed that a few vehicles and a pedestrian are detected as moving objects (yellow rectangle) correctly. Moreover, the lane (green) and the curb (blue) are detected from the ground data and reflection intensity map.

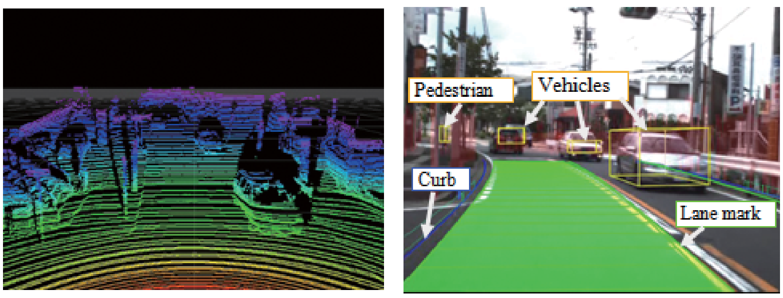
Since precise 3D range data are acquired from LIDAR, accurate motion estimation and collision determination against vehicles and pedestrians, which are important roles of ADAS, can be achieved.

4. Precise Pedestrian Recognition Using LIDAR

In this section, we propose a novel precise pedestrian recognition method by high-definition LIDAR. Compared with a camera, the spatial resolution is inferior even when using high-definition LIDAR. As described in section 3, moving objects such as pedestrians and vehicles can be detected. However, if pedestrians are standing and not moving, conventional algorithms are not able to detect them because appearance information, which is usually validated for pattern recognition described in section 2, cannot be acquired from LIDAR. In contrast, since dense range data is acquired from high-definition LIDAR, the rich information can be exploited for pedestrian recognition. Moreover, as accurate range information can be provided by LIDAR, precise motion estimation of pedestrians is expected in comparison with a camera.

The literature[15] proposed a method for tracking pedestrians in 3D range data from Velodyne. The 3D point cloud of the target is divided into three parts corresponding to the legs and trunk of a pedestrian, and the variances of the 3D points contained in each part are utilized to discriminate the pedestrians. In addition, they represent a 3D pedestrian shape by 2D histograms on two principal planes. This approach has the advantage of a low computational cost since the feature extraction is very simple. However, the detection performance is very sensitive to the target distance. We extend previous methods[15,16] to employ two novel features in order to achieve high-performance pedestrian recognition even at a long range.

The algorithm for segmenting obstacles from range data is the same as shown in Fig. 7. From the segmented obstacles, the clusters corresponding to vehicles and pedestrians are generated by a distance-based clustering algorithm. To reduce computational cost, the clustering process is carried out on the occupancy grid using the labeling technique for image processing. If the distance between two points is within 0.5 m, these points are integrated into the same cluster. A rectangular parallelepiped is applied to each cluster by the rotation calipers method.[17] Then, the pedestrian



(a) 3D range data from the driver's viewpoint

(b) Recognition result from the 3D range data shown in (a)

Fig. 8 Example of the recognition results for vehicles, a pedestrian, curbs and lane marks.

candidates are extracted on the basis of the size of the clusters. The conditions are as follows:

$$0.8 \le h \le 2.0, \quad \cdots\cdots (1)$$

$$w \le 1.2, \quad \cdots\cdots (2)$$

$$l \le 1.2, \quad \cdots\cdots (3)$$

where h, w and l denote the height, the width and the length of the cluster, respectively.

A feature vector is computed from the 3D point cloud of each pedestrian candidate. **Table 2** lists all nine features used in the proposed method. The set of feature values of each candidate C_j forms a vector $f_j = (f_1, \ldots, f_9)$. Features f_1 and f_2 are introduced by the Premebida method.[16] Features f_3 to f_7 are proposed using the Navarro-Serment method.[15] To improve the classification performance, the proposed method adds two more features.

The first feature is the slice feature as shown in **Fig. 9**. A rough profile from the head to the legs is utilized to represent the 3D shape of the pedestrians. Three principal axes for the pedestrian candidates are calculated by principal component analysis (PCA). We assume that most pedestrians are in an upright position; thus, the principal eigenvector is expected to be vertically aligned with the pedestrian's body. 3D points in the cluster are divided into N blocks of the same size along the principal eigenvector, as shown in the left image of Fig. 9. As a result of the division, a common feature can be extracted from pedestrians with different heights, such as an adult and a child. Then, the 3D points in each block are projected onto a plane orthogonal to the principal eigenvector, and two widths along the other eigenvectors are computed as the features. The number of blocks is 10. Eventually, the feature vector, referred to as the "slice feature" in this paper, is represented as follows:

$$f_8 = \{w_{00}, w_{01}, \ldots, w_{j0}, w_{j1}, \ldots, w_{N0}, w_{N1}\}. \quad \cdots (4)$$

The second feature is computed from the reflection intensities of the 3D points, which are expected to depend on the materials of the targets. The following three values are computed from the reflection intensities contained in the pedestrian candidates; i) mean intensity, ii) standard deviation of the intensities and iii) normalized histogram where the number of bins is 25 and the range of the intensities is divided at equal intervals.

On the basis of the calculated nine features as shown in Table 2, SVM is applied to classify the pedestrians from the candidates.

An example of the experimental result is shown in **Fig. 10**. We evaluated the recognition performance using a dataset which includes 7,865 positive samples and 8,055 negative samples.[11] The measurement range is up to 50 m. Figure 10 shows the classification result for the overall samples. The true positive rate of the proposed method is approximately 0.1 higher than that of the Navarro-Serment method at the point where the false positive rate is 0.01. To evaluate the performance at different ranges, all samples are divided into four range groups: 10 to 20 m, 20 to 30 m, 30 to 40 m, and 40 to 50 m. **Figure 11** shows the result of the evaluation at the different ranges. The proposed method generally shows higher performances than the Navarro-Serment method,

Table 2 Features for pedestrian classification. f_8 and f_9 are the proposed features for precise classification.

No.	Description	Dimension
f_1	Number of points including the cluster	1
f_2	The minimum distance to the cluster	1
f_3	3D covariance matrix of the cluster	6
f_4	The normalized moment of inertia tensor	6
f_5	2D covariance matrix in three zones	9
f_6	2D histogram for the main plane	98
f_7	2D histogram for the secondary plane	45
f_8	Slice feature for the cluster	20
f_9	Reflectivity distribution for the cluster	27

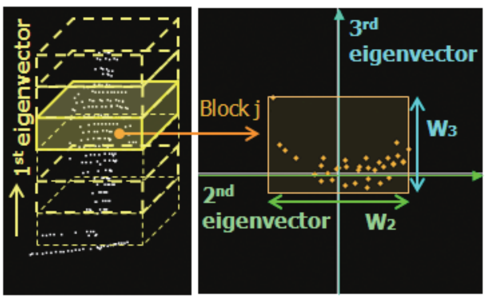


Fig. 9 Proposed slice feature. Left image shows a pedestrian candidate composed of a 3D point cloud with N blocks. Right image shows a profile of Block j for calculating the feature vector.

especially at the range between 30 to 50 m. The two additional features work effectively for the long range, where the spatial resolution of the measurement decreases. These quantitative evaluations confirm that the proposed method improves the classification ability, especially at a range of more than 30 m.

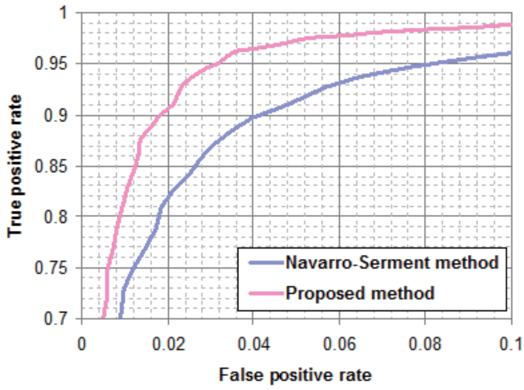


Fig. 10 Recognition performance for the overall pedestrians in our dataset with the Navrro-Sermet method.

5. Conclusion

Some research achievements for recognizing road environment were presented in this paper. The proposed high-dimensional image feature, FIND, enables competent performance of pedestrian detection from captured camera images. Our proposed method for pedestrian recognition can also estimate the direction of the pedestrian, suggesting that the developed algorithm can be applied to the prediction of pedestrian behavior. In addition, studies utilizing a range sensor called LIDAR were presented. The road environment including vehicles, pedestrians, curbs and lanes can be comprehensively recognized by high-definition LIDAR. Moreover, the alternative novel method for pedestrian detection applying rich 3D point clouds and reflection intensity information showed high performance from some experiments.

For future research, we will combine functions attained by a camera and LIDAR to recognize the road environment robustly and precisely. For example, sensor integration will be expected to realize intelligent inferences for a collision between pedestrians and vehicles. We confirmed that our future research will contribute to “Zeronize” the number of traffic accidents and “Maximize” the positive aspects of vehicles.

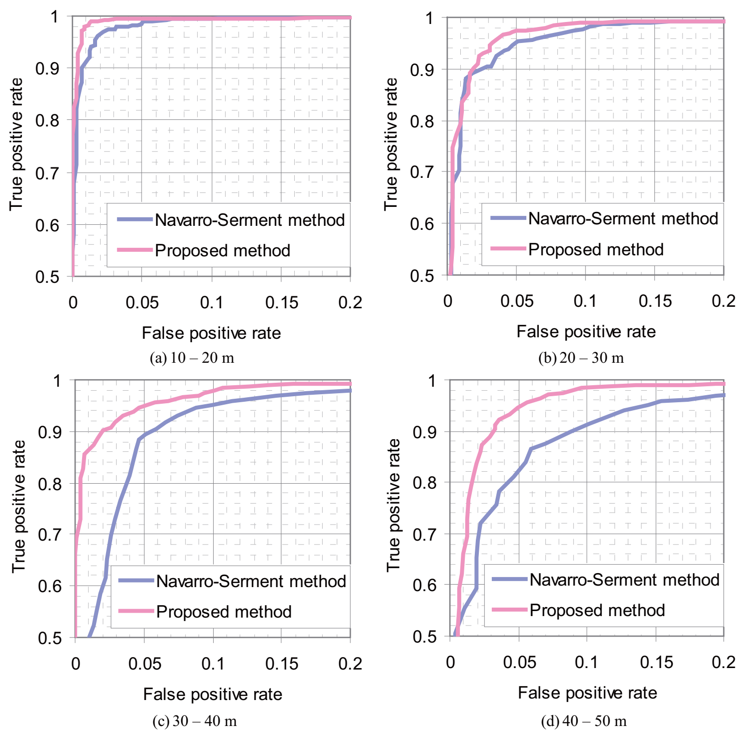


Fig. 11 Evaluation results of pedestrian recognition according to distances.

References

(1) World Health Organization, *Global Status Report on Road Safety: Time for Action* (2009), World Health Organization.

(2) “Active Safety”, <http://www.toyota-global.com/innovation/safety_technology_quality/safety_technology/active_safety/>, (accessed 2012-02-20).

(3) “Pre-crash Safety System”, <http://www.toyota-global.com/innovation/safety_technology_quality/safety_technology/pre-crash_safety/>, (accessed 2012-02-20).

(4) “TOYOTA's ITS Vision”, <http://www.toyota-global.com/innovation/intelligent_transport_systems/toyota_ap/vision/>, (accessed 2012-02-20).

(5) Geronimo, D., Lopez, A. M., Sappa, A. D. and Graf, T., “Survey on Pedestrian Detection for Advanced Driver Assistance Systems”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.32, No.7 (2010), pp.1239-1258.

(6) Cao, H., Yamaguchi, K., Naito, T. and Ninomiya, Y., “Pedestrian Recognition Using Second-order HOG Feature”, *Proc. of 9th Asian Conference on Computer Vision* (2009), pp.628-634.

(7) Dalal, N. and Triggs, B., “Histograms of Oriented Gradients for Human Detection”, *Proc. of IEEE Conference on Computer Vision and Pattern Recognition*, Vol.1 (2005), pp.886-893.

(8) Goto, K., Kidono, K., Kimura, Y. and Naito, T. “Pedestrian Detection and Direction Estimation by Cascade Detector with Multi-classifiers Utilizing Feature Interaction Descriptor”, *Proc. 2011 IEEE Intelligent Vehicles Symposium* (2011), pp.224-229.

(9) Fan, R., Chang, K., Hsieh, C., Wang, X. and Lin, C., “LIBLINEAR: A Library for Large Linear Classification”, *Journal of Machine Learning Research*, Vol.9 (2008), pp.1871-1874.

(10) “Urban Challenge”, <http://archive.darpa.mil/grandchallenge/index.asp>, (accessed 2012-02-20).

(11) Miyasaka, T., Ohama, Y. and Ninomiya, Y., “Ego-motion Estimation and Moving Object Tracking Using Multi-layer LIDAR”, *Proc. 2009 IEEE Intelligent Vehicles Symposium* (2009), pp.151-156.

(12) Kidono, K., Miyasaka, T., Watanabe, A., Naito T. and Miura, J., “Pedestrian Recognition Using High-definition LIDAR”, *Proc. 2011 IEEE Intelligent Vehicles Symposium* (2011), pp.405-410.

(13) Thrun, S., Burgard, W. and Fox, D., *Probabilistic Robotics* (2005), The MIT Press.

(14) Besl, P. and McKay, N., “A Method for Registration and 3D Shape”, *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol.12, No.2 (1992).

(15) Navarro-Serment, L. E., Mertz, C. and Hebert, M., “Pedestrian Detection and Tracking Using Three-dimensional LADAR Data”, *Proc. Int. Conf. on Field and Service Robotics* (2009).

(16) Premebida, C., Ludwing, O. and Nunes, U., “Exploiting LIDAR-based Features on Pedestrian Detection in Urban Scenarios”, *Proc. 12th Int. IEEE Conf. on Intelligent Transportation Systems* (2009).

(17) Toussaint, G., “Solving Geometric Problems with the Rotating Calipers”, *Proc. IEEE MELECON*, Vol.9 (1983), pp.1-8.

Takashi Naito

Research Fields:
- Image Processing
- Pattern Recognition
- Computer Vision
- ADAS

Academic Societies:
- The Robotics Society of Japan
- The Institute of Electronics, Information and Communication Engineers

Kunihiro Goto

Research Fields:
- Image Processing
- Pattern Recognition
- Computer Vision
- ADAS

Academic Degree: Dr. Eng.

Academic Society:
- The Institute of Electrical Engineers of Japan

Kiyosumi Kidono

Research Fields:
- Computer Vision
- Robotics
- ITS

Academic Societies:
- The Robotics Society of Japan
- The Institute of Electronics, Information and Communication Engineers
- The Society of Instrument and Control Engineers