Fig. 5. Experimental vehicle

IV. Experiments

To confirm the effectiveness of the proposed method, experiments on scene-wise reliability estimation were conducted by using actual point clouds measured using a LiDAR as input.

A. Experimental Conditions

To acquire point clouds, we prepared a vehicle that mounted Velodyne LiDAR HDL-64e on the roof as shown in Fig. 5. The vehicle traveled through urban areas during the day time, and obtained 17 sequences composed of 3,871 frames. Each experiment was carried out with a leave-one-out cross-validation scheme where one sequence was used for evaluation and the others for learning.

In this experiment, pedestrians with no occlusion within approximately 40 m from the sensor were targets for the detection, which were annotated manually.

B. Pedestrian Detector

As pre-processing, road plane removal and clustering using Euclidean distance were performed for candidates extraction. Next, these candidates were classified whether they were pedestrian or not by a classifier. The pedestrian classifier used in this experiment was constructed with Real AdaBoost [12]. The following 225 dimensions feature was used for this.

Features proposed by Kidono et al. [3]

213-dimensional features.

Features proposed by Liu et al. [13]

Seven-dimensional features based on eigenvalues.

Additional features

Five-dimensional features. (The height of the cluster, the maximum and minimum of the Z coordinate of each point, the area in the X-Y plane, the median of the intensity of each point.)

C. Validity of the Proposed Reliabilities

Here we present an experiment to confirm the validity of the proposed reliabilities described in Section II.

The proposed method defined "threshold where oversights or false detections are minimized" as reliabilities. For evaluation of these criteria, the thresholds in each frame were calculated with the pedestrian detector constructed in Section IV-B. Note that the pedestrian detector applied to each sequence was trained with other sequences. In addition, clusters extracted from the pre-processing were used for evaluating oversights and false detections.

As an example, the result from sequence #14 is shown in Fig. 6. The threshold of the detector is in the range of [0.00, 1.00]. Figure 6(a) shows the maximum thresholds without oversights defined in (1) for each frame. As mentioned in Section II-A, we can say that outputs from scenes with higher threshold values can be trusted. On the other hand, Figure 6(b) shows the minimum thresholds without false detections defined in (3) for each frame. As mentioned in Section II-B, we can say that outputs from scenes with lower threshold values can be trusted. Note that blanks in Fig. 6(a) are frames with no pedestrian to be detected.

From Fig. 6, we can see that the threshold changes across frames in accordance with the change of scene following the travelling of the vehicle. A frame indicated by a solid circle in Fig. 6(a) is shown in Fig. 7(a) as an example of a low reliability scene considering oversights. A pedestrian existing near a wall, so it was difficult to detect. Similarly, a frame indicated by a dotted circle in Fig. 6(a) is shown in Fig. 7(b) as an example of a high reliability scene considering oversights. A pedestrian existing in an open intersection, so it was easy to detect.

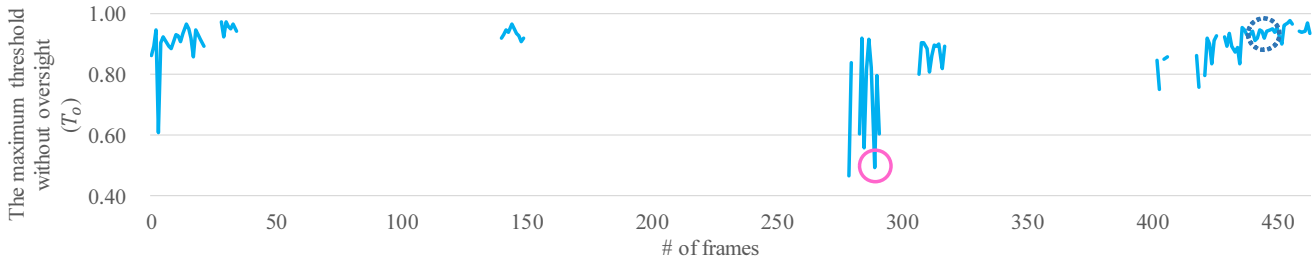
Meanwhile, a frame indicated by a solid circle in Fig. 6(b) is shown in Fig. 7(c) as an example of a low reliability scene considering false detections. Since there were many structures similar to pedestrians such as poles, trees, and so on, false detections were likely to occur. Similarly, a frame indicated by a dotted circle in Fig. 6(b) is shown in Fig. 7(d) as an example of a high reliability scene considering false detections. Although there are trees along the street, they were easy to be separated from pedestrians because they were few and thin.

As described above, we confirmed the validity of the proposed reliabilities. In the remaining part of the experiments, the reliabilities calculated here are used as ground truths.

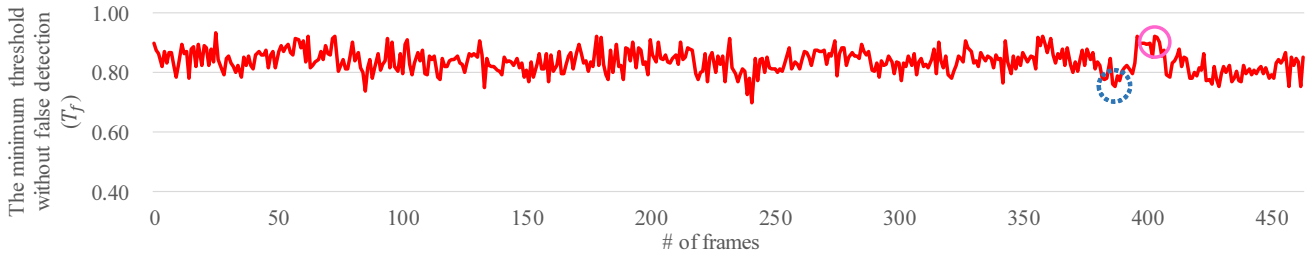
D. Effectiveness of the Proposed Reliability Estimation

Here we present an experiment to confirm the effectiveness of the proposed reliability estimator described in Section III. The estimators were constructed with the reliabilities calculated in Section IV-C as ground truths. Note that the estimators for each sequence were learned with the reliabilities of other sequences.

As a result, the proposed estimator could estimate the reliability considering oversights with 0.055 mean absolute error, and that considering false detections with 0.033 mean absolute error. As an example, the estimation result for sequence #14 is shown in Fig. 8. From Fig. 8(b), we can see that the reliability estimator considering false detections was able to estimate quite well. Meanwhile, from Fig. 8(a), we can see that the reliability estimator considering oversights failed to estimate well when the reliability decreased drastically. The following two reasons could be the cause for this.



(a) Considering oversights.



(b) Considering false detections.

Fig. 6. Thresholds where oversights or false detections are minimized.

Lack of features

Features leading to oversights such as integration of clusters were lacking.

Insufficient amount of training data

The reliability considering oversights is calculated from only frames including pedestrians, so the size of the training data was insufficient.

V. Summary

In this paper, we focused on the problem that outputs from pedestrian detectors can not be fully trusted in real environments. Considering this, we defined reliabilities of detectors for a given scene, and proposed a method to construct a scene-wise reliability estimator. The proposed method defined the "threshold where oversights or false detections are minimized" as scene-wise reliabilities.

To demonstrate the effectiveness of the proposed method, experiments were conducted by applying the proposed method to actual point clouds measured from a LiDAR. The experimental results showed that the definitions of reliabilities were reasonable, and that the proposed estimator can estimate the reliabilities properly.

For future work, we need to extend features of the reliability estimator for more accurate estimation. For example, rather than outputting a global reliability for a scene, we may calculate estimation features and the reliability of local regions. Furthermore, we should also design a method that considers the final detection results considering the reliability.

ACKNOWLEDGMENT

Parts of this research were supported by the Center of Innovation Program from Japan Science and Technology Agency, JST.

References

[1] H. Yoshida, D. Suzuo, D. Deguchi, I. Ide, H. Murase, T. Machida, and Y. Kojima, "Pedestrian detection based on deep convolutional neural network with ensemble inference network," in *Proc. 2013 IEEE Intelligent Vehicles Symposium*, Jun. 2013, pp. 654–659.

[2] H. Fukui, T. Yamashita, Y. Yamauchi, H. Fujiyoshi, and H. Murase, "Pedestrian detection based on deep convolutional neural network with ensemble inference network," in *Proc. 2015 IEEE Intelligent Vehicles Symposium*, Jul. 2015, pp. 223–228.

[3] K. Kidono, T. Miyasaka, A. Watanabe, T. Naito, and J. Miura, "Pedestrian recognition using high-definition LIDAR," in *Proc. 2011 IEEE Intelligent Vehicles Symposium*, Jun. 2011, pp. 405–410.

[4] T. Ogawa, H. Sakai, Y. Suzuki, K. Takagi, and K. Morikawa, "Pedestrian detection and tracking using in-vehicle LIDAR for automotive application," in *Proc. 2011 IEEE Intelligent Vehicles Symposium*, Jun. 2011, pp. 734–739.

[5] M. Engelcke, D. Rao, D. Z. Wang, C. H. Tong, and I. Posner, "Vote3Deep: Fast object detection in 3D point clouds using efficient convolutional neural networks," in *Proc. 2017 Int. Conf. on Robotics and Automation*, May 2017, pp. 1355–1361.

[6] C. R. Qi, L. Yi, H. Su, and L. J. Guibas, "PointNet++: Deep hierarchical feature learning on point sets in a metric space," *Advances in Neural Information Processing Systems 30*, pp. 5099–5108, 2017.

[7] T. Yamamoto, F. Shinmura, D. Deguchi, Y. Kawanishi, I. Ide, and H. Murase, "Efficient pedestrian scanning by active scan LIDAR," in *Proc. 2018 Int. Workshop on Advanced Image Technology*, no. C4-2, Jan. 2018, pp. 1–4.

[8] Y. Tatebe, D. Deguchi, Y. Kawanishi, I. Ide, H. Murase, and U. Sakai, "Pedestrian detection from sparse point-cloud using 3DCNN," in *Proc. 2018 Int. Workshop on Advanced Image Technology*, no. C4-4, Jan. 2018, pp. 1–4.

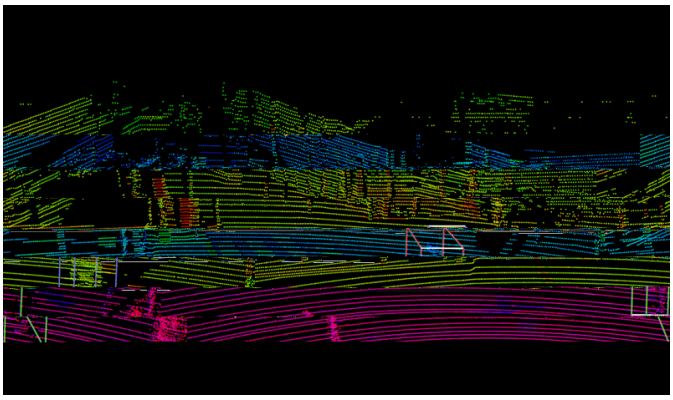
[9] A.-L. Jousselme, A.-C. Boury-Brisset, B. Debaque, and D. Prevost, "Characterization of hard and soft sources of information: A practical illustration," in *Proc. 2014 Int. Conf. on Information Fusion*, Oct. 2014, pp. 1–8.

[10] N.-E. E. Faouzi and L. A. Klein, "Data fusion for ITS: Techniques and research needs," *Transportation Research Procedia*, vol. 15, pp. 495–512, Jun. 2016.

[11] A. R. Hilal, "Context-aware source reliability estimation for multi-sensor management," in *Proc. 2017 Annual Int. Systems Conf.*, Apr. 2017, pp. 1–4.

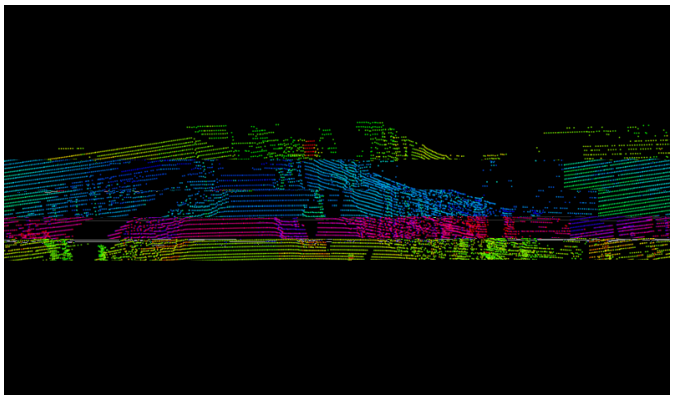
[12] R. E. Schapire and Y. Singer, "Improved boosting algorithms using confidence-rated predictions," *Machine Learning*, vol. 37, no. 3, pp. 297–336, Dec. 1999.

[13] X. Liu, G. Zhao, J. Yao, and C. Qi, "Background subtraction based on low-rank and structured sparse decomposition," *IEEE Trans. on Image Processing*, vol. 24, no. 8, pp. 2502–2514, Apr. 2015.



(a) Low reliability scene considering oversights due to large structures.

(b) High reliability scene considering oversights due to open space.

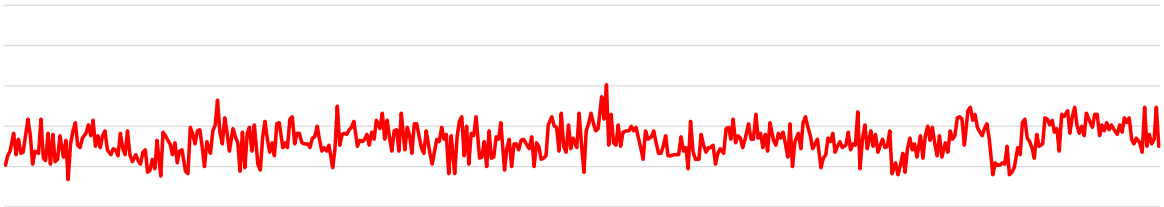


(c) Low reliability scene considering misdetections due to existence of similar objects.

(d) High reliability scene considering misdetections due to inexistence of similar objects.

Fig. 7. Examples of scenes with low and high reliabilities.

(a) Considering oversights.



(b) Considering false detections.

Fig. 8. Estimated reliabilities.