Ambient Environment Recognition Algorithm Fusing Vision and LiDAR Sensors for Robust Multi-channel V2X System

Gyu Ho Lee[1] , Ki Hoon Kwon[1], Min Young Kim[1,2]

[1]Department of Electronics Engineering, Kyungpook National University, Daegu, Korea

[2] Research Center for Neurosurgical Robotic System, Daegu, Korea

leegh623@naver.com , kwons149@naver.com , minykim@knu.ac.kr

Abstract— Recently, 5G commercialization issues and standardization of WAVE communication have been done on V2X communication of autonomous vehicles. In this paper, we propose an algorithm for predicting the communication performance of multi-channel V2X with a vision sensor. The proposed method recognizes the surrounding environment as a vision sensor and provides information to the TCU board to select optimal parameters. The sensing system integrates camera and LiDAR sensor data into a single data set. We applied the CNN-based object detection algorithm to the fusion sensor and recognized the driving environment. We defined the situation that affects the communication performance and measured the accuracy by perceiving the scenario. This algorithm can optimize the communication channel and select the communication channel suitable for the driving environment in advance to improve the communication stability.

Keywords— Multi-Channel V2X; LiDAR; Vision; CNN; Object Detection

I. INTRODUCTION

There are two types of autonomous navigation technologies: sensor type and communication type. The sensor system using the camera, radar, and LiDAR detects the surrounding obstacles directly. However, it has a relatively small detection area and a limitation when obstacles obstruct the sensor. The V2X communication technology can complement the limitations of sensors. V2X refers to a technology in which a vehicle exchanges information with other vehicles, roads, and other infrastructure through a wireless network. However, there is also a problem with V2X communication technology. Its infrastructure should be built on the road and has a problem of leakage of personal information. In addition, if a hacker attacks the network, the vehicle network is paralyzed and the safety can be threatened.

In this paper, we try to solve the problem of communication performance deterioration due to the surrounding environment of the V2X communication vehicle. Vehicle communication quality changes dynamically due to low antenna position or mobility. Especially, a vehicle running at high speed may lose a signal, and the channel is saturated in a crowded environment. Currently, most vehicles use only one channel of WAVE or LTE using Telecommunication Control Unit (TCU). When one channel quality deteriorates, it is difficult to provide a stable communication link. In a vehicle communication service, functions such as urgent message transmission are associated with safety, so that minimization of delay time and reliability of information transmission should be guaranteed.

In this paper, we propose a vision sensing system and environment awareness algorithm for multi-channel V2X performance analysis. This algorithm can improve the V2X communication performance by intelligently selecting the channel and optimizing the parameters by using the environment information of the V2X communication vehicle using the multi-channel communication. Currently, most of the detection algorithms use an image-based deep learning algorithm to detect nearby objects. The main objects are pedestrians, cars, bicycles and so on. However, we need a new approach because we recognize the objects and environment that affect the V2X communication from the proposed algorithm. Therefore, we determine the objects that affect the communication performance and the driving situation. First, the objects that affect communication performance are relatively huge objects such as trucks, buses, buildings, and so on. Next, the driving environment that is fatal to communication passes through tunnels, uphill roads, under bridge piers, and so on. In this paper, we define the scenarios that affect the communication performance and evaluate the performance of the algorithm[10].

II. METHOD

A. Scenario

Conventional ambient awareness algorithm detect objects on the road. It mainly detects people, vehicles, bicycles, etc. [1-3]. The final goal is to detect and classify objects or create paths. Recently, 3D point cloud information is applied to deep learning algorithm along with image to estimate the direction of object [4-5]. However, in this paper, it is aimed to recognize the situation that affects communication performance. So we have defined a few scenarios. The following four scenarios affect communication.

- The environment in which the truck or bus is going higher than the antenna position of the vehicle ahead.

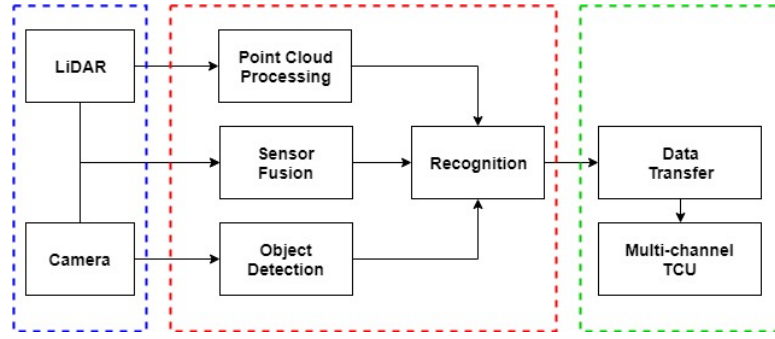- Environment with high buildings nearby.

Fig. 1.   This flowchart is our entire system. The blue area is a sensing system and used-two LiDAR and once one camera. The red area is the ambient environment recognition algorithm proposed in this paper. In the green region, the output of the algorithm is transfered to the TCU. The proposed method uses a sensing system to allow the TCU to select a channel.

- An environment where there are tunnels, piers, and soundproof walls on roads.
- A situation where a vehicle climbs uphill

Fig.1. is a flow chart of the entire system. It uses LiDAR and camera for obtaining the surrounding data. Then the proposed algorithm recognizes objects located in the surrounding environment. Based on the perceived information, TCU boards using multi-channel transmit appropriate parameter values to select the channel. The information of the object delivered by the proposed algorithm is as follows.

- Position; x, y, z, distance.
- Object ID; object class.
- State of object; Inactive or active.
- 3d box; width, length and height of box.
- Lane; Number of lane and driving lane in current.
- Path history; Position of x-frame, relative heading & speed.

B. Sensing System

In this paper, we used a camera and two 16-channel Velodyne LiDAR. The camera sensor provides multiple RGB data, but the detection area is relatively small. The LiDAR sensor provides point cloud data and has the 360 degrees detection area. However, there is a disadvantage in that the number of point clouds is small and it is difficult to identify an object at a distance. However, it provides accurate distance information and reflectance values. We fused the camera and the Lidar sensor into a single sensor. By integrating the coordinate systems of the sensors into one, we took advantage of each.

Fig. 1.   Algorithm experimental vehicle equipped with two VLP16 Puck-LITE and Camera. The camera is attached to the inside of the vehicle to acquire the front scene.
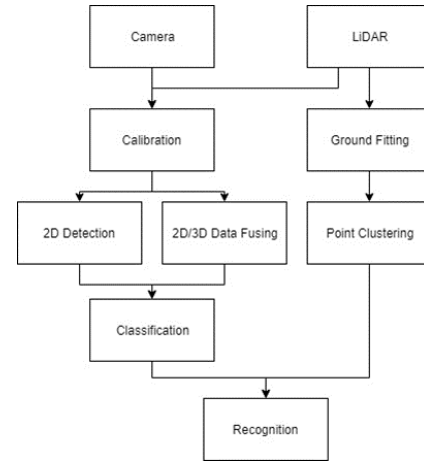
III.   IMPLEMETATION

Fig. 2.   Architecture of Recognition algorithm fusing Vision and LiDAR. This is the blue and red areas in Fig. 1.

Fig.2. shows the operation sequence of this algorithm. First, the coordinate system of the two sensors is integrated through the camera and LiDAR calibration. Second, the surrounding objects are detected using the 2D image information. The

distance and location information of the detected object is obtained using the fused information. Third, the data is processed by removing the ground from the point cloud. Then the clustering method is applied to classify each object. However, point clustering cannot determine the class of an object. Therefore, it is possible to extract objects that cannot be detected by the camera using the information extracted from the previously merged information. The extracted object is not detected due to the FOV of the camera. It is mainly a huge object such as a truck or a building. Finally, the class, position, and distance information of the object detected in the surrounding environment are used to predict communication performance.
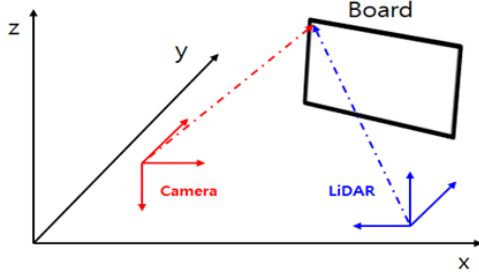
A. Calibration

Fig. 3. Camera and LiDAR coordinate system.

In order to utilize the two sensors as a single fusion system, a calibration process of matching the coordinate system between the sensors is required. There are many ways to integrate the LiDAR and camera sensors. The conventional method is to use the board to obtain the corresponding point [6-8]. Recently, a method of extracting feature values and matching them to fuse has been studied. In this paper, feature points are extracted from the data of each sensor using a calibration board from Fig. 3. The extracted feature points are analyzed to match corresponding points [9]. As a result, the transformation matrix can be obtained by using the relationship between the extracted corresponding points. The obtained transformation matrix projects LiDAR data to the pixel coordinates of the camera and uses the sensed data in one coordinate system. Equation (1) shows the relationship between the pixel coordinate matrix A and the transformation matrix M of the LiDAR coordinate matrix B.

$$\begin{pmatrix} u_i \\ v_i \\ 1 \end{pmatrix} = \begin{bmatrix} m_{11} & m_{12} & m_{13} & m_{14} \\ m_{21} & m_{22} & m_{23} & m_{24} \\ m_{31} & m_{32} & m_{33} & m_{34} \end{bmatrix} \begin{pmatrix} X_i \\ Y_i \\ Z_i \\ 1 \end{pmatrix} \qquad (1)$$

Fig. 4. Projection of a LiDAR point into the pixel coordinate system. The red dot is LiDAR point projected onto the image.

B. Detection

Object detection utilizes an image of a camera sensor that provides a large amount of information. In this paper, we applied the CNN (Convolutional Neural Network) based deep - learning algorithm which is specialized in image processing [5]. CNN-based detection algorithms are divided into two types, one is a one-stage detector and the other is a two-stage detector algorithm [1-3]. Object detection and recognition are distinguished according to whether they are performed in one neural network at a time or whether neural networks for object detection and neural networks for object recognition are separated. In this paper, we use a faster one-stage algorithm in terms of processing speed to detect nearby objects [1].

IV. RESULT

The experiment was conducted on two vehicles with LiDAR and a camera. The PC environment is a notebook equipped with an i7 CPU and 1070 GPU. The final frame is 25 frames and the maximum detection distance is 50m. The resulting value is the object information, distance and location. Image-based detection performance is the same as that of CNN-based detection algorithm.
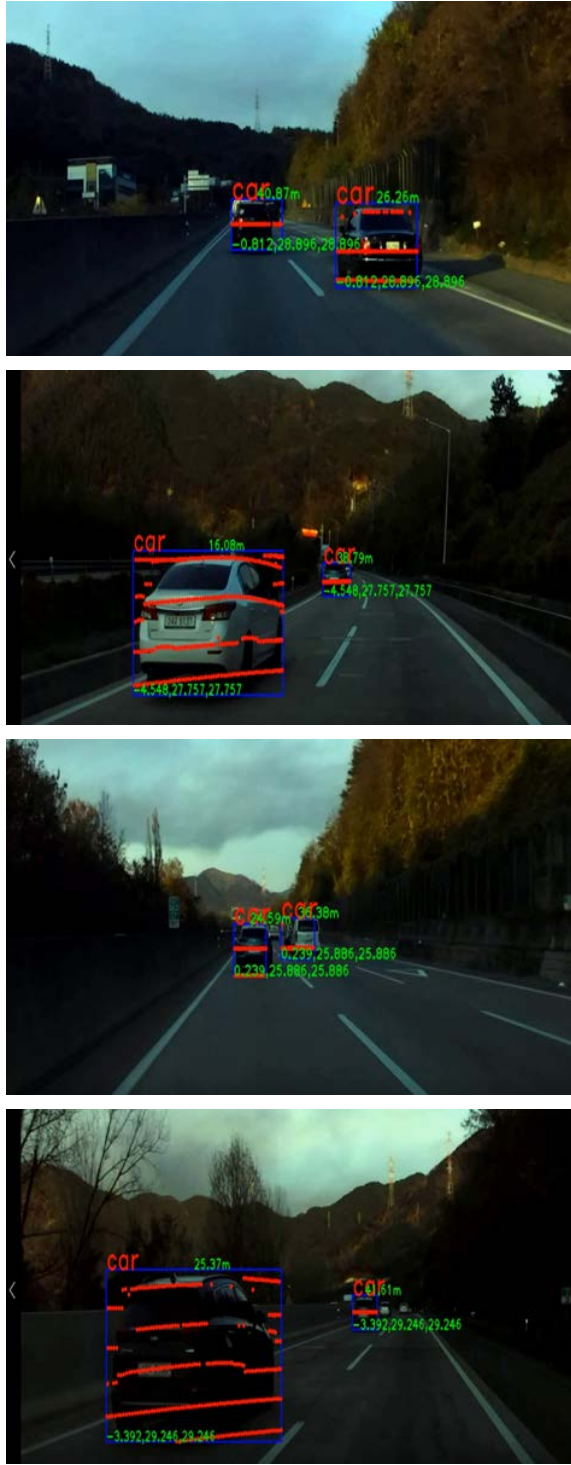
Fig. 5. Surrounding Environment Recognition Result. The output is the distance and coordinates of the objects.

V. CONCLUSION

The proposed method uses a camera and a LiDAR sensor as a single sensor. Fusion sensor detects objects that affect V2X communication. The output is the coordinates, distance and classification of the object. It improves communication performance by adjusting communication parameter based on object information. The proposed method can improve the stability and performance of multi-channel TCU board [10].

REFERENCES

[1] Redmon J., et al. "You Only Look Once: Unified, Real-Time Object Detection" Proceedings of the IEEE conference on computer vision and pattern recognition, 2016.

[2] S. Ren, K. He, R. Girshick, and J. Sun, "Faster r-cnn: Towards realtime object detection with region proposal networks," in Advances in Neural Information Processing Systems 28, pp. 91–99, 2015.

[3] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, and S. Reed, "SSD: Single shot multibox detector," arXiv:1512.02325, 2015.

[4] Jason Ku, Melissa Mozifian, Jungwook Lee, Ali Harakeh, and Steven Waslander. Joint 3d proposal generation and object detection from view aggregation. arXiv preprint arXiv:1712.02294, 2017.

[5] K. He and J. Sun. Convolutional neural networks at constrained time cost. In CVPR, 2015.

[6] X. Chen, H. Ma, J. Wan, B. Li, and T. Xia. Multi-view 3d object detection network for autonomous driving. In IEEE CVPR, 2017.

[7] G. Pandey, J. R. McBride, S. Savarese, and R. M. Eustice, "Automatic targetless extrinsic calibration of a 3D LiDAR and camera by maximizing mutual information," in Proc. 26th AAAI Conf. Artif. Intell., Toronto, ON, Canada, pp. 2053–2059, 2012.

[8] W. Wang, K. Sakurada, and N. Kawaguchi, "Reflectance intensity assisted automatic and accurate extrinsic calibration of 3d lidar and panoramic camera using a printed chessboard," Remote Sensing, vol. 9, no. 8, p. 851, 2017.

[9] Q. Zhang and R. Pless, "Extrinsic calibration of a camera and laser range finder (improves camera calibration)," in Intelligent Robots and Systems, 2004.(IROS 2004). Proceedings. 2004 IEEE/RSJ International Conference on, vol. 3, pp. 2301–2306, 2004.

[10] Hong-Jong Jeong and Sutaek Oh, "Development of Web-based V2X Field Test & Analysis System using Big Data Solution", 22nd ITS World Congress, ITS-2311, 2015.