IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received March 10, 2020, accepted March 22, 2020, date of current version May 5, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2985075

Camera and Radar Sensor Fusion for Robust Vehicle Localization via Vehicle Part Localization

DAEJUN KANG[1,2] AND DONGSUK KUM[1], (Member, IEEE)
[1]Korea Advanced Institute of Science and Technology, Daejeon 34141, South Korea
[2]Korea Automotive Technology Institute, Cheonan-si 31214, South Korea

Corresponding author: Dongsuk Kum (dskum@kaist.ac.kr)

This work was supported in part by the Industrial Core Technology Development Program and the Technology Innovation Program under grant 10070276, 10083646, funded by the Ministry of Trade, Industry and Energy.

ABSTRACT Many production vehicles are now equipped with both cameras and radar in order to provide various driver-assistance systems (DAS) with position information of surrounding objects. These sensors, however, cannot provide position information accurate enough to realize highly automated driving functions and other advanced driver-assistance systems (ADAS). Sensor fusion methods were proposed to overcome these limitations, but they tend to show limited detection performance gain in terms of accuracy and robustness. In this study, we propose a camera-radar sensor fusion framework for robust vehicle localization based on vehicle part (rear corner) detection and localization. The main idea of the proposed method is to reinforce the azimuth angle accuracy of the radar information by detecting and localizing the rear corner part of the target vehicle from an image. This part-based fusion approach enables accurate vehicle localization as well as robust performance with respect to occlusions. For efficient part detection, several candidate points are generated around the initial radar point. Then, a widely adopted deep learning approach is used to detect and localize the left and right corners of target vehicles. The corner detection network outputs their reliability score based on the localization uncertainty of the center point in corner parts. Using these position reliability scores along with a particle filter, the most probable rear corner positions are estimated. Estimated positions (pixel coordinate) are translated into angular data, and the surrounding vehicle is localized with respect to the ego-vehicle by combining the angular data of the rear corner and the radar's range data in the lateral and longitudinal direction. The experimental test results show that the proposed method provides significantly better localization performance in the lateral direction, with greatly reduced maximum errors (radar: 3.02m, proposed method: 0.66m) and root mean squared errors (radar: 0.57m, proposed method: 0.18m).

INDEX TERMS Bird's-eye view, vehicle localization, sensor fusion, lane change prediction, trajectory prediction, particle filter.

I. INTRODUCTION

A key enabler of recent DAS technologies is the drastically improved perception technologies owing to vision and radar sensors. These sensors have their own strengths and weaknesses. Table 1 summarizes the strengths and weaknesses of three popular sensors in terms of performance, cost, and robustness to environmental effects. For mass-production vehicles, cost is usually the dominant factor behind the sensor choice, and thus Lidar is rarely used for production cars despite its superior performance. Instead, radar and camera sensors are the two most widely used sensor choices. These two sensors work adequately when they are applied in a well-conditioned environment for low-level automated driving systems. However, in order to develop highly-advanced DAS such as a highly automated driving system, accurate position information is essential to predict the future motions of surrounding vehicles and to take appropriate actions while avoiding faulty ones [1], [2]. Unfortunately, radar and camera sensors do not provide sufficiently accurate position information when used individually.

The research on cameras in DAS has focused mainly on vehicle detection with a lighter calculation load and greater robustness to external environmental changes [3]–[12]. A few recent studies investigated the distance estimation problem [13]–[20]. These studies use a method that estimates

The associate editor coordinating the review of this manuscript and approving it for publication was Mohammad Ayoub Khan.

TABLE 1. Comparison of sensor characteristic.

Sensor	Range accuracy		Cost	Environmental affect
	Radial	Lateral		
Radar	○	□	△	△
Camera	□	○	○	□
Lidar	○	○	□	□

○: Good, △: Normal, □: Bad

the distance between a vehicle with ADAS installed, called an ''ego-vehicle'', and surrounding vehicles through geometric relationships. When using this method, a distance estimation error often occurs because the vehicle detection bounding box cannot always be fitted suitably. A detected vehicle is represented as a bounding box in an image, and the contact point is defined as the point that overlaps both the bounding box and the road plane in the image. The distance is then estimated using the geometric relationship between this contact point and a camera installed on the vehicle. However, since a camera has low resolution in the longitudinal direction, a slight position error of the bounding box causes large longitudinal distance variation. Additionally, due to the variation of the camera optical axis as the vehicle is driven, estimation errors based on geometric relationships are inevitable. Consequently, existing distance estimation methods using only a camera cannot represent the position of the vehicle precisely.

Radar also cannot always accurately provide precise position information. As summarized in Table 1, radar shows good range accuracy in the radial direction but poor performance in the lateral direction. Some studies on radar related to improving accuracy in lateral direction have been conducted [21]–[29]. However, limitations due to inherent characteristics still exist. Basically, a radar uses measured phase difference from more than two Rx antennas to estimate the angle of arrival of the object. The relationship between the phase and angle is nonlinear, and the sensitivity of the phase to angle degrades as the angle increases. Accordingly, when the angle is estimated using only radar, an error tends to occur as the angle increases. And since the position of peak reflectivity of the object is continuously changing in the tracking process, the resulting angle error also exists. Also, as shown in Fig. 1, radar is still limited when classifying the vehicle's corner part, which is needed to accurately localize the vehicle onto bird's-eye-view coordinates.

In order to complement the limitations of radar and cameras, the radar and camera sensor fusion algorithm has been studied. However, since most sensor fusion studies have focused on improving detection performance through cross-validation or reducing computation load, the weaknesses of each sensor are still not completely solved [30]–[37].

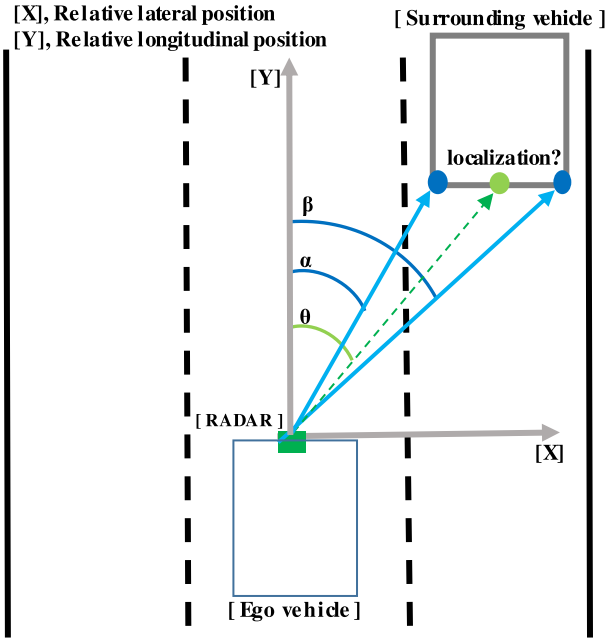

FIGURE 1. Problem of radar regarding vehicle localization in a bird's-eye view.

To achieve better localization performance, a few sensor fusion techniques have been proposed [34], [38]–[40]. In a prior study, a simple radar and camera coordinate transformation calibration method was proposed as a type of sensor fusion and has served as preliminary work for radar and camera sensor fusion studies [34]. In another study, to represent the position of the vehicle, the position information from each sensor is displayed on a grid cell in a bird's-eye-view and the vehicle position is derived by superimposing this information [38]. Because the vehicle position is displayed over several adjacent grids, uncertainty arises. In order to reduce uncertainty in vehicle localization and maintain continuity of the localization results, a Kalman filter is used to fuse both the radar and camera data [39]. This results in better performance, effectively reducing the noisy distance estimation data of each sensor in the radial direction. Moreover, other work sought to complement the high lateral position variance due to radar using the symmetry of the rear parts of the vehicle [40]. This approach can provide a relatively accurate lateral position in the case of a vehicle moving forward. However, it is limited when the vehicle is partially visible, as it relies on the symmetry of the rear contour of the vehicle. In a close-up case, unlike when a vehicle is moving forward, it is also difficult to apply symmetry detection due to the varying viewpoints. Overall, existing sensor fusion methods are effectively used only as the cross validation to increase detection accuracy, but there is a limitation to robustly localizing a vehicle in diverse driving environments including occlusion.

Therefore, in this study, we present a sensor fusion method which reinforces the azimuth angle accuracy of the radar data by localizing a vehicle's rear corner part using a camera. For vehicle part localization, the center position of the vehicle's

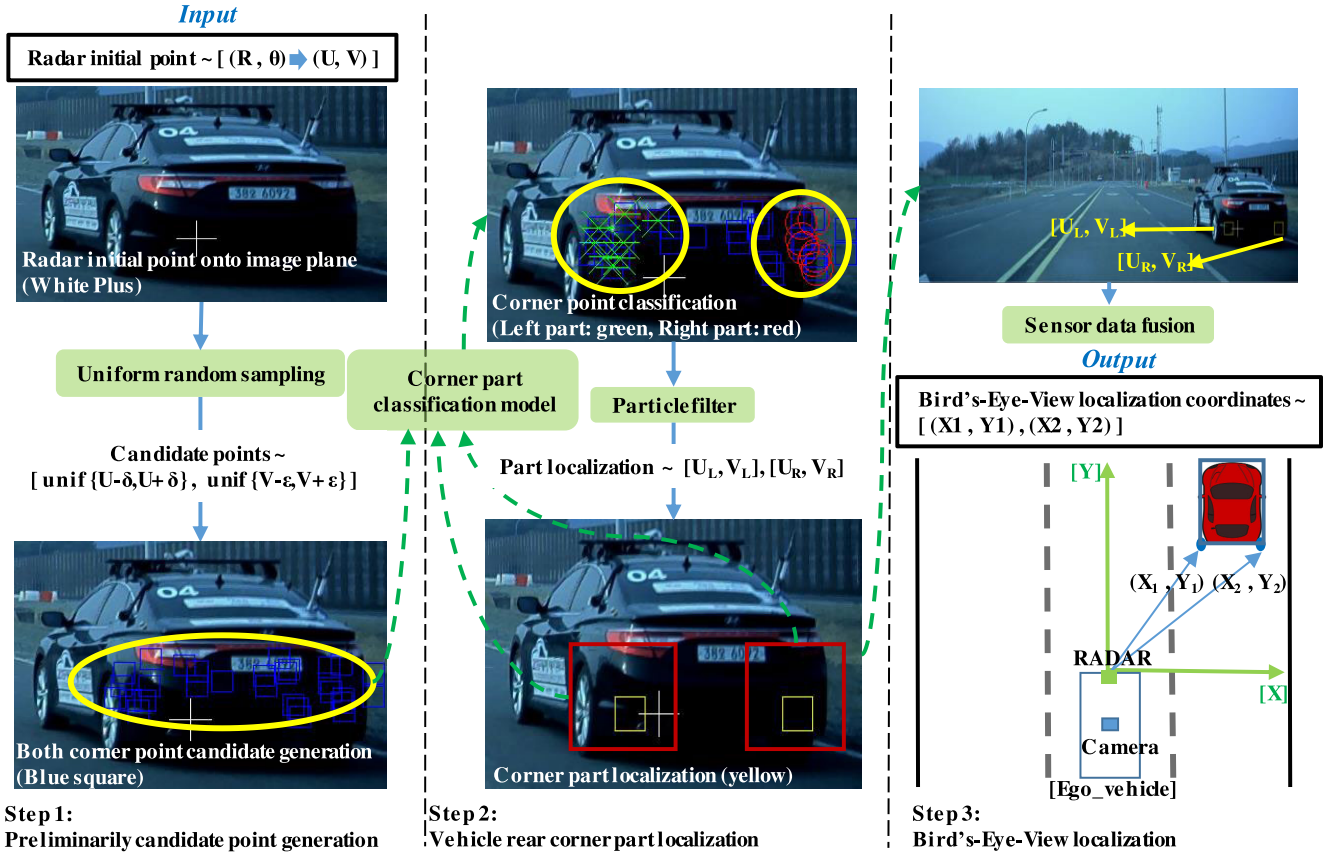

FIGURE 2. Overall work flow of this study.

rear corner is estimated and tracked based on a particle filter framework. Thus, through this part-based fusion approach, we can provide accurate vehicle localization as well as robust performance with respect to occlusions. In this way, the shortcomings of each sensor are compensated for and surrounding vehicles are localized accurately onto a bird's-eye view in actual driving environments.

The main contributions in this study can be summarized as follows:

First, our method shows accurate localization performance by estimating a vehicle's rear corner part robustly. The vehicle's rear corner part is classified robustly via a proposed vehicle part classification model and using these classified corner parts, the most probable position of the vehicle's rear corner is estimated accurately through a particle filter.

Second, our method shows reliable results when localizing the relative positions of surrounding vehicles under diverse driving conditions. Through tracking each vehicle's rear corner part separately based on particle filter framework, our method copes well with sudden partial observation driving situations such as cut-in, cut-out driving.

The overall workflow of this study is described in Fig. 2. and consists of the following three steps.

Step 1: Preliminarily candidate point generation

The radar data is translated to the image plane and preliminarily candidate points for classifying a vehicle's rear corner part are generated around the initially translated radar data.

Step 2: Vehicle rear corner part localization

Each corner part of the vehicle is classified from the candidate points, after which the most probable corner part position is estimated using the classified candidate points. Once the corner part position is estimated, the position is continuously tracked around the prior estimated position through an iterative process.

Step 3: Bird's-eye-view localization

Both estimated rear corner part positions in an image plane are combined with the range and angle data from the radar system, which is then translated to the coordinates of the vehicle's rear corner part in a bird's-eye-view.

The remaining sections covered in this paper are as follows. The vehicle's rear corner part localization on the image plane is described in Section II. In this section, an overview of the corner part localization method, the vehicle rear corner classification model, candidate point generation and the particle-filter–based framework for the extraction and tracking of the vehicle rear corner part position are explained in detail. In Section III, adding vehicle localization into a bird's-eye view is explained using both radar data and camera data. Section IV describes the test environment and presents

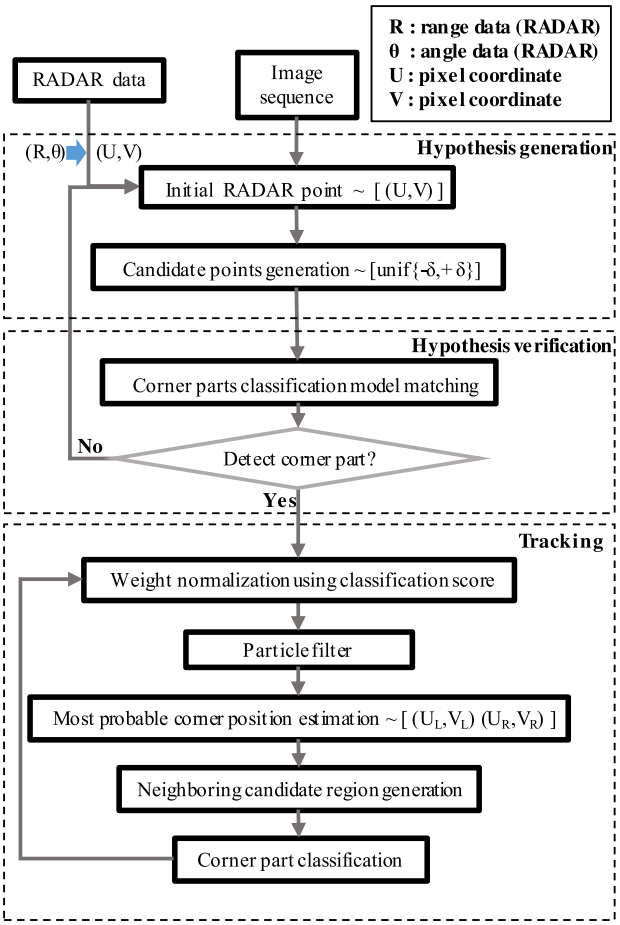

FIGURE 3. Both rear corner part localization sequence.

experimental results on the accuracy of the surrounding vehicle localization process. Finally, the conclusion and future works are presented in Section V.

II. VEHICLE REAR CORNER PART LOCALIZATION

A. OVERVIEW

During the second step shown in Fig. 2, both rear corner part positions are estimated and tracked on an image plane in order to localize the vehicle in the lateral direction with respect to the relative coordinates of the ego-vehicle. Even when either the left or right corner part is not visible due to occlusion, one corner part is continuously tracked to realize a close approximation of the position of the vehicle. For part localization, transformation of the radar data to the image plane should take precedence. Subsequently, the candidate points are generated around the initial radar point with a uniform distribution to classify both rear side parts of the vehicle, efficiently reducing the processing time. From among the candidate points, the vehicle's rear corner parts on the left and right are identified through a classification model pre-trained on both corner parts, with the classified points having their weights assigned from their corresponding classification scores. Using the position and weight of the classified points, the most probable left and right side

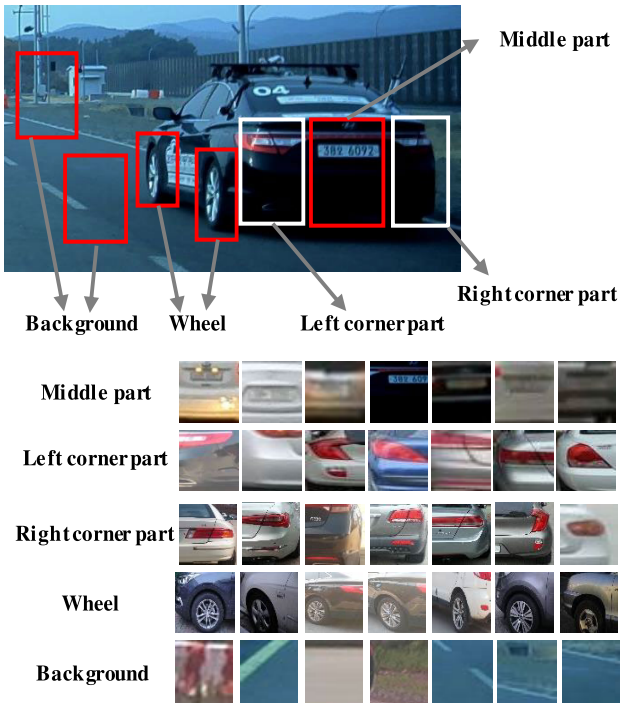

FIGURE 4. Training image samples in this study.

positions are estimated. Once the probable position is estimated, the candidate points are regenerated around the prior estimated position to track both rear corner parts of a vehicle, and each corner part is tracked through an iterative process. The surrounding vehicles then localized onto a bird's eye view based on the rear corner parts of the tracked vehicle. The workflow to track the most probable left and right side positions of the surrounding vehicles in the image sequences is described in Fig. 3.

B. DEEP LEARNING-BASED VEHICLE PART CLASSIFICATION MODEL

1) TRAINING DATASET

In order to classify a vehicle's rear corner part, we must establish a dataset which is composed of various classes. Thus, we define five classes including other parts of the vehicle in the vicinity of the corner. Each of these classes is defined in Fig 4. In total, 29,185 images of vehicles are included in the dataset: some of them from Stanford which contains 16,185 images of vehicles [41]. Using those images, we define the class of the vehicle part and train it to classify the rear corner.

2) DEEP LEARNING MODEL FOR REAR CORNER CLASSIFICATION

A pre-trained model is used to train a vehicle rear corner part classification model. In this study, we use the VGG16 network [42] as a pre-trained model, which is simple but shows good performance in many applications of image classification. Since our defined dataset has five classes, we have

TABLE 2. Confusion matrix for test dataset.

Target / Output	Class 1 (Middle part)	Class 2 (Left corner part)	Class 3 (Right corner part)	Class 4 (Wheel)	Class 5 (Background)	Accuracy
Class 1 (Middle part)	425	4	3	2	0	97.9
Class 2 (Left corner part)	0	296	0	4	0	98.7
Class 3 (Right corner part)	5	3	235	1	0	96.3
Class 4 (Wheel)	0	2	2	198	2	97.1
Class 5 (Background)	0	0	0	0	212	100
Accuracy	98.8	97	97.9	96.6	99.1	98

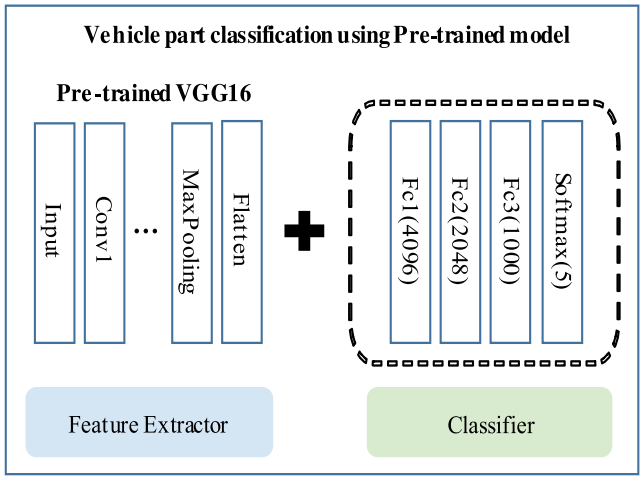

FIGURE 5. Pre-trained classification model for vehicle rear corner classification.

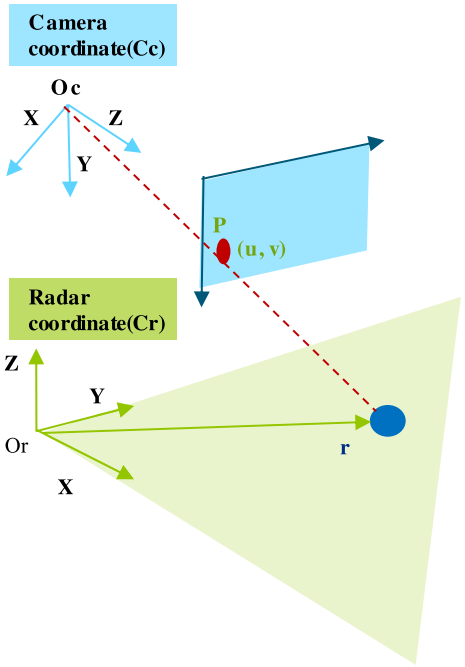

FIGURE 6. The relationship between the camera coordinates and radar coordinates.

to replace the output layer of VGG16 to customize it to the dataset. Therefore, for customizing we replace the classifier of the VGG16 network to the newly added classifier as shown in Fig 5. Finally, to confirm the performance of the trained detection model, classification is carried out throughout the test dataset. The total simulation result shows 98 percent classification accuracy for the test dataset. The accuracy of each class classification as a confusion matrix in the test dataset is shown in Table 2.

C. CANDIDATE POINT GENERATION FOR REAR CORNER POINTS BASED ON RADAR DATA

1) DATA TRANSFORMATION

a: GENERALIZED DIRECT LINEAR TRANSFORM

In order to use the radar data in an image, appropriate coordinate transformation must be done initially. Thus, the direct linear transform (DLT) method is employed to translate the radar point to the corresponding point on the image plane. The radar and image coordinate systems are defined in Fig. 6, where the transformation matrix between two coordinate systems is defined as (1).

$$\begin{bmatrix} xi \\ yi \\ wi \end{bmatrix} = \begin{bmatrix} h11 & h12 & h13 & h14 \\ h21 & h22 & h23 & h24 \\ h31 & h32 & h33 & h34 \end{bmatrix} \begin{bmatrix} Xi \\ Yi \\ Zi \\ 1 \end{bmatrix} \quad (1)$$

where $[x_i, y_i, w_i]$ is the set of homogeneous coordinates for the image point; $[h_{11}, \ldots, h_{34}]$ is an element of the camera matrix; $[X_i, Y_i, Z_i, 1]$ is the set of homogeneous coordinates for the radar point. Assuming that all radar points are on the plane ($Z = 0$), the transformation matrix can be represented

as projective transformation (H) as (2).

$$\begin{bmatrix} xi \\ yi \\ wi \end{bmatrix} = \begin{bmatrix} h11 & h12 & h13 \\ h21 & h22 & h23 \\ h31 & h32 & h33 \end{bmatrix} \begin{bmatrix} Xi \\ Yi \\ 1 \end{bmatrix} \tag{2}$$

This can be rearranged as (3) for the pixel coordinates u and v.

$$\begin{aligned} ui &= \frac{xi}{wi} = \frac{h11 * Xi + h12 * Yi + h13}{h31 * Xi + h32 * Yi + h33} \\ vi &= \frac{xi}{wi} = \frac{h11 * Xi + h12 * Yi + h13}{h31 * Xi + h32 * Yi + h33} \end{aligned} \tag{3}$$

For each point i, (3) can be rewritten as a polynomial with respect to h, as shown in (4).

$$\begin{aligned} &[-X_i \ -Y_i \ -1 \ 0 \ 0 \ 0 \ u_iX_i \ u_iY_i \ u_i] \begin{bmatrix} h_{11} \\ h_{12} \\ h_{13} \\ h_{21} \\ h_{22} \\ h_{23} \\ h_{31} \\ h_{32} \\ h_{33} \end{bmatrix} \rightarrow A_i^T h = 0 \\ &[0 \ 0 \ 0 \ -X_i \ -Y_i \ -1 \ v_iX_i \ v_iY_i \ v_i] \begin{bmatrix} h_{11} \\ h_{12} \\ h_{13} \\ h_{21} \\ h_{22} \\ h_{23} \\ h_{31} \\ h_{32} \\ h_{33} \end{bmatrix} \rightarrow B_i^T h = 0 \end{aligned} \tag{4}$$

In conclusion, to obtain each element h of the homography matrix, the equation can be defined as a least square problem, as shown in (5).

$$\begin{aligned} &\begin{bmatrix} A_i^T \\ B_i^T \\ \dots \\ A_k^T \\ B_k^T \end{bmatrix} h = Ph = w, \quad For\ each\ data\ point(i \dots k) \\ &h = \arg\min_h w^T w \end{aligned} \tag{5}$$

b: CALIBRATION DATASET

In order to solve the least square problem for the homography matrix defined as (5), the distance from the radar sensor to the object (corner reflector) was measured and the data was matched manually with the pixel coordinates of the object in an image. A test sample image for the calibration dataset is shown in Fig. 7. Like Fig. 7, data was collected from 19 separate points which are arranged in Table 3.

FIGURE 7. Sample image of the calibration dataset.

TABLE 3. Radar and camera calibration dataset.

Sample point	X lateral direction(m)	Y longitudinal direction(m)	U pixel coordinate	V pixel coordinate
1	-3.113	6.227	86	570
2	-1.121	6.351	435	567
…	…	…	…	…
18	3.238	32.255	776	445
19	7.597	32.006	964	454

2) REAR CORNER CANDIDATE POINT GENERATION

The initial candidate points for classifying the rear corner parts of the surrounding vehicles are generated around the initial radar points. Each initial radar point is converted into pixel coordinates in an image through the homography matrix, calibrated using the process described above. The candidate points are then generated around the pixel coordinates of the radar points with a uniform distribution. In each case, the mean of the uniform distribution is the pixel coordinate of the radar point and the size of the interval in the uniform distribution is dependent on pixel coordinate v in the image plane, as defined in (6). In accordance with this process, generated candidate points around the initial radar point are shown in Fig. 8.

$$\begin{aligned} S &= \{S1, S2, \dots, Sn\} \\ S(i) &= \begin{cases} S(i) - k, & \text{if } S(i) > size\ of\ image\ frame \\ unif(m-\sigma, m+\sigma), & otherwise \end{cases} \\ \sigma &= \alpha \cdot v + \beta \end{aligned} \tag{6}$$

where, $S(i)$ is a candidate point; m is the mean of the uniform distribution; σ is the interval size of the uniform distribution which is set differently according to the v pixel coordinate; α, β and k are hyper parameters which are adjusted to the image frame size; The small n represents sample numbers that is defined as 30 in this study.

D. REAR CORNER PART CLASSIFICATION FROM CANDIDATE POINTS

Once the candidate points are generated according to the initial radar point, the trained vehicle rear corner part classification model is allocated to a window box with a rectangular shape centered on the candidate points. Because the closer vehicle is located in the lower region of an image, the size of

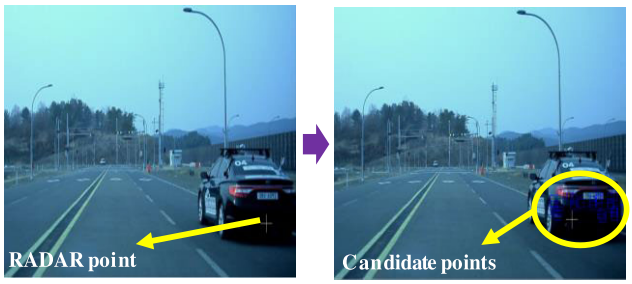

FIGURE 8. Candidate point generation using radar points: (a) initial radar point, and (b) generated candidate points around the initial radar point.

FIGURE 9. Classification result for both vehicle rear corner parts.

the window box for classification depends on the position of the candidate points. The classification score for the window box is then calculated through the allocated classification model to find the rear left and right corner parts of the vehicle. Despite the fact that these parts may be classified as a corner part initially, they are regarded as false positives if the classification score is lower than a predefined threshold value. Through the procedures described above, vehicle rear corner parts are classified for the right and left part in each respective case among the candidate points. Figure 9 shows the classified rear left and right corner parts of the vehicle. Among the generated candidate points, the green 'star' represents the points classified as left corner parts and the red 'circle' represents the points classified as right corner parts.

E. TRACKING THE REAR CORNER POINT USING A PARTICLE FILTER

In order to extract the best possible position of the rear corner among the classified rear corner parts, a particle filter framework is employed. A point regarded as the left or right corner among all candidate points has its weight assigned from the classification score based on the classification model. The weights are normalized and an arbitrary probability distribution is generated based on these values. Thus, the most suitable left and right corner part positions are estimated based on the generated arbitrary probability distribution. In order to keep track of more suitable positions in consecutive frames, the search points are regenerated around the estimated position of the prior frame and the best possible position is estimated through an iterative process among the customized regions. The position estimation and tracking process based on the particle filter framework for vehicle rear corner parts is described in Algorithm 1.

Algorithm 1: Rear corner part position tracking algorithm

Input: $X = \{< x^{[j]}, w^{[j]} >\}$ (Paricle set)
Output: $\overline{P}$ (*Estimated vehicle rear side part position*), X
1: $\overline{P} = \emptyset$
2: *for* $j = 1$ *to* J *do*
3: $\quad p_t^{[j]} = \frac{w_i^{[j]}}{\left\| w_i^{[j]} \right\|}$
4: $\quad \overline{P}_t = \overline{P}_t + \sum c_t^{[j]} \cdot p_t^{[j]}$
5: *endfor*
6: *return* $\overline{P}_t$, $X_t = \emptyset$
7: *for* $k = 1$ *to* K *do*
8: $\quad$ *sample* ${x_{t+1}}^{[k]} \sim N(\overline{P}_t, \sigma^2)$
9: $\quad {w_{t+1}}^{[k]} \sim$ *classification score of* k^{th} *sample*
10: $\quad$ *If* ${w_{t+1}}^{[k]}$ *to* ${X_{t+1}}^{[k]}$
11: $\quad\quad$ *draw* k *with probability* ${w_{t+1}}^{[k]}$
12: $\quad\quad$ *add* ${x_{t+1}}^{[k]}$ *to* X_{t+1}
13: $\quad$ *endif*
14: $\quad X_{t+1} = X_{t+1} + < {x_{t+1}}^{[k]}, {w_{t+1}}^{[k]} >$
15: *endfor*
16: *return* X_{t+1}

Through Algorithm 1, the positions $[(U_L, V_L), (U_R, V_R)]$ of both vehicle rear corner parts can be extracted and tracked, as shown in Fig. 10. By tracking the position of each corner part of a vehicle in this way, it becomes possible to estimate the position of the vehicle even when only a part of the vehicle is visible due to the camera's field of view or occlusion. Moreover, by rapidly reducing the search regions to the estimated position, the computing load can be greatly reduced. If there is no region detected as the left and right-side parts in the initial candidate group, it is initialized as the first time and is searched again in the candidate regions generated based on the next radar data input.

III. VEHICLE LOCALIZATION

A. ASSUMPTION IN THE RADAR DATA

Figure 11 demonstrates the limitation of the commercial radar output data. Commercial radar transmits the output data in the form of points, including a radial range and angle information; however, it does not include information about where the point is on the vehicle, as shown in Fig. 11. Therefore, in this study we assume that the point lies between the left corner and the right corner of the vehicle rear part. In this case, even if the radar angle shows a large deviation, the vehicle position in the longitudinal direction is kept constant ($\boldsymbol{R} * \boldsymbol{cos}\,(\theta)$).

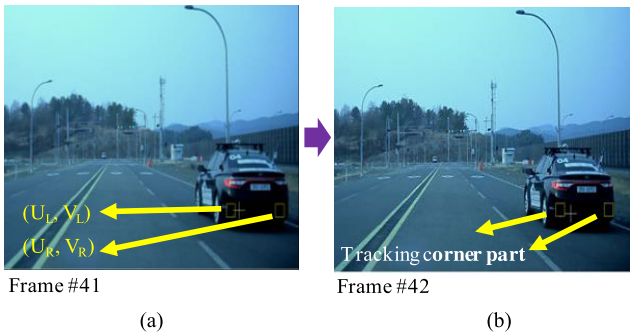

FIGURE 10. Most probable position estimation and tracking: (a) pixel coordinates of rear left & right side parts, and (b) tracking for the rear left and right corner parts.

Where, R is the relative distance between an ego-vehicle and surrounding vehicles with respect to the radial direction; θ is the relative angle between an ego-vehicle and surrounding vehicles with reference to the center front of the ego-vehicle.

B. LOCALIZATION USING RADAR AND CAMERA DATA

The vehicle position with respect to the ego-vehicle is calculated based on both the radar output data and the spatial resolution of the camera. The vehicle position in the longitudinal direction is calculated from the radar output data directly, and is combined with the angular data of the rear corner part to calculate the vehicle's lateral position.

The pixel coordinates from the estimation of the vehicle rear corner part position on the image plane are translated into angular data with respect to the optical axis, as shown in Fig. 11(a). Because the radar and camera are aligned in the same direction, the longitudinal position of the vehicle with respect to the camera can be calculated simply by adding the difference in the length (T) between the radar and the camera, as shown in Fig. 11. Accordingly, the vehicle position in the lateral direction is calculated as shown in Fig. 11(b).

IV. PERFORMANCE EVALUATION

A. TEST SCENARIO

In this study, the performance of the proposed method is evaluated in terms of the accuracy of the relative position between the ego-vehicle and the target vehicle. The radar system used as the main part of the sensor system in a typical ADAS offers mainly poor performance, especially regarding the 2D reconstruction of all edges of the surrounding vehicles. In order to solve this problem, we developed a relative position estimation method for the surrounding vehicles using sensor fusion based on mono-vision and radar. To evaluate the proposed method, datasets were generated through test driving. In the test datasets, surrounding vehicles were positioned under diverse conditions, such as in the left lane, middle lane, and right lane, to reflect the relative locations of the surrounding vehicles sufficiently in an actual driving environment. An occlusion case is also included which consists

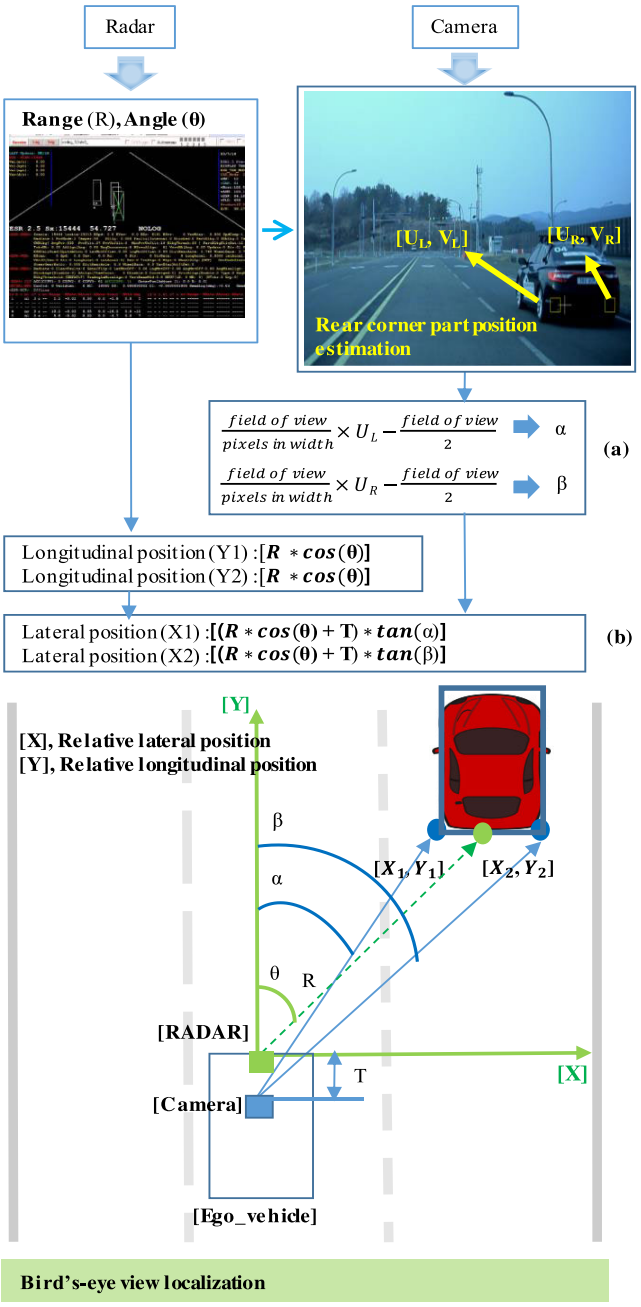

FIGURE 11. Bird's-eye-view localization using radar and a camera: (a) translation pixel coordinates to angular data, and (b) calculation of the lateral position.

of observations of partial occlusions due to the camera's field of view and other vehicles.

B. TEST ENVIRONMENT CONFIGURATION

In order to obtain the reference data (ground truth) for a performance evaluation of the proposed method, three vehicles (one ego-vehicle and two target vehicles) with systems installed (RT3002 and RT-Range (Oxford Technology System)) were set up, as shown in Fig. 12. The data from RT was calibrated based on the center of the front bumper of the ego-vehicle, which is the origin coordinate of the radar system as

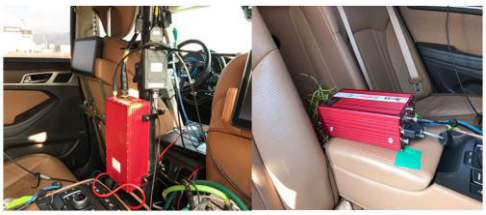

FIGURE 12. Test vehicle with RT3002 and RT-Range.

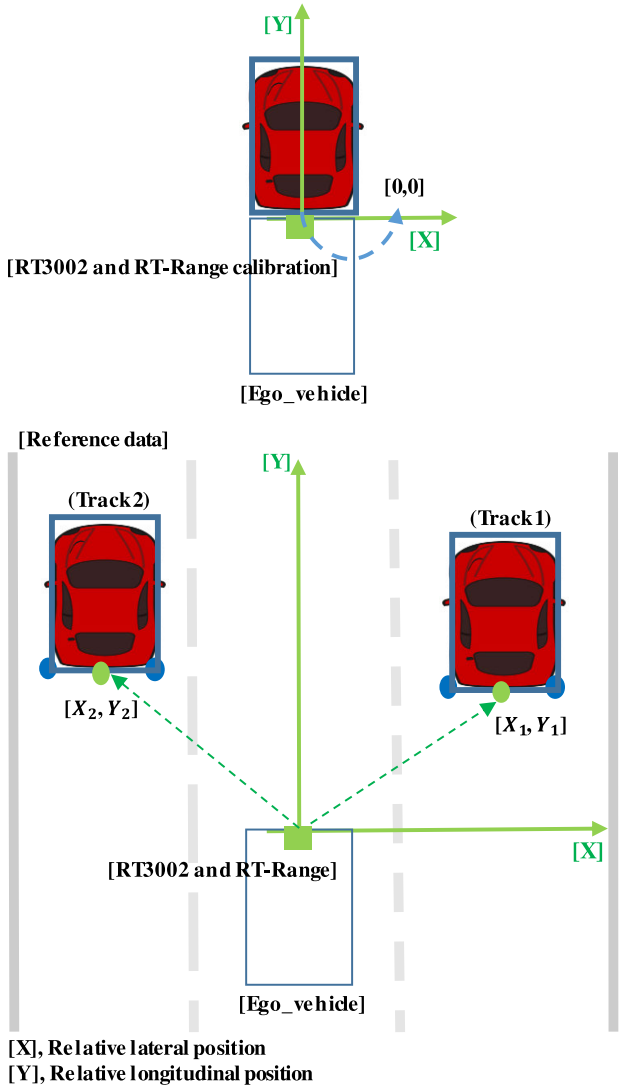

FIGURE 13. The reference data calibration and acquisition from RT equipment.

described in Fig. 13. The data accuracy of the RT equipment regarding the relative position of the target vehicle is 0.03m in the longitudinal, lateral range. Through the RT3002 and RT-Range systems, the relative position data between the ego-vehicle and the target vehicles can be measured as shown in Fig. 13. In order to evaluate the proposed method compared to reference data, the ego-vehicle was equipped with a radar system (Delphi ESR), a camera and the RT-equipment as

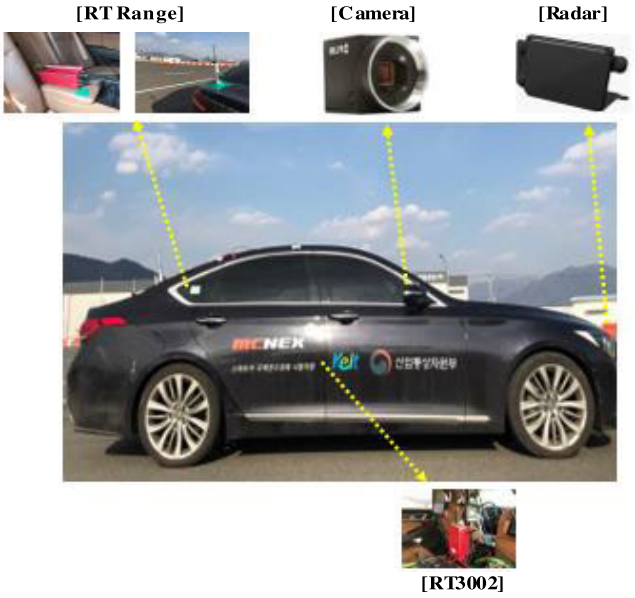

FIGURE 14. Ego-vehicle sensor configuration.

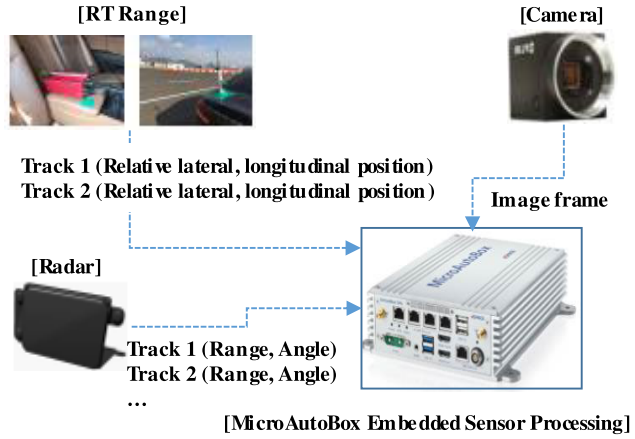

FIGURE 15. The data interface of each sensor.

shown in Fig. 14. The data interface of each sensor and sensor data processing unit are described in Fig. 15.

For data synchronization between heterogeneous devices with different data update and output times, the default data logging time is set to 20Hz (50ms), which is the minimum update interval value of radar. Because the RT signal is received via CAN communication, it is matched to the radar signal through the receiving period. For the vision data, the time reference is transmitted to the image processing board through the CAN signal every 50ms, which is the logging time of the radar data, and data synchronization is performed.

C. SURROUNDING VEHICLES LOCALIZATION PERFORMANCE

The surrounding vehicles are shown as track 1 and track 2 and are classified as the left lane, middle lane, and right lane

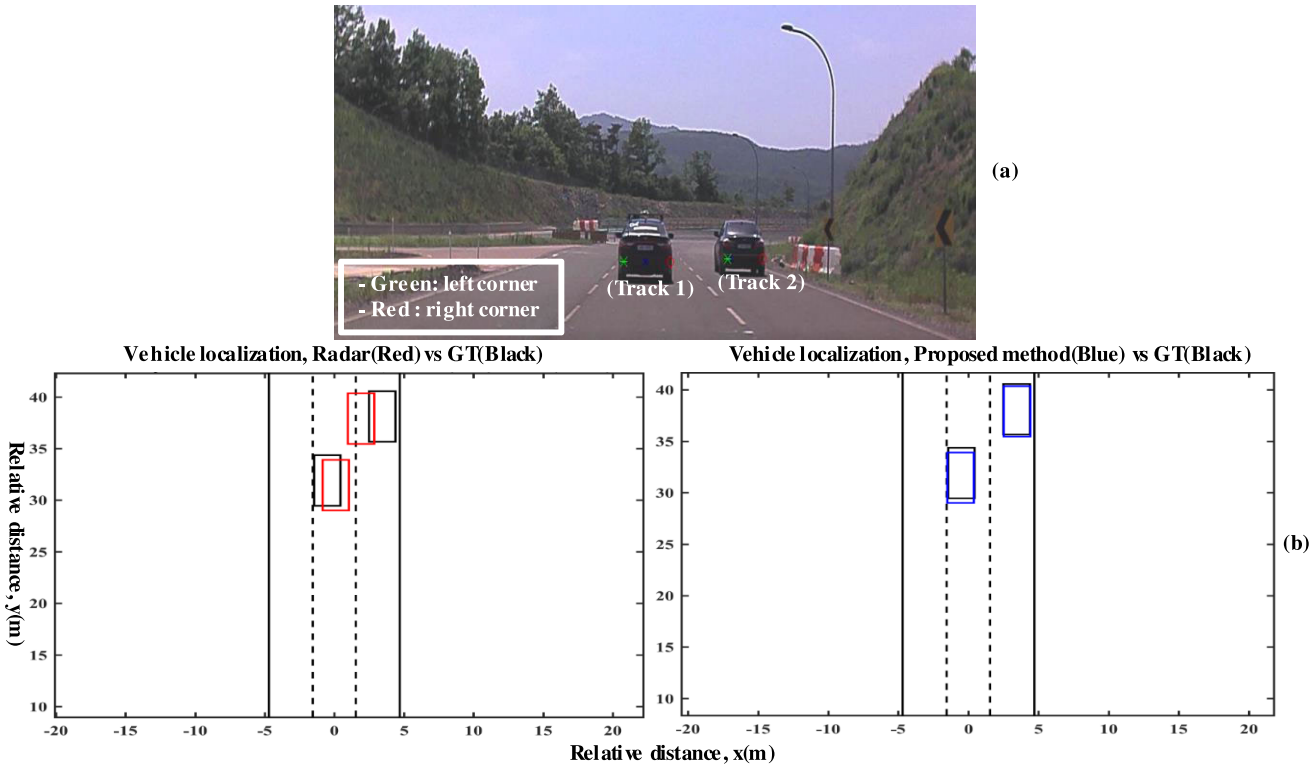

FIGURE 16. Vehicles are in the middle and right lanes: (a) both left and right corner part tracking, (b) bird's-eye view vehicle localization (radar vs proposed method) (left: radar (red), right: proposed method (blue)), ground truth (black).

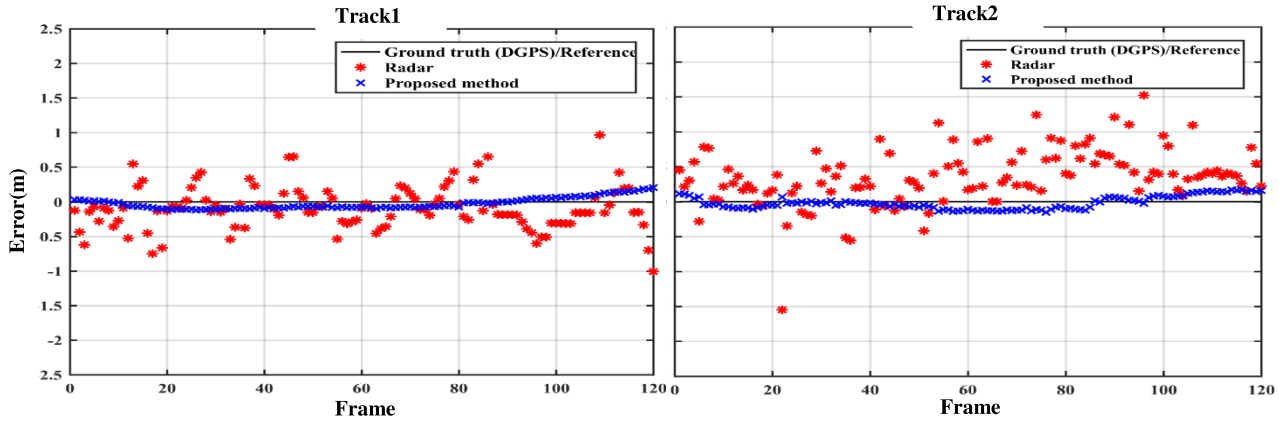

FIGURE 17. Distance error in lateral direction in 1) (Middle lane, Right lane), (Red: Radar, Blue: Proposed method).

depending on their position relative to the ego-vehicle. In each test case, bird's-eye view localization outcomes using both the proposed method and the radar tracking data are compared with the ground truth. Moreover, each lateral position error with respect to the ground truth is shown through bird's-eye view localization. When localizing surrounding vehicles onto the bird's-eye view, it is assumed that the radar tracking data is centered on the rear part of the vehicle and that the width of the vehicle is known.

The test results of 1), 2), and 3) below have been uploaded to YouTube [43].

1) MIDDLE LANE, RIGHT LANE (NORMAL DRIVING SITUATION)

In a normal driving situation without vehicle occlusion, it is confirmed that the proposed method tracks both corner parts of the vehicle, as shown in Fig. 14(a). Bird's-eye view vehicle localization based on this rear corner part tracking significantly reduces lateral position errors compared to the tracking data of the radar system. In Fig. 15, the radar tracking data (track 2) shows a maximum error of 1.5m, which could lead to a serious malfunction in an ADAS.

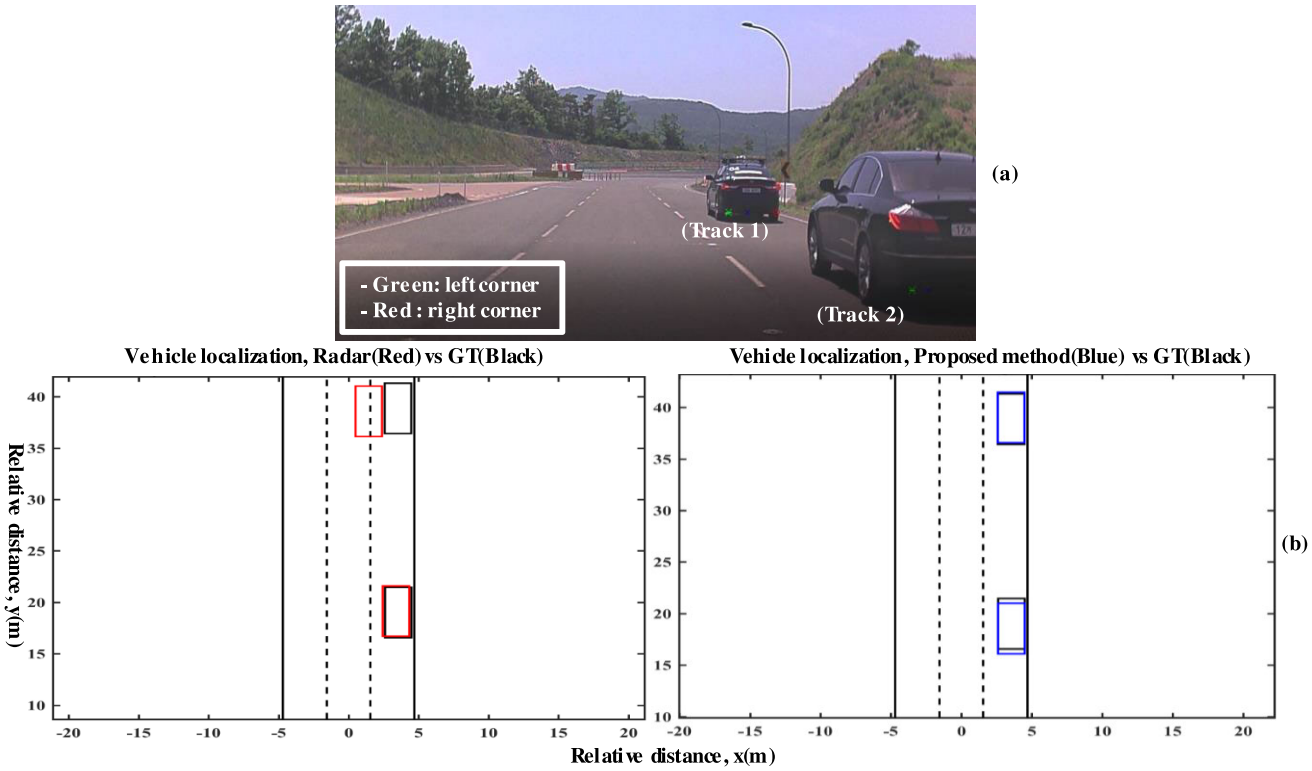

FIGURE 18. Vehicles are in the right lane: (a) both left and right corner part tracking, and (b) bird's-eye view vehicle localization (radar vs proposed method) (left: radar (red), right: proposed method (blue)), ground truth (black).

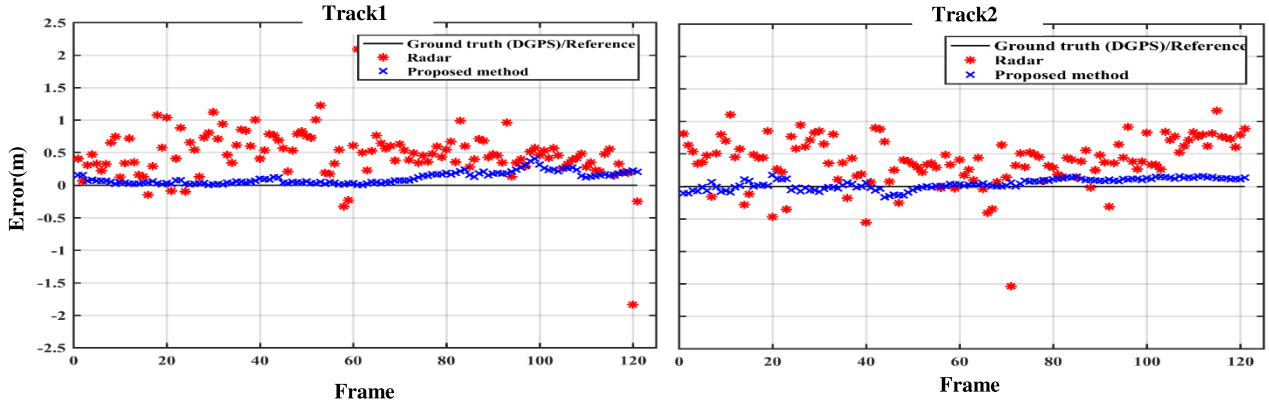

FIGURE 19. Distance error in the lateral direction in 2) (right lane), (red: radar, blue: proposed method).

2) RIGHT LANE (OCCLUSION DUE TO THE CAMERA FIELD OF VIEW)

For an adjacent vehicle in the right lane or left lane, it must be localized more precisely for an ADAS to make the correct decision. However, as shown in Fig. 17, the radar data for an adjacent vehicle (track 2) shows a large variation about −1.5m to 1m in the lateral direction. This large variation makes the ego-vehicle unable to distinguish which lane the adjacent vehicle is in. On the other hand, the proposed method shows a small variation regarding an adjacent vehicle (track 2) of about -0.2m to 0.1m in the lateral direction. It works well for either left or right corner part tracking even when the vehicle is partially visible due to the camera's field of view, as shown in Fig. 16(a). In this case, the left corner part is tracked, as represented by the green 'star' in Fig. 16(a). Using this corner part localization, bird's-eye view vehicle localization is carried out. Because the left and right corner parts are classified during the tracking process, surrounding vehicles can be localized precisely onto the bird's-eye view via one corner part localization.

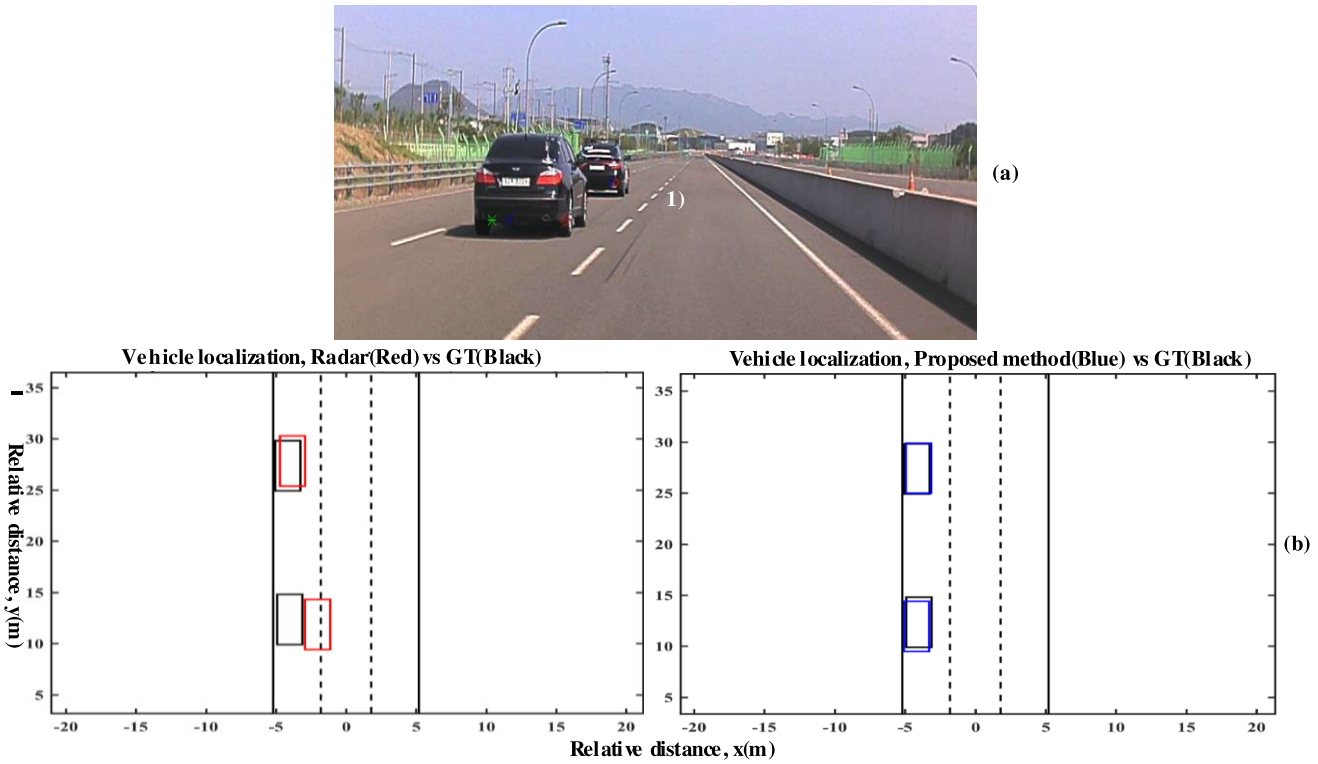

FIGURE 20. Vehicles are in the left lane: (a) both left and right corner part tracking, and (b) bird's-eye view vehicle localization (radar vs proposed method) (left: radar (red), right: proposed method (blue)), ground truth (black).

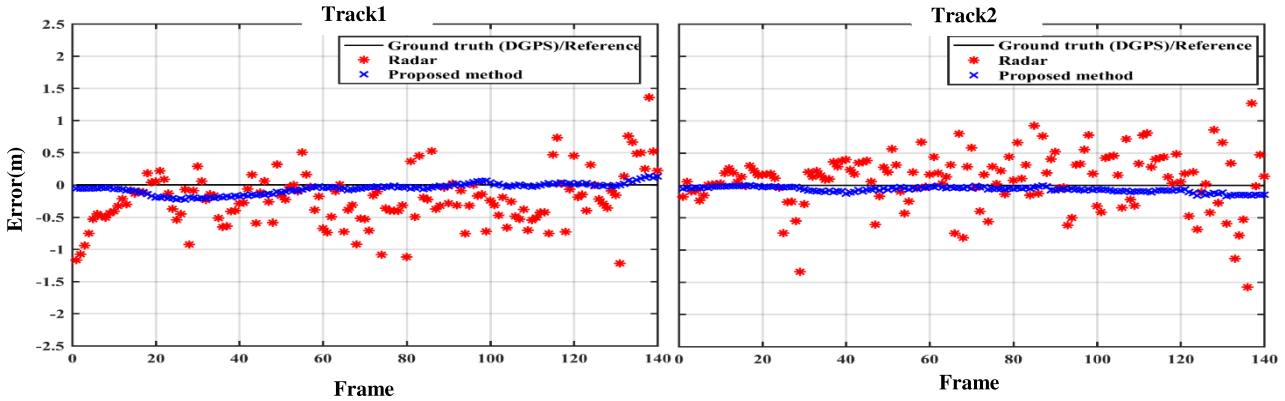

FIGURE 21. Distance error in the lateral direction in 3) (left lane), (red: radar, blue: proposed method).

3) LEFT LANE (OCCLUSION DUE TO OTHER VEHICLE)

Similarly, for the surrounding vehicles in the left lane, as described in the previous case, the right corner part is classified and tracked for the occluded vehicle (track 1), as represented by the red 'circle' in Fig. 18(a), even when the vehicle is partially visible due to other vehicles. The occluded vehicle (track 1) is localized precisely onto the bird's-eye view via the localization of this one corner part. In Fig. 19, the proposed method shows an estimation distance error for track 1 of approximately −0.15m to 0.1m in the lateral direction. However, the radar tracking data shows a large variation for track 1 of about −1.25m to 1.4m in the lateral direction.

4) RESULTS OF AN ANALYSIS ACCORDING TO THE RELATIVE POSITIONS OF SURROUNDING VEHICLES

In order to analyze the position estimation results according to the relative position of the surrounding vehicles, the relative positions of the surrounding vehicles in the test cases are summarized by lane, with the lateral position accuracy results depending on the lane, as shown in Fig. 20 and Table 4. As indicated by the test results, the radar data with a maximum lateral position error of 2.3∼3.02m may cause a crucial fault because it cannot detect the lane of the surrounding vehicle. Also, as shown by the total Root Mean Square Error (RMSE, 0.5773), application is difficult to the lateral control system in the ADAS due to the consecutive position

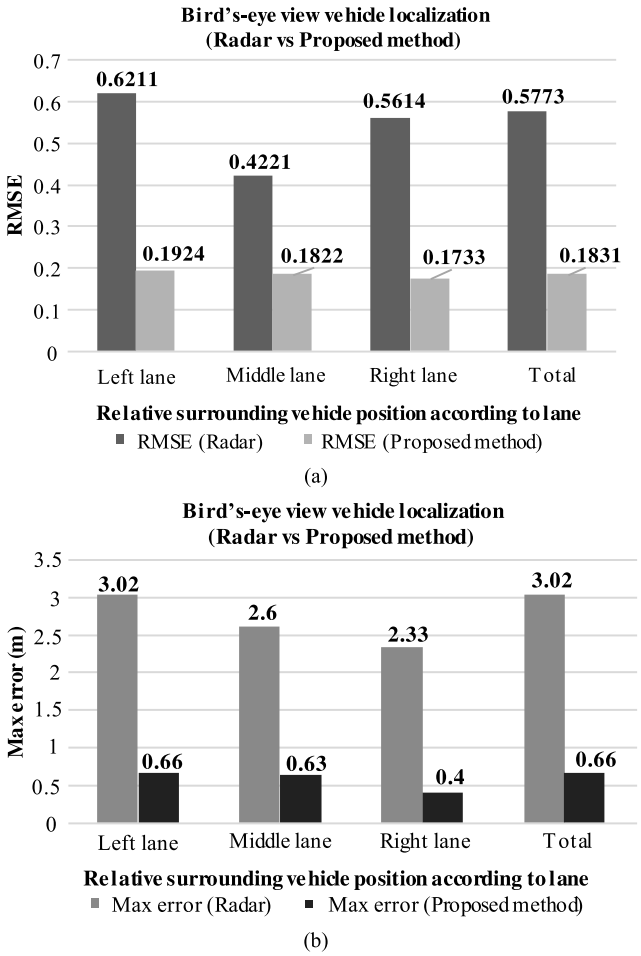

FIGURE 22. Lateral position accuracy during bird's-eye view vehicle localization according to the lane: (a) RMSE, and (b) max error (m).

TABLE 4. Lateral position accuracy in Bird's-eye view vehicle localization according to the lane.

Case	Bird's-eye view localization method	RMSE	Max error(m)
Left lane	Radar	0.6211	3.02
	Proposed method	0.1924	0.66
Middle lane	Radar	0.4221	2.6
	Proposed method	0.1822	0.63
Right lane	Radar	0.5614	2.33
	Proposed method	0.1733	0.4
Total	Radar	0.5773	3.02
	Proposed method	0.1831	0.66

fluctuations of the surrounding vehicles. The maximum error and total RMSE of the proposed method are 0.66m and 0.1831, respectively. It is confirmed that the lanes of the surrounding vehicles can be accurately discriminated regardless of the relative positions of the surrounding vehicles. In addition, because the rear left and right corners of the vehicle are classified and separately extracted through the proposed method, it is possible to estimate the positions regardless of partial occlusion when driving.

V. CONCLUSION AND FUTURE WORKS

The relative position estimation of surrounding vehicles around an ego-vehicle are important for the safety of an ADAS and automated vehicles. In order to develop a highly automated driving system, not a semi-automated driving system, accurate localization of surrounding vehicles is essential. Accurate relative position estimations can allow motion predictions of the surrounding vehicles to help an automated vehicle prevent accidents due to unexpected situations. However, a sensor system using radar and camera has limitations when accurately localizing surrounding vehicles onto a bird's-eye view in the actual driving environment. In that sense, the proposed method shows accurate and reliable localization performance with diverse driving test condition considering actual driving environment, not the processed virtual dataset from a conventional dataset. The outcomes here can also be sufficiently extended to solve problems concerning motion predictions of surrounding vehicles and the coordinated signal control system using it [44]. However, additional research on relative position estimations of surrounding vehicles should be carried out for robust estimations. Issues related sensor fusion should also be assessed regarding the occurrence of radar range data errors. If there is an error of radar range data, the localization performance corresponding to both longitudinal and lateral direction will be degraded. Therefore, additional research is needed to cope with this error.

REFERENCES

[1] J.-H. Kim and D.-S. Kum, "Threat prediction algorithm based on local path candidates and surrounding vehicle trajectory predictions for automated driving vehicles," in *Proc. IEEE Intell. Vehicles Symp. (IV)*, Jun. 2015, pp. 1220–1225.
[2] S. Yoon and D. Kum, "The multilayer perceptron approach to lateral motion prediction of surrounding vehicles for autonomous vehicles," in *Proc. IEEE Intell. Vehicles Symp. (IV)*, Jun. 2016, pp. 1307–1312.
[3] H. Zhu, K.-V. Yuen, L. Mihaylova, and H. Leung, "Overview of environment perception for intelligent vehicles," *IEEE Trans. Intell. Transp. Syst.*, vol. 18, no. 10, pp. 2584–2601, Oct. 2017.
[4] A. Mukhtar, L. Xia, and T. B. Tang, "Vehicle detection techniques for collision avoidance systems: A review," *IEEE Trans. Intell. Transp. Syst.*, vol. 16, no. 5, pp. 2318–2338, Oct. 2015.
[5] S. Sivaraman and M. M. Trivedi, "Looking at vehicles on the road: A survey of vision-based vehicle detection, tracking, and behavior analysis," *IEEE Trans. Intell. Transp. Syst.*, vol. 14, no. 4, pp. 1773–1795, Dec. 2013.
[6] Z. Sun, G. Bebis, and R. Miller, "On-road vehicle detection: A review," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, no. 5, pp. 694–711, May 2006.
[7] J. Zhou, J. Duan, and H. Yu, "Machine-vision based preceding vehicle detection algorithm: A review," in *Proc. 10th World Congr. Intell. Control Autom.*, Jul. 2012, pp. 4617–4622.
[8] C.-C. Chiu, M.-L. Chung, and W.-C. Chen, "Real-time front detection algorithm for an asynchronous binocular system," *J. Inf. Sci. Eng.*, vol. 26, pp. 735–752, Sep. 2010.
[9] B. Tian, B. T. Morris, M. Tang, Y. Liu, Y. Yao, C. Gou, D. Shen, and S. Tang, "Hierarchical and networked vehicle surveillance in ITS: A survey," *IEEE Trans. Intell. Transp. Syst.*, vol. 18, no. 1, pp. 25–48, Jan. 2017.
[10] R. Labayrade, D. Aubert, and J.-P. Tarel, "Real time obstacle detection in stereovision on non flat road geometry through 'v-disparity' representation," in *Proc. Intell. Vehicle Symp.*, Jun. 2002, pp. 646–651.

[11] S. Sivaraman and M. M. Trivedi, "A general active-learning framework for on-road vehicle recognition and tracking," *IEEE Trans. Intell. Transp. Syst.*, vol. 11, no. 2, pp. 267–276, Jun. 2010.
[12] S. S. Teoh and T. Bräunl, "Symmetry-based monocular vehicle detection system," *Mach. Vis. Appl.*, vol. 23, no. 5, pp. 831–842, Sep. 2012.
[13] L.-C. Liu, C.-Y. Fang, and S.-W. Chen, "A novel distance estimation method leading a forward collision avoidance assist system for vehicles on highways," *IEEE Trans. Intell. Transp. Syst.*, vol. 18, no. 4, pp. 937–949, Apr. 2017.
[14] C.-C. Lin and M.-S. Wang, "A vision based top-view transformation model for a vehicle parking assistant," *Sensors*, vol. 12, no. 4, pp. 4431–4446, 2012.
[15] H.-K. Cheung, W.-C. Siu, S. Lee, L. Poon, and C.-S. Ng, "Accurate distance estimation using camera orientation compensation technique for vehicle driver assistance system," in *Proc. IEEE Int. Conf. Consum. Electron. (ICCE)*, Jan. 2012, pp. 227–228.
[16] A. Joglekar, D. Joshi, R. Khemani, S. Nair, and S. Sahare, "Depth Estimation Using Monocular Camera," *Int. J. Comput. Sci. Inf. Technol.*, vol. 2, no. 4, pp. 1758–1763, 2011.
[17] S. Lessmann, M. Meuter, D. Muller, and J. Pauli, "Probabilistic distance estimation for vehicle tracking application in monocular vision," in *Proc. IEEE Intell. Vehicles Symp. (IV)*, Jun. 2016, pp. 1199–1204.
[18] K.-Y. Park and S.-Y. Hwang, "Robust range estimation with a monocular camera for vision-based forward collision warning system," *Sci. World J.*, vol. 2014, Jan. 2014, Art. no. 923632.
[19] M. Rezaei, M. Terauchi, and R. Klette, "Robust vehicle detection and distance estimation under challenging lighting conditions," *IEEE Trans. Intell. Transp. Syst.*, vol. 16, no. 5, pp. 2723–2743, Oct. 2015.
[20] F. Gökçe, G. Üçoluk, E. Şahin, and S. Kalkan, "Vision-based detection and distance estimation of micro unmanned aerial vehicles," *Sensors*, vol. 15, no. 9, pp. 23805–23846, 2015.
[21] J. Shi, G. Hu, X. Zhang, F. Sun, and H. Zhou, "Sparsity-based two-dimensional DOA estimation for coprime array: From sum–difference coarray viewpoint," *IEEE Trans. Signal Process.*, vol. 65, no. 21, pp. 5591–5604, Nov. 2017.
[22] J. Shi, G. Hu, X. Zhang, and F. Sun, "Sparsity-based DOA estimation of coherent and uncorrelated targets with flexible MIMO radar," *IEEE Trans. Veh. Technol.*, vol. 68, no. 6, pp. 5835–5848, Jun. 2019.
[23] F. Wen, X. Zhang, F. Yang, and Z. Zhang, "Direction finding in bistatic MIMO radar with unknown spatially colored noise," in *Proc. IEEE 23rd Int. Conf. Digit. Signal Process. (DSP)*, Nov. 2018, pp. 1–13.
[24] F. Wen, J. Shi, and Z. Zhang, "Joint 2D-DOD, 2D-DOA, and polarization angles estimation for bistatic EMVS-MIMO radar via PARAFAC analysis," *IEEE Trans. Veh. Technol.*, vol. 69, no. 2, pp. 1626–1638, Feb. 2020.
[25] J.-J. Lin, Y.-P. Li, W.-C. Hsu, and T.-S. Lee, "Design of an FMCW radar baseband signal processing system for automotive application," *SpringerPlus*, vol. 5, no. 1, p. 42, Dec. 2016.
[26] S. Patole, M. Torlak, D. Wang, and M. Ali, "Signal processing for smart vehicle technologies: Part 2," *IEEE Signal Process. Mag.*, vol. 33, no. 6, p. 22, Mar. 2017.
[27] N. A. Baig and A. Hussain, "RADAR signal processing for target range, Doppler and DoA estimation," in *Proc. 14th Int. Bhurban Conf. Appl. Sci. Technol. (IBCAST)*, Jan. 2017, pp. 820–825.
[28] W. Yurong, W. Congling, L. Xunbo, and Z. Chen, "Study on automotive anti-collision radar system and its signal processing algorithm," in *Proc. Int. Forum Inf. Technol. Appl.*, May 2009, pp. 586–590.
[29] P. Gupta and S. P. Kar, "MUSIC and improved MUSIC algorithm to estimate direction of arrival," in *Proc. Int. Conf. Commun. Signal Process. (ICCSP)*, Apr. 2015, pp. 757–761.
[30] X. Wang, L. Xu, H. Sun, J. Xin, and N. Zheng, "On-road vehicle detection and tracking using MMW radar and monovision fusion," *IEEE Trans. Intell. Transp. Syst.*, vol. 17, no. 7, pp. 2075–2084, Jul. 2016.
[31] H.-T. Kim and B. Song, "Vehicle recognition based on radar and vision sensor fusion for automatic emergency braking," in *Proc. 13th Int. Conf. Control, Autom. Syst. (ICCAS)*, Oct. 2013, pp. 1342–1346.
[32] M. Bertozzi, L. Bombini, P. Cerri, P. Medici, P. C. Antonello, and M. Miglietta, "Obstacle detection and classification fusing radar and vision," in *Proc. IEEE Intell. Vehicles Symp.*, Jun. 2008, pp. 608–613.
[33] S. Sugimoto, H. Tateda, H. Takahashi, and M. Okutomi, "Obstacle detection using millimeter-wave radar and its visualization on image sequence," in *Proc. 17th Int. Conf. Pattern Recognit. (ICPR)*, Aug. 2004, pp. 342–345.
[34] S. Han, X. Wang, L. Xu, H. Sun, and N. Zheng, "Frontal object perception for intelligent vehicles based on radar and camera fusion," in *Proc. 35th Chin. Control Conf. (CCC)*, Jul. 2016, pp. 4003–4008.
[35] A. Haselhoff, A. Kummert, and G. Schneider, "Radar-vision fusion for vehicle detection by means of improved Haar-like feature and Adaboost approach," in *Proc. 15th Eur. Signal Process. Conf.*, Sep. 2007, pp. 2070–2074.
[36] Z. Ji, M. Luciw, J. Weng, and S. Zeng, "Incremental online object learning in a vehicular radar-vision fusion framework," *IEEE Trans. Intell. Transp. Syst.*, vol. 12, no. 2, pp. 402–411, Jun. 2011.
[37] L. Bombini, P. Cerri, P. Medici, and G. Alessandretti, "Radar-vision fusion for vehicle detection," in *Proc. Int. Workshop Intell. Transp.*, vol. 70, Mar. 2006, p. 65.
[38] D. Gao, J. Duan, X. Yang, and B. Zheng, "A method of spatial calibration for camera and radar," in *Proc. 8th World Congr. Intell. Control Autom.*, Jul. 2010, pp. 6211–6215.
[39] Y. Feng *et al.*, "Distance estimation by fusing radar and monocular camera with Kalman filter," SAE Tech. Paper 2017-01-1978, 2017, doi: 10.4271/2017-01-1978.
[40] M. Nishigaki, S. Rebhan, and N. Einecke, "Vision-based lateral position improvement of RADAR detections," in *Proc. 15th Int. IEEE Conf. Intell. Transp. Syst.*, Sep. 2012, pp. 90–97.
[41] J. Krause, M. Stark, J. Deng, and L. Fei-Fei, "3D object representations for fine-grained categorization," in *Proc. IEEE Int. Conf. Comput. Vis. Workshops*, Dec. 2013, pp. 554–561.
[42] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," 2014, *arXiv:1409.1556*. [Online]. Available: http://arxiv.org/abs/1409.1556
[43] *Test Result for Robust Vehicle Localization*. Accessed: Jan. 2, 2019. [Online]. Available: https://www.youtube.com/channel/UCVvHzDhJOqZ4AdPa0BK30Gg
[44] C. Ma, W. Hao, A. Wang, and H. Zhao, "Developing a coordinated signal control system for urban ring road under the vehicle-infrastructure connected environment," *IEEE Access*, vol. 6, pp. 52471–52478, 2018.

DAEJUN KANG received the B.S. degree in industrial engineering and the M.S. degree in mechanical engineering from the Korea Advanced Institute of Science and Technology (KAIST), Daejeon, South Korea, in 2006 and 2009, respectively, and the Ph.D. degree from the Graduate School for Green Transportation, KAIST, in 2015. He is currently a Senior Researcher with the Korea Automotive Technology Institute (KATECH). His research interests are sensor fusion, motion prediction, and collision avoidance in autonomous driving vehicles.

DONGSUK KUM (Member, IEEE) received the Ph.D. degree in mechanical engineering from the University of Michigan, Ann Arbor, MI, USA, in 2010.

He had worked with the General Motors Research and Development Propulsion Systems Research Laboratory, Warren, MI, USA, as a Visiting Research Scientist. He is currently an Associate Professor with the Cho Chun Shik Graduate School for Green Transportation, Korea Advanced Institute of Science and Technology (KAIST), and the Director of the Vehicular Systems Design and Control Laboratory. His also works at General Motors focused on advanced propulsion system technologies including hybrid electric vehicles, flywheel hybrid, and waste heat recovery systems. His research centers on the modeling, control, and design of advanced vehicular systems with particular interests in hybrid electric vehicles and autonomous vehicles.