

Research Article

In-Lane Localization and Ego-Lane Identification Method Based on Highway Lane Endpoints

Kyoungtaek Choi ¹, Jae Kyu Suhr ¹ and Ho Gi Jung ²

¹School of Intelligent Mechatronics Engineering, Sejong University, 209 Neungdong-ro, Gwangjin-gu, Seoul 05006, Republic of Korea

²Department of Electronic Engineering, Korea National University of Transportation, 50 Daehak-ro, Chungju-si, Chungbuk 27469, Republic of Korea

Correspondence should be addressed to Ho Gi Jung; hogijung@ut.ac.kr

Received 15 July 2019; Revised 2 December 2019; Accepted 28 December 2019; Published 1 February 2020

Academic Editor: Hocine Imine

Copyright © 2020 Kyoungtaek Choi et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The low-cost global navigation satellite systems combined with an inertial navigation system (GNSS/INS) used most frequently for vehicle localization show errors up to 10 m, approximately, even in open-sky environments such as highways. To reduce this error on highways, this paper proposes a localization method based on lane endpoints. Since a lane endpoint frequently appears on a road and can be detected in close proximity even by a low-cost monocular camera, it is a very useful landmark for precise localization. However, the lane width is generally less than 3.5 m, and the localization error from the GNSS is about 10 m. Therefore, if an ego-lane is not identified, the lane endpoints detected in an ego-lane can be falsely corresponded to the lane endpoints in the other lane of a map. This paper proposes an in-lane localization method that uses lane endpoints, the relation between a camera and a road, and the estimated vehicle's orientation from a map. In addition, this paper proposes an ego-lane identification method that generates a hypothesis about an ego vehicle position per lane by using the proposed in-lane localization method and verifies each hypothesis by the projection of lane endpoints and an additional landmark such as a road sign. The average error of the proposed in-lane localization method is 0.248 m on highways. The success rate of the proposed ego-lane identification method is 99.28% by one trial and reaches 100% by fusing the results.

1. Introduction

Vehicle localization is used to estimate the current global position of a vehicle and is one of the core components in autonomous driving [1]. The most widely adapted localization systems are global navigation satellite systems (GNSSs) which estimate their position through multilateration with satellite signals [2]. The precision of a GNSS can be degraded by the diffused reflection of signals on skyscrapers, signal blocking in tunnels, or atmospheric signal distortion. Cooperative positioning (CP) robust to atmospheric signal distortion was proposed [3]. CP compensates for signal distortion by sharing the signal distortion information between several receivers. The representative systems of the CP approach are differential GNSS (DGNSS),

satellite-based augmentation systems (SBASs), ground-based augmentation systems (GBASs), and real-time kinematic (RTK). And to reduce the localization error caused by signal diffused reflection or signal blocking, a GNSS/INS that combines GNSS and an inertial navigation system (INS) has been researched [4]. The low-cost GNSSs generally installed on mass-produced vehicles have localization errors of up to 10 m even in an open-sky environment such as a highway. However, more expensive GNSS/INS with very precise INS and RTK GPS can keep their localization errors to less than the width of a lane even in urban areas where the diffused reflection of signals or signal blocking often occurs [5]. However, this type of system is not affordable for mass-produced vehicles because of its high price, and even precise INS cannot avoid cumulative errors, too. Moreover, an

autonomous driving vehicle requires a localization system whose precision is under tens of centimeters rather than tens of meters [6].

To improve localization precision up to tens of centimeters by compensating the localization error from a GNSS/INS, map matching-based localization systems have been broadly researched [7–19]. A map matching-based localization system recognizes a landmark such as a building or a road facility through perception sensors and estimates the vehicle's global position by finding the correspondence of the landmark in a digital map. As a perception sensor, sensors such as Lidar, radar, and stereo cameras are mainly used to capture 3D information and a monocular camera for capturing 2D information. The map-matching based localization system can be categorized into a feature point-based and road facility-based approach according to the types of landmarks used. The feature point-based approach utilizes feature points extracted on any surrounding static objects as landmarks [7–14]. The road facility-based approach recognizes a road facility such as a lane marking and a road sign as a landmark [15–19]. The former can be applied in any place but increases tremendously the volume of a digital map to store a lot of feature points according to the size of place. Because of this, this approach is mainly used in limited areas such as an indoor place or parking lot. The road facility-based approach can be utilized only on a road, but the volume of a digital map is relatively small because the map stores only the information of road facilities. Most road facilities such as lane markings, road signs, and traffic signs are strictly maintained by transport authorities to be easily recognized by a driver, and since their shapes are standardized, the essential amount of information to be stored (ex: four vertices of a road sign and a center position of a traffic sign) is small. For these reasons, the road facility-based approach is effective for vehicle localization on a road.

This paper deals with the road facility-based vehicle localization approach on highways. In order to recognize road facilities, this paper utilizes a forward-looking monocular camera installed in most advanced driver assistance systems (ADASs) because of its low cost and high usability. The major road facility used in this paper is the endpoint of a dashed lane marking, as shown in Figure 1. The lane endpoint frequently appears on roads, and their usability to compensate the localization error of a GNSS/INS is high. In particular, it appears over almost the entire section of a highway and, unlike lane markings, provides information on both lateral and longitudinal positions [20]. Moreover, since it can be captured at a close distance, its position accuracy is guaranteed. According to the driving direction of a road, the lane endpoint is classified into a lane starting point and a lane ending point that are shown as a green and a red point, respectively, in Figure 1. According to Korean regulations, on highways, the lane width is 3.5 m, the length of a dashed lane marking is 8.0 m, and the longitudinal distance between two dashed lane markings is 12.0 m, as shown in Figure 1. Therefore, the same type of endpoints is at least 20 m apart in the longitudinal direction and at least 3.5 m apart in the lateral direction. Since the localization error of a low-cost GPS on a highway is under 10 m, there is no possibility that

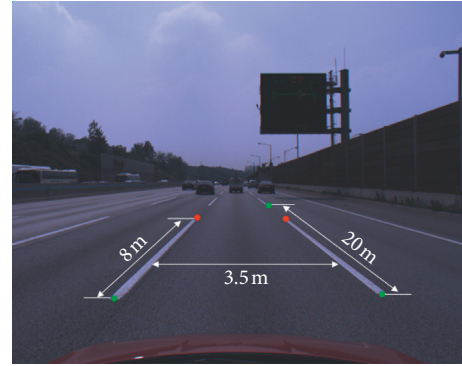


FIGURE 1: Lane endpoints and the printed regulation of dashed lane markings on a highway.

the detected endpoint falsely corresponds with the endpoint stored in a map in the longitudinal direction. However, if the ego-lane is not identified on a map, a detected endpoint can falsely correspond with the endpoint stored on a map in the lateral direction. Therefore, this paper proposes the framework to solve the vehicle localization in highways by dividing the problem into in-lane localization and ego-lane identification. In the proposed framework, in-lane localization using lane endpoints reduces the m-level localization uncertainty from GNSS/INS to the cm-level uncertainty on each lane by generating the camera hypothesis on each lane. And the hypothesis on each lane has been kept by the localization filter such as Kalman or particle filter until the ego-lane is identified by using an additional landmark.

The major contributions of this paper are as follows:

- (1) It proposes an efficient framework that divides the localization problem into the in-lane localization and the ego-lane identification
- (2) The proposed method achieves an average of 24.8 cm localization precision in highway
- (3) The proposed method is practical enough to be applied to a mass-produced vehicle since it requires a low-volume digital map and uses only low-cost sensors: a monocular camera and a low-cost GNSS/INS

Since the localization filter has been well introduced in other references, this paper describes only the system overview, the in-lane localization, and the ego-lane identification. The rest of this paper is organized as follows. Section 2 explains the previous works related to the ego-lane identification. Section 3 describes the proposed in-lane localization and the proposed ego-lane identification. Section 4 presents the experimental results and analyses. Finally, this paper concludes with future works in Section 5.

2. Previous Ego-Lane Identification Research

The previous works for ego-lane identification can be categorized into the vehicle to vehicle (V2V) communication-based approach, perception sensor-based approach, and map-guided perception sensor-based approach. The V2V

communication approach identifies the ego-lane by exchanging the relative positions between an ego-vehicle and others through the communication line [21, 22]. This approach can identify the ego-lane of all vehicles equipped with a communication device. However, there should be surrounding vehicles with a communication device on every lane. Therefore, as V2V communication is not in common usage, this system is suitable only for vehicle platooning.

Most of the perception sensor-based approaches detect road boundaries and identify the ego-lane by measuring the lateral offset of an ego-vehicle from the road boundary [23–25]. Lee et al. proposed a method that extracts multiple road characteristics (lane marking type, lane marking color, number of lanes, and lateral offset of each lane from a road boundary) and identifies an ego-lane by feeding these characteristics into the Bayesian network (BN) [23]. There is also a method that estimates lane width and the ego-lane lateral offset from a road boundary through a Kalman filter [24]. Kim and Park assume that a traffic sign is close to a road boundary. Their proposed method detects a traffic sign with a stereo camera to estimate its 3D position and identifies an ego-lane by estimating the lateral offset from the traffic sign to an ego-vehicle [25]. There is a multisensor fusion method that combines the surrounding vehicles detected by a Lidar and lane markings detected by a camera [26]. The perception sensor-based approach is easy to apply to mass-produced vehicles because it does not require any communication device or digital map. However, when the overall width of a road is too large to be within the sensor detection range or the road is too crowded for its lane markings to be observed, it is hard to recognize the road characteristics such as the number of lanes and the road boundary. Moreover, the methods to infer an ego-lane with the relative position from surrounding vehicles like the V2V communication approach have the limitation that there should be a vehicle in each lane [26].

The map-guided perception sensor-based approach utilizes both a perception sensor and a digital map to identify an ego-lane. Ballardini et al. define the difference between the detected lane number and the lane number recorded on a map as a sensor failure and identify an ego-lane through hidden Markov model (HMM) considering this sensor failure [27]. Nedeveschi et al. proposed an ego-lane identification method around an intersection [16]. This method detects lane markings, road markings, and surrounding vehicles with a stereo camera. Then, it estimates the lateral position by identifying an ego-lane through BN. This method also estimates the longitudinal position by detecting a stop line and finding its correspondence on a map. Svensson and Sorstedt estimate the probability that each lane is an ego-lane based on the Bayesian formula [28]. In order to estimate the probability, they utilize the information from a perception sensor (lane markings and surrounding vehicles), the vehicle estimated position from a low-cost GPS, and the information recorded in a map (number of lanes, lane width, and lane curvature). They also estimate the vehicle position within a lane through a localization filter such as a particle filter. Most map-guided perception sensor-based methods pay attention to estimate directly the order of the ego-lane with the information from a perception sensor and a map. Unlike the

previous map-guided perception sensor-based methods, Cao et al. generate many hypotheses of an ego-vehicle position through Monte Carlo simulation and select a reasonable one among the hypotheses to the observation results [29]. To do this, they generate a virtual image by projecting a 3D precise map into a virtual camera based on each hypothesis and each hypothesis is verified by comparing its virtual image to a real image. Their method consumes high computational power to verify a hypothesis, and in order to increase the localization precision, the method must generate many hypotheses.

Our proposed method belongs to the map-guided perception sensor-based approach and the proposed method is similar to Cao's method in that an ego-lane is identified by selecting a reasonable hypothesis to the current observation. While Cao's method requires high computational power to generate many hypotheses, the proposed method generates only one hypothesis per lane since the proposed method can estimate a vehicle position within a lane by using the endpoints of lane markings. Moreover, while Cao's method requires a large volume 3D precise map storing the information about the overall shape of a road and the shape of static objects on a road, the proposed method requires a relatively small-sized map since the proposed method utilizes the essential shape information of road facilities such as endpoints of lane markings and vertices of a road sign. And also, instead of generating a virtual image, since the proposed method only checks how the essential shape information of a detected road facility is reasonable to each hypothesis, the amount of computation is very small. In summary, the proposed ego-lane identification method projects the shape information of a road facility based on a camera hypothesis and compares the information to the detected facility. The proposed method is very effective for the following reasons. First, since the camera hypothesis generated by the in-lane localization is very precise, and only one camera hypothesis on each lane is needed. Second, the information of a road sign to be projected is very small, such as the 3D coordinates of road sign vertices.

The previous works that utilize a road boundary can identify an ego-lane in any section of a road without road facilities such as a lane endpoint and a road sign. However, on highways, the cases that the lane change of an autonomous driving vehicle is necessary are mainly collision avoidance and the entering of an intersection (IC) or a junction (JC). For collision avoidance, ego-lane identification is not necessary, and only the recognition whether the adjacent lane for a vehicle to move into is free or not is needed. In order to enter into an intersection (IC) or a junction (JC), ego-lane identification is necessary. Fortunately, there are always road signs to notify in advance the entrance of an IC or a JC on a highway, and also the type of lane marking is a dashed line to allow a lane change. Therefore, the proposed method that utilizes a lane endpoint and a road sign is suitable for ego-lane identification on a highway.

3. Proposed Method

3.1. System Overview. The flowchart of the proposed method is shown in Figure 2. In Figure 2, the initial vehicle position

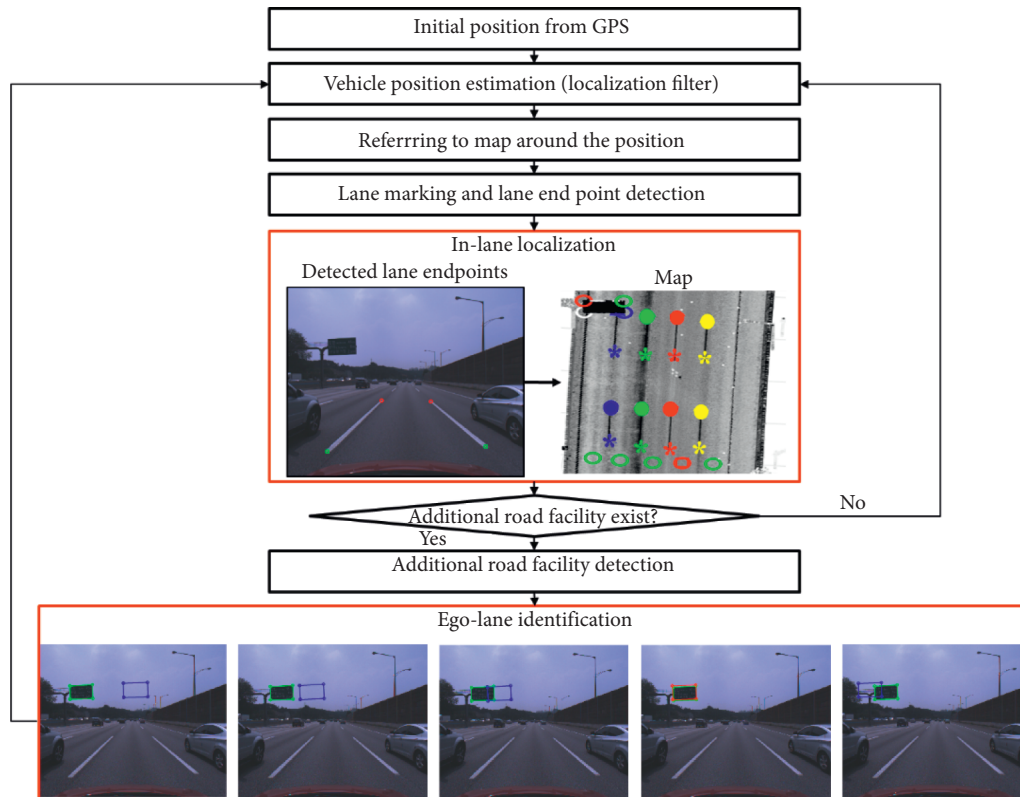


FIGURE 2: System flowchart.

is given from GPS. After the initial position is given, the vehicle position is estimated by the localization filter such as Kalman or particle filter. The localization filter exploits two inputs: the outputs from GPS and IMU (inertial measurement unit) and the vehicle position estimated by matching the detected lane endpoints to the corresponding ones in a map. After the ego-vehicle position is predicted by a localization filter, the digital map around the predicted position is referred. Lane markings and lane endpoints are detected in an image, and the in-lane localization estimates a vehicle position within a lane by matching the detected lane endpoints to the corresponding ones in a map. When the ego lane is not known yet, lane's endpoints on a map corresponding to the detected lane endpoints cannot be determined. Therefore, the in-lane localization generates vehicle position hypotheses for every lane, as shown in the right picture of the in-lane localization block of Figure 2. In the right picture, the hollow circles on a road indicate the generated hypotheses and the red circle among them indicates the hypothesis on an ego-lane. If there is an additional landmark such as a traffic sign or a road sign on a map, the additional landmark is detected. Otherwise, the vehicle position hypotheses generated in the in-lane localization are inputted into the localization filter. The ego-lane is then identified by verifying the vehicle position hypothesis per lane with the help of the additional landmark. In the ego-lane identification block of Figure 2, the green box indicates the detected additional landmark and the blue boxes indicate the projected additional landmarks stored in a map by using the false vehicle position hypotheses. The red box indicates the projected additional landmark by using

the hypothesis on an ego-lane. The ego-lane is identified by comparing the position difference between the detected landmark and the projected landmark. Before the ego-lane is identified, the localization filter is updated with the vehicle position hypotheses for every lane and keeps the possibility of the vehicle position within each lane. When the particle filter is used as the localization filter, the particles are weighted according to their distances to their nearest vehicle position hypothesis. The output of the localization filter is the mean of the particles. When the Kalman filter is used, a Kalman filter is generated for each vehicle position hypothesis. Therefore, there are multiple outputs from multiple Kalman filters. In this case, the map around the mean of the multiple outputs is referred. After the ego-lane identification, the localization filter is updated only with the vehicle position hypothesis on the ego-lane. After the ego-lane is identified once, the localization filter has tracked the vehicle position. Therefore, it is not necessary to identify the ego-lane in every frame. In this paper, the localization filter is assumed to be a particle filter. The detection methods for lane markings, lane endpoints, and an additional landmark such a road sign are referred to in our previous research [20, 30, 31]. Since this paper proposes in-lane localization and the ego-lane identification depicted as red blocks in Figure 2, the following sections describe only the in-lane localization and the ego-lane identification in detail.

3.2. In-Lane Localization. In-lane localization is to estimate a plausible position of a camera within each lane in a map. If there are N lanes in a map, a camera position per lane is

estimated with the correspondences between the lane endpoints of the lane in a map and the detected lane endpoints. The lane endpoints of a lane in a map and the detected lane endpoints are sorted according to their distances from the camera position estimated by GNSS/INS and from the origin in the camera coordinate system, respectively. And then the endpoints whose types are the same and distances are similar are associated. There may be several solutions to estimate a camera position with the correspondences between global points and the points in the camera coordinate system. When the number of detected lane endpoints are more than four, the perspective n points (PnP) solution, homography-based method, and perspective transformation (PT) based method are applicable. When the number of points is less than four, the perspective transformation-based method with the vehicle direction assumed from a map is a feasible solution. By comparison of localization precisions for the PnP solution, the homography-based method, and PT-based method, this paper suggests that the PT-based method is most appropriate for in-lane localization using lane endpoints.

3.2.1. PNP Method. The localization of the vehicle rigidly coupled with a perception sensor is equal to the perception sensor localization. When N correspondences between global coordinates and image coordinates are given, the camera pose estimation in the global coordinate system is known to be a PnP problem. This paper compares the efficient PnP (EPnP) method best known among many other methods available to solve the PnP problem [32]. The EPnP method takes less computation than other PnP methods that find a solution iteratively. When the solution from EPnP is used as an initial solution for other iterative methods, the more precise solution is quickly estimated. In the EPnP method, a point (\mathbf{p}_i^w) in the global coordinate system is expressed as the linear combination of four control points (\mathbf{c}_j^c) in the same coordinate system as follows:

$$\mathbf{p}_i^w = \sum_{j=1}^4 \alpha_{ij} \mathbf{c}_j^w, \text{ with } \sum_{j=1}^4 \alpha_{ij} = 1, \quad (1)$$

and, an image point (\mathbf{u}_i) is expressed as a projection of the linear combination of four control points (\mathbf{c}_j^c) in the camera coordinate system as follows:

$$\begin{bmatrix} \mathbf{u}_i \\ 1 \end{bmatrix} = \mathbf{K} \sum_{j=1}^4 \alpha_{ij} \mathbf{c}_j^c, \quad (2)$$

where \mathbf{K} is a camera intrinsic matrix. After the four global and camera control points are calculated, respectively, the rigid transformation between global control points and camera control points are simply estimated to find the camera pose [33]. With this solution as an initial one, the final solution to minimize the differences between image points and the projections of global points is precisely estimated through an optimization algorithm, such as Levenberg–Marquardt (LM) [34].

3.2.2. Homography Method. If the global points are on a single plane, the relation between the global points (X_i, Y_i, Z_i) and their image correspondences (x_i, y_i) is defined as homography transformation. Therefore, the camera projection matrix in equation (3) is reduced to the homography matrix shown in equation (4). The homography matrix \mathbf{H} is estimated with more than four correspondences between global points and image points [35]. The camera intrinsic matrix \mathbf{K} is known by the camera calibration in advance. Therefore, the rotation matrix \mathbf{R} and the translation vector \mathbf{t} between the camera coordinate system and the global coordinate system are calculated by equations (5) and (6), respectively. The camera position in the global coordinate system is calculated from the matrix \mathbf{R} and the vector \mathbf{t} :

$$\begin{bmatrix} x_i \\ y_i \\ 1 \end{bmatrix} = s \cdot \mathbf{K} \cdot \mathbf{M} \cdot \begin{bmatrix} X_i \\ Y_i \\ Z_i \\ 1 \end{bmatrix} = s \cdot \begin{bmatrix} fs_x & 0 & c_x \\ 0 & fs_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \quad (3)$$

$$\cdot \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \end{bmatrix} \cdot \begin{bmatrix} X_i \\ Y_i \\ Z_i \\ 1 \end{bmatrix},$$

$$\begin{bmatrix} x_i \\ y_i \\ 1 \end{bmatrix} = \mathbf{H} \cdot \begin{bmatrix} X_i \\ Y_i \\ 1 \end{bmatrix} = [\mathbf{h}_1 \ \mathbf{h}_2 \ \mathbf{h}_3] \cdot \begin{bmatrix} X_i \\ Y_i \\ 1 \end{bmatrix} \quad (4)$$

$$= s \cdot \begin{bmatrix} fs_x & 0 & c_x \\ 0 & fs_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} r_{11} & r_{12} & t_x \\ r_{21} & r_{22} & t_y \\ r_{31} & r_{32} & t_z \end{bmatrix} \cdot \begin{bmatrix} X_i \\ Y_i \\ 1 \end{bmatrix},$$

$$\mathbf{R} = [\mathbf{r}_1 \ \mathbf{r}_2 \ \mathbf{r}_3] = \left[\frac{1}{s} \mathbf{K}^{-1} \mathbf{h}_1 \ \frac{1}{s} \mathbf{K}^{-1} \mathbf{h}_2 \ \mathbf{r}_1 \times \mathbf{r}_2 \right], \quad (5)$$

$$\mathbf{t} = \frac{1}{s} \mathbf{K}^{-1} \mathbf{h}_3, \quad (6)$$

$$s = \frac{\|\mathbf{K}^{-1} \mathbf{h}_1\| + \|\mathbf{K}^{-1} \mathbf{h}_2\|}{2}. \quad (7)$$

3.2.3. PT-Based Method. The PnP and the homography methods described above did not fully utilize the prior knowledge. The lane endpoints exist only on a road surface, and the road surface is almost flat, but the PnP method does not utilize this knowledge. On the contrary, the homography method utilizes the assumption that the lane endpoints exist on the plane. However, this method does not utilize the relationship between the road surface and the camera which is known in advance through the calibration. The PT-based

method utilizes this relationship which can be known in advance by the calibration with the pattern, as shown in Figure 3.

In the PT-based method, the virtual camera whose image plane is perpendicular to the road surface is created by using the calibration result, as shown in Figure 4. The transform from the real camera coordinates \vec{C}_r to the virtual camera coordinates \vec{C}_v is calculated as follows:

$$\vec{C}_v = KR^tK^{-1}\vec{C}_r, \quad (8)$$

where K is the intrinsic camera matrix and R is the rotational matrix from the extrinsic calibration. Then, the coordinates of lane endpoints in the virtual camera coordinate system are calculated. The coordinate of a lane endpoints p_1 in the virtual camera coordinate system is $(X_1, Y_1, -Z_c)$, as shown in Figure 4, and Z_c is equal to the camera installation height from the road surface and is known by the camera calibration in advance. Y_1 is calculated as follows by the similarity in triangles shown in Figure 5(a):

$$Y_1 = \frac{-f \cdot Z_c}{(v_1 - o_v)}, \quad (9)$$

where f is a focal length, v_1 is the vertical image coordinate of lane endpoint p_1 , and the image coordinates of the principal point is (o_u, o_v) . X_1 is calculated as follows by the similarity in triangles shown in Figure 5(b):

$$X_1 = \frac{(u_1 - o_u) \cdot Y_1}{f}, \quad (10)$$

where u_1 is the horizontal image coordinate of lane endpoint p_1 . When more than four lane endpoints are detected and their corresponding global points are given, the rotation matrix and the translation vector between the virtual camera coordinate system and the global coordinate system can be estimated from these correspondences [33]. The camera global position is calculated from the rotation matrix and the translation vector.

When both of the left and right lane markings beside an ego-vehicle are dashed lines, there is high probability that more than four lane endpoints are detected. However, in the first lane or the lane next to the road boundary, only one of the lane markings is a dashed line. But, if the vehicle travelling direction is given, even in the case where only a single lane endpoint is detected, the vehicle position can be estimated by the PT-based method. Generally, one of the global coordinate axes is parallel to the gravity direction vector. When the global coordinates of a lane endpoint p_1 is (X_1^w, Y_1^w, Z_1^w) and Z_1^w is the value of the gravity directional axis, if the camera is not far from p_1 , the value of the gravity directional axis for the camera can be approximated to $Z_1^w + Z_c$. By this approximation, the camera position is on the circle whose center and radius are $p_1' = (X_1^w, Y_1^w, Z_1^w + Z_c)$ and $\gamma = \sqrt{X_1^w{}^2 + Y_1^w{}^2}$, as shown in Figure 6. The circle is on the plane at $Z = Z_1^w + Z_c$. Under the assumption that a vehicle follows a lane most of the time, if the travelling direction \vec{v}_c^w of a vehicle is approximated to the difference between two different type lane endpoints (p_1 and p_2) as in



FIGURE 3: Camera extrinsic parameter calibration.

equation (11), the optical axis of only the camera colored with red in Figure 6 is matched to \vec{v}_c^w . The x axis in the camera coordinate system is equal to the vector \vec{v}_R^w that is the cross product between \vec{v}_c^w and the gravity directional axis $[0, 0, 1]$ in the global coordinate system as in equation (12). The camera position p_c in the global coordinate system is estimated as in equation (13):

$$\vec{v}_c^w = \frac{\mathbf{P}_2 - \mathbf{P}_1}{\|\mathbf{P}_2 - \mathbf{P}_1\|}, \quad (11)$$

$$\vec{v}_R^w = \vec{v}_c^w \times \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \quad (12)$$

$$\mathbf{p}_c = \mathbf{p}_1 - \left(Y_1 \cdot \vec{v}_c^w + X_1 \cdot \vec{v}_R^w \right) + Z_c \cdot \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}. \quad (13)$$

3.3. Ego-Lane Identification. Mentioned in the introduction, the longitudinal and the lateral intervals between the same type lane endpoints on a highway are 20 m and 3.5 m, respectively. Therefore, in the case that the maximum GNSS/INS error is about 10 m, it is unknown which lane the detected lane endpoints belong to. That is, in order to estimate a vehicle position from a lane endpoint, an ego-lane identification is necessary.

The proposed ego-lane identification method is to select the one among the camera position hypotheses generated in the in-lane localization most reasonable to the additionally detected landmarks. In order to associate the additionally detected landmarks and the landmarks in a map, after the detected landmarks and the landmarks in a map are ordered, respectively, from left to right with respect to the optical axis in the camera coordinate system and in the global coordinate system, a detected landmark is associated one by one with the landmark in a map. For each camera position hypothesis, the camera rotation matrix is finely recalculated by using the additionally detected landmark and lane endpoints with their corresponding ones in a map through the PnP method. The reason why the camera rotation matrix is recalculated is that if the camera position hypothesis is the correct one, the camera rotation matrix can be more precisely estimated,

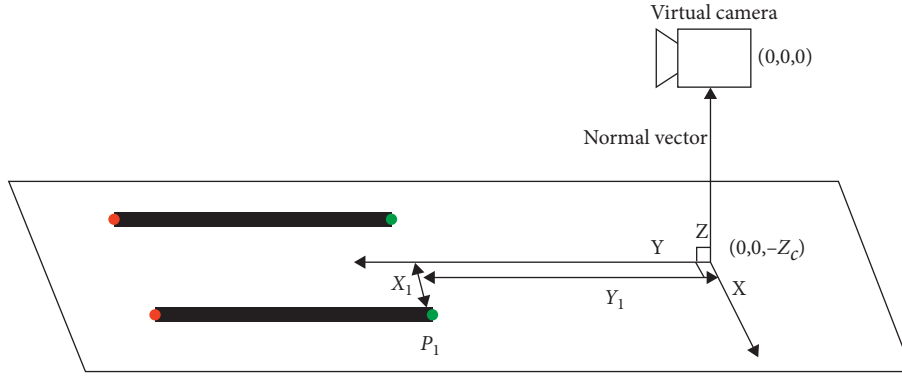


FIGURE 4: Relationship between lane endpoints and a virtual camera.

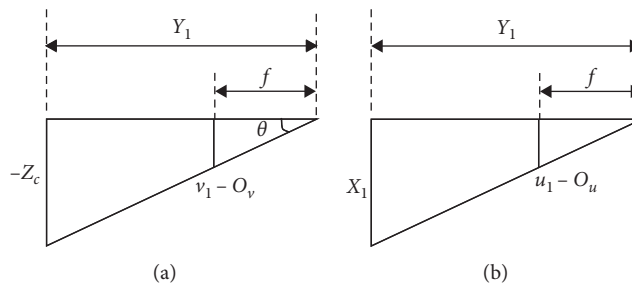


FIGURE 5: Relationship between a lane endpoint in the camera coordinate system and the corresponding point in the image coordinate system. (a) similarity in triangles for Y_1 ; (b) similarity in triangles for X_1 .

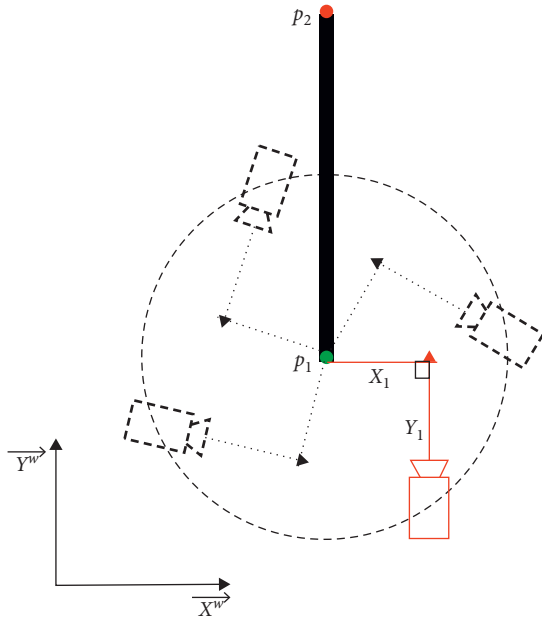


FIGURE 6: Relation between a lane endpoint and the estimated camera position.

thanks to the additional landmarks; otherwise, the camera rotation matrix will be incorrectly estimated to increase more the projection error. The projection error is the position difference in an image coordinate system between

detected landmarks and their corresponding ones from a map. After the camera position hypothesis and the camera rotation matrix are given, the global points for lane endpoints and additionally detected landmarks are acquired from a map and the global points are projected into the camera located in the camera position hypothesis. For each hypothesis, the projection error is calculated. The probability $P(T/L_i)$ that i_{th} among N hypotheses is true, as follows:

$$\max_{L_i} P(T/L_i) = \max_{L_i} \frac{1 - \left(PE(L_i) / \sum_j^N PE(L_j) \right)}{\sum_j^N \left(1 - \left(PE(L_i) / \sum_j^N PE(L_j) \right) \right)} \quad (14)$$

where $PE(L_i)$ is the projection error when lane L_i is assumed to be an ego-lane. The ego-lane is identified by selecting the hypothesis having the maximum probability. That is, the ego-lane is identified by selecting the hypothesis having the minimum projection error among the generated hypotheses. If only a single image is given, the ego-lane can be identified by equation (14). However, as the number of observed images increases, the performance of the ego-lane identification can be improved. Since the ego-lane identification result of each image is not independent, in this paper, the probability of each hypothesis from each image is simply fused as a summation:

$$\max_{L_i} \sum_{t=1}^T P_t(T/L_i). \quad (15)$$

As a camera nears a landmark, the angle between the camera optical axis and the landmark becomes large and the projection error of the incorrect hypothesis becomes larger. Therefore, we can fuse the sequential ego-lane identification result through the weighted sum by giving the larger weight in the reverse time order, but we fuse the probability for each image by a simple summation in equation (15). Since the proposed method generates a hypothesis per lane and projects only the vertices of a landmark to an image, its computation is very small. For reference, the additional landmarks in the experiment of this paper are road signs including a milestone, as shown in Figure 7. Only the global positions of vertices on a road sign are utilized.

4. Experimental Results

4.1. Experimental Database. In this paper, the experiments with the database collected on a highway show the comparison results of the three in-lane localization methods and show the performance analysis of the proposed ego-lane identification method. The high definition map and image database in this paper are constructed by a mobile mapping system (MMS). The MMS in this paper equips the high precision Lidar as a range sensor [36] and utilizes the positioning sensor that combines a RTK-GPS, a high precision inertial measurement unit (IMU), and a distance measurement indicator (DMI) [37]. The specification of the MMS is shown in Table 1.

The camera for the landmark detection is rigidly coupled to the MMS, as shown in Figure 8. The rigid transformation between the MMS and the camera is known by the off-line calibration and the camera is synchronized by receiving the trigger signal from the MMS. Therefore, the precise global position of the camera in the capturing time is calculated from the MMS, and the global position is used as the ground truth. The red cross in a red square in Figure 9(a) depicts the camera global position calculated from the MMS on the Lidar point clouds. The asterisks and the circles in Figure 9(a) depict the global positions of the lane starting points and lane ending points, respectively. The hollow circles depict the global positions of the road sign vertices. The global positions of lane endpoints and road sign vertices are projected to the camera on the ground truth position, as shown in Figure 9(b). The result that the projected positions of lane endpoints and road sign vertices are almost matched to the true positions, as shown in Figure 9(b), proves that the camera global position calculated from the MMS is very precise.

The database was collected within about a 42 km range from the Seoul toll gate to the Hobeop junction of the Yeongdong highway in South Korea, as shown in Figure 10. The highway around the Seoul toll gate is very wide and has 10 lanes in both directions. Table 2 summarizes the information about the experimental DB. Of a total of 24,289 images where a lane endpoint is detected, 6016 images have more than four lane endpoints. The number of images where a lane endpoint and a road sign are detected together for ego-lane identification is 419.

4.2. In-Lane Localization Result. For in-lane localization, the correspondences between the detected lane endpoints in an image and their global points stored on a map should be given. For this, first, under the assumption that the ego-lane is already identified, the corresponding lane markings of the left and right lane markings detected in an image are found in a map. The lane endpoints on a map and an image are, respectively, sorted in the ascending order of the distance from the camera with respect to the driving direction of a road. The endpoints in a map are corresponded sequentially to the same type endpoints in an image. For the 6016 images that have more than four detected lane endpoints, the in-lane localization comparisons of the PnP method, the homography method, and the PT-based method are shown in Table 3.

Table 3 shows that, as the prior knowledge is added, the localization error tends to be decreased. In particular, the PnP method using no prior knowledge is severely affected by the position error of the detected lane endpoints. Figure 11 shows an example of the in-lane localization results of three methods.

Figure 11(a) shows the detected lane endpoints, and Figure 11(b) shows the lane endpoints on a map and the localization results. In Figure 11(b), the red rectangle depicts the ground-truth camera position. The red, blue, and green hollow circles depict the estimated camera position of the PT-based method, the homography method, and the PnP method, respectively. Although the lane endpoints are detected accurately, as shown in Figure 11(a), because of the high sensitivity to detection error and the small number of detected points, the estimated position of the PnP can often differ greatly to its ground truth, as shown in Figure 11(b). However, the camera position estimated by the PT-based is almost overlapped over the ground truth. Most cases where the localization errors from the PT-based are large are caused by falsely detected lane endpoints, as shown in Figure 12. In Figure 12(a), there is one falsely detected lane endpoint. Because of this, the localization errors of the PT-based method, the homography method, and the PnP method become 1.7 m, 1.8 m, and 5.4 m, respectively.

Table 4 shows the average localization error of the PT-based method according to the number of lane endpoints. The average localization error of the PT-based method is very small, 0.284 m, even with one detected lane endpoint. As the number of detected lane endpoints increases, the average localization error naturally decreases. However, the maximum localization error does not decrease since the error is affected by a falsely detected lane endpoint.

In Table 4, when the number of detected lane endpoints increases from three to four, the localization error decreases significantly. This is caused by a slight change of the estimation method according to the number of lane endpoints. When the number of detected lane endpoints is not over three, the camera position is calculated as the average of the position estimated from each detected lane endpoint. On the contrary, when over four, the rigid transformation between the camera coordinate system and the global coordinate system is estimated by using all lane endpoints and the camera position is calculated from the transformation.

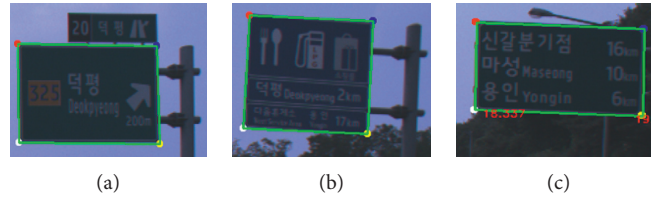


FIGURE 7: Examples of an additional road facility.

TABLE 1: Sensor specifications of the MMS.

Sensor	Model	Specification	
Ranging sensor	Velodyne HDL32E	Measurement range	5 cm~100 m
		Field of view	Horizontal 360°, vertical 40°
		Accuracy	<2 cm
		Angular resolution	Vertical 1.25°
		Acquisition frequency	5~20 Hz
Positioning sensor	Applanix POS LV 210	X and Y position	0.020 m (RMS)
		Z position	0.050 m (RMS)
		Roll and pitch	0.020° (RMS)
		Heading	0.050° (RMS)
		Acquisition frequency	100 Hz

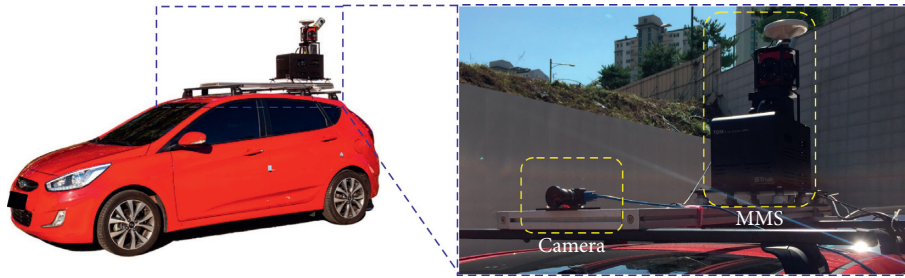


FIGURE 8: MMS and camera used in the experiment.

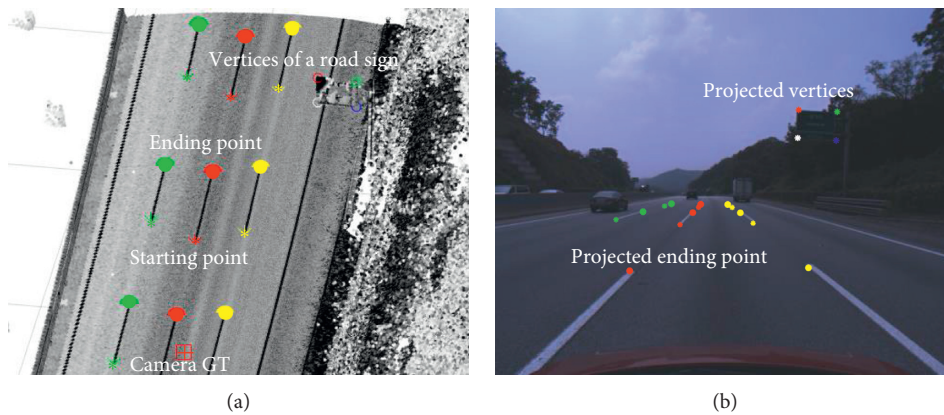


FIGURE 9: The camera global position calculated from the MMS and the projected global points of lane endpoints and road sign vertices: (a) lane endpoints and road sign vertices on the Lidar cloud points; (b) lane endpoints and road sign vertices projected on an image.

Figure 13 shows the localization example of the PT-based method using one detected lane endpoint. Figure 13(a) shows the detected lane endpoint. Figure 13(b) shows the lane endpoints stored in a map, the estimated camera position, and the camera position GT. In Figure 13, the difference between the estimated camera position and

the camera position GT is as small as 5 cm. Figure 14 shows the example where the error of the camera position estimated from one lane endpoint is as large as 1.7 m. Although a lane endpoint is correctly detected in Figure 14(a), the estimated camera position is far from its ground truth in Figure 14(b). This error is caused by the fact that the

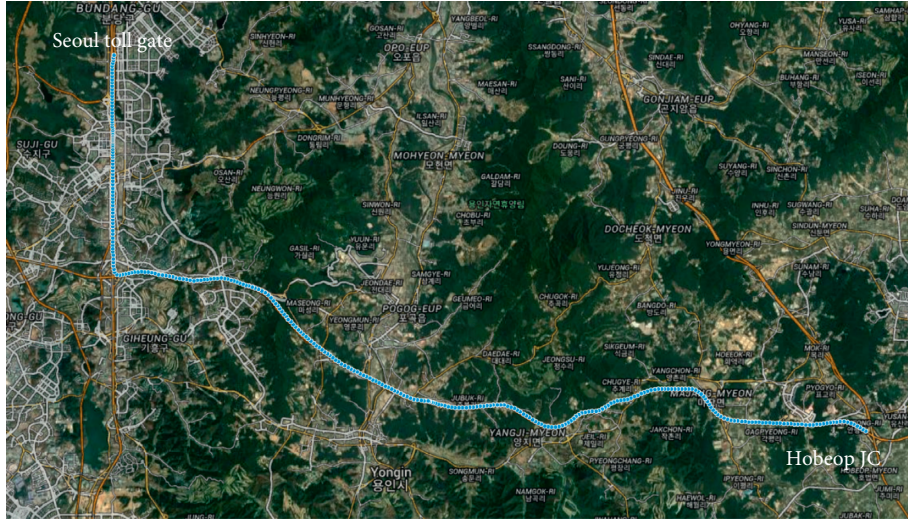


FIGURE 10: Route of the vehicle collecting the database.

TABLE 2: Summary of experimental DB.

Number of frame	Number of frames <4 endpoints	Number of frames ≥ 4 endpoints	Endpoints and road signs
24,289	18,273	6,016	419

TABLE 3: Comparison of in-lane localizations with more than four lane endpoints.

Error (m)	Methods		
	PT-based	Homography	PnP
Average	0.194	0.457	7.657
Maximum	1.960	21.982	6504.599

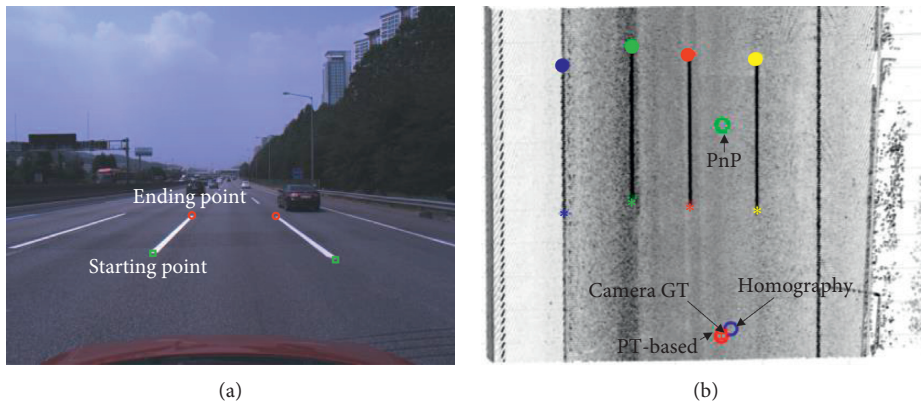


FIGURE 11: Comparison of the estimated camera positions: (a) detected lane endpoints; (b) estimated camera positions.

repainted lane markings are not yet updated in a map. In Figure 14(b), the lane endpoint stored in a map is different to the lane endpoint on the Lidar points cloud.

4.3. Ego-Lane Identification Results. The proposed ego-lane identification can be executed when lane endpoints and an additional landmark such as a road sign are detected in a single image. Therefore, the ego-lane identification is tested

with 419 images. Table 5 shows the performance of the proposed ego-lane identification method according to the number of images.

In the case of processing just a single image, only three among 419 ego-lane identification tests failed. In the case of processing two images, the number of tests is reduced from 419 to 387 since the first image where a road sign is detected is excluded in the test count. When the number of processing images increases up to four, all tests are successful.

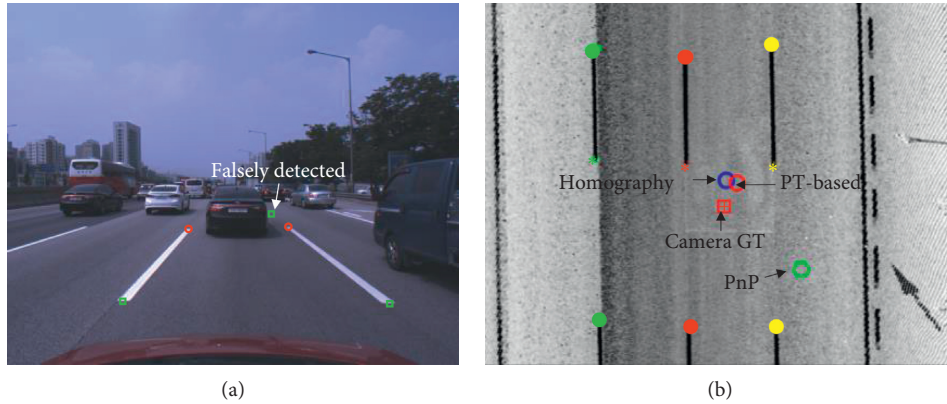


FIGURE 12: Example of estimated camera position error caused by a falsely detected lane endpoint: (a) a falsely detected lane endpoint; (b) estimated camera positions.

TABLE 4: Localization error of the PT-based method.

Number of detected lane endpoints	Number of frames	Average error (m)	Maximum error (m)
1	4758	0.284	1.727
2	10971	0.259	1.998
3	2544	0.258	1.948
More than 4	6016	0.194	1.960
Total	24289	0.248	1.998

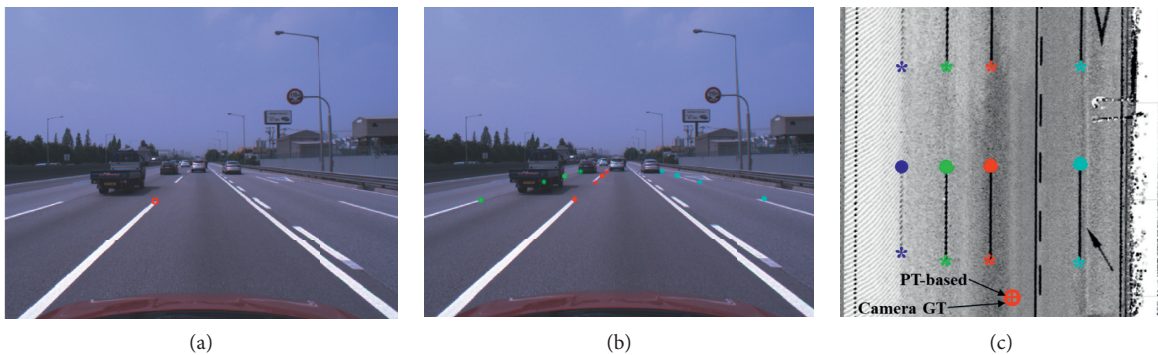


FIGURE 13: Example of estimated camera position with one detected lane endpoint: (a) detected lane endpoint; (b) projected lane endpoints from a map; (c) estimated camera position.

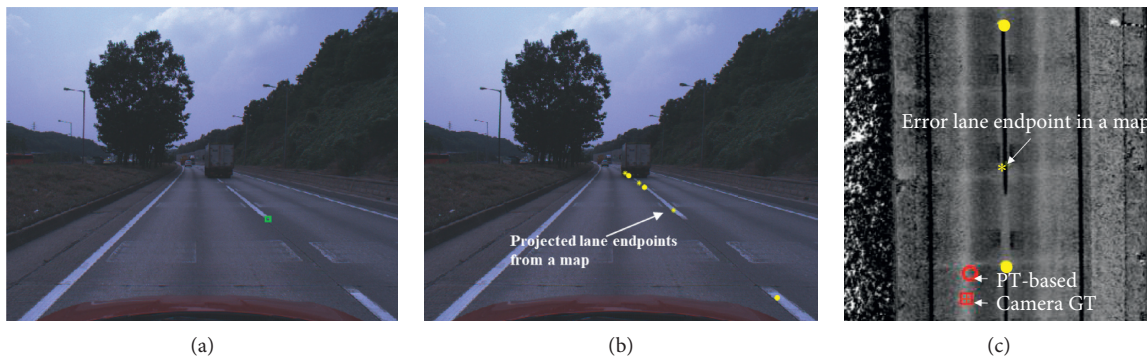


FIGURE 14: Example of large camera position error estimated with one detected lane endpoint: (a) detected lane endpoint; (b) projected lane endpoints from a map; (c) estimated camera position.

TABLE 5: Ego-lane identification results according to the number of processed images.

Number of processed images	Number of test	Accuracy (%)	Number of error
1	419	99.28	3
2	387	99.22	3
3	355	99.72	1
4	323	100	0

TABLE 6: Vehicle localization results of the proposed method in a five-lane road.


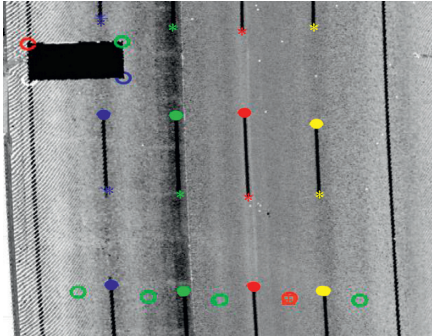
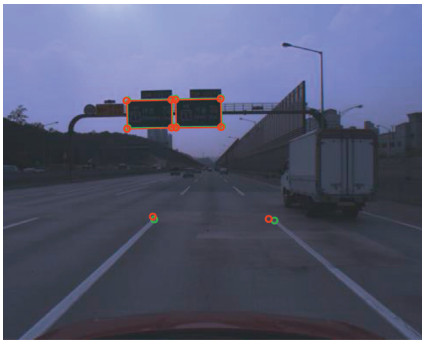
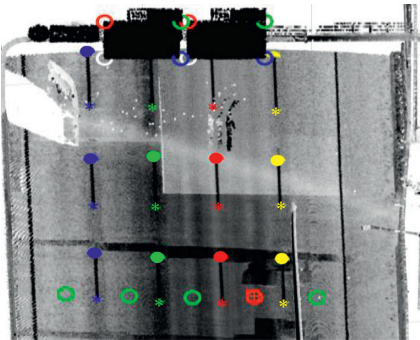

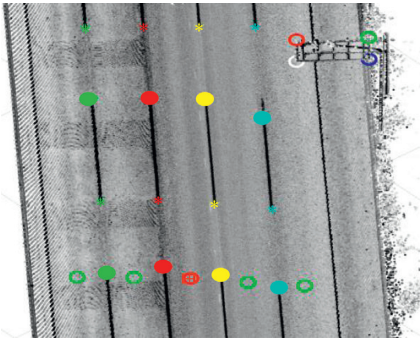
Description	Detected landmarks and projected landmarks	Localization results
<p>Sign position: left Projection error: 17.1 pixel Vehicle position error: 0.21 m</p>		
<p>Sign position: center Projection error: 4.8 pixel Vehicle position error: 0.37 m</p>		
<p>Sign position: right Projection error: 13.1 pixel Vehicle position error: 0.17 m</p>		

Table 6 shows the localization results by the proposed method from one image when a road sign is on the left, center, or right side of the road which has five lanes in one direction. The localization results in Table 6 are obtained by performing in-lane localization for each lane using the PT-based method and selecting an ego-lane by the proposed ego-lane identification method. The first column in Table 6

describes the position (left, center, and right) of a road sign on a road, the projection error, and the localization error. The second column in Table 6 shows the detected landmarks (a road sign and a lane end point) and the projections of the corresponding landmarks stored in a map. In the second column, the detected landmarks are depicted as green circles and the projections of the corresponding landmarks are

TABLE 7: Vehicle localization results of the proposed method in various situations.

Description	Detected landmarks and projected landmarks	Localization results
Number of lanes: 4 Sign pos: R Projection error: 4.3 pixel Vehicle position error: 0.07 m		
Number of lanes: 4 Curved road Projection error: 18.9 pixel Vehicle position error: 0.53 m		
Number of lanes: 4 Curved road Projection error: 20.6 pixel Vehicle position error: 0.12 m		
Number of lanes: 5 A road sign missed Projection error: 3.6 pixel Vehicle position error: 0.11 m		

depicted as red circles. The last column shows the Lidar points cloud where the road sign vertices and lane endpoints stored in a map are displayed. In the images of the last column, the hollow circles on a road depicts the vehicle position hypothesis per lane generated by the in-lane localization, and the red hollow circle among them is the hypothesis on the identified ego-lane. The camera ground truth position is depicted as a red cross within a square. Table 6 shows that the proposed method can correctly identify an ego-lane in the broad road. In particular, since the accurate in-lane localization results are acquired by the PT-based method using lane endpoints and the ego-lane is easily identified by verifying a handful of hypotheses with a map, the proposed method can acquire the accurate localization results with a small amount of computation.

Table 7 shows the localization results by the proposed method from one image in various situations. The 2nd and 3rd rows in Table 7 show the localization results in the curved section. When the number of detected lane endpoints is below four, as shown in the 4th row, the vehicle position can be precisely estimated. In a curved road, the camera optical axis approximation by the lane mark direction makes the localization error relatively large. However, even on a curved road, when the number of detected lane endpoints is more than four, as shown in the 3rd row, the localization error is not large due to not using the optical axis approximation. In this paper, we do not use the yaw rate sensor already installed in a vehicle to estimate the camera optical axis, but if it is used, it is expected to estimate the axis more precisely even in a curved road. The last row of Table 7 shows the case that one of two road signs is not detected. In this case, since the detected road sign can be corresponded to both road signs on a map, the difference between the projection errors of a true hypothesis and a false one can become small. However, even in this case, there is little possibility that the projection error of a false hypothesis is smaller than the error of a true hypothesis. In the last row of Table 7, the projection error of the true hypothesis is 3.6 pixels, and the error of the hypothesis for the left lane to the ego-lane is 30.1 pixels. Since the road where the experimental database is collected is highway, its curvature is not large, but some parts of the road include the curved section, as shown in Figure 10. In order to make the curvature of the road in the 3rd row in Table 7 noticeable, the expanded LIDAR map of the road of the 3rd row is shown in Figure 15.

The major reason for the false ego-lane identification is the false detection of a road sign or lane endpoints. Figures 16 and 17 show the false ego-lane identification cases. In Figure 16(a), two right vertices of a road sign are falsely detected. Therefore, the rightmost lane is falsely identified as an ego-lane, as shown in Figure 16(b). Figure 17 shows the example of a falsely detected lane endpoint. In order to find the correspondences between the detected lane endpoints and those stored on a map, instead of considering the distance between lane endpoints, only the type of lane endpoint is checked after lane endpoints are sorted in the ascending order of the distance from a camera. Therefore, if there is a false one among the detected lane endpoints, the detected points can be falsely corresponded so that the error

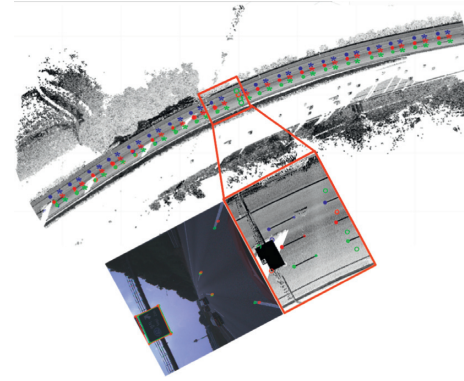
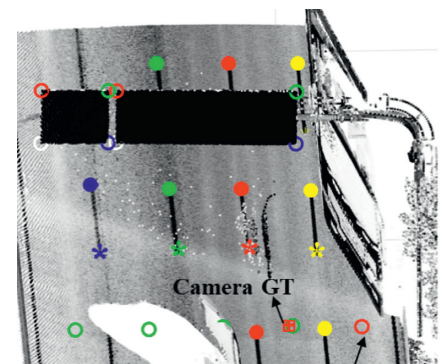


FIGURE 15: Example of camera hypotheses and ego-lane identification result in a curved road.



(a)



False identification

(b)

FIGURE 16: Ego-lane identification failure by a falsely detected road sign: (a) falsely detected road sign; (b) localization result.

of the in-lane localization per lane becomes large, as shown in Figure 17(b). This problem can be solved through the removal of falsely detected lane endpoints by checking the type change of the sequential lane endpoints and the distance between them. Also, this problem can be solved by using a robust data association method such as Hungarian algorithm [38]. As mentioned above, although the ego-lane identification error can be reduced by removing falsely detected road signs or lane endpoints, these false detections

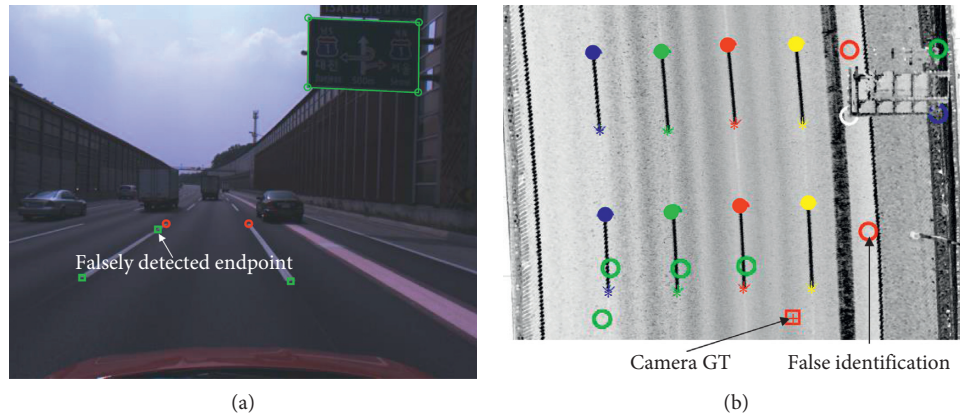


FIGURE 17: Ego-lane identification failure by a falsely detected lane endpoint: (a) falsely detected lane endpoint; (b) localization result.

do not continuously happen. Therefore, by simply fusing the identification results from sequential images, the ego-lane identification error can be reduced, as shown in Table 5.

5. Conclusion and Future Work

This paper proposes a localization method that estimates a vehicle position within a lane and identifies an ego-lane by using lane endpoints and an additional landmark. The average error of the PT based in-lane localization to use lane endpoints is 0.248 m, and an error level is low enough to satisfy the requirement specification of autonomous driving. The proposed ego-lane identification method perfectly identified an ego-lane by fusing the identification results from four images even when there are false detections of lane endpoints or a road sign.

The in-lane localization in this paper simply approximates the normal vector of a road surface as the gravity direction vector and the camera optical axis as the vector from a starting point to an ending point. However, we expect to improve the localization precision through the normal vector estimation with lane endpoints in a map and the optical axis estimation in the global coordinate system with the angle between a detected lane marking and the optical axis in the camera coordinate system. The proposed method does not consider the false detections of a landmark such as lane endpoints or a road sign. The false detection of a road sign as shown in Figure 16(a) can be easily removed by checking the aspect ratio difference between the detected road sign and the road sign stored in a map. Also, the case where the same type lane endpoints on a lane marking are continuously detected as shown in Figure 17(a) is caused by a false positive or a false negative. This case can be removed by checking the order of the lane endpoint type and the distance between endpoints stored in a map. The performance of the proposed method will be improved by removing these false cases. Finally, the performance of the proposed method was evaluated under the assumption that the coarse localization given by a GPS is synchronized to camera capturing. In the future, we plan to develop a final localization system by combining the proposed method with a localization filter such as a Kalman filter or particle filter.

Data Availability

It is difficult to share the experimental data and source code because the research result will be used for the commercial product.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This research was supported by a grant (code 19TLRP-B101406-05) from “Development of Cooperative Automated Driving Highway Systems” Program funded by the Ministry of Land, Infrastructure and Transport of Korean government.

References

- [1] J. Ziegler, P. Bender, M. Schreiber et al., “Making bertha drive an autonomous journey on a historic route,” *IEEE Intelligent Transportation Systems Magazine*, vol. 6, no. 2, pp. 8–20, 2014.
- [2] I. Skog and P. Handel, “In-car positioning and navigation technologies-A survey,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 10, no. 1, pp. 4–21, 2009.
- [3] N. Alam and A. G. Dempster, “Cooperative positioning for vehicular networks: facts and future,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 14, no. 4, pp. 1708–1717, 2013.
- [4] D. S. Yun and H. S. Yu, “Development of the optimized autonomous navigation algorithm for the unmanned vehicle using extended kalman filter,” *Transactions of KSAE*, vol. 16, no. 3, pp. 7–14, 2008.
- [5] A. Pirti, K. Gumus, H. Erkaya, and R. G. Hosbas, “Evaluating repeatability of RTK GPS/GLONASS near/under forest environment,” *Croatian Journal of Forest Engineering*, vol. 31, no. 1, pp. 23–33, 2010.
- [6] S. Nedeveschi, V. Popescu, and R. Danescu, “Accurate ego-vehicle global localization at intersections through alignment,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 14, no. 2, pp. 1–15, 2012.
- [7] Y. Li, Z. Hu, Z. Li, M. A. Sotelo, and Y. Ma, “Multiscale site matching for vision-only self-localization of intelligent

- vehicles,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 10, no. 3, pp. 170–183, 2018.
- [8] A. Y. Hata and D. F. Wolf, “Feature detection for vehicle localization in urban environments using a multilayer LIDAR,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 17, no. 2, pp. 420–429, 2016.
- [9] C. Valgren and A. J. Lilienthal, “SIFT, SURF & seasons: appearance-based long-term localization in outdoor environments,” *Robotics and Autonomous Systems*, vol. 58, no. 2, pp. 149–156, 2010.
- [10] A. Schlichting and C. Brenner, “Localization using automotive laser scanners and local pattern matching,” *IEEE IV2003 Intelligent Vehicles Symposium*, vol. 10, pp. 414–419, 2014.
- [11] K. Yoneda, H. Tehrani, T. Ogawa, N. Hukuyama, and S. Mita, “Lidar scan feature for localization with highly precise 3-D map,” *IEEE IV2003 Intelligent Vehicles Symposium*, vol. 10, pp. 1345–1350, 2014.
- [12] H. Lategahn and C. Stiller, “City GPS using stereo vision,” *IEEE Transactions on Aerospace and Electronic Systems*, vol. 11, pp. 1–6, 2011.
- [13] J. Ziegler, H. Lategahn, M. Schreiber et al., “Video based localization for bertha,” *IEEE IV2003 Intelligent Vehicles Symposium*, vol. 12, pp. 1231–1238, 2014.
- [14] M. Stübler, J. Wiest, and K. Dietmayer, “Feature-based mapping and self-localization for road vehicles using a single grayscale camera,” *IEEE IV2003 Intelligent Vehicles Symposium*, vol. 15, pp. 267–272, 2015.
- [15] M. Schreiber, C. Knöppel, and U. Franke, “LaneLoc: lane marking based localization using highly accurate maps,” *IEEE IV2003 Intelligent Vehicles Symposium*, vol. 17, pp. 449–454, 2013.
- [16] S. Nedeveschi, V. Popescu, R. Danescu, T. Marita, and F. Oniga, “Accurate ego-vehicle global localization at intersections through alignment of visual data with digital map,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 14, no. 2, pp. 673–687, 2013.
- [17] K. Jo, Y. Jo, J. K. Suhr, H. G. Jung, and M. Sunwoo, “Precise localization of an autonomous car based on probabilistic noise models of road surface marker features using multiple cameras,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 16, no. 6, pp. 3377–3392, 2015.
- [18] J. K. Suhr, J. Jang, D. Min, and H. G. Jung, “Sensor fusion-based low-cost vehicle localization system for complex urban environments,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 18, no. 5, pp. 1078–1086, 2017.
- [19] D. Gruyer, R. Belaroussi, and M. Revilloud, “Accurate lateral positioning from map data and road marking detection,” *Expert Systems with Applications*, vol. 43, pp. 1–8, 2016.
- [20] E. S. Jang, J. K. Suhr, and H. G. Jung, “Lane endpoint detection and position accuracy evaluation for sensor fusion-based vehicle localization on highways,” *Sensors*, vol. 18, no. 12, pp. 1–23, 2018.
- [21] N. Houdali, T. Ditchi, E. Géron, J. Lucas, and S. Holé, “Road-vehicle cooperation for lateral guidance,” *Progress in Electromagnetics Research*, vol. 10, pp. 791–795, 2013.
- [22] T.-S. Dao, K. Y. K. Leung, C. M. Clark, J. P. Huissoon, and J. P. Huissoon, “Markov-based lane positioning using inter-vehicle communication,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 8, no. 4, pp. 641–650, 2007.
- [23] S. Lee, S. W. Kim, and S. W. Seo, “Accurate ego-lane recognition utilizing multiple road characteristics in a Bayesian network framework,” *IEEE IV2003 Intelligent Vehicles Symposium*, vol. 10, pp. 543–548, 2015.
- [24] Y. W. Seo and R. Rajkumar, “Tracking and estimation of ego-vehicle’s state for lateral localization,” in *Proceedings of the 2014 17th International Conference on Intelligent Transportation Systems*, Qingdao, China, 2014.
- [25] S. Kim and S.-Y. Park, “Lane-level positioning based on 3D tracking path of traffic signs,” *Proceedings of the 11th Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications*, vol. 3, pp. 642–648, 2016.
- [26] H. Weigel, P. Lindner, and G. Wanielik, “Vehicle tracking with lane assignment by camera and lidar sensor fusion,” *IEEE IV2003 Intelligent Vehicles Symposium*, vol. 10, pp. 513–520, 2009.
- [27] A. L. Ballardini, D. Cattaneo, R. Izquierdo, I. Parra, M. A. Sotelo, and D. G. Sorrenti, “Ego-lane estimation by modeling lanes and sensor failures,” in *Proceedings of the 2017 IEEE 20th International Conference on Intelligent Transportation Systems (ITSC)*, Yokohama, Japan, 2017.
- [28] D. Svensson and J. Sorstedt, “Ego lane estimation using vehicle observations and map information,” in *Proceedings of the 2016 IEEE Intelligent Vehicles Symposium (IV)*, Gothenburg, Sweden, 2016.
- [29] G. Cao, F. Damerow, B. Flade, M. Helmling, and J. Eggert, “Camera to map alignment for accurate low-cost lane-level scene interpretation,” in *Proceedings of the 2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC)*, Rio de Janeiro, Brazil, 2016.
- [30] K. Choi, J. K. Suhr, and H. G. Jung, “FAST pre-filtering-based real time road sign detection for low-cost vehicle localization,” *Sensors*, vol. 18, no. 10, pp. 1–18, 2018.
- [31] K. Choi, J. K. Suhr, and H. G. Jung, “Map-matching based cascade landmark detection and vehicle localization,” *IEEE Access*, vol. 7, pp. 127874–127894, 2019.
- [32] V. Lepetit, F. Moreno-Noguer, and P. Fua, “EPnP: An accurate O (n) solution to the PnP problem,” *International Journal of Computer Vision*, vol. 81, no. 2, 2009.
- [33] S. Umeyama, “Least-squares estimation of transformation parameters between two point patterns,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 13, no. 4, pp. 376–380, 1991.
- [34] P. R. Gill, W. Murray, and M. H. Wright, *The Levenberg-Marquardt Method*, Academic Press, London, UK, 1981.
- [35] The MathWorks Inc., “Compute location of calibrated camera,” <https://kr.mathworks.com/help/vision/ref/extrinsics.html>.
- [36] Velodyne HDL-32E., 2018, <https://velodynelidar.com/hdl-32e.html>.
- [37] Applanix POS LV 210, 2018, <https://www.applanix.com/products/poslv.htm>.
- [38] J. Munkres, “Algorithms for the assignment and transportation problems,” *Journal of the Society for Industrial and Applied Mathematics*, vol. 5, no. 1, pp. 32–38, 1957.