Dense Stereo-Based Robust Vertical Road Profile Estimation Using Hough Transform and Dynamic Programming

Jae Kyu Suhr, *Member, IEEE*, and Ho Gi Jung, *Senior Member, IEEE*

***Abstract*—This paper proposes a dense stereo-based robust vertical road profile estimation method. The vertical road profile is modeled by a cubic B-spline curve, which is known to be accurate and flexible but difficult to estimate under a large proportion of outliers. To robustly estimate a cubic B-spline curve, the proposed method utilizes a two-step strategy that initially estimates a piecewise linear function and then obtains a cubic B-spline curve based on the initial estimation result. A Hough transform and dynamic programming are utilized for estimating a piecewise linear function to achieve robustness against outliers and guarantee optimal parameters. In the experiment, a performance evaluation and comparison were conducted using three publicly available databases. The result shows that the proposed method outperforms three previous methods in all databases. In particular, its performance is superior to the others in the cases of a large proportion of outliers and road surfaces distant from the ego-vehicle.**

***Index Terms*—B-spline curve, dense stereo, dynamic programming, Hough transform, piecewise linear function, road profile, road surface, stereo vision.**

I. Introduction

STEREO cameras have been widely used for driving environment perception since they can provide both image and depth information. Road surface estimation is one of the most important procedures for stereo vision systems because an obstacle is generally defined as an object located above the road surface. Although a road is a 3-D curved surface in the real world, mostly it has been approximated to a 2-D slice capturing the vertical road height along the heading direction of the ego-vehicle because its lateral variation is negligible in on-road situations. This approximation is beneficial since it reduces a computational burden and increases robustness against outliers produced by obstacles or stereo matching errors. The 2-D slice representing a road surface is often referred to as a vertical road profile.

A vertical road profile has been modeled as a polynomial function [1]–[4] and a piecewise linear function [5]–[7]. A B-spline curve has been presented as a generalization of polynomial and piecewise linear functions [8]–[12]. Among these popular vertical road profile models, a cubic B-spline curve has been known as the most accurate and flexible model. However, this curve is difficult to estimate under a large proportion of outliers produced by obstacle or stereo matching errors due to its high degree of freedom (DOF). To alleviate this problem, Wedel *et al.* [8], [9] proposed a Kalman filter-based estimation method. Keller *et al.* [10], [11] utilized a region of interest (ROI)-based outlier removal approach. Schauwecker and Klette [12] presented an M-estimator-based region growing method. Despite all these efforts, the previous methods are unable to work properly in cases of non-Gaussian noises, increasing triangulation errors at the distant road surface, and a larger portion of outliers.

To overcome the drawbacks of the previous methods, this paper proposes a novel vertical road profile estimation method based on a cubic B-spline curve. The proposed method utilizes a two-step strategy because it is difficult to directly estimate a function with a high DOF from the realistic data contaminated by outliers. First, the road surface is divided into several subregions, and Hough transform is applied to the 3-D points in each subregion to obtain their voting results. Dynamic programming is utilized to estimate the parameters of the piecewise linear function by treating the voting results as a data term and the changes of slopes and locations of adjacent linear functions as a smoothness term. Second, the vertical road profile with a cubic B-spline curve is estimated by utilizing the points uniformly sampled from the estimation result of the piecewise linear function. Fig. 1 shows the flowchart of the proposed method.

This paper has the following contributions.

1) A two-step strategy is proposed in order to reliably estimate a cubic B-spline curve from the realistic data contaminated by outliers with a reasonable computational cost.
2) A method that combines a Hough transform and dynamic programming is proposed to robustly find the optimal parameters of a piecewise linear function under a large proportion of outliers.

The proposed method was tested using three publicly available databases and compared with three previous methods that utilize robust estimators. Experimental results reveal that the proposed method outperforms the three previous methods in all three databases. In particular, its performance is superior to the others in the cases of a large proportion of outliers and road surfaces distant from the ego-vehicle.

Manuscript received December 18, 2013; revised May 7, 2014, July 18, 2014, and September 16, 2014; accepted October 26, 2014. Date of publication December 4, 2014; date of current version May 29, 2015. This work was supported in part by Hyundai Mobis and in part by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (2012-0002467). The Associate Editor for this paper was S. Birchfield.

J. K. Suhr is with the Research Institute of Automotive Control and Electronics, Hanyang University, Seoul 133-791, Korea.

H. G. Jung is with the Department of Automotive Engineering, Hanyang University, Seoul 133-791, Korea (e-mail: hogijung@hanyang.ac.kr).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TITS.2014.2369002

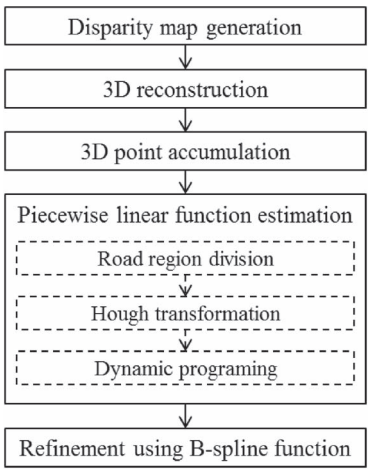

Fig. 1. Flowchart of the proposed method.

II. Related Research

Reliable terrain estimation is an important task for intelligent vehicle applications. Terrain estimation using stereo vision starts with generating 3-D points from a disparity map. Since an off-road environment usually has complex and arbitrary terrain, various types of grid maps (2-D, 2.5-D, or 3-D) have been utilized for off-road terrain estimation. A 2-D grid map stores occupancy information (traversable/nontraversable) at each cell and is often called an occupancy grid [13]. Although the 2-D grid map is effective in avoiding simple obstacles on flat terrain, a detailed grid map is required to handle complex situations. A 2.5-D grid map, a so-called elevation map, contains height information at each cell [14]–[17]. This map is useful for terrain with a single surface layer, but cannot deal with multiple surface layers such as bridges, tunnels, or underpasses. A 3-D grid map is able to overcome this problem by representing terrain using stacks of cubes with occupancy information [18], [19]. This map is often called a voxel map and can be efficiently described by a hierarchical data structure for spatial subdivision such as an octree [20]. Since dealing with a 3-D grid map usually requires a high computational cost, elevation maps that compromise accuracy and computational cost are more popularly used for intelligent vehicle applications.

Unlike the off-road environment, terrain in on-road situations consists of a man-made paved road with a fairly simple shape. Terrain estimation of on-road situations is often called a road surface estimation. Since a road is a 3-D curved surface, 3-D models have been used for the road surface estimation. Sappa *et al.* [21], [22] modeled a road surface with a 3-D plane. This method selects 3-D points that are likely to be a part of the road surface using the YZ-plane accumulation. The road surface is estimated by applying a random sample consensus (RANSAC)-based 3-D plane estimator to the selected 3-D points. Suhr *et al.* [23] also presented a method similar to this in rear-view situations. Danescu *et al.* [24] and Oniga and Nedevschi [25] utilized a 3-D quadratic surface as a road surface model. This method estimates the initial solution via RANSAC and refines it using a region growing process.

Although a road is a 3-D curved surface in the real world, a small number of previous methods utilize 3-D road models. This is because its lateral variation in on-road situations is negligible. Thus, many previous methods approximate the road surface to a 2-D slice by considering accuracy, robustness, and computational cost. The 2-D slice representing a road surface is often referred to as a vertical road profile. Labayrade and Aubert [1] and Benenson *et al.* [2] modeled a vertical road profile with a straight line. These methods generate a v-disparity image using a sparse disparity map or stereo matching cost and apply a robust estimator to find a straight line. Since the straight line model is only suitable for a flat road, various methods have been proposed to handle nonflat road situations. Nedevschi *et al.* [3] utilized a clothoid curve model. This method accumulates 3-D points on the YZ-plane and estimates the curve using an angle histogram of the near 3-D points and a curvature histogram of the far 3-D points. Sappa *et al.* [4] compared various polynomial-based vertical road profile models (linear, quadratic, and cubic) in both v-disparity and YZ-plane domains using synthetic scenes. Labayrade *et al.* [5], Suhr *et al.* [6], and Suhr and Jung [7] used a piecewise linear function model. Labayrade *et al.* [5] applied Hough transform to a whole v-disparity image and estimated a piecewise linear function by selecting several straight lines that have large accumulation values and compose an upper or a lower envelope. Suhr *et al.* [6] divided a v-disparity image into several parts and applied Hough transform to the consecutive parts to find a piecewise linear function. Suhr and Jung [7] sampled 3-D points using the YZ-plane accumulation and estimated a piecewise linear function by applying RANSAC to the points sampled in consecutive regions.

Meanwhile, a B-spline curve has been presented as a generalization of polynomial and piecewise linear functions. Among the popular vertical road profile models discussed above, a cubic B-spline curve has been known as the most accurate and flexible model because of its high DOF [8]–[12]. Wedel *et al.* [8], [9] estimated a cubic B-spline curve-based vertical road profile using a Kalman filter. This method properly works if the 3-D points provided to the Kalman filter are from the actual road surface and are only contaminated by Gaussian noise. However, it is known that this method gives inaccurate results if the 3-D points include outliers such as points produced by obstacles or stereo matching errors [10], [11]. Keller *et al.* [10], [11] proposed an ROI-based outlier removal approach. This method selects ROIs according to the predicted driving corridor, variance of triangulation error, height distribution of 3-D points, and empirically designed fixed area. This method is able to stabilize the estimation results of the B-spline curve by reducing the number of outliers. However, it is difficult for the empirically selected ROIs to reject most of the outliers while retaining inliers under various road conditions. In particular, in the case of the distant road surface, it is hard to accurately select ROIs because errors of the driving corridor prediction and triangulation dramatically increase. This might be the reason why Keller *et al.* [10], [11] utilized this method only for road surfaces up to 30–40 m. Schauwecker and Klette [12] presented a robust estimator-based method. This method iteratively performs least squares estimation of a B-spline curve by growing the road surface region participating in the estimation process. It can be regarded as an M-estimator-based method with an

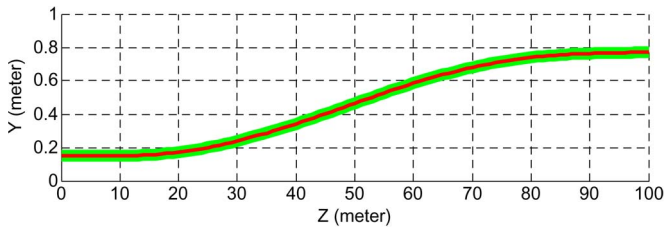

Fig. 2. Example of the function approximation. Thick green and thin red lines indicate a cubic B-spline curve and its approximation using a piecewise linear function, respectively.

empirical binary weighting function. The M-estimator-based method can robustly estimate a vertical road profile if the 3-D points are contaminated by a small proportion of outliers. However, it might converge to a local minimum when a large proportion of outliers are presented. There are mainly two reasons for this drawback. One is that the performance of the M-estimator is highly dependent on the initial estimate, which is difficult to correctly establish in the case of a large proportion of outliers [26]. The other is that the larger the proportion of outliers, the higher the possibility that the empirical binary weighting function selects the wrong 3-D points [12]. This situation frequently occurs when a road surface is severely occluded by obstacles close to the ego-vehicle or the proportion of outliers increases on the distant road surface due to the perspective projection.

III. Proposed Method

A. Two-Step Strategy

A cubic B-spline curve with a high DOF and nonlinearity is hard to estimate from the realistic data contaminated by outliers. This paper utilizes a two-step strategy that initially estimates a piecewise linear function and then obtains a cubic B-spline curve from the initial estimation result. In this paper, a road surface up to 100 m is modeled by a cubic B-spline curve with control points every 20 m, and this curve is approximated to 20 linear functions with 5-m interval. Fig. 2 shows an example of the function approximation. In this figure, thick green and thin red lines indicate a cubic B-spline curve and its approximation using a piecewise linear function, respectively. It can be observed that a cubic B-spline curve can be precisely approximated by a piecewise linear function.

This approximation can achieve robustness against outliers. Robust estimation of a function with a high DOF requires a large amount of computational complexity. If RANSAC [27] is utilized for estimating parameters of the two functions shown in Fig. 2, a cubic B-spline curve whose DOF is 8 requires $\log(1-p)/\log(1-w^8)$ iterations, but a piecewise linear function needs only $20 \times \log(1-p)/\log(1-w^2)$ iterations by assuming that adjacent linear functions are independent. p and w are the probability that at least one of the random samples is free from outliers and the proportion of inliers, respectively. If Hough transform [28] is utilized instead, the former requires N^8 voting bins, but the latter needs only $20 \times N^2$ voting bins. N is the number of voting bins in each parameter space. The decrease of computational complexity not only reduces the computing time but also raises the possibility of correct estimation. Since the piecewise linear function representing a vertical road profile is continuous, there are dependencies among the adjacent linear functions. However, they are temporarily ignored in this stage to achieve computational efficiency. These dependencies are restored during the parameter estimation stage by the smoothness term of dynamic programming. It will be discussed later in Section III-C.

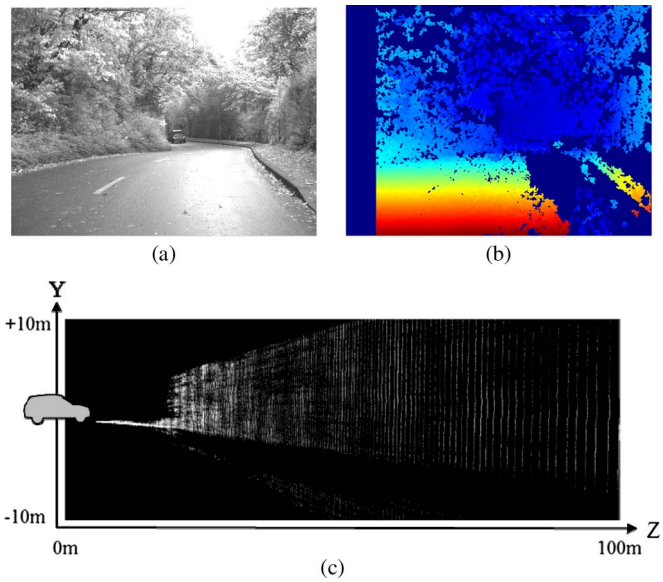

Fig. 3. (a) Left image of stereo camera. (b) Dense disparity map. (c) Three-dimensional point accumulation result in logarithmic scale.

B. Three-Dimensional Point Accumulation

A dense disparity map is utilized for vertical road profile estimation. Fig. 3(a) and (b) shows the left image of a stereo camera and the corresponding dense disparity map calculated by Semiglobal Matching (SGM) [29] implemented in [30], respectively. The road in this figure is sloped downward. All pixels in a dense disparity map are three-dimensionally reconstructed and accumulated on the YZ-plane after compensating the camera tilting angle obtained by an offline precalibration. The tilt angle compensation is for removing the effect caused by camera installation variation. Since each public database provides this value, this paper simply used it. This approach can achieve computational efficiency since numerous 3-D points are able to be represented by limited number of bins, and it has been also used in [3], [7], [21], and [22]. Fig. 3(c) shows an accumulation result of 3-D points on the YZ-plane. In the figure, the accumulation is displayed in logarithmic scale for better visualization. The horizontal axis indicates Z-coordinates from 0.0 m (left) to 100.0 m (right), and the vertical axis indicates Y-coordinates from -10.0 m (bottom) to $+10.0$ m (top). Bin size is set to 0.1 m $\times$ 0.1 m. For better understanding, a picture of a small car is depicted at the approximated ego-vehicle location. Due to the perspective projection, an image of a distant object appears smaller than that of a close object, although they are equal in metric size. This causes an accumulation inequality among objects located at different distances. Approximately,

an object size in an image is proportional to focal length, f, and inversely proportional to depth, Z. Thus, accumulation inequality due to the perspective projection can be moderated by multiplying the accumulation result by Z/f as

$$\bar{A}(j,i) = A(j,i)\frac{Z(i)}{f} \tag{1}$$

where $A(j,i)$ and $\bar{A}(j,i)$ are the original and moderated accumulation results, respectively; and $Z(i)$ and f are a depth corresponding to $A(j,i)$ and the focal length in pixel units, respectively. i and j are indexes of the accumulation result in horizontal and vertical directions, respectively. The origin of i and j is located at the top left corner in Fig. 3(c).

C. Parameter Estimation

A piecewise linear function has been used for representing a vertical road profile [5]–[7]. Labayrade *et al.* [5] estimated a piecewise linear function in v-disparity using Hough transform. This method estimates a vertical road profile by calculating an upper or a lower envelope of multiple straight lines that have the highest accumulation values. However, it can only handle slope changes in one direction. Suhr *et al.* [6] (our previous method) also utilized v-disparity image and Hough transform. This method divides a v-disparity image into several parts and applied Hough transform to the consecutive parts to find a piecewise linear function. Suhr and Jung [7] (our previous method) estimated a piecewise linear function in the YZ-plane using RANSAC. This method first samples the 3-D points expected to compose a road surface and estimates a piecewise linear function by adaptively changing the interval. Since the estimation procedure of the methods in [6] and [7] is sequentially conducted, a linear function in a subregion is influenced only by the linear function in the previous subregion. This may cause an estimation error in the near subregion to be propagated to the far subregions.

For robust and accurate estimation of a piecewise linear function, this paper proposes a method that utilizes Hough transform [28] and dynamic programming [31]. This combination is suited for this problem, since Hough transform produces voting results (fitness values) for all discrete parameter combinations within a search range and dynamic programming can directly use them as the data term for optimization. In addition, dependence of adjacent linear functions scarified for computational efficiency can be restored as the smoothness term. Unlike the previous methods [5]–[7], this method can produce more reliable estimation results and handle slope changes in both directions because it selects the most appropriate combinations of linear functions via global optimization. In this paper, Hough transform does not require a high computational cost because a small search range with rough resolution is sufficient for the given problem.

A piecewise linear function in the YZ-plane can be represented as

$$Y(Z) = \tan(\theta_n)\cdot Z + \rho_n, \qquad Z_{n-1} \le Z < Z_n \tag{2}$$

where θ_n and ρ_n are the slope angle and y-intersection at the start point of the nth subregion $Y(Z_{n-1})$, respectively. $Z_{n-1} \le Z < Z_n$ indicates the range of the nth subregion. Since θ_n means the slope of the road at the nth subregion, this method can handle the slanted roads. Dynamic programming is utilized to estimate the piecewise linear function. Cost of dynamic programming $C(\theta_n, \rho_n, \theta_{n+1}, \rho_{n+1})$ consists of two terms as

$$C(\theta_n, \rho_n, \theta_{n+1}, \rho_{n+1}) = D(\theta_n, \rho_n) + S(\theta_n, \rho_n, \theta_{n+1}, \rho_{n+1}) \tag{3}$$

where $D(\theta_n, \rho_n)$ and $S(\theta_n, \rho_n, \theta_{n+1}, \rho_{n+1})$ are the data and smoothness terms, respectively.

The data term is calculated by applying Hough transform to the 3-D point accumulation result shown in Fig. 3(c). Before applying Hough transform, the 3-D point accumulation result is manipulated. Since there is ideally no 3-D point underneath the road surface, a certain location that has a large accumulation value underneath it is unlikely to be a part of the road surface. Based on this, the accumulation result is manipulated by subtracting the maximum accumulation value below a certain location from the accumulation value of that location as (4). This procedure suppresses the accumulation values generated by obstacles above the road surface.

$$\underline{A}(j,i) = \overline{A}(j,i) - \max\left\{\overline{A}(j+1,i), \overline{A}(j+2,i), \ldots, \overline{A}(j_{\max},1)\right\} \tag{4}$$

where $j_{\max}$ indicates the maximum of row index j. After this manipulation, the road surface is divided into 20 subregions, and Hough transform is applied to the manipulated accumulation result ($\underline{A}$) of each subregion. During Hough transform, the range and interval of θ_n are set to -4.0° to $+4.0^\circ$ and 1.0°, respectively; and those of ρ_n are set to -5.0 m to $+5.0$ m and 0.1 m, respectively. Fig. 4(a) and (b) shows the divided subregions and corresponding Hough transform results, respectively. Horizontal and vertical axes of the Hough transform result are θ_n and ρ_n, respectively. In this figure, only 19 Hough transform results are presented because the region up to 5.0 m that is out of the stereo camera's field of view is omitted. These Hough transform results are used for the data term of dynamic programming as

$$D(\theta_n, \rho_n) = -H(\theta_n, \rho_n) \tag{5}$$

where $H(\theta_n, \rho_n)$ denotes the accumulation value of the voting bin at (θ_n, ρ_n) calculated by applying Hough transform to the nth subregion. A negative sign is added because dynamic programming is designed to solve a minimization problem.

The smoothness term consists of two subterms related with two properties of the piecewise linear function as

$$S(\theta_n, \rho_n, \theta_{n+1}, \rho_{n+1}) = S_1(\theta_n, \rho_n, \theta_{n+1}, \rho_{n+1}) + S_2(\theta_n, \theta_{n+1}). \tag{6}$$

The first subterm of (6) reflects the property that adjacent linear functions should meet at the region boundaries as

$$S_1(\theta_n, \rho_n, \theta_{n+1}, \rho_{n+1}) = \begin{cases} 0, & \text{if } R(\theta_n, \rho_n, \theta_{n+1}, \rho_{n+1}) \le T \\ \infty, & \text{otherwise} \end{cases} \tag{7}$$

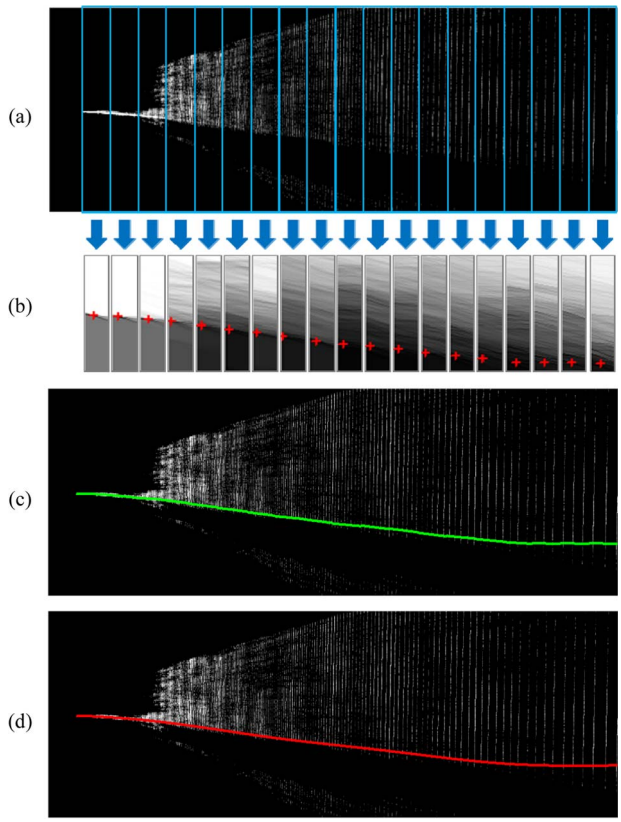

Fig. 4. Procedures of the proposed vertical road profile estimation method. (a) Divided sub-regions. (b) Hough transform results and estimated linear function parameters (red crosses). (c) Estimation result of the piecewise linear function (green line). (d) Estimation result of the cubic B-spline curve (red line).

where $R(\theta_n, \rho_n, \theta_{n+1}, \rho_{n+1})$ indicates the distance in Y-coordinates between adjacent linear functions at the region boundary, Z_n, as

$$R(\theta_n, \rho_n, \theta_{n+1}, \rho_{n+1}) = |\{\tan(\theta_n) \cdot Z_n + \rho_n\} - \{\tan(\theta_{n+1}) \cdot Z_n + \rho_{n+1}\}|. \quad (8)$$

If this distance is less than or equal to a predetermined threshold, T, the first subterm is set to 0, and otherwise, it is set to infinity. Thus, it provides a strong constraint that makes the adjacent linear functions connected to each other at the region boundary. T is set to the same value as the bin size used for the 3-D point accumulation (0.1 m) to cope with the quantization error caused by the accumulation process. The second subterm of (6) reflects the property that the slopes of adjacent linear functions are gradually changing as

$$S_2(\theta_n, \theta_{n+1}) = |\{(Z_{n+1} - Z_n) \cdot \tan(\theta_n)\} - \{(Z_{n+1} - Z_n) \cdot \tan(\theta_{n+1})\}|. \quad (9)$$

This term calculates the slope difference between adjacent linear functions in Y-coordinates within the range of the subregion $(Z_{n+1} - Z_n)$. To make the cost of dynamic programming well balanced, both the data term in (5) and the smoothness term in (6) are designed to have metric values. Fig. 4(b) and (c) shows the estimation result of the piecewise linear function via the proposed method. Red crosses on the Hough transform results in Fig. 4(b) indicate the parameters of linear functions for corresponding subregions, and a green line in Fig. 4(c) is the vertical road profile represented by the estimated piecewise linear function.

Since the piecewise linear function has a drawback in that slope changes are abrupt [8], [9], the vertical road profile is refined by estimating a cubic B-spline curve based on the estimation result of the piecewise linear function. To this end, points are uniformly sampled with intervals of 0.1 m on the estimated piecewise linear function, and the cubic B-spline curve is obtained by applying the least squares estimator to the sampled points. A red line in Fig. 4(d) shows the final vertical road profile estimation result with the cubic B-spline curve.

TABLE I
DATABASE DESCRIPTION

Database	Number of images	Stereo matching algorithm	Image resolution (pixels)
6D-VISION DB	500	SGM (implemented by Daimler)	1024 × 440
KITTI DB	287	SGM (implemented in [30])	1242 × 375 (slightly varying)
DAIMLER DB	306	SGM (implemented in [30])	640 × 480

IV. EXPERIMENTS

A. Database Description

Three publicly available databases were used for performance evaluation: 1) 6D-Vision Scene Labeling Database (6D-VISION DB) [32]; 2) KITTI Vision Benchmark Suite - Road Database (KITTI DB) [33]; and 3) Daimler Stereo Pedestrian Detection Benchmark Data Set (DAIMLER DB) [34]. The 6D-VISION DB and the KITTI DB were chosen because they provide the designation results of road surface region, and the DAIMLER DB was chosen because it includes various urban situations. The 6D-VISION DB and the KITTI DB provide 500 and 287 images with the road surface designation results, respectively. In the case of the DAIMLER DB, we selected uniformly distributed 306 images and manually designated their road surface regions. Since dense disparity maps are included only in the 6D-VISION DB, those of the KITTI DB and the DAIMLER DB were obtained by SGM implemented in [30]. Table I describes the test databases in detail. Fig. 5 shows examples of three databases. Fig. 5(a) shows the designated road surface region (red line) on a left stereo camera image, and Fig. 5(b) shows the corresponding dense disparity map. Examples in the first, second, and third rows are from the 6D-VISION DB, the KITTI DB, and the DAIMLER DB, respectively.

B. Performance Evaluation and Comparison

Performance evaluation was conducted by comparing the ground truth of the vertical road profile with the estimated vertical road profile. The ground truth of the vertical road profile was manually generated via the following steps.

1) Disparities inside the manually designated road surface region are selected.
2) Miscalculated disparities are manually removed.

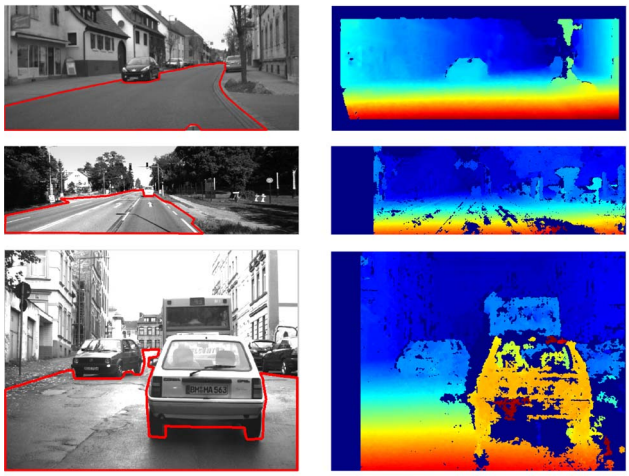

Fig. 5. Examples of test databases. (a) Designated road surface region (red line) on a left stereo camera image. (b) Corresponding dense disparity map. Examples in the first, second, and third rows are from the 6D-VISION DB, the KITTI DB, and the DAIMLER DB, respectively.

3) Retained disparities are three-dimensionally reconstructed.
4) Maximum distance of the road surface region visible in the current image is manually designated.
5) Ground truth of vertical road profile is obtained by estimating a cubic B-spline curve using the reconstructed 3-D points within the maximum distance via the least squares estimator.

We believe that this curve can be considered as the most accurate vertical road profile that can be estimated from the given dense disparity map and a cubic B-spline curve model. This procedure is a slightly modified version of [12] and [35].

For quantitative evaluation, this paper suggests a criterion called mean absolute vertical difference (MAVD). If G and E are the areas under the ground truth and estimated vertical road profiles, MAVD is calculated by dividing the symmetric difference of G and E $(G\Delta E)$ within the maximum distance of the road surface region (L) by L as

$$\text{MAVD} = \frac{G\Delta E}{L} = \frac{(G \cup E) - (G \cap E)}{L}. \tag{10}$$

Since MAVD means the average absolute difference between the ground truth and estimated vertical road profiles in Y-coordinates, a lower MAVD means better performance.

The proposed method was compared with three previous methods: M-estimator-based region-growing method [12], RANSAC-based sequential estimation method (our previous method) [7], and Hough transform and v-disparity-based method [5]. In the case of the method in [12], our implementation might be slightly different from the original one because detailed descriptions were omitted in their paper. Since the method in [7] was originally proposed to estimate a piecewise linear function, a cubic B-spline curve is estimated using the inliers that support the estimated piecewise linear function.

Table II shows the MAVDs of the four vertical road profile estimation methods in three different databases. It can be easily noticed that the proposed method outperforms the three previous methods in all three databases. On average, the proposed method gives MAVD of 9.6 cm, and the methods in [5], [7], and [12] give 18.6, 19.5, and 20.3 cm, respectively. The performance of the method in [12] is dramatically degraded in the DAIMLER DB. This is because the 6D-VISION DB and the KITTI DB mostly consist of situations where forward obstacles are distant from the ego-vehicle and the road surface is not much occluded, as shown in the first and second rows in Fig. 5(a). On the contrary, the DAIMLER DB includes many situations where a large proportion of the road surface is occluded by forward obstacles close to the ego-vehicle, as shown in the last row in Fig. 5(a). It reveals that the method in [12] is more sensitive to these situations compared with the others. This will be discussed in detail later. All four methods show the best results in the 6D-VISION DB. This is because the dense disparity maps included in this database are more reliable compared with those calculated by SGM implemented in [30], as shown in Fig. 5(b).

TABLE II
MAVDs OF THE FOUR METHODS

Database	Method in [12]	Method in [7]	Method in [5]	Proposed method
6D-VISION DB	13.6 cm	15.9 cm	13.4 cm	7.8 cm
KITTI DB	18.3 cm	24.1 cm	19.9 cm	11.1 cm
DAIMLER DB	32.9 cm	20.9 cm	25.7 cm	11.2 cm
Total	**20.3 cm**	**19.5 cm**	**18.6 cm**	**9.6 cm**

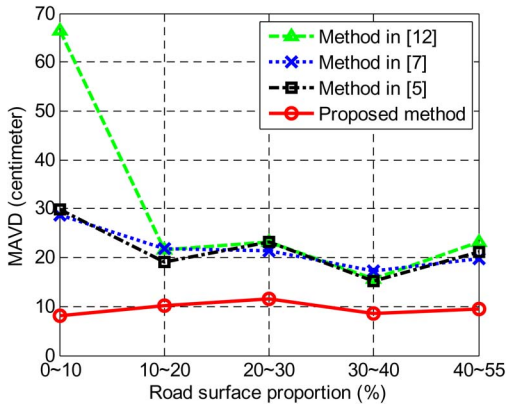

Fig. 6. Changes in MAVD with different road surface proportions.

There are mainly two reasons that make the proposed method outperform the previous methods. The first reason is that the proposed method is more robust against the situations with a large proportion of outliers compared with the previous methods. Fig. 6 shows the changes in MAVD with different road surface proportions. In this figure, black dash–dot, blue dotted, green dashed, and red solid lines indicate the MAVDs of the methods in [5], [7], and [12] and the proposed method, respectively. Since the road surface proportion is the ratio between the image of the road surface region and the whole image, a lower road surface proportion means a higher proportion of outliers. In Fig. 6, it can be easily observed that the proposed

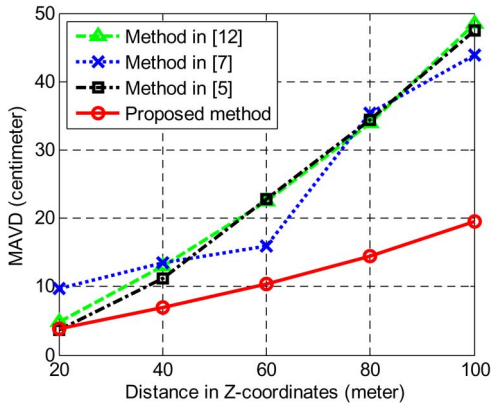

Fig. 7. Changes in MAVD with different distances in Z-coordinates.

method gives similar MAVDs with different road surface proportions, but the MAVDs of the other three methods are increased, as the road surface proportion is getting lower. Particularly, the MAVD of the method in [12] is more rapidly increased compared with the other methods. This is because the M-estimator of the method in [12] is more sensitive to outliers (3-D points not produced from the road surface) compared with the RANSAC of the method in [7] and the Hough transform of the method in [5] and the proposed method. This is a well-known drawback of the M-estimator [26]. Furthermore, the proposed method can more efficiently restrict the parameter space using Hough space and smoothness term of dynamic programming compared with the method in [12].

The second reason is that the MAVD of the proposed method gradually grows when the distance in Z-coordinates increases. Fig. 7 shows the changes in MAVD with different distances in Z-coordinates. In this figure, black dash–dot, blue dotted, green dashed, and red solid lines indicate the MAVDs of the methods in [5], [7], and [12] and the proposed method, respectively. It is obvious that the MAVD of the proposed method is more gradually increased compared with the MAVDs of the previous methods. This is because the proposed method estimates a vertical road profile using a global optimization via dynamic programming. That is, a vertical road profile in each subregion is determined by considering the vertical road profiles of all subregions. However, in the cases of the methods in [7] and [12], the vertical road profile in a subregion is estimated by considering only the previous subregion since they sequentially estimate a vertical road profile from near to far subregions. Due to this, the estimation error in the near subregion can be propagated to the far subregions. This causes the MAVDs of the methods in [7] and [12] to more rapidly increase compared with the proposed method. In the case of the method in [5], the estimation result highly depends on the road surface near the ego-vehicle because the contribution of this region to the v-disparity accumulation is dominant due to the perspective projection. This makes the MAVD of the method in [5] become larger when the distance in Z-coordinates increases. In addition, this result can be analyzed in terms of the outlier proportion. The proportion of road surface is decreased in an image when the distance in Z-coordinates is getting farther because of the perspective projection. This means that the 3-D

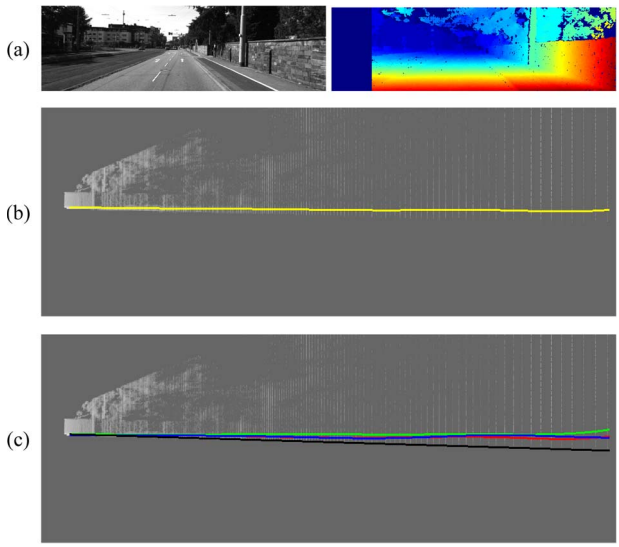

Fig. 8. Vertical road profile estimation result from the KITTI DB. (a) Left image of stereo camera and corresponding dense disparity map. (b) Ground truth vertical road profile (yellow line). (c) Vertical road profile estimation results of the method in [5] (black line), the method in [7] (blue line), the method in [12] (green line), and the proposed method (red line).

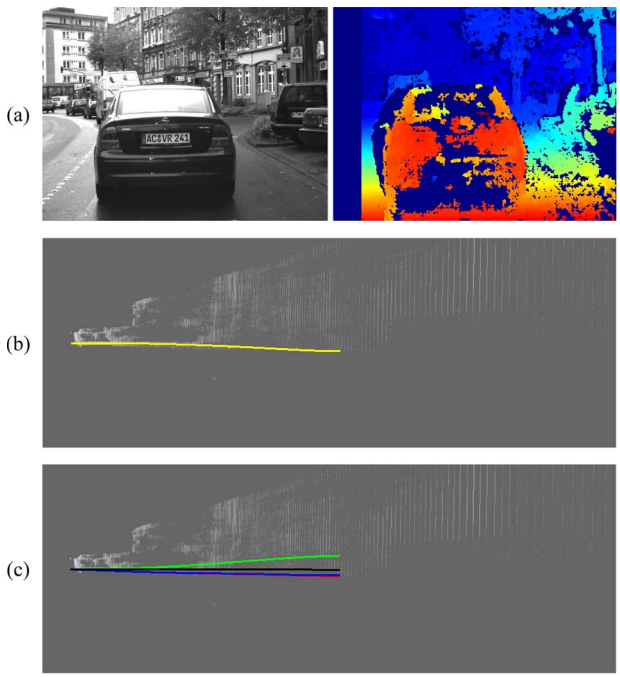

Fig. 9. Vertical road profile estimation result from the DAIMLER DB. (a) Left image of stereo camera and corresponding dense disparity map. (b) Ground truth vertical road profile (yellow line). (c) Vertical road profile estimation results of the method in [5] (black line), the method in [7] (blue line), the method in [12] (green line), and the proposed method (red line).

points produced from a distant subregion include a large proportion of outliers compared with those produced from a subregion close to the ego-vehicle. Since the proposed method is superior to the others in terms of robustness against outliers, as already discussed with Fig. 6, the proposed method can more accurately estimate the distant vertical road profile.

Figs. 8–10 show the vertical road profile estimation results of the four methods. Figs. 8(a), 9(a), and 10(a) show left images

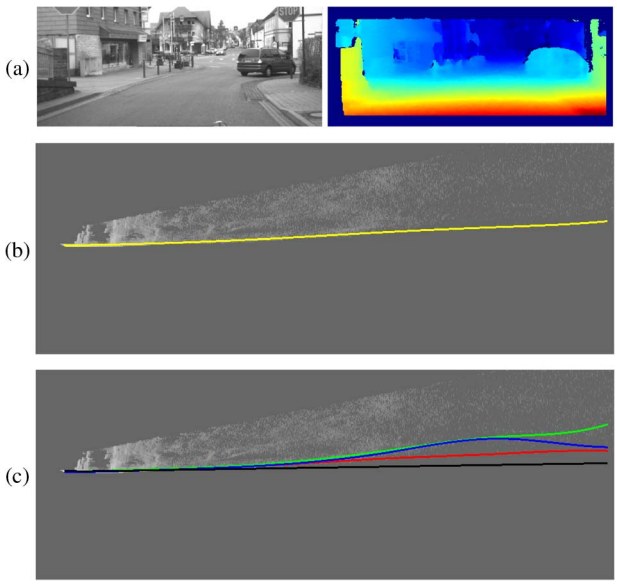

Fig. 10. Vertical road profile estimation result from the 6D-VISION DB. (a) Left image of stereo camera and corresponding dense disparity map. (b) Ground truth vertical road profile (yellow line). (c) Vertical road profile estimation results of the method in [5] (black line), the method in [7] (blue line), the method in [12] (green line), and the proposed method (red line).

of the stereo camera and corresponding dense disparity map, and Figs. 8(b) and (c), 9(b) and (c), and 10(b) and (c) show the ground truth vertical road profiles (yellow line) and the vertical road profile estimation results, respectively. Black, blue, green, and red lines in Figs. 8(c), 9(c), and 10(c) indicate the estimation results of the methods in [5], [7], and [12] and the proposed method, respectively. Fig. 8 shows the vertical road profile estimation result from the KITTI DB. In this example, a wide road surface is presented without occlusions caused by obstacles close to the ego-vehicle. The proposed method and the methods in [7] and [12] give reliable results in this case. However, the result of the method in [5] becomes unreliable when the distance in Z-coordinates increases. This is mainly due to its high dependence on the road surface close to the ego-vehicle. Fig. 9 shows the vertical road profile estimation result from the DAIMLER DB. A small proportion of road surface is visible due to the forward vehicle close to the ego-vehicle. The proposed method and the method in [7] successfully estimate the vertical road profile, and the method in [5] gives a result with a small error. However, the method in [12] fails due to its sensitivity to the outlier proportion. This result is consistent with the analysis in Fig. 6. Fig. 10 shows the vertical road profile estimation result from the 6D-VISION DB. In this example, the proportion of the road surface is dramatically decreased due to the perspective projection and the surrounding obstacles (buildings, vehicles, etc.), whereas the distance in Z-coordinates is increased. The proposed method accurately estimates the vertical road profile, whereas the three previous methods are getting more erroneous after the location of approximately 50 m. This result is consistent with the analysis in Fig. 7.

Some 3-D points are located under the ground truth, although there must be no 3-D points under the road surface in an ideal situation. This is because of the triangulation error presented in a real situation. This error causes the 3-D points from the road surface to become more scattered while its distance increases.

TABLE III
COMPUTATION TIMES OF THE FOUR METHODS

Database	Method in [12]	Method in [7]	Method in [5]	Proposed method	
	Matlab	Matlab	Matlab	Matlab	C
6D-VISION DB	1.72 s	0.94 s	0.38 s	1.10 s	0.042 s
KITTI DB	1.56 s	0.79 s	0.32 s	1.06 s	0.042 s
DAIMLER DB	1.37 s	0.59 s	0.41 s	0.99 s	0.038 s

C. Execution Time

All methods were implemented in MATLAB, and their execution times were measured on a 3.40 GHz Intel Core i7-2600 CPU with 4 G RAM. Table III shows the computation times of the four methods. The computation time of the method in [12] is higher than those of the other methods because it directly utilizes 3-D points, but the other methods utilize 2-D accumulation results. The proposed method was also implemented in C language, and it requires 38 ms in the DAIMLER DB and 42 ms in the 6D-VISION DB and the KITTI DB. Although the pixel resolutions of the 6D-VISION DB and the KITTI DB are almost twice that of the DAIMLER DB, there is a slight difference in execution time (only 4 ms). This is because the computational cost of the proposed method is consistent regardless of pixel resolutions once 3-D points are reconstructed and accumulated onto the YZ-plane. This is one of the advantages of using the YZ-plane accumulation. The computation time presented in Table III does not include stereo matching procedure.

V. SUMMARY AND FUTURE WORK

This paper has proposed a dense stereo-based vertical road profile estimation method. The proposed method utilizes a two-step strategy and combination of Hough transform and dynamic programming. The two-step strategy enables a cubic B-spline curve to be reliably estimated from the realistic data contaminated by outliers. The combination of Hough transform and dynamic programming makes the estimation procedure robust against a large proportion of outliers and guarantee optimal parameters. Experimental results show that the proposed method produces reliable vertical road profiles even in the cases of a large proportion of outliers and road surfaces distant from the ego-vehicle. The proposed method cannot properly work in cases of a severe roll angle and complete occlusion of the road surface. Thus, we are planning to overcome these drawbacks by adding an online roll angle calibration and road profile tracking in the future. In addition, we will improve the proposed method by compensating the increasing sparseness of the 3-D point accumulation considering the sensor error model.

REFERENCES

[1] R. Labayrade and D. Aubert, "Robust and fast stereovision based road obstacles detection for driving safety assistance," in *Proc. IAPR Workshop Mach. Vis. Appl.*, 2002, pp. 624–627.

[2] R. Benenson, R. Timofte, and L. V. Gool, "Stixels estimation without depth map computation," in *Proc. Int. Conf. Comput. Vis. Workshops*, 2011, pp. 2010–2017.

[3] S. Nedevschi *et al.*, "High accuracy stereovision approach for obstacle detection on non-planar roads," in *Proc. Intell. Eng. Syst.*, 2004, pp. 211–216.

[4] A. D. Sappa, R. Herrero, F. Dornaika, D. Gerónimo, and A. López, "Road approximation in Euclidean and V-disparity space: A comparative study," in *Proc. Lect. Notes Comput. Sci.*, 2007, vol. 4739, pp. 1105–1112.

[5] 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. IEEE Intell. Veh. Symp.*, 2002, pp. 646–651.

[6] J. K. Suhr, H. M. Kang, and H. G. Jung, "Dense stereo based critical area detection for an active pedestrian protection system," *Electron. Lett.*, vol. 48, no. 19, pp. 1199–1201, Aug. 2012.

[7] J. K. Suhr and H. G. Jung, "Noise-resilient road surface and free space estimation using dense stereo," in *Proc. IEEE Intell. Veh. Symp.*, 2013, pp. 461–466.

[8] A. Wedel, U. Franke, H. Badino, and D. Cremers, "B-spline modeling of road surfaces for freespace estimation," in *Proc. IEEE Intell. Veh. Symp.*, 2008, pp. 828–833.

[9] A. Wedel *et al.*, "B-spline modeling of road surfaces with an application to free-space estimation," *IEEE Trans. Intell. Transp. Syst.*, vol. 10, no. 4, pp. 572–583, Dec. 2009.

[10] C. G. Keller, D. F. Llorca, and D. M. Gavrila, "Dense stereo-based ROI generation for pedestrian detection," in *Proc. DAGM Symp. Pattern Recognit.*, 2009, pp. 81–90.

[11] C. G. Keller *et al.*, "The benefits of dense stereo for pedestrian detection," *IEEE Trans. Intell. Transp. Syst.*, vol. 12, no. 4, pp. 1096–1106, Dec. 2011.

[12] K. Schauwecker and R. Klette, "A comparative study of two vertical road modeling techniques," in *Proc. Asian Conf. Comput. Vis. Workshop*, 2010, pp. 174–183.

[13] S. Thrun, W. Burgard, and D. Fox, *Probabilistic Robotics*. Cambridge, MA, USA: MIT Press, 2005.

[14] P. F. Malartre, T. Feraud, C. Debain, and R. Chapuis, "Digital elevation map estimation by vision-lidar fusion," in *Proc. Int. Conf. Robot. Biomim.*, 2009, pp. 523–528.

[15] A. Broggi, E. Cardarelli, S. Cattani, and M. Sabbatelli, "Terrain mapping for off-road autonomous ground vehicles using rational B-spline surfaces and stereo vision," in *Proc. IEEE Intell. Veh. Symp.*, 2013, pp. 648–653.

[16] R. Hadsell, J. A. Bagnell, D. Huber, and M. Hebert, "Accurate rough terrain estimation with space-carving kernels," in *Proc. Robot. Sci. Syst.*, 2009, pp. 1–8.

[17] R. Hadsell, J. A. Bagnell, D. Huber, and M. Hebert, "Space-carving kernels for accurate rough terrain estimation," *Int. J. Robot. Res.*, vol. 29, no. 8, pp. 981–996, Jul. 2010.

[18] J. Ryde and H. Hu, "3D mapping with multi-resolution occupied voxel lists," *Auton. Robot.*, vol. 28, no. 2, pp. 169–185, Feb. 2010.

[19] A. Broggi, S. Cattani, M. Patander, M. Sabbatelli, and P. Zani, "A full-3D Voxel-based dynamic obstacle detection for urban scenario using stereo vision," in *Proc. Int. IEEE Conf. Intell. Transp. Syst.*, Oct. 2013, pp. 71–76.

[20] K. M. Wurm, A. Hornung, M. Bennewitz, C. Stachniss, and W. Burgard, "OctoMap: A probabilistic, flexible, and compact 3D map representation for robotic systems," in *Proc. Int. Conf. Robot. Autom. Workshop*, 2010, pp. 1–8.

[21] A. D. Sappa, D. Gerónimo, F. Dornaika, and A. López, "On-board camera extrinsic parameter estimation," *Electron. Lett.*, vol. 42, no. 13, pp. 745–747, Jun. 2006.

[22] A. D. Sappa, F. Dornaika, D. Ponsa, D. Gerónimo, and A. López, "An efficient approach to onboard stereo vision system pose estimation," *IEEE Trans. Intell. Transp. Syst.*, vol. 9, no. 3, pp. 476–490, Sep. 2008.

[23] J. K. Suhr, H. G. Jung, K. Bae, and J. Kim, "Automatic free parking space detection by using motion stereo-based 3D reconstruction," *Mach. Vis. Appl.*, vol. 21, no. 2, pp. 163–176, Feb. 2010.

[24] R. Danescu, F. Oniga, S. Nedevschi, and M.-M. Meinecke, "Tracking multiple objects using particle filters and digital elevation maps," in *Proc. IEEE Intell. Veh. Symp.*, 2009, pp. 88–93.

[25] F. Oniga and S. Nedevschi, "Processing dense stereo data using elevation maps: Road surface, traffic isle, and obstacle detection," *IEEE Trans. Veh. Technol.*, vol. 59, no. 3, pp. 1172–1182, Mar. 2010.

[26] P. Torr and D. Murray, "The development and comparison of robust methods for estimating the fundamental matrix," *Int. J. Comput. Vis.*, vol. 24, no. 3, pp. 271–300, Sep. 1997.

[27] M. Fischler and R. Bolles, "Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography," *Commun. ACM*, vol. 24, no. 6, pp. 381–395, Jun. 1981.

[28] R. O. Duda and P. E. Hart, "Use of the Hough transformation to detect lines and curves in pictures," *Comm. ACM*, vol. 15, no. 1, pp. 11–15, Jan. 1972.

[29] H. Hirschmüller, "Stereo processing by semi-global matching and mutual information," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 30, no. 2, pp. 328–341, Feb. 2008.

[30] OpenCV 2.1, 2013. [Online]. Available: http://opencv.willowgarage.com/

[31] R. Bellman, *Dynamic Programming*. Princeton, NJ, USA: Princeton Univ. Press, 1957.

[32] T. Scharwächter, M. Enzweiler, S. Roth, and U. Franke, "Efficient multi-cue scene segmentation," in *Proc. German Conf. Pattern Recog.*, 2013, pp. 435–445.

[33] J. Fritsch, T. Kuehnl, and A. Geiger, "A new performance measure and evaluation benchmark for road detection algorithms," in *Proc. Int. IEEE Conf. Intell. Transp. Syst.*, 2013, pp. 1693–1700.

[34] C. Keller, M. Enzweiler, and D. M. Gavrila, "A new benchmark for stereo-based pedestrian detection," in *Proc. IEEE Intell. Veh. Symp.*, 2011, pp. 691–696.

[35] K. Schauwecker, S. Morales, S. Hermann, and R. Klette, "A comparative study of stereo-matching algorithms for road-modeling in the presence of windscreen wipers," in *Proc. IEEE Intell. Veh. Symp.*, 2011, pp. 7–12.

Jae Kyu Suhr (M'12) received the B.S. degree in electronic engineering from Inha University, Incheon, Korea, in 2005 and the M.S. and Ph.D. degrees in electrical and electronic engineering from Yonsei University, Seoul, Korea, in 2007 and 2011, respectively.

He is a Research Assistant Professor with the Research Institute of Automotive Control and Electronics, Hanyang University, Seoul. His research interests include computer vision, image analysis, and pattern recognition for intelligent vehicle and visual surveillance.

Ho Gi Jung (M'05–SM'10) received the B.E., M.E., and Ph.D. degrees from Yonsei University, Seoul, Korea, in 1995, 1997, and 2008, respectively, all in electronic engineering.

From 1997 to April 2009, he was with MANDO Corporation Global Research and Development Headquarters. He developed environmental recognition systems for an intelligent parking assist system, a collision warning and avoidance system, and an active pedestrian protection system. From May 2009 to February 2011, he was a Full-Time Researcher and a Research Professor with Yonsei University. He researched computer vision applications for intelligent surveillance systems and biometric systems. Since March 2011, he has been an Assistant Professor with Hanyang University, Seoul. He is researching recognition systems for intelligent vehicles. His research interests include recognition system for intelligent vehicle, next-generation vehicle, computer vision applications, and pattern recognition applications.