International Journal of Innovative
Computing, Information and Control
Volume **16**, Number **4**, August **2020**

 ISSN 1349-4198
pp. 1225–1235

DEPTH-BASED REAR-OBSTACLE DETECTION APPROACH FOR DRIVING IN THE REVERSE GEAR

Po-Yuan Huang[1], Huei-Yung Lin[1] and Chin-Chen Chang[2,*]

[1]Department of Electrical Engineering
National Chung Cheng University
No. 168, Sec. 1, University Road, Minhsiung, Chiayi 621301, Taiwan

[2]Department of Computer Science and Information Engineering
National United University
No. 2, Lienda Road, Miaoli 36063, Taiwan
*Corresponding author: chinchen.chang@gmail.com

Received February 2020; revised June 2020

Abstract. *A vehicle is mainly driven in the forward direction and occasionally driven in the reverse gear. Blind spots for vehicles refer to the areas behind the vehicle that cannot be seen by a driver while driving. These spots are danger zones while driving a vehicle in the reverse gear. Therefore, vehicle safety in these spots is critical. In this paper, a rear-obstacle detection approach was proposed for drivers through stereo vision techniques. A binocular camera system based on parallax maps was used for rear obstacle detection. A stixel representation was developed for an outdoor traffic scene in a columnar manner. Also, disparity maps were used to estimate the ground region in the representation, identify free space in the representation, and segment the obstacles present in the image based on their height. Moreover, a road-edge-tracking step was incorporated in the proposed approach for stabilizing the approach. The experimental results reveal that the proposed approach exhibits high performance for rear obstacle detection.*
Keywords: Obstacle detection, Reverse driving, Stereo vision, Disparity map

1. **Introduction.** From a driver's viewpoint, accidents often occur in the areas behind vehicles. A vehicle's blind spot represents an area that cannot be directly viewed by drivers from their seats. The lines of sight of a driver seated in a normal position are obstructed by the body of the vehicle driven. In this case, errors can be easily made by a driver due to the blind spots, such as misjudgment and operation errors, thus leading to accidents [9]. Therefore, it is crucial for developing techniques for detecting an obstacle present behind a vehicle and for sending a warning signal to a driver while driving in reverse gear [21]. Such techniques are essential to reduce the errors made by drivers due to misjudgment, operation method, and poor driving skills [12].

Reliable sensors are crucial for obstacle detection. The different types of sensors used in vehicular applications include ultrasonic, microwave radar, laser rangefinder, and complementary metal oxide semiconductor (CMOS) or charge coupled device (CCD) image sensors. The measurement distance of ultrasonic sensors is short. Moreover, these sensors are easily influenced by weather conditions and generate additional noise. Microwave radar sensors are adopted for conducting long distance measurements and are generally not influenced by the weather. However, the price of these sensors is relatively high. The cost of the sensors used in this study to develop safety equipment for general vehicles is a critical factor. CMOS or CCD image sensors can be integrated with the processing units

DOI: 10.24507/ijicic.16.04.1225

on the back-end of line of a single chip. This integration can increase the computation speed and provide real-time performance. Thus, these sensors are core technologies for many low-cost applications [16]. In this study, we used stereo vision with a dual-camera system to detect the obstacles behind a vehicle.

In this paper, we presented a rear-obstacle detection approach for drivers through stereo vision techniques. The proposed approach can provide in-depth information of the obstacles and provide better measurement accuracy and efficiency. A technique for simplifying dual-camera stereoscopic parallax information was adopted in the image processing stage. This technique was used to convert a complex disparity map to a columnar pixel scene. We incorporated a ground-edge-tracking algorithm to improve the robustness of the camera system. Experiments were conducted using real scene images for demonstrating the feasibility of the proposed approach.

2. **Related Works.** The existing obstacle detection methods are very diverse [1-4,6,11, 14,15,17-19]. The obstacle detection methods that use images can be mainly categorized based on their usage of monocular or stereo vision systems for image acquisition and scene analysis.

The obstacle detection methods with monocular vision systems can be mainly categorized into three methods: feature-based tracking, model-based tracking, and optical-flow-based estimation [22]. Feature-based methods use vehicle characteristics, such as the aspect ratio of the license plate and the geometric symmetry of the vehicle, to reduce the amount of data that must be processed. The features are continuously tracked and compared in this method to determine the status of a moving vehicle [15]. However, the accuracy may be decreased for short-range detection of vehicles. An optical flow is mainly generated due to the relative motion between the camera and object. Thus, the intensities of the continuously obtained images change in terms of the time and spatial displacements. A foreground object can be separated from the background based on the difference between the moving speed of a vehicle and the background [17]. Optical-flow-based estimation methods can be classified into four categories based on their rules: differential-based, correlation-based, energy-based, and phase-based methods. Most studies are based on two optical-flow-based estimation methods, such as differential methods or associative methods. For example, edge information can be used to separate moving objects from the background by using different speeds of the moving objects and the background change [11]. However, this approach has a drawback that the motion boundary cannot be precisely detected for some situations.

Stereo vision approaches include four primary models for providing a scene representation: probabilistic occupancy map generation, digital elevation map generation, scene flow segmentation, and geometry-based clustering. The probability occupancy map displays a scene in a grid format [1], and the results of each grid are segregated into free, occupied, and unknown regions. After the distance values of the regions are calculated, probability is used to evaluate the state of each grid. Different coordinate systems are used for preparing a probability occupancy map. The first one is the Cartesian coordinate system. In this system, each grid is a square and is evenly distributed on a map. The other two coordinate systems use the columns of an image as the abscissa, and the ordinates represent the disparity and depth. However, the lateral space of the free space will be narrow for driving applications.

Badino et al. [2] converted a stereo image pair in a columnar pixel (stixel) representation through a probability occupancy map. They divided all the elements in the scene into three categories: the ground, vertical objects on the ground, and the sky. The ground is

extended infinitely in the occupancy map. Moreover, the objects that are placed perpendicular to the ground are represented using columns. This study adopted a dimensionality reduction scene representation. For example, if the columnar stixel width is set to five pixels, then an image with a width of 640 pixels only has 128 stixels. This simplification method considerably improves the post processing result required for object detection. However, heights of some objects cannot be accurately estimated such as traffic signs.

A numerical elevation map can be used to convert disparity maps or three-dimensional (3D) values to a digital elevation map (DEM) [13]. Density- and road-based obstacle classifiers were adopted. The density-based obstacle classifier uses the density of 3D points for marking a DEM unit as a road or an obstacle. Conversely, the road-based classifier employs the random sample consensus (RANSAC) method to detect obstacles present in front of the vehicle. However, this approach has a disadvantage of the robustness of detection due to poor road texture.

Optical flow methods are mainly used in scene stream segmentation to estimate motion and depth information between two frames captured using the same camera at different time instants. The principle of this segmentation method is that pixel-level segmentation is conducted on an image of a scene to detect an obstacle in the scene. Taylor et al. [20] calculated a scene flow through parallax and optical flow. A method known as "6D Vision" was proposed that considers vision as a process of space and time [7,14]. This method combines depth and motion information obtained from a sequence of images captured by mobile observers. The 3D position and 3D motion of many image points are simultaneously estimated using a Kalman filter. However, this approach has much more robust estimations. Due to the recent development of convolutional neural networks (CNN), an optical flow can be completed more quickly and accurately through training the CNNs [21]. The use of CNNs in obstacle detection becomes popular.

Although these studies have provided several techniques for obstacle detection, there are still some problems. In this paper, we presented a rear-obstacle detection approach for drivers through stereo vision techniques. The proposed approach can provide in-depth information of the obstacles and provide better measurement accuracy and efficiency.

3. **Proposed Approach.**

In this paper, we proposed a depth-based rear obstacle detection approach for detecting obstacles while driving in the reverse gear. In the proposed approach, two cameras that face backward were installed at the rear side of a vehicle. The proposed approach comprises three parts: a) free-space computation, b) road-edge tracking, and c) height-based segmentation.

In free-space computation, the proposed approach computes a disparity map with depth information. In conventional free-space computation methods, only a disparity map is usually employed, and the information obtained between consecutive images is not used. While conducting road-edge tracking, the system stability can be maintained appropriately by tracking the free space over time through the boundaries between the obstacles and free regions. In height segmentation, obstacles are segmented based on their heights. Moreover, the objects closest to the vehicle are specifically marked in the image. Then, the entire scene is precisely described for driving in the reverse gear. Figure 1 presents a flowchart of the proposed approach.

3.1. **Free space computation.**

A basic requirement while driving a vehicle is that the surrounding environment should be appropriately examined in all directions to check whether there are any obstacles in the path, and the distance between the obstacle and the vehicle should be estimated. Without including any traffic rules, the space between

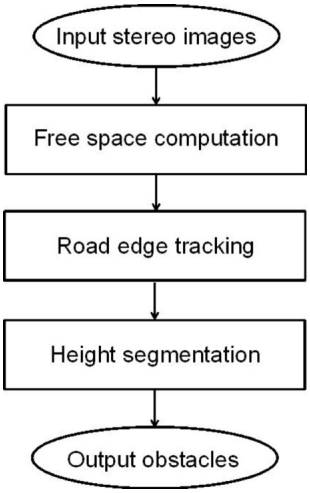

FIGURE 1. Flowchart of the proposed approach

an obstacle and a vehicle in which the vehicle can move freely is known as the free space in our system.

A semi-global block matching (SGBM) algorithm is used on an image pair acquired from a stereo camera rig to develop a disparity map with depth information [8]. To determine the road surface in the image pair, the disparity map with depth information is first statistically analyzed in the v direction to develop a v-disparity histogram [10]. The RANSAC method is used to find the slope that corresponds to the road surface in the v-disparity map [5]. The parameters used to represent a mathematical model in the RANSAC algorithm are estimated in an iterative manner by using an obtained data set containing outliers. The basic assumption of this method is that the data can be categorized into an "inner group" and "outlier group". The inner group contains the parameters pertaining to several sets of models, whereas the outliers are not suitable for the model. In this method, the road surface data are included in the inner group.

The disparity value on the slope pertains to the road surface. The parallax points are first eliminated from the original disparity map. Subsequently, dynamic programming (DP) is conducted to find the upper edge of the road surface. This process resembles the construction of a directed acyclic graph. Each pixel in the graph is a vertex, and any two pixels pertaining to adjacent columns are connected through edges. The weight of a vertex is the reciprocal of its placeholder value. This value is processed according to the difference in the number of rows equal to the predetermined value. The borders between the free space and obstacles are considered as the road edges.

3.2. **Road edge tracking.** A stability mechanism was incorporated for road-surface-edge tracking. The relationship between the consecutive images was considered so that the free-space calculation does not encounter a significant number of changes and that the free-space region can be derived on a pixel-by-pixel basis for each image.

A road point was set after every five pixels on the edge of the road to represent the road edge. The set decision points can only move upward and downward but cannot move in the horizontal direction. Here, each point independently represents a column. The advantage of such a setup is that only the calculation between columns is required, and the calculation rate can be increased by not considering the information between the rows. The operation methodology is the same as the process used for computing an obstacle's height. The system used for the road-surface tracking is not in the initial state when the operation is initiated. The edge point of the current image is determined through

the edge details of the road surface as the center of a searching region. A 5×5 pixel mask is used to compute the Euclidean distance between the edge point in the previous image and the edge point within a 10-pixel range in the current image. The edge point that has the minimum Euclidean distance value which is obtained from the 5×5 pixel mask, is used to update the decision point.

Finally, the road-surface edge point of the current image is determined by comparing the decision point location and the associated disparity in the current image. The average value is used when the absolute difference between the tracking change ΔI and the disparity change Δd is less than a preset threshold. A large $|\Delta I - \Delta d|$ value implies that the change on the road surface is significant. Moreover, the probability of the disparity computation being incorrect is very high. In this case, the image-tracking result obtained using ΔI is used as the final result. The advantage of this approach is that if the road surface cannot be correctly identified by the disparity map, then the image-tracking result can be adopted instead.

3.3. **Height segmentation.** After the free space in the images is calculated and eliminated from the disparity map, the remaining disparity values correspond to obstacles. The bottom portion of an obstacle that is closest to the ground is considered to be related to the road surface. This piece of information can be used to determine the obstacles that are closest to the vehicle. In the next step of the reverse-gear driving scenario, the height of an obstacle is calculated. Many problems may occur while determining the upper edge of an obstacle through a fixed threshold value for segmentation because we cannot select an appropriate threshold for the different disparity values.

Thus, the membership functions are combined to obtain a cost function for calculating the height of an obstacle. The height segmentation is defined by stixels [2], and each columnar pixel has the same disparity value. Thus, the disparity map of the obstacle on the ground described previously must be replanned. The primary steps involved in this replanning are as follows: pixel membership and cost function computations.

a) Membership Function:

The membership function is defined as follows:

$$M_d(u, v) = 2^{1-\left(\frac{d-d_u}{\Delta D_u}\right)^2} - 1, \tag{1}$$

where d represents the disparity value of the area to be calculated, ΔD_u is the calculation parameter, and d_u denotes the initial disparity value of an obstacle obtained from the free space, that is, the bottom disparity value of the obstacle object in the uth column. Figure 2 shows the membership function.

$$\Delta D_u = d_u - f_d(Z_u + \Delta Z_u), \quad f_d(z) = \frac{b \cdot f_x}{z}. \tag{2}$$

Here, $f_d(z)$ is the parallax corresponding to the depth z, b represents the stereo baseline, and f_x is the focal length. The result of the exponential membership function is used to approximate the Boolean function. The membership function is attributed an object or a background when $M_d(u, v) > 0$ or $M_d(u, v) \leq 0$, respectively.

b) Cost Function:

The cost function is calculated after the pixel membership degree is derived for each point in the disparity map. The larger a cost function is, the higher is the probability that the function is associated with the edge of an obstacle's height. The cost function is defined as follows:

$$C(u, v) = \sum_{f=v}^{v_b} M(u, f) - \sum_{b=0}^{v-1} M(u, b), \tag{3}$$

where v_b is the initial disparity of the foreground and v is the currently calculated point. Equation (3) is used to accumulate the current points from the bottom region of an obstacle and subtract those points from the points pertaining to the top portion of the obstacle to obtain the cost. The higher the cost value is, the higher is the probability that the point is at the edge.

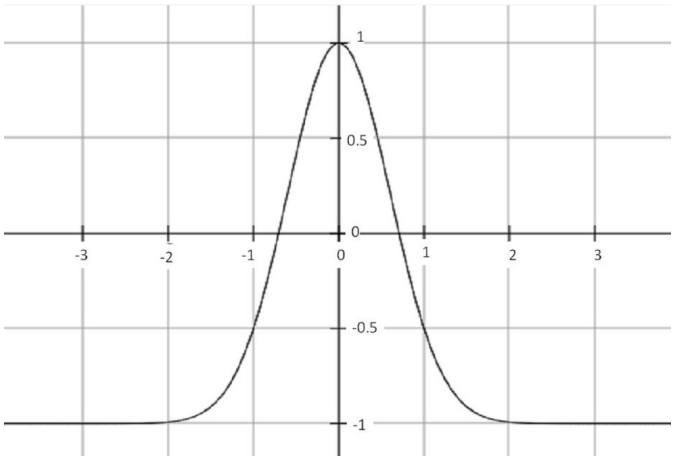

FIGURE 2. Membership function

DP is conducted after developing the cost map of the image to find the segmentation of the obstacle's height. After the pixels corresponding to these regions are identified, an obstacle is represented by its bottom portion and height. An obstacle is considered to contain five pixels in terms of its width for the representation in this study. Each group of column pixels is arranged based on the obstacle height by considering the median value to reduce noise in the disparity map. In the final stage, the disparity and height of the columnar pixels are assessed. If both disparity value and height are small, then the pixel is eliminated to avoid the presence of noise associated with the ground.

4. **Experimental Results.** The stereo vision system used in our experiments comprises two industrial cameras (acA2500-14gc, Basler), and the stereo baseline is set at 14 cm. The system was mounted near the rear windshield of a vehicle at approximately 1.2 m from the ground. The system can capture images at a resolution of 860×644 pixels. This resolution was reduced to 430×322 pixels to remove unnecessary areas and increase the computation speed of disparity maps. A desktop computer with Windows 10, Intel i7 8700 (3.2 GHz), and 16 GB memory was used as the computation platform. Microsoft Visual Studio 2013 and OpenCV library 2.4.13 were used in the implementation with C and C++. The total execution time for an image is approximately 360 ms. The major part of the computation is the generation of disparity through stereo matching by using SGBM, and this step requires approximately 230 ms. The time required for object-height-segmentation and road-tracking calculations is approximately 50 and 80 ms, respectively.

For free space computation, the ground region cannot be correctly measured using the v-disparity map while calculating the free space due to some reasons. Errors are caused in the free-space computation if the disparity map created is incomplete and contains blank areas. Refer to Figure 3 as an illustration. For height segmentation, the cost is first calculated and the obstacle boundary is identified by the height. The obstacles are then segmented from the free space. Figure 4 displays the segmentation results of the obstacles.

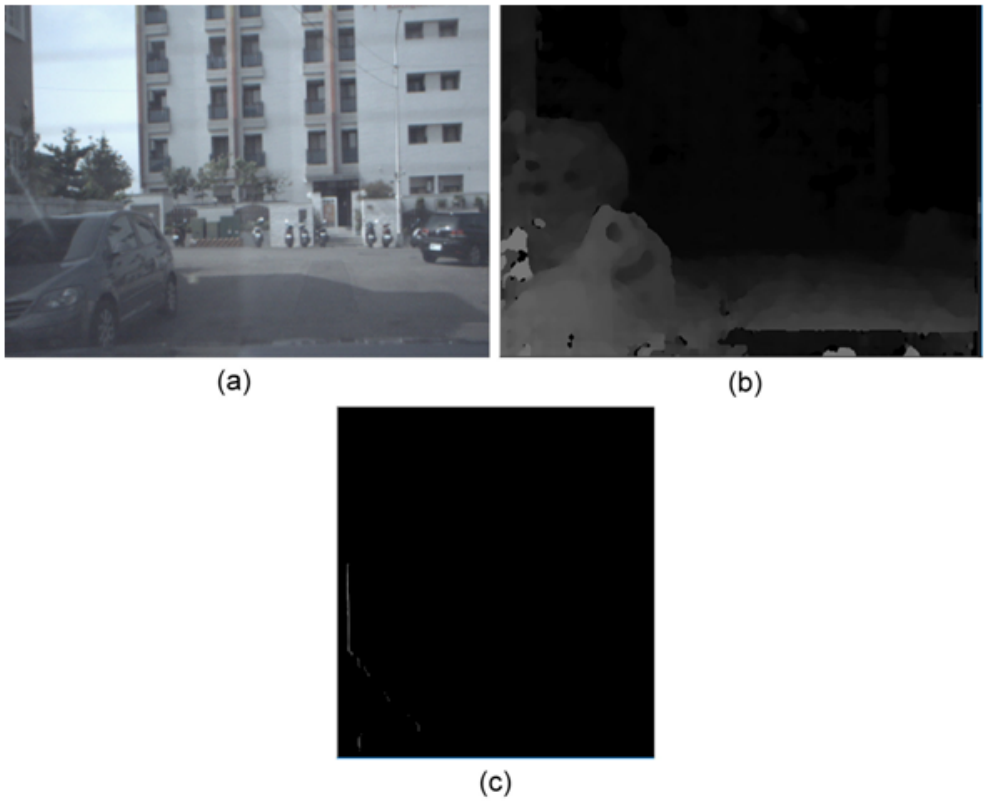

FIGURE 3. (a) Original image, (b) incomplete disparity map, and (c) corresponding v-disparity

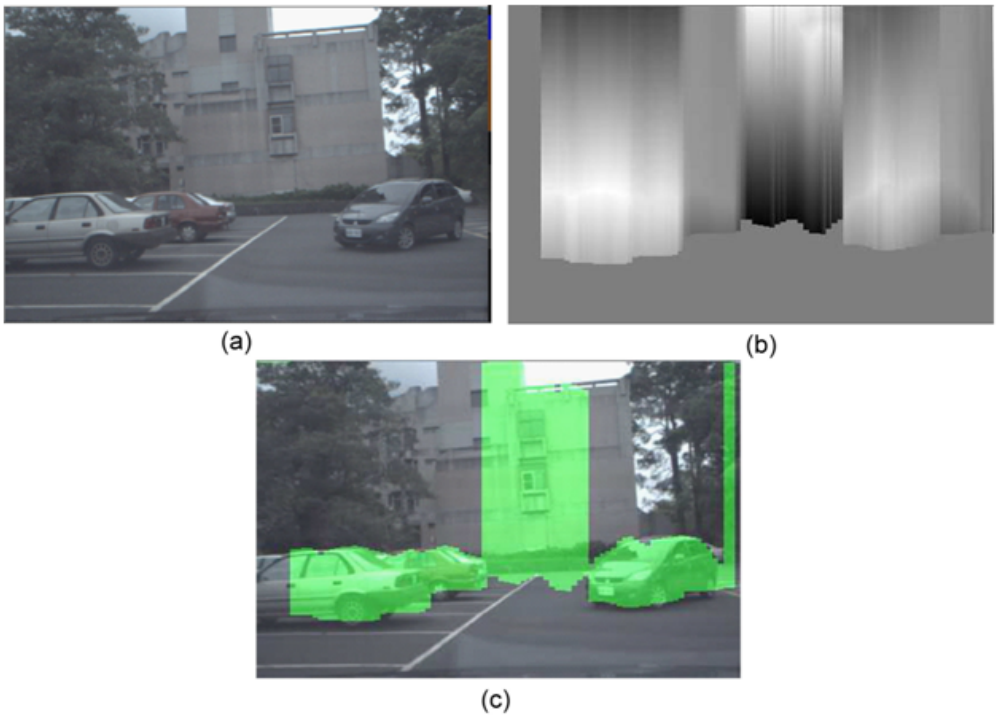

FIGURE 4. (color online) (a) Original image, (b) cost presented in an image, and (c) obstacle segmentation result

Four image sequences were tested with all scenes captured when the vehicle is moving in the reverse direction. Each sequence comprises approximately 100 to 200 image frames. All the stixels in the images were examined. A stixel is considered incorrect if a selected area is less than half of the height of an obstacle or higher than the height of the obstacle. The quantitative evaluation is displayed in Table 1. The input images and segmentation outputs are displayed in Figure 5. The free space (road), vehicles, and other obstacles are marked in different colors.

TABLE 1. Results of three image sequences and the quantitative evaluation

	Number of images	Stixels	Error stixels	Stixels accuracy
Scene 1	146	11388	3065	73.09%
Scene 2	97	7663	2453	67.99%
Scene 3	117	6849	914	86.66%

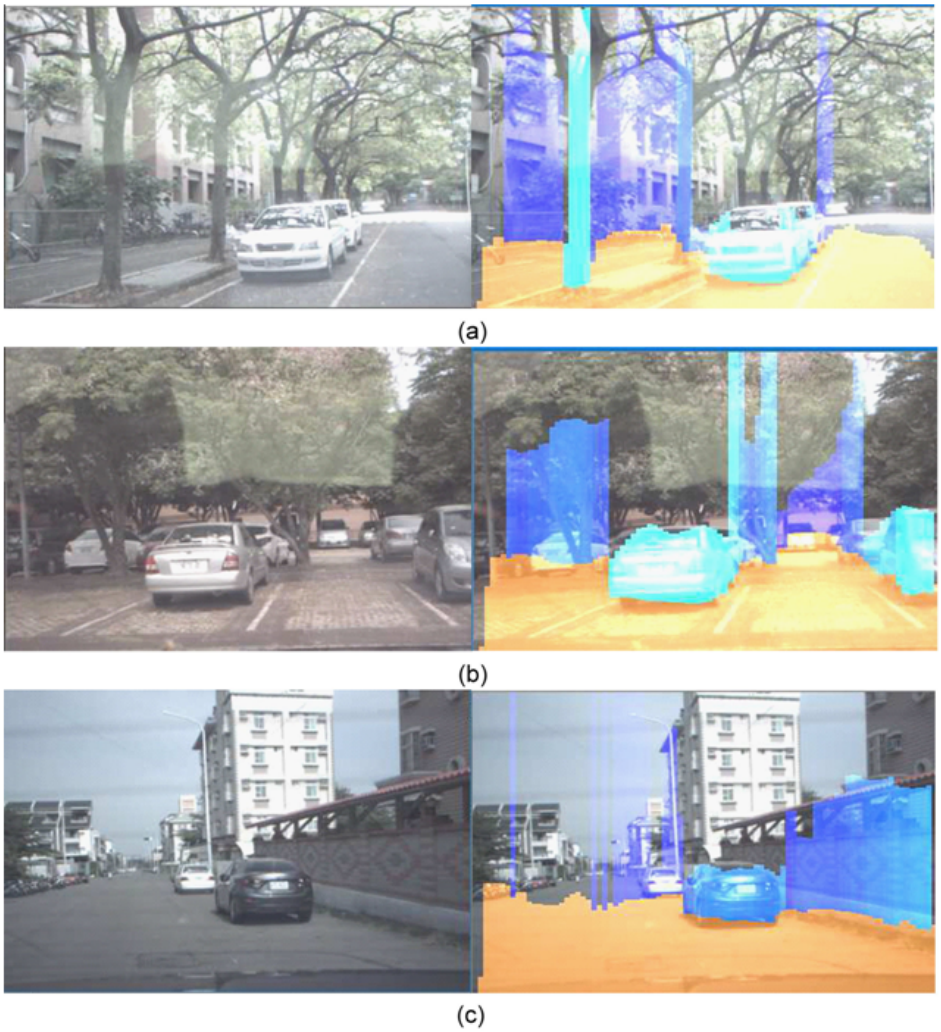

FIGURE 5. (color online) Input images and segmentation results: (a) Scene 1, (b) Scene 2, and (c) Scene 3

The main error occurrences in the results can be divided into three categories: free-space computation error, road-surface-tracking deviation due to road segmentation error, and height segmentation with excessive sensitivity or dullness. If RANSAC finds that

the slope of the road surface in the v-disparity histogram does not fit appropriately with the real road surface during the free-space calculation, then the free space is considered incorrect and should be marked as an obstacle. If an error always occurs in the road-surface tracking and the parallax, then the cumulative deviation of the edge tracking causes the road segmentation to fail after a period of time. The error observed in the height segmentation calculation occurs due to two reasons. The first reason is that the disparity map has many blank regions and the segmentation height is small. Figure 6 shows an illustration. The second reason is that it occurs far away and the disparity map is not obvious. Hence, errors can easily occur in calculations. Moreover, because this segmentation method is based on the height of the obstacle closest to the car, overlapping information may sometimes be ignored. As illustrated by Figure 7(a), the bush at the lower right corner is more prominent in the disparity map 7(b). Hence, for the height segmentation, only the height of the bush was segmented, and the existence of the iron sheet at the back is ignored in Figure 7(c). In such conditions, the driver's judgment may sometimes be affected.

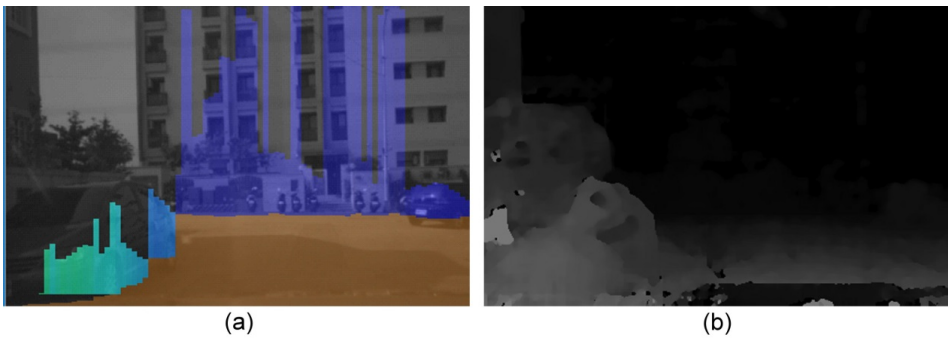

FIGURE 6. (color online) (a) Small height of the original segmentation, and (b) disparity map with many blank regions

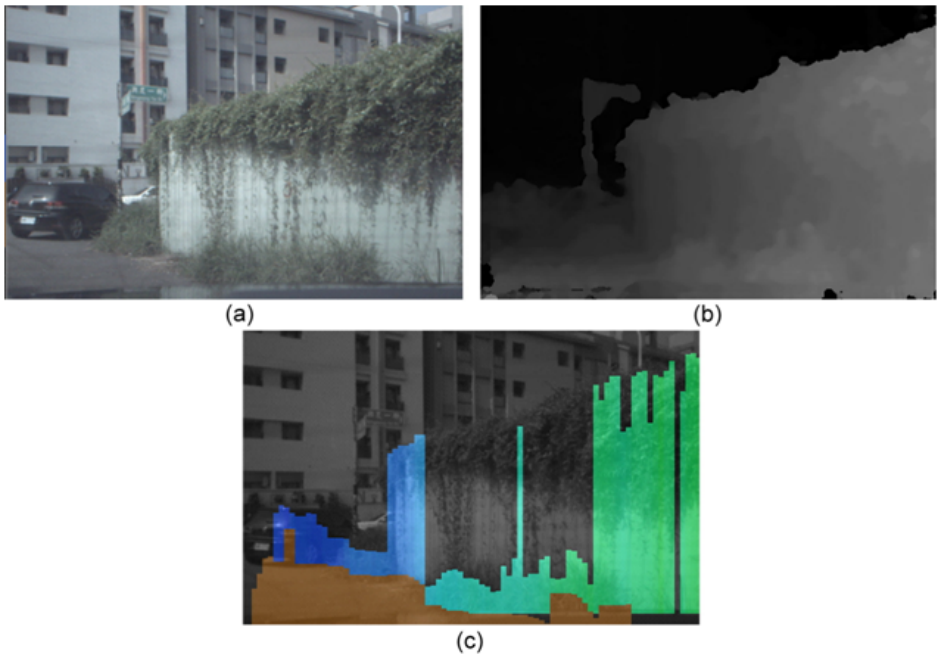

FIGURE 7. (color online) (a) Original image, (b) disparity map, and (c) segmented results

Here, we analyzed the proposed approach and theoretically compared it with previous methods. For free space computation, we added an extra technique to stabilize the calculation of free space. For previous methods, the calculation of free space required only information from the disparity map, and no connections were formed between consecutive images, resulting in system instability. Therefore, we track the intersecting edges between the free space and the obstacle to link the information from the previous and subsequent images. For road edge tracking, we added a road edge tracking stabilization mechanism for the information between the previous and subsequent images; this prevents wide fluctuations in the calculation of free space, stabilizes the calculation, and ultimately ensures that a free space can be identified for every image. For previous methods, free space calculation errors occur for various reasons, such as when the ground region cannot be correctly measured using the v-disparity map or when the calculation matching result is a broken disparity map. For height segmentation, we combine membership functions into cost functions to calculate obstacle height. For previous methods, when determining the upper edges of obstacles, numerous problems occur if a fixed threshold value is directly used for segmentation, and a suitable threshold value cannot be chosen among various disparity values. In the proposed approach, we combine membership functions into cost functions to calculate obstacle height.

5. **Conclusions.** We have presented an obstacle detection approach to provide information pertaining to rear obstacles to drivers while driving in the reverse gear. A stereo vision system was used in the proposed method and was installed near the rear windshield of a vehicle. Stereo matching was employed to derive disparity maps for a sequence of images. These maps were then used to identify the free space and the height of an object to guide drives while driving in reverse gear. Results reveal that the effectiveness of the proposed method is presented using real scene images.

In the future, because the calculations for this method are based on the depth disparity map, the overall accuracy of the system can be enhanced if learning methods are used to synthesize the disparity map with higher levels of accuracy. Moreover, we will add a route-planning component to establish a complete system for aiding car reversal.

Acknowledgment. This paper is a revised and expanded version of a paper entitled, *Rear Obstacle Warning for Reverse Driving Using Stereo Vision Techniques*, presented at 2019 IEEE International Conference on Systems, Man, and Cybernetics (SMC 2019), 2019.

REFERENCES

[1] H. Badino, U. Franke and R. Mester, Free space computation using stochastic occupancy grids and dynamic programming, *Proc. of International Conference on Computer Vision, Workshop on Dynamical Vision*, 2007.

[2] H. Badino, U. Franke and D. Pfeiffer, The stixel world – A compact medium level representation of the 3D-world, *Proc. of the 31st DAGM Symposium on Pattern Recognition*, pp.51-60, 2009.

[3] H. Bingo and H. Ohtake, Two-dimensional obstacle avoidance behavior based on three-dimensional environment map for a ground vehicle, *ICIC Express Letters, Part B: Applications*, vol.9, no.6, pp.485-492, 2018.

[4] A. Broggi, M. Buzzoni, M. Felisa and P. Zani, Stereo obstacle detection in challenging environments: The VIAC experience, *Proc. of 2011 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp.1599-1604, 2011.

[5] M. A. Fischler and R. C. Bolles, Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography, *Communications of the ACM*, vol.24, no.6, pp.381-395, 1981.

[6] A. M. Fouad, R. M. Sharkawy and A. Onsy, Fixed obstacle detection for autonomous vehicle, *Proc. of 2019 IEEE Conference on Power Electronics and Renewable Energy (CPERE)*, 2019.
[7] U. Franke, C. Rabe, H. Badino and S. Gehrig, 6D-vision: Fusion of stereo and motion for robust environment perception, *Proc. of the 27th DAGM Symposium on Pattern Recognition*, pp.216-223, 2005.
[8] H. Hirschmuller, Accurate and efficient stereo processing by semi-global matching and mutual information, *Proc. of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2005)*, vol.2, pp.807-814, 2005.
[9] M. Keall, B. Fildes and S. Newstead, Real-world evaluation of the effectiveness of reversing camera and parking sensor technologies in preventing backover pedestrian injuries, *Accident Analysis & Prevention*, vol.99, pp.39-43, 2017.
[10] R. Labayrade, D. Aubert and J. Tarel, Real time obstacle detection in stereovision on nonflat road geometry through "*v*-disparity" representation, *Proc. of IEEE Intelligent Vehicle Symposium*, vol.2, pp.646-651, 2002.
[11] Y. Mae, Y. Shirai, J. Miura and Y. Kuno, Object tracking in cluttered background based on optical flow and edges, *Proc. of the 13th International Conference on Pattern Recognition*, vol.1, pp.196-200, 1996.
[12] M. M. Monwar and B. V. K. Vijaya Kumar, Vision-based potential collision detection for reversing vehicle, *Proc. of 2013 IEEE Intelligent Vehicles Symposium Workshops*, pp.81-86, 2013.
[13] F. Oniga and S. Nedevschi, Processing dense stereo data using elevation maps: Road surface, traffic isle, and obstacle detection, *IEEE Trans. Vehicular Technology*, vol.59, no.3, pp.1172-1182, 2010.
[14] C. Rabe, U. Franke and S. Gehrig, Fast detection of moving objects in complex scenarios, *2007 IEEE Intelligent Vehicles Symposium*, Istanbul, pp.398-403, 2007.
[15] B. Ran and H. Liu, Development of vision-based vehicle detection and recognition system for intelligent vehicles, *Transportation Research Record: Journal of the Transportation Research Board*, vol.1679, pp.130-138, 1999.
[16] B. Ranft and C. Stiller, The role of machine vision for intelligent vehicles, *IEEE Trans. Intelligent Vehicles*, vol.1, no.1, pp.8-19, 2016.
[17] A. Schaub, D. Baumgartner and D. Burschka, Reactive obstacle avoidance for highly maneuverable vehicles based on a two-stage optical flow clustering, *IEEE Trans. Intelligent Transportation Systems*, vol.18, no.8, pp.2137-2152, 2017.
[18] M. Shimakawa, I. Taguchi, C. Okuma and K. Kiyota, Smartphone application program of obstacle detection for visually impaired people, *ICIC Express Letters, Part B: Applications*, vol.10, no.3, pp.219-226, 2019.
[19] A. Talukder, R. Manduchi, A. Rankin and L. Matthies, Fast and reliable obstacle detection and segmentation for cross-country navigation, *Proc. of IEEE Intelligent Vehicle Symposium*, vol.2, pp.610-618, 2002.
[20] L. A. Taylor, B. W. Eakins, K. S. Carignan, R. R. Warnken, D. C. Schoolcraft, T. S. Sazonova and G. Sharman, *Digital Elevation Model of Dutch Harbor, Alaska: Procedures, Data Sources and Analysis*, 2008.
[21] L. Wu and H. Lin, Overtaking vehicle detection techniques based on optical flow and convolutional neural network, *Proc. of the 4th International Conference on Vehicle Technology and Intelligent Transport Systems*, pp.133-140, 2018.
[22] H. Zhu, K. Yuen, L. Mihaylova and H. Leung, Overview of environment perception for intelligent vehicles, *IEEE Trans. Intelligent Transportation Systems*, vol.18, no.10, pp.2584-2601, 2017.