

Improved Vehicle Detection Algorithm in Heavy Traffic for Intelligent Vehicle

Chung-Hee Lee¹, Young-Chul Lim¹, Dongyoung Kim¹, Kyu-Ik Sohng²

¹Daegu Gyeongbuk Institute of Science & Technology(DGIST),

50-1 Sang-Ri, Hyeonpung-Myeon, Dalseong-Gun, Daegu, Republic of Korea

²School of Electrical Engineering and Computer Science, Kyungpook National University,

1370 Sankyug-dong, Buk-gu, Daegu, Republic of Korea

chlee@dgist.ac.kr

Abstract—Despite significant progress in vehicle detection over the last few decades, vehicle detection performance in heavy traffic is still inadequate. In this paper, we propose a new algorithm for vehicle detection in heavy traffic to improve detection performance. It uses two proposed segmentation methods, namely, the disparity map-based bird's-eye-view mapping segmentation method and the edge distance weighted conditional random field (CRF)-based segmentation method. Our experimental results show that the proposed algorithm outperforms conventional algorithms. The improvements in performance range from 10.8 % to 20.5 % increase in F-measure.

Index Terms—Intelligent vehicle, image segmentation, object recognition, stereo image processing.

I. INTRODUCTION

Forward vehicle detection is one of the most important technologies in an intelligent vehicle, which is closely related to autonomous driving, collision prevention, etc [1]. Many researchers have proposed vehicle detection methods using various sensors [2], [3]. Camera sensors in particular have many advantages, such as possession of a large amount of information and low cost compared to other sensors. Thus, the vehicle detection algorithms based on images have been studied for a long time [4]. Widely studied methods include the AdaBoost classifier [5] using intensity information and u/v-disparity [6] or column detection [7], [8] using a disparity map. The Adaboost classifier can cause many false alarms while detecting vehicles in occluded situations, such as heavy urban traffic, since the classifier solely uses intensity-based features. A disparity map-based vehicle detection method can minimize the problems by using three-dimensional information. However, it is also difficult to detect vehicles accurately in heavy traffic because of a stereo matching error and the precision of the disparity map.

In this paper, we focus on the improvement of vehicle detection performance in heavy traffic based on the proposed segmentation methods. In most cases, extracting exact obstacle areas from images is a very important step for improving vehicle detection performance. Although a

disparity map can provide approximate boundaries of obstacles and backgrounds, there are still many limitations in relying on only the disparity information. Intensity and edges are other important cues for segmentation in small obstacle areas, namely results of disparity map-based obstacle detection and segmentation. Thus, two proposed segmentation methods, disparity map-based bird's-eye-view mapping segmentation and edge distance weighted conditional random field (CRF)-based segmentation are proposed here for accurate vehicle detection. The segmentation method which is suitable for on-road vehicle detection, has not been proposed yet.

The remainder of this paper is organized as follows. Section II presents the proposed disparity map-based obstacle detection and segmentation algorithm. Section III describes the proposed intensity-based obstacle segmentation algorithm. In Section VI, we describe the area selection and verification method. The experimental results are presented in Section V. Finally, conclusions of this paper are given in Section VI.

II. DISPARITY MAP-BASED OBSTACLE DETECTION AND BIRD'S-EYE-VIEW MAPPING SEGMENTATION

Our proposed algorithm is described in detail according to steps and a block diagram is shown in Fig. 1. First, all road obstacles are detected by using road feature information extracted by a previously proposed method [7]. In our previous work [7], we found that it was very important to extract road feature information robustly to improve detection performance in various traffic situations. However, there are still many obstacles and backgrounds in the detected areas. Thus, each area needs to be segmented more accurately.

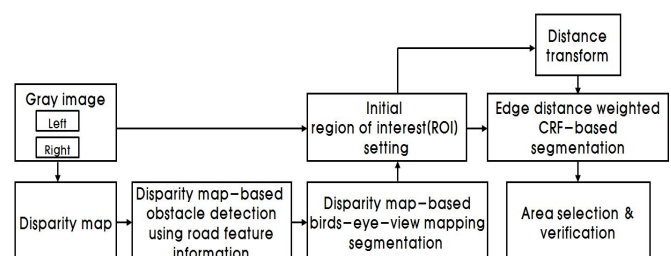


Fig. 1. Block diagram of vehicle detection algorithm.

An automotive stereo vision system model is shown in

Fig. 2. Two cameras are located h from the ground plane in the world coordinate system $(R_w(X_w, Y_w, Z_w))$, and b and θ imply the baseline and pitch, respectively. Each camera coordinate system is defined as R_{cl} (the left camera) and R_{cr} (the right camera), and the image plane is also defined as $I(u, v)$. The detected areas in the disparity map are mapped into a bird's-eye-view using stereo vision modelling equations and projection of the X - Z plane, which can be calculated by:

$$X = (u_l - u_r - 2u_0) \frac{\{(Y + h) \sin \theta + Z \cos \theta\}}{2fm}, \quad (1)$$

$$Y = \frac{(v_l - v_r)b \cos \theta + fmb \sin \theta - dh}{d}, \quad (2)$$

$$Z = \frac{b\{fm \cos \theta - (v_l - v_r) \sin \theta\}}{d}, \quad (3)$$

where $P(X, Y, Z, 1)^T$ is a point in the world coordinates, namely, the lateral, vertical, and longitudinal positions, respectively; u_l and u_r are the horizontal positions of the point in the left and right images, respectively; and v is the vertical position of the point. Here, (u_0, v_0) is the center of the image, and d is the disparity, which can be expressed by $u_l - u_r$; f is the focal length of the camera; and m is the number of pixels per unit distance in the image plane.

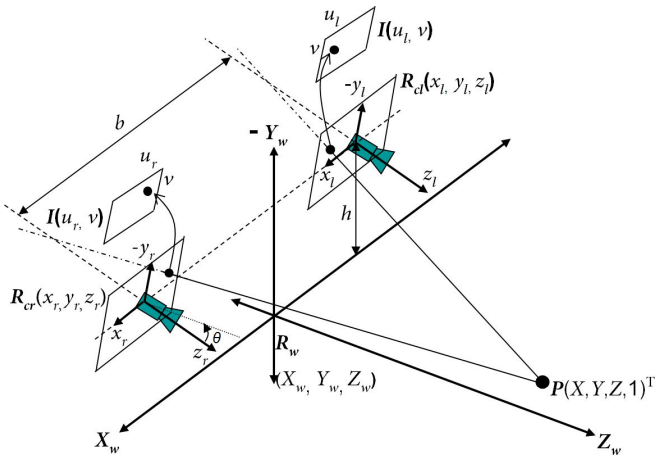


Fig. 2. Automotive stereo vision system model.

The positions of obstacles are identified very well, because the results of (1)–(3) can be represented on a flat plane by projecting them on an X - Z plane. Thus, the obstacles can be segmented more accurately and easily with disparity map-based bird's-eye-view mapping. Segmentation in the bird's-eye view consists of two steps, row segmentation and post-segmentation. Row segmentation is performed by detecting peaks and valleys in each row of the bird's-eye view. Post-segmentation is then performed additionally, because the obstacles are divided excessively in the case of long obstacles such as guide rails, median barriers, etc. When the centre position and disparity value of the segmented obstacles are changed regularly, the obstacles are regarded as long obstacles and remerged in the post-segmentation. Thus, very long obstacles can be identified and removed. We can also remove an obstacle whose height is too tall or too short using the disparity information. From (2), the obstacle height

$(v_1 - v_2)$ in the image can be expressed by

$$v_1 - v_2 = (Y_1 - Y_2) \frac{d}{b \cos \theta}, \quad (4)$$

where $Y_1 - Y_2$ is the obstacle height in the world coordinate. Thus, if we predefine a vehicle's height ($Y_1 - Y_2$), an obstacle whose height is longer or shorter than the vehicle's height can be removed in the disparity map-based segmentation.

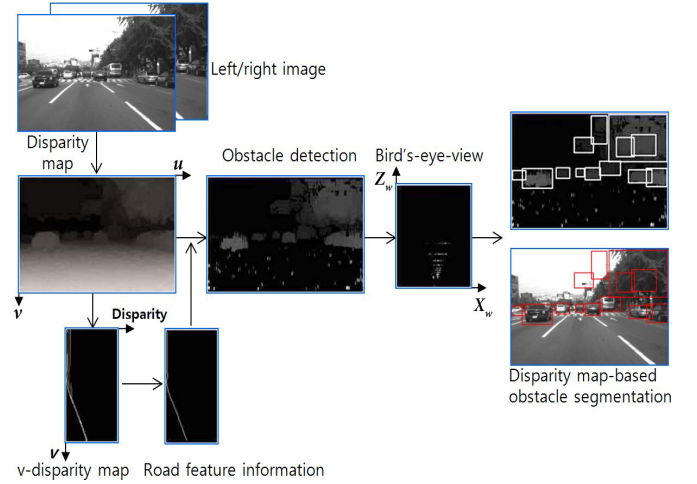


Fig. 3. Disparity map-based obstacle detection and segmentation.

The results segmented in the bird's-eye-view are reconverted into the disparity map. These disparity map-based segmentations improve detection performance in heavy traffic thanks to accurate identification of obstacle position and removal of unnecessary obstacles. However, additional segmentation is needed due to the limitation of the disparity map. The entire procedure of disparity map-based obstacle detection and segmentation is presented in Fig. 3.

III. EDGE DISTANCE WEIGHTED CRF-BASED SEGMENTATION

In this paper, edge distance weighted CRF-based segmentation using a new pairwise potential function is proposed to obtain more accurate results from the above results. Let $S = \{1, \dots, n\}$ be an index in the image lattice, where n is the number of pixels in the observed image. Let $X = \{x_i | i \in S\}$ be a label, where x_i is one of the labels in $L = \{1, \dots, c\}$ at location i , and c is the number of labels. Let $Y = \{y_i | i \in S\}$ be the observed image, where y_i is the known pixel intensity at location i . CRF-based calculation of the label distribution for segmentation [9] can be performed by

$$P(X|Y) = \frac{1}{z} \exp\{-\mathbb{E}(x, Y) - \sum_{i \in N_i} \mathbb{E}(x, x_i, Y)\}, \quad (5)$$

where z is the partition function, $\mathbb{E}_i(x_i, Y)$ and $\mathbb{E}_{ij}(x_i, x_j, Y)$ are unary and pairwise potentials, respectively, and N_i denotes the neighbours of pixel i .

The unary potential is defined as the Gaussian likelihood

$$\mathbb{E}(x, Y) = -\frac{(y - \tilde{y})}{2\uparrow}, \quad (6)$$

where \tilde{y}_{x_i} and \uparrow_{x_i} are the mean and standard deviation of intensity associated with the class indicated by label x at location i , respectively. In the case of pairwise potential, a novel potential for accurate vehicle detection is proposed, which is one of the major contributions of this paper. The edge and intensity are important cues for segmentation in small obstacle areas, and the vertical edge in particular is a good standard for separating the vehicle from the background or other vehicles in close proximity. The proposed pairwise potential is defined as:

$$\mathbb{E}(x, x, Y) = \chi w \exp\left(-\frac{(y - \tilde{y})}{2\uparrow}\right) D(i, j) u, \quad (7)$$

$$w = \frac{D + D}{2}, \quad (8)$$

where D (or D) = $\frac{d$ (or d)}, χ is the model parameter, w_{ij} is the weighting parameter, \uparrow_{N_i} is the standard deviation of intensity in the neighbours including the current pixel, $D_{inv}(i, j)$ is the inverse of the Euclidean distance between i and j , and u_{x_i, x_j} is the Kronecker delta. Here, D_i (or D_j) is the normalized distance transform value, d_i (or d_j) is the distance transform value [10] of the pixel, and $\mathbb{E}_{ij}(x_i, x_j, Y)$ represents the label relationship between the current pixel and the neighbours, and it uses the distance transform as the weighting (w_{ij}).

Generally, the probability that the current pixel has the same label as its neighbours is high in regions that do not have an edge. On the contrary, the probability is low in regions that have an edge. In other words, the farther away from the edge the current pixel is, the more important the label information of the neighbours, such as pairwise potential, is. On the other hand, the closer to the edge the current pixel is, the more important the feature data (intensity), such as unary potential, is. Thus, the distance information between the current pixel and the edge is utilized in the proposed CRF. The relative importance of pairwise and unary potential is determined by the edge distance. The distance information is obtained from the distance transformation using a vertical edge image extracted from each detected area. Each edge distance is normalized, and the mean is utilized as the final edge distance.

Note that $\mathbb{E}_{ij}(x_i, x_j, Y)$ also uses other information, such as the spatial distance and the difference of intensity between two pixels. The shorter the pixel distance and the smaller the difference of intensity are, the higher the probability is that the current pixel has the same label as its neighbours. The Euclidean distance and the Gaussian model are utilized to measure the spatial distance and the difference of intensity between two pixels, respectively.

The model parameter χ is estimated by using the pseudo-likelihood method and is defined as

$$\hat{\chi} = \arg \max_x \prod_{k=1}^M \prod_{i \in S} P(x_i^k | x_{N_i}^k, Y^k, \chi), \quad (9)$$

where $\hat{\chi}$ is the estimated parameter, M is the number of training images, and x_{N_i} is the label of the neighbour of pixel i . Other parameters, \tilde{y}_{x_i} and \uparrow_{x_i} are first calculated directly from the same labelled pixels from the initial label status. However, these parameters are updated each time during the inference. The iterated conditional modes (ICM) algorithm [11] is utilized to optimize the proposed model. Our proposed local conditional probability $P(x_i | x_{N_i}, Y)$ for ICM can be written as

$$P(x | x, Y) \propto \exp\left\{-\frac{(y - \tilde{y})}{2\uparrow} + \chi w \exp\left(-\frac{(y - \tilde{y})}{2\uparrow}\right) D(i, j) u\right\}. \quad (10)$$

Thus, the label of each pixel \hat{x}_i is estimated using

$$\hat{x}_i = \arg \max_x P(x | x, Y). \quad (11)$$

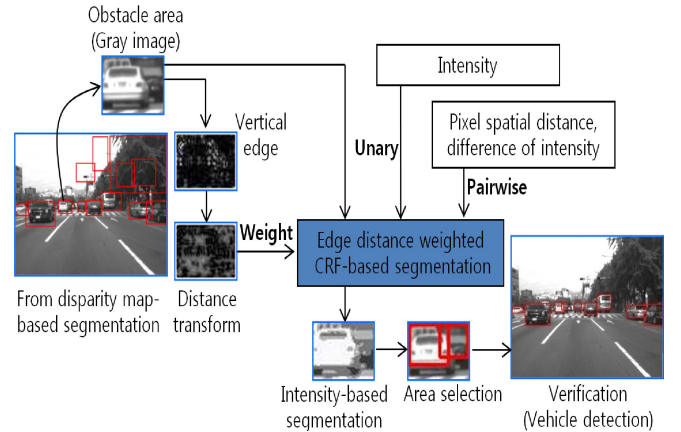


Fig. 4. Intensity-based obstacle segmentation and verification.

For quick convergence, \tilde{y}_{x_i} and \uparrow_{x_i} are updated every time, and the neighbour pixels' labels obtained in the current iteration are utilized immediately when the pairwise potential of the current pixel is calculated. In other words, the current estimated data are updated instantly to be utilized when the label of the next pixel is determined.

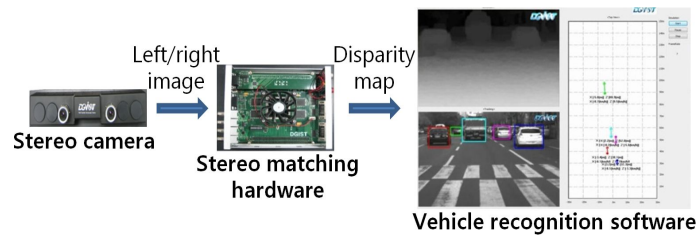
IV. AREA SELECTION AND VERIFICATION

After segmentation processing, the isolated small areas are removed by image processing, and size-based area selection is performed. The largest area is selected because a vehicle has almost unvarying intensity and has the largest size in the segmented results. However, if an area whose size is greater than half size of the largest area exists, the area is also selected. Finally, the selected area is verified with respect to whether it is a vehicle or not [7]. The entire process of intensity-based obstacle segmentation and verification is presented in Fig. 4.

V. EXPERIMENTAL RESULTS

The proposed vehicle detection algorithm was evaluated by application to heavy traffic images captured by our stereo vision system. Our system architecture and a photo of the

system installed in the test vehicle are presented in Fig. 5. The belief propagation algorithm is utilized to make the disparity map, which is implemented on FPGA hardware for real-time processing [12]. Our algorithms are implemented using vehicle recognition software with C++.

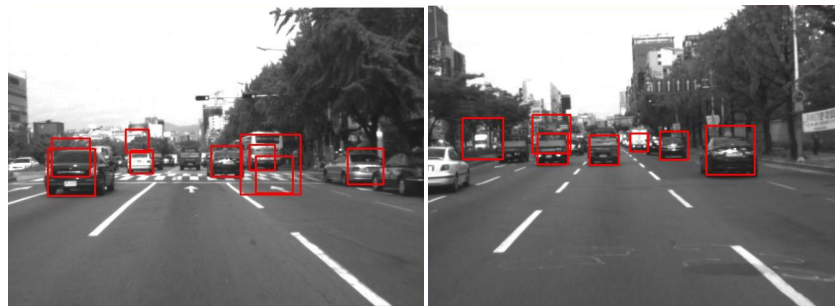


(a)

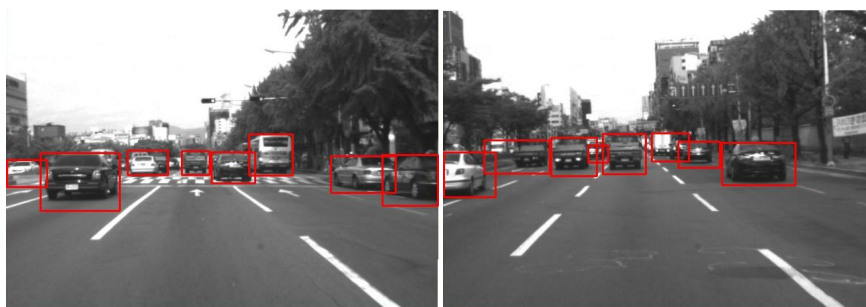


(b)

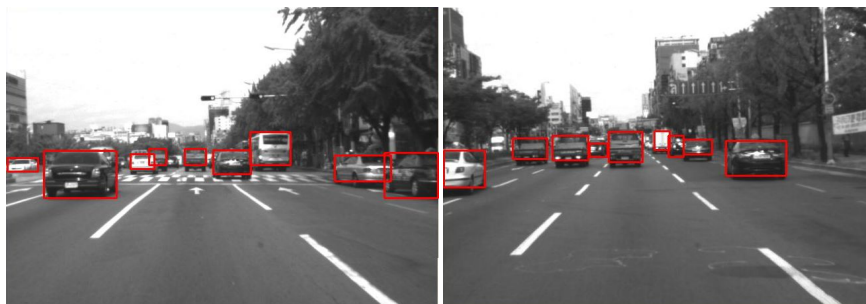
Fig. 5. Vehicle detection system for intelligent vehicle: (a) system architecture; (b) system installed in test vehicle.



(a) Adaboost classifier [5]



(b) Column detection [7]



(c) Our method

Fig. 6. Vehicle detection results in heavy traffic: (a) adaboost classifier [5]; (b) column detection [7]; (c) our method.

We set the number of labels to 4 because the average size of the segmented results was small and the vehicle had approximately the same intensity in the results. The model parameter X was estimated to be 23 by the pseudo-likelihood method. We compared our results with those obtained with the Adaboost classifier [5] and disparity map-based column detection [7] in Fig. 6. As seen in Fig. 6 and Table I, the proposed algorithm provides better detection performance than the other methods. The improvements in performance range from 10.8 % to 20.5 % increase in F-measure. Furthermore, vehicles are detected more accurately owing to segmentation, which is a good characteristic to track vehicles.

TABLE I. VEHICLE DETECTION PERFORMANCE.

	Adaboost classifier [5]	Column detection [7]	Our method
Recall	0.70	0.75	0.85
Precision	0.67	0.74	0.79
F-measure	0.68	0.74	0.82

VI. CONCLUSIONS

A new vehicle detection algorithm using disparity and intensity-based segmentation suitable for on-road vehicle detection was proposed. Though accurate vehicle detection in heavy traffic is a very difficult task in general, we effectively addressed this issue using powerful segmentation algorithms, namely, disparity map-based bird's-eye-view mapping segmentation and edge distance weighted CRF-based segmentation. We verified this approach by applying it to real heavy traffic images. The results showed that the proposed method worked very well in heavy traffic compared to other methods. Generally, it is not easy to accurately recognize vehicles in real traffic situations. However, thanks to improved performance in heavy traffic, our algorithm can be used in intelligent vehicles as a core technology of various important functions, such as forward collision warning or mitigation, adaptive cruise control, road environment recognition system, etc.

In the future we will use more sophisticated potentials

based on vehicle shape model and other features, such as colour. We believe that such potentials will yield better performance.

REFERENCES

- [1] K. D. Kusano, H. C. Gabler, "Safety benefits of forward collision warning, brake assist, and autonomous braking systems in rear-end collisions", *IEEE Trans. Intelligent Transportation Systems*, vol. 13, no. 4, pp. 1546–1555, 2012. [Online]. Available: <http://dx.doi.org/10.1109/TITS.2012.2191542>
- [2] C. Lundquist, U. Orguner, T. B. Schon, "Tracking stationary extended objects for road mapping using radar measurements", *IEEE Intelligent Vehicles Symposium*, 2009, pp. 405–410.
- [3] S. Eigo, S. Morito, S. Shigeru, H. Norio, T. Tomonobu, T. Masatoshi, "Processing vehicle detection using stereo image and non-scanning millimeter-wave radar", *IEICE Trans. Inf. Syst.*, pp. 2101–2108, 2006.
- [4] S. Zehang, B. George, M. Ronald, "On-road vehicle detection: a review", *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 28, no. 5, pp. 694–711, 2006. [Online]. Available: <http://dx.doi.org/10.1109/TPAMI.2006.104>
- [5] P. Viola, M. Jones, "Rapid object detection using a boosted cascade of simple features", *IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, USA, pp. 511–518, 2001.
- [6] R. Labayrade, D. Aubert, "Robust and fast stereovision based obstacles detection for driving safety assistance", *IEICE Trans. Inf. Syst.*, pp. 80–88, 2004.
- [7] C. H. Lee, Y. C. Lim, S. Kwon, J. H. Lee, "Stereo vision-based vehicle detection using a road feature and disparity histogram", *Opt. Eng.*, vol. 50, no. 2, 2011. [Online]. Available: <http://dx.doi.org/10.1117/1.3535590>
- [8] B. M. Collins, A. L. Kornhauser, "Stereo vision for obstacle detection in autonomous navigation", *DARPA Grand Challenge Princeton university technical paper*, 2006.
- [9] P. Kohli, J. Rihan, M. Bray, P. H. S. Torr, "Simultaneous segmentation and pose estimation of humans using dynamic graph cuts", *Int. J. Comput. Vis.*, vol. 79, no. 3, pp. 285–298, 2008. [Online]. Available: <http://dx.doi.org/10.1007/s11263-007-0120-6>
- [10] R. Fabbri, L. D. F. Costa, J. C. Torelli, O. M. Bruno, "2D Euclidean distance transform algorithms: a comparative survey", *ACM. Comput. Surv.*, vol. 40, no. 1, 2008. [Online]. Available: <http://dx.doi.org/10.1145/1322432.1322434>
- [11] S. Nayak, S. Sarkar, B. Loeding, "Automated extraction of signs from continuous sign language sentence using iterated conditional modes", *IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, USA, 2009, pp. 2583–2590.
- [12] S. Kwon, C. H. Lee, Y. C. Lim, J. H. Lee, "A sliced synchronous iteration architecture for real-time global stereo matching", *Proc. SPIE*, vol. 7543, 2010.