International Journal of Automotive Technology, Vol. 8, No. 4, pp. 493–504 (2007)

Copyright © 2007 KSAE
1229-9138/2007/035-12

STEREO VISION-BASED FORWARD OBSTACLE DETECTION

H. G. JUNG[1),2)*], Y. H. LEE[1)], B. J. KIM[1)], P. J. YOON[1)] and J. H. KIM[2)]

[1)]Mando Global R&D H.Q., 413-5 Gomae-dong, Giheung-gu, Yongin-si, Gyeonggi 449-901, Korea
[2)]School of Electrical and Electronic Engineering, Yonsei University, Seoul 120-749, Korea

(Received 12 January 2006; Revised 8 June 2007)

ABSTRACT–This paper proposes a stereo vision-based forward obstacle detection and distance measurement method. In general, stereo vision-based obstacle detection methods in automotive applications can be classified into two categories: IPM (Inverse Perspective Mapping)-based and disparity histogram-based. The existing disparity histogram-based method was developed for stop-and-go applications. The proposed method extends the scope of the disparity histogram-based method to highway applications by 1) replacing the fixed rectangular ROI (Region Of Interest) with the traveling lane-based ROI, and 2) replacing the peak detection with a constant threshold with peak detection using the threshold-line and peakness evaluation. In order to increase the true positive rate while decreasing the false positive rate, multiple candidate peaks were generated and then verified by the edge feature correlation method. By testing the proposed method with images captured on the highway, it was shown that the proposed method was able to overcome problems in previous implementations while being applied successfully to highway collision warning/avoidance conditions. In addition, comparisons with laser radar showed that vision sensors with a wider FOV (Field Of View) provided faster responses to cutting-in vehicles. Finally, we integrated the proposed method into a longitudinal collision avoidance system. Experimental results showed that activated braking by risk assessment using the state of the ego-vehicle and measuring the distance to upcoming obstacles could successfully prevent collisions.

KEY WORDS : Active safety vehicle, Obstacle detection, Stereo vision system, Disparity histogram

1. INTRODUCTION

Collision Avoidance Systems in general consist of four components: environment recognition, electronically controlled vehicles, collision avoidance controllers, and HMIs (Human Machine Interfaces) (Blum and Escandarian, 2002). Environment recognition refers to fixed environment information such as roadway geometry, weather conditions, and geometrical locations, as well as target vehicle information provided to collision avoidance controllers. Electronically controlled vehicles refer to motion sensor outputs used for the estimation of vehicle states, such as wheel speed sensors, steering angle sensors, yaw-rate sensors, and acceleration sensors. At the same time, electronically-controlled vehicles implement steering, braking, and acceleration commands transferred via the CAN (Controller Area Network) using X-by-Wire systems. Collision avoidance controllers fulfill risk assessment, path planning for collision avoidance, and command generation/transmission for path tracking. HMIs receive driver intension or operation commands and inform drivers of risk assessment results and control situations.

*Corresponding author. e-mail: hgjung@mando.com

Vehicle detection methods can be categorized as either or passive. Active sensor-based methods include millimeter wave radar-based (Park *et al.*, 2003), Lidar-based (Light Detection and Ranging) (Wang *et al.*, 2003; Hancock *et al.*, 1997) and acoustic-based (Chellappa *et al.*, 2004) methods. Optical sensors, such as normal cameras, are usually referred to as passive sensors because they acquire data in a non-intrusive way. Generally, passive sensors offer some advantages over active sensors: low cost, wide spatial coverage, high spatial resolution, fast scanning speed, and no interference that might be caused when a large number of vehicles move simultaneously in the same direction using the same type of sensor (Herbert, 2000).

Sun's recent review of vision-based vehicle detection effectively shows the state of the art in terms of the passive sensor-based method (Sun *et al.*, 2006). In this research, the vehicle detection procedure was separated into two steps: HG (Hypothesis Generation) and HV (Hypothesis Verification).

Approaches to the HG step can be classified mainly into three categories: knowledge-based, motion-based, and stereo-based approaches. The objective of the HG step is to quickly find candidate vehicle locations in an

image. Knowledge-based methods employ *a priori* knowledge to estimate vehicle locations in an image using techniques such as symmetry, color, shadow, corners, texture, vehicle light, and vertical/horizontal edges. Motion-based methods employ relative motion via the calculation of optical flow (Giachetti *et al.*, 1998). There are two types of stereo-based methods: disparity map-based and IPM (Inverse Perspective Mapping)-based methods.

The input to the HV step is the set of hypothesized locations obtained during the HG step. Approaches to the HV step can be classified as template-based or appearance-based approaches. Template-based approaches use predefined patterns and perform correlations between the given image and the template (Ito *et al.*, 1995). Appearance-based approaches are further divided into a two-class pattern classification problem: vehicles versus non-vehicles. Appearance-based approaches discover the characteristics of the appearance of any given vehicle by using a set of training images to capture the variability of the vehicle class (Sun *et al.*, 2005).

The major interest of this paper is stereo vision-based vehicle detection for longitudinal collision warning/avoidance via braking. Although dense stereo matching algorithms are progressing continuously and are able to recognize the shapes and judge the distances of objects precisely and minutely, they are not useful in automotive collision warning/avoidance because the development direction is not aimed at real-time computation and target selection (Brown *et al.*, 2003; Scharstein and Szeliski, 2002; Kastrinaki *et al.*, 2003). This may be a reason why Sun's review mentions only two kinds of stereo vision-based methods.

The IPM-based method generates two virtual images respectively for left and right images by assuming that every object is located on a flat ground plane. Then, the difference between the remapped left and right images reveals the location of obstacles, because anything located above the road makes large clusters of nonzero pixels in the difference image (Bertozzi and Broggi, 1998). Because of the flat ground plane assumption, the performance of the IPM-based method is degraded severely when the tilt angle of the camera changes on uneven roadways.

The difference between corresponding pixels in the left and right images is called disparity. The disparities of all the image points form the disparity-map. Once the disparity map is available, all the pixels within the depth of interest, according to a disparity interval, are determined and accumulated in a disparity histogram. If an obstacle is present within the depth of interest, then a peak point will occur at the corresponding histogram bin (Franke and Kutzbch, 1996). Because Franke's implementation was developed for stop-and-go applications, it used a fixed rectangular ROI and peak detection by constant threshold and then simply recognized the closest object, i.e. the largest disparity, as the control target. This approach could not be applied directly to highway collision warning/avoidance because multiple preceding vehicles exist over a wide distance range.

This paper proposes a way to measure the distance between the ego-vehicle and target vehicles. The HG step of this paper consists of ROI establishment using recognized lane information, edge feature-based stereo matching, disparity histogram generation, peak detection in the disparity histogram using the threshold-line, and candidate validation by peakness evaluation. The HV step of this paper verifies candidate locations by using the correlation of the left and right edge images with the candidate disparity, which is similar to the template matching-based HV. The main contribution of this paper is extending the disparity map-based method to highway applications by establishing the ROI using recognized lane information, and by compensating for the variation of peak height using the threshold-line. All of the adopted operations proved to be suitable for real-time operations and parallel implementations by VHDL (Very-high-speed-integrated-circuit Hardware Description Language). Experimental results showed that the proposed method was able to successfully measure the distance to the preceding vehicle and was able to overcome curved road problems and road surface noise factors. Thanks to the wider FOV (Field Of View) of the camera, the vision-based method detected cutting-in vehicles faster than a laser radar. Experiments using a test vehicle with the ESP (Electronic Stability Program) showed that the longitudinal collision avoidance system (incorporating the proposed method) was able to successfully stop vehicles before collisions.

2. STEREO VISION-BASED LONGITUDINAL COLLISION AVOIDANCE SYSTEM

This system consists of five main components: a stereo vision-based obstacle distance measurement, a dynamic model of the ego-vehicle, the collision avoidance algorithm, active braking, and the HMI (Human Machine Interface). This dynamic model of the ego-vehicle estimates required state variables for the ego-vehicle by utilizing sensors already installed on the vehicle such as wheel speed, lateral acceleration, yaw-rate, and steering angle. The collision avoidance algorithm measures the risk of collision using current vehicle states and obstacle states, and then sends the required braking commands to active braking systems via CAN in order to avoid upcoming collisions. The active braking system should be able to generate the required braking force without requiring the driver to push the pedal. In our system, the Mando MGH-series ESP was used as an active braking actuator. Figure

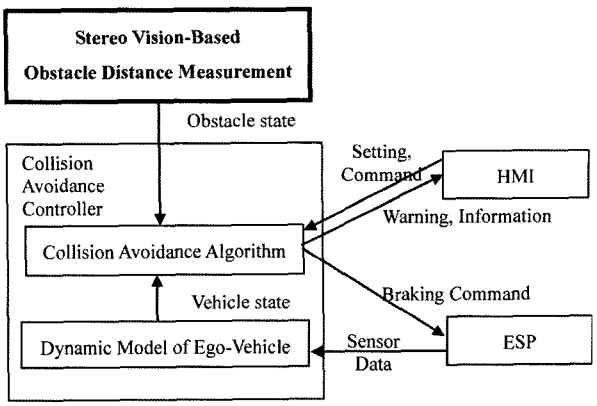


Figure 1. The architecture of the longitudinal collision avoidance system.

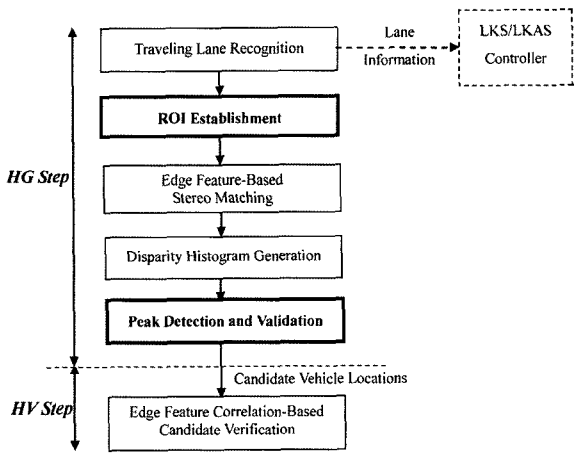


Figure 2. Vision system architecture.

1 shows the brief architecture of the system.

Stereo vision-based obstacle distance measurement consists of six phases as shown in Figure 2: traveling lane recognition, ROI establishment, edge feature-based stereo matching, disparity histogram generation, peak detection/validation in the disparity histogram, and edge feature correlation-based candidate verification. The proposed method is basically an extension or modification of Franke's disparity histogram-based obstacle distance measurement (Franke and Kutzbch, 1996). It was assumed that the preceding vehicle will make a definite peak in the disparity histogram and peak detection would measure the distance. In order to make the disparity histogram emphasize the preceding vehicle and ignore potential disturbances effectively, the fixed rectangular ROI was replaced with an adaptive ROI according to the recognized traveling lane region, i.e. a traveling lane-based ROI. Peak height variation in the disparity histogram with respect to distance was compensated for by replacing constant threshold-based peak detection with threshold-line-based peak detection and peakness evaluation-based validation. In addition, the recognized traveling lane was utilized by the LKS (Lane Keeping System)/LKAS (Lane Keeping Assist System) controller.

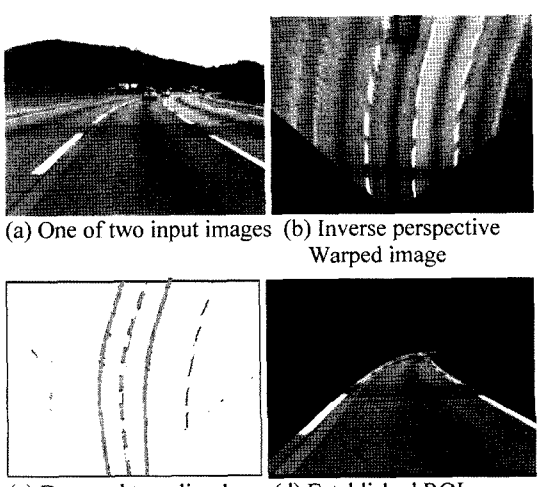
3. ENHANCED DISPARITY HISTOGRAM-BASED HYPOTHESIS GENERATION

3.1. Travelling Lane-Based ROI Establishment

The ROI for the following operations was established according to recognized traveling lanes. In general, ROI establishment is important in two aspects: to reduce computational load and to improve the robustness of systems. Furthermore, if an established ROI in general contains sufficient image portions to detect the distance to the preceding vehicle, explicit object boundary detection is unnecessary.

McCall's recent in-depth survey shows the state of the art in vision-based lane detection and tracking technology, which has developed over the last 15 to 20 years (McCall and Trivedi, 2006). Although there have been a lot of advanced lane detection technologies, we implemented a simple lane detection method similar to Bertozzi's approach (Bertozzi, 1998) to verify the feasibility of our main proposal, i.e. traveling lane-based ROI and threshold-line based peak detection in the disparity histogram.

The implemented lane detection method consists of four steps: inverse perspective warped image generation, lane marking template matching, lane marking grouping, and curve fitting. Some assumptions about the structured nature of road surfaces include: (1) the road/lane texture is consistent; (2) the road/lane width is locally constant; (3) road markings follow strict rules for appearance or placement; and (4) the road is a flat plane or follows a



(a) One of two input images (b) Inverse perspective Warped image

(c) Detected traveling lane (d) Established ROI

Figure 3. Traveling lane-based ROI establishment.

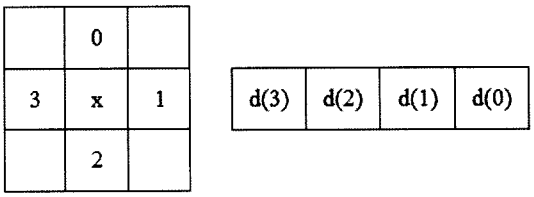
strict model for elevation change (McCall and Trivedi, 2006). It is noteworthy that the absence of flat ground in lane detection is not as serious as in obstacle distance measurement.

An inverse perspective warped image, or bird's eye view image, was obtained by using the homography between the image plane and the ground plane derived from the camera height and the tilt angle (Batavia *et al.*, 1997). Figure 3(a) is an input image and Figure 3(b) is the inverse perspective warped image corresponding to the input image. The inverse perspective warped image eliminates the effect of perspective distortion. Lane marking template-matching uses the fact that lane marking appears as a black-white-black transition with almost the same width in the horizontal direction. Lane marking grouping collects detected lane markings in a vertical direction, and lane marking groups with proper spacing, i.e. roadway width, can be recognized as left/right lane markings. Figure 3(c) shows the detected traveling lane information. Left lane marking and right lane marking were modeled as a quadratic polynomial using LS (Least Squared error)-based curve fitting. Finally, the recognized traveling lane was remapped onto the input image as shown in Figure 3(d). This traveling lane-based ROI takes the place of a fixed rectangular ROI in order to cope with highway applications.

3.2. Edge Feature-Based Stereo Matching

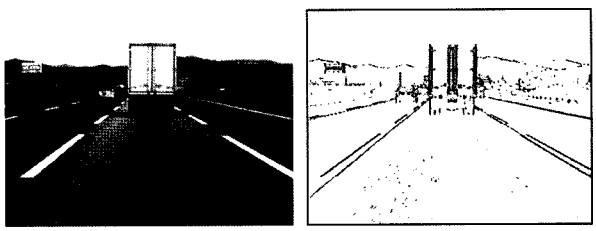
Feature-based stereo matching was employed to construct the sparse disparity map of the input images. In automotive vision, it is well known that vertical edges are sufficient to detect noticeable objects (Franke and Kutzbch, 1996). Consequently, stereo matching using only vertical edges can drastically reduce the computational load. The feature-based stereo matching process consisted of pixel classification and similarity-based matching.

In general, pixel classification investigates the intensity differences between a pixel and four directly connected neighbors so as to assign a class value to the pixel reflecting the intensity configuration. It is known that feature-based stereo matching with pixel class is fast and robust to noise (Franke and Kutzbch, 1996; Jung *et al.*, 2006). Equation (1) shows how the relationship between $g(x)$ and $g(i)$ was encoded, where $g(x)$ represents the grey value of the investigated pixel and $g(i)$ represents the gray value of the neighboring pixel with index i, which was designated as shown in Figure 4(a). Consecutively, Equation (2) and Figure 4(b) show how the class value $Class(x)$ was constructed by concatenating four $d(i)$ values. A smooth pixel was classified in the zero class and a pixel with more edges was classified in the non-zero class. The threshold T was adaptively modified to maintain the ratio of non-zero pixels within a reasonable range, e.g. 5~10%. Figure 5(a) shows the original image



(a) Neighboring pixel index (b) Class encoding

Figure 4. Neighboring pixel indexing and class encoding.



(a) Original image (b) Classification results

Figure 5. Pixel classification results.

and Figure 5(b) shows the pixel classification results. Only 6.4% of the total pixels were classified in the non-zero class.

$$d(i)=\begin{cases} 1 & if \quad g(i)-g(x)>+T \\ 2 & if \quad g(i)-g(x)<-T \\ 0 & else \end{cases} \tag{1}$$

$$Class(x)=\sum_{i=0}^{3}(d(i)<<(2\cdot i))=\sum_{i=0}^{3}4^{i}\cdot d(i) \tag{2}$$

Stereo matching was performed only on pixels with vertical edges. Furthermore, stereo matching was composed of step-by-step test sequences: class comparison, class similarity, color similarity, and maximum similarity detection (Jung *et al.*, 2006). Only correspondence candidates passing the previous test steps were investigated in the next test step. Assuming that the vertical alignment of the stereo rig was calibrated, the search range of a pixel was limited to a horizontal line with a fixed displacement. First, a correspondence test was performed on the pixels in the same class as the investigated pixel. The class similarity of the investigated pixel (x, y) with respect to the disparity d defined by Equation (3) showed how the candidate pixel was similar to the investigated pixel in the sense of a 3×3 class window. The color similarity of the investigated pixel (x,y) with respect to disparity d defined by Equation (4) showed how the candidate pixel was similar to the investigated pixel in the sense of the 5×5 color window. The total similarity defined by Equation (5) was the product of the class similarity and the color similarity. If the highest total

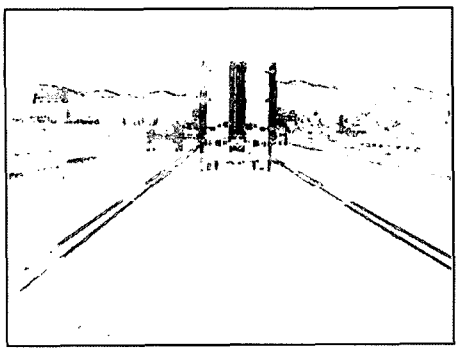


Figure 6. Stereo matching result (disparity map).

similarity was lower than a certain threshold, the investigated pixel failed to find the corresponding point and was ignored. A pixel with disparity d receiving the highest total similarity was determined as the corresponding point of the investigated pixel.

$$ClassSimilarity(x, y, d) = \frac{1}{3\times3}\sum_{u=-1}^{1}\sum_{v=-1}^{1} f(Class_{left}(x+u,y+v), Class_{left}(x+u+d,y+v))$$

$$\text{where, } f(Class_{left}, Class_{right}) = \begin{cases} 0, & Class_{left} \neq Class_{right} \\ 1, & Class_{left} = Class_{right} \end{cases}$$

$$ColorSimilarity(x, y, d) = 1 - \frac{1}{256}\sqrt{\frac{ColorSSD(x,y,d)}{5\times5}} \quad (4)$$

where, $ColorSSD(x,y,d) =$

$$\sum_{u=-2}^{2}\sum_{v=-2}^{2}\left\{\frac{1}{3}\times\begin{pmatrix}(R_{left}(x+u,y+v)-R_{right}(x+u+d,y+v))^2+\\(G_{left}(x+u,y+v)-G_{right}(x+u+d,y+v))^2+\\(B_{left}(x+u,y+v)-B_{right}(x+u+d,y+v))^2\end{pmatrix}\right\}$$

$$Similarity(x, y, d) = ClassSimilarity(x, y, d)\times ColorSimilarity(x, y, d) \quad (5)$$

3.3. Threshold-Line-Based Peak Detection and Peakness Evaluation-Based Validation

In general, the candidate locations of the preceding vehicle are determined by peak detection in a disparity histogram. Candidate generation consists of three steps: disparity histogram construction, threshold-line-based peak detection and peakness evaluation-based validation.

The preceding vehicle is supposed to form a peak in the disparity histogram (Franke and Kutzbch, 1996). The disparity histogram measures how many pixels have a certain disparity value and are implemented as an accumulator array. While investigating the disparity value of all feature pixels within the established ROI, a histogram bin corresponding to each disparity value increases. Figure 7 shows an example of a disparity histogram.

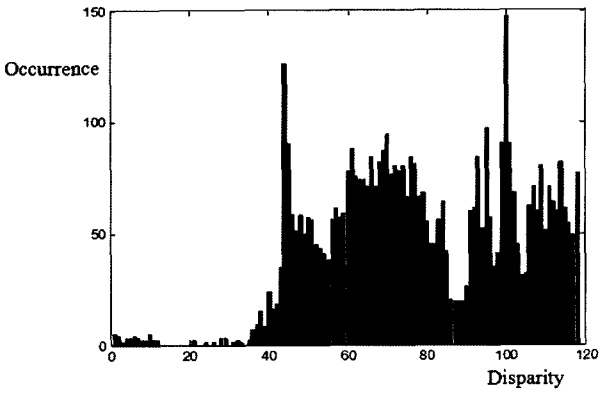


Figure 7. Disparity histogram.

Although the preceding vehicle certainly produces a peak in the disparity histogram, the height of the peak varies with respect to its disparity value, as shown in Figure 8. This can be naturally derived from the fact that distant objects appear small and near objects appear large. Because near objects appear big, the probability of the occurrence of vertical edge pixels is high. A near distance means large disparity. Therefore, near objects are expected to generate high peaks and large disparity values. Inversely, because distant objects appear small, the probability of the occurrence of vertical edge pixels is low. Distant objects are expected to generate low peaks and small disparity values.

Peak detection when using a constant threshold cannot reflect the relationship between disparity and peak height. If the constant threshold is too low, the false detection rate at large disparity values increases. If threshold is set too high to avoid false detection, the preceding vehicle might be missed. The threshold-line with respect to disparity is expected to overcome the drawback of the con-

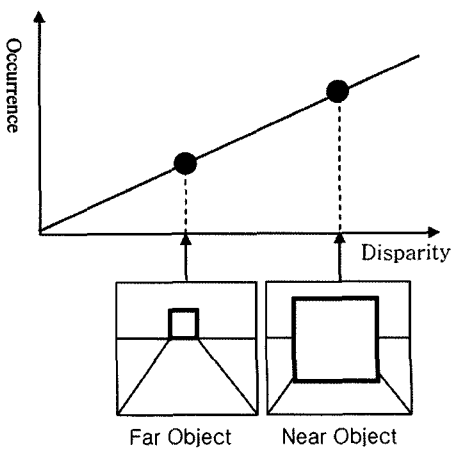


Figure 8. Threshold line compensating for perspective distortion.

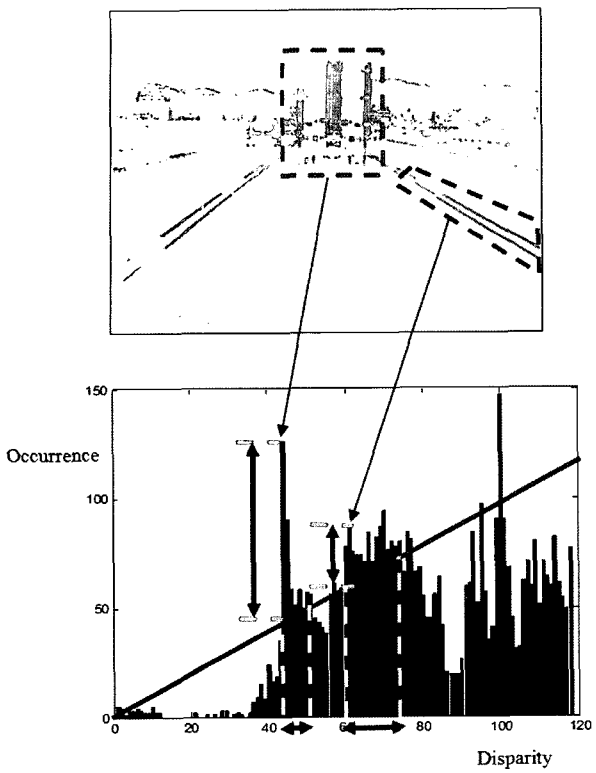


Figure 9. Different disparity distribution of preceding vehicle and lane marking.

stant threshold. The threshold-line is a line passing through the origin with its slope calibrated empirically. The constant slope of the threshold line is justified by feature ratio maintenance in the feature detection phase. Peakness evaluation and the HV step will be explained below.

In general, objects attached to the ground contribute to a wide range of disparities because they can exist in a wide range of distances from the camera. Contrarily, objects located above the ground, including the preceding vehicle, tend to contribute to a narrow range of disparities because they exist within a narrow range of distances from the camera, as shown in Figure 9. In other words, the camera can observe the backs of the preceding vehicles since they are parallel to the camera image plane. Vertical edges caused by the backs of preceding vehicles are supposed to exist almost at the same distance.

If peak $\mathcal{P}$ is defined as a group of continuous disparities whose histogram values are above the threshold-line, the peak width is the difference between the maximum and minimum disparity values in the group and the peak height is the maximum difference between the histogram value and the threshold value in the group. Peakness evaluation encodes *a priori* knowledge such as: 1) a valid peak should be sharp and the ratio of the peak height to the peak width should be high, and 2) the higher the peak height, the more likely the peak is to be of correct disparity; the ratio of the peak height to the threshold should be high. Peakness evaluation with respect to disparity d is defined in Equation (6). The first term reflects the first kind of knowledge and the second term reflects the second kind of knowledge, where $h(d)$ denotes the disparity histogram value of disparity d and θ denotes the slope of the threshold-line. After evaluation of the candidate peaks, any peak below the predefined threshold is rejected.

$$Peakness(\mathcal{P}) = \frac{\max_{d\in\mathcal{P}}\left(h(d)-\theta\cdot d\right)}{\max_{d\in\mathcal{P}} d - \min_{d\in\mathcal{P}} d} \times \frac{\max_{d\in\mathcal{P}}\left(h(d)-\theta\cdot d\right)}{\theta\cdot \operatorname{median}_{d\in\mathcal{P}} d} \tag{6}$$

4. EDGE FEATURE CORRELATION-BASED HYPOTHESIS VERIFICATION

At times, peak detection can find several peaks above the threshold-line. These peaks represent candidates for the preceding vehicle. Among these candidates, the best one is selected by edge feature correlation-based verification.

Because the correct disparity value refers to the displacement between the left and right pixels belonging to the preceding vehicle, the right image shifted by the correct disparity value will be exactly overlapped on the left image in the region of the preceding vehicle, as shown in Figure 10. Therefore, checking the overlapping quality of each candidate disparity can determine the correct disparity value. The overlapping quality can be measured within the potential vehicle boundary, a small rectangle with the same width and half the height of a general vehicle, and stands at the distance calculated by candidate disparity and perspective geometry, as shown

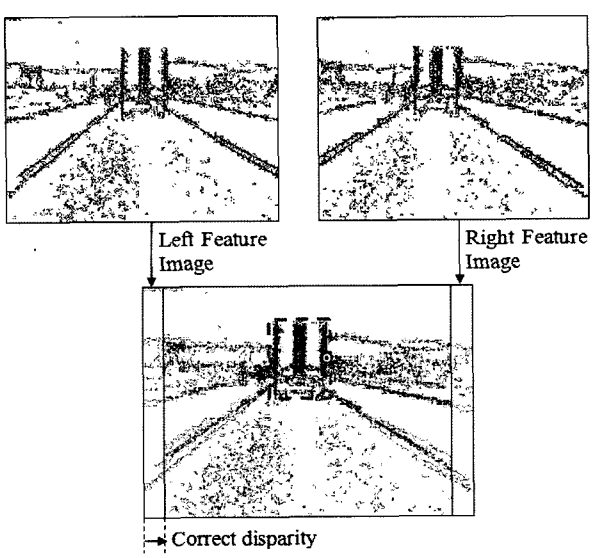
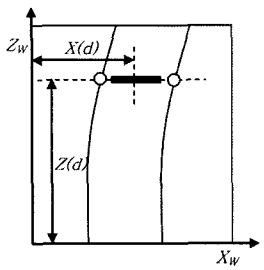
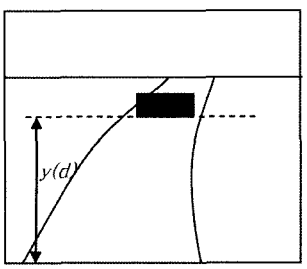


Figure 10. Preceding vehicle will be overlapped with the correct disparity value.



(a) Inverse perspective warped view



(b) Perspective view

Figure 11. Potential vehicle boundary corresponding to disparity d.

in Figure 11. The distance to the bottom line of the potential vehicle boundary, $Z(d)$, was calculated by the relationship between the disparity and the distance, as shown in Equation (7), where f denotes the focal length of the camera and B denotes the baseline of the stereo camera. The center of the bottom line, $X(d)$, was calculated as the center of two cross-points between the ROI boundary and the horizontal line at $Z(d)$. The upper line of the potential vehicle boundary was calculated by adding half the vehicle height in the Y_W coordinates. Once the coordinates of the four corners were determined in the world coordinate system, (X_W, Y_W, Z_W), the corresponding coordinates in the image coordinate system, (x,y), were calculated by perspective geometry. Alternatively, the bottom line of the potential vehicle boundaries in image coordinate systems, $y(d)$, was directly calculated by Equation (8) (Franke, 1996), where H represents the height of the camera position and B represents the baseline of stereo vision. fx and fy represent the focal length of the x-axis and the y-axis respectively, and α denotes the tilt angle of the camera.

$$Z(d)=\frac{f \cdot B}{d} \tag{7}$$

$$y(d)=\frac{f_y}{\cos\alpha}\left(\frac{1}{f_x}\frac{H}{B} \cdot d-\sin\alpha\right) \tag{8}$$

The edge similarity of disparity d is defined as in Equation (9) and measures the number of edge pixels that are overlapped if shifted by the disparity. Edge similarity is a normalized way of having a value between 0 and 1, where $\mathcal{E}$ is the set of non-zero class pixels in the left image within the potential vehicle boundary corresponding to disparity d, and N is the size of $\mathcal{E}$. The edge color similarity of disparity d is defined in Equation (10) and measures how similar corresponding edge pixels are in the sense of color code values (R, G, B refer to red, green, and blue values). Edge color similarity is also a normalized way of having a value between 0 and 1. The overlapping quality of disparity d is the product of edge similarity and edge color similarity. Consequently, candidate disparity with the highest overlapping quality is recognized as the disparity of the preceding vehicle. Because disparity histograms might represent a kind of dimension reduction transformation, impatient determination of the best peak would lead to a higher false detection rate. Acceptance of multiple candidates and re-examination of the detection principle not only increases the true positive rate but also decreases the false positive rate.

$$EdgeSimilarity(d) = \frac{1}{N}\sum_{(x,y)\in\mathcal{E}} f(Class_{left}(x,y), Class_{right}(x+d,y)) \tag{9}$$

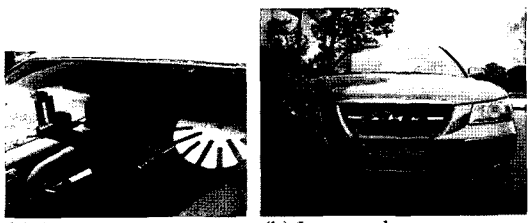
$$\text{where, } f(Class_{left}, Class_{right}) = \begin{cases} 0, & Class_{left} \neq Class_{right} \\ 1, & Class_{left} = Class_{right} \end{cases}$$

$$EdgeColorSimilarity(d) = \frac{1}{N}\sum_{(x,y)\in\mathcal{E}}\left(1-\frac{1}{256}\sqrt{ColorSquaredError(x,y,d)}\right) \tag{10}$$

$$\text{where, } ColorSquaredError(x,y,d)=\frac{1}{3}\times\begin{Bmatrix}(R_{left}(x,y)-R_{right}(x+d,y))^2+\\(G_{left}(x,y)-G_{right}(x+d,y))^2+\\(B_{left}(x,y)-B_{right}(x+d,y))^2\end{Bmatrix}$$

5. EXPERIMENTAL RESULTS

The proposed system was validated by in-vehicle tests. The traveling lane-based ROI establishment was validated by comparison with the fixed ROI. Threshold-line based peak detection and peakness evaluation-based validation were validated by comparison with the constant threshold. The distance measured by the proposed method was compared with the distance measured by the laser radar. Within the distance range of longitudinal collision warning/avoidance, the proposed method showed reasonable accuracy and shorter response time in terms of cutting-in vehicle detection. Finally, we implemented the proposed method with a longitudinal collision avoidance system. By experiments with a balloon vehicle on a test track, we confirmed that the proposed method



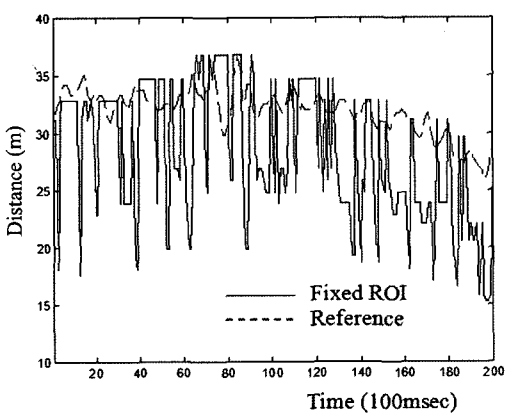
(a) Stereo camera (b) Laser radar

Figure 12. Forward-looking sensors on a test vehicle.

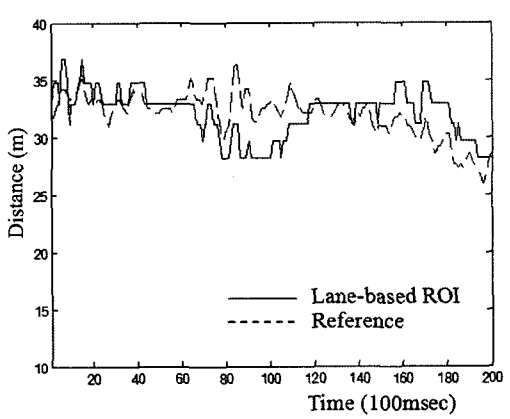
successfully stopped the test vehicle before a collision.

5.1. Experimental Settings

The stereo camera used in the experiments was made with two off-the-shelf CMOS cameras and its baseline set to 30 cm in order to be able to detect distant objects. Figure 12(a) shows the stereo camera module that was installed on the windshield of the test vehicle. A Caltech calibration toolbox was used for stereo camera calibration and rectification (Bouguet, 2006). Images used for tests and evaluations were acquired on the highway.

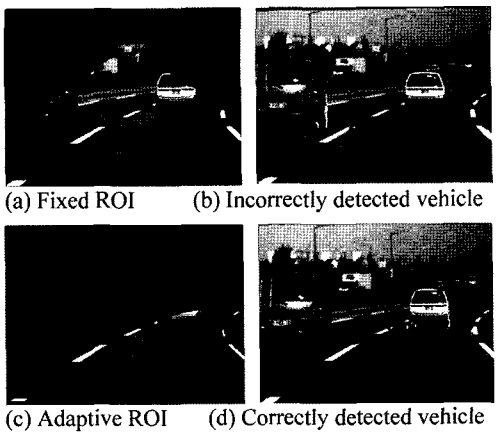


(a) Distances calculated by the fixed ROI



(b) Distances calculated by the proposed ROI

Figure 13. Fixed ROI vs. traveling lane-based ROI.



(a) Fixed ROI (b) Incorrectly detected vehicle

(c) Adaptive ROI (d) Correctly detected vehicle

Figure 14. Comparison of ROI methods in a curved road situation.

During the experiments, one laser radar was installed on a test vehicle as shown in Figure 12(b) in order to record the reference distance.

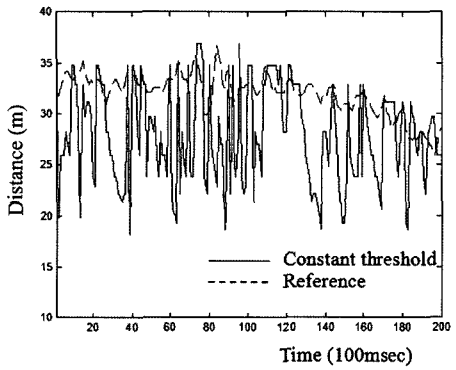
5.2. Effect of Traveling Lane-Based ROI

Distances of the image sequence were calculated by two different methods: the fixed ROI and the traveling lane-based ROI. Figure 13(a) shows the distances calculated by the fixed ROI-based method. The distances contained many large noise factors. Figure 13(b) shows the distances calculated by the traveling lane-based ROI method and Table 1 shows the comparison of the two cases. The traveling lane-based ROI method produced small average errors compared to the fixed ROI method.

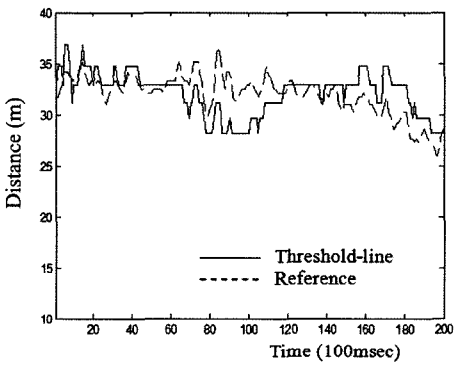
By investigating the situation on a curved road, it was shown how the proposed traveling lane-based ROI improved the performance of preceding vehicle detection. Figure 14(a) shows an example of the fixed ROI and Figure 14(b) shows pixels with the same disparity values as the output of preceding vehicle detection. Because the road was curved, a vehicle on the adjacent lane was detected. The correct preceding vehicle, which should have been used for longitudinal control, was rejected because of its comparatively smaller peak height. Figure 14(c) shows the ROI established according to the recognized traveling lane. It was shown that the established ROI successfully captured the image of the preceding vehicle while ignoring adjacent vehicles. Figure 14(d) shows the correctly-detected preceding vehicle.

Table 1. Error mean and variance of two ROI methods.

	Fixed ROI	Lane-based ROI
Error mean	5.1738	1.8509
Error variance	15.8154	1.8453

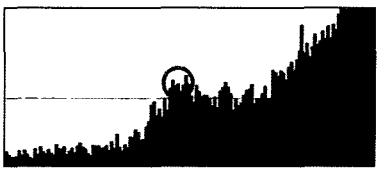


(a) Distances calculated by the constant threshold

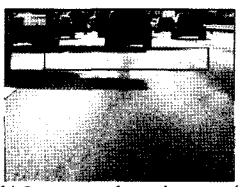


(b) Distances calculated by the .threshold-line and peakness evaluation

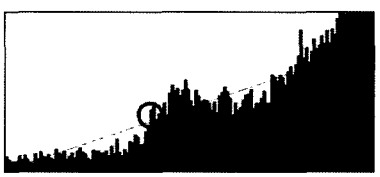
Figure 15. Constant threshold vs. the threshold-line and peakness evaluaiton.



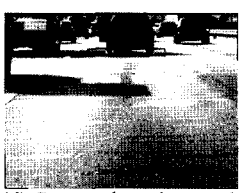
(a) Constant threshold and detected peaks



(b) Incorrect detection result



(c) Threshold-line and detected peaks



(d) Corrent detection result

Figure 16. Comparison of the two peak detection methods.

5.3. Effect of Peak Detection by the Threshold-Line and Peakness Evaluation

Preceding vehicle detection was tested with two different peak detection methods: the constant threshold-based, and threshold-line and peakness evaluation-based methods. Figure 15(a) shows the distances calculated by the constant threshold. The distances contained many large noise factors. Figure 15(b) shows the distances calculated by the threshold-line and peakness evaluation. Table 2 shows that the proposed method produced smaller average errors than the constant threshold-based method. During the experiments, the constant threshold was small so the candidate peaks never failed to include correct ones. Pixels belonging to the background portion, which generally form small peaks at small disparity values, were detected as preceding vehicles. Errors in the constant threshold-based method were caused by background pixels and traffic markings.

By investigating critical situations, it was shown that

Table 2. Error mean and variance of two threshold methods.

	Constant threshold	Threshold-line and Peakness Evaluation
Error mean	4.1874	1.8509
Error variance	9.4403	1.8453

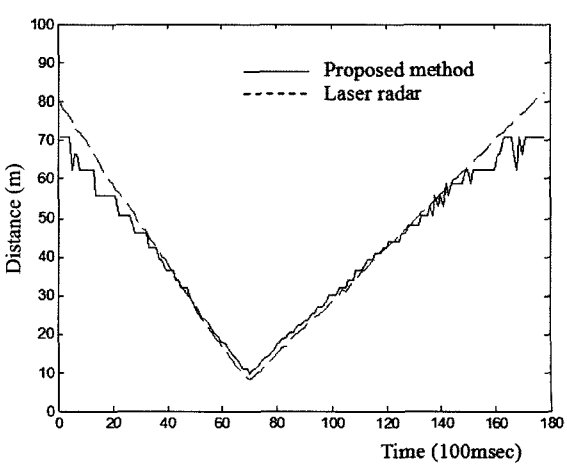
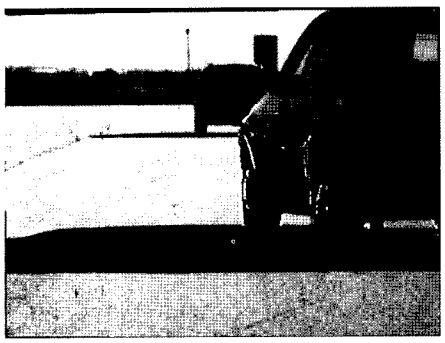
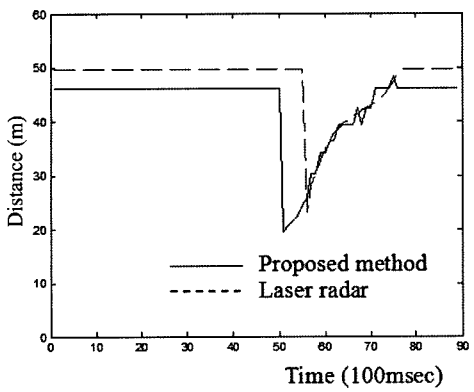


Figure 17. Measured distances with a laser radar and the proposed method.



(a) Detected cutting-in vehicle



(b) Faster response of vison-based system

Figure 18. Effect of the vision system's wider FOV.

the proposed threshold-line and peakness evaluation improved peak detection performance. Figure 16(a) is an example of the constant threshold and the detected peaks. The peak in the circle denotes a peak recognized as a preceding vehicle. In Figure 16(b), the traffic sign on the ground surface corresponds to a detected peak. Figure 16(c) is an example of the threshold-line and the detected peaks. The circled peak also denotes a peak recognized as a preceding vehicle. In this case, thanks to threshold-line and peakness evaluation, the correct peak, (not detected by the constant threshold-based method) was successfully detected. Figure 16(d) shows that the detected peak did in fact correspond to a preceding vehicle.

5.4. Comparison with a Laser Radar

The proposed method was verified by comparing its output with the distance measured by a laser radar. It was confirmed that the larger FOV of the vision system improved response times with respect to cutting-in vehicles when compared with laser radars.

In an open space, a test vehicle approached a vehicle standing still and then returned to the initial position. For the sake of safety, we performed the experiment at a test track and did not use the traveling lane-based ROI method. Figure 17 shows the two distance sequences measured by the proposed system and a laser radar. The proposed system measured the distance correctly within the range of 10-60 m. Although the measured distance was not perfect, it proved sufficient for collision warning and avoidance. Furthermore, while comparing the proposed method with laser radars, the vision-based system was shown to simultaneously detect traveling lanes, which is necessary for lane-keeping and target resolutions on curved roads.

Another major difference between stereo vision and laser radars is the FOV. Laser radars use a narrow FOV to cover far distances due to the TOF (Time Of Flight) principle. By contrast, the vision system uses a wider FOV because it measures the bearing angle of objects. Furthermore, the traveling-lane based ROI eliminates the disturbance of vehicles on adjacent lanes. Figure 18 shows the distance measurements when a vehicle was about to cut-in. The proposed vision system detected cutting-in vehicles faster than the laser radar. This rapid response time is crucial for the successful management of cutting-in vehicles.

5.5. Vehicle Test of Longitudinal Collision Avoidance System

Using the proposed stereo vision-based obstacle distance measurement, we implemented a longitudinal collision avoidance system by braking. The collision avoidance controller continuously calculated the required braking distance by considering the ego-vehicle states and the obstacle distance. Once the required braking distance was smaller than the measured distance, active braking was activated to prevent predicted collisions.

Active braking was implemented by using the Mando MGH-25 ESP system. The collision avoidance controller sent a deceleration command to the braking controller, and then the braking controller implemented deceleration

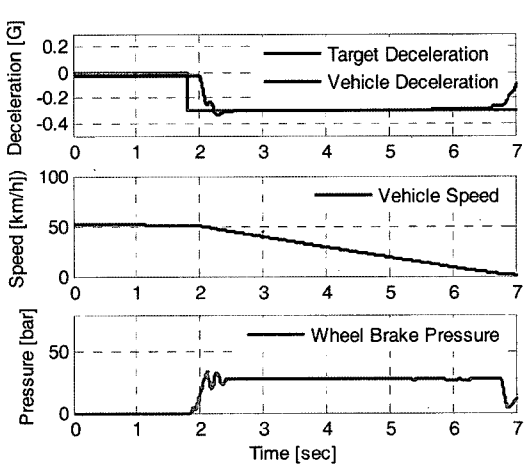


Figure 19. Deceleration control results by ESP pressure control.

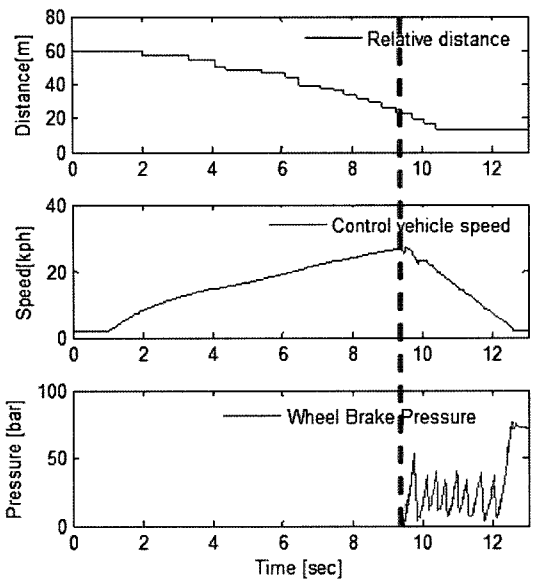


Figure 20. Longitudinal collision avoidance by the proposed distance measurement and active braking.

by controlling the ESP hydraulic actuator. The first graph of Figure 19 shows the target deceleration and the measured deceleration. The third graph of Figure 19 shows the controlled wheel pressure and the second graph shows the speed of the ego-vehicle.

The required braking distance $D_{required}$ was calculated using Equation (11). The first term represents the distance before the actual braking and the second term represents the distance during deceleration with constant deceleration control. T_{delay} denotes the time delay from recognition to actuation and $V_{relative}$ and $a_{relative}$ denote the relative velocity and deceleration between the ego-vehicle and the obstacle, respectively. Once the required braking distance was much smaller than the measured between-distance, deceleration by active braking was activated. Figure 20 shows the result of the longitudinal collision avoidance system with a balloon target vehicle on snow in winter. The vertical dotted line designates a warning alert and active braking activation. The ego-vehicle was stopped before collision with the obstacle.

$$D_{required}=\left(V_{relative}\cdot T_{delay}+\frac{1}{2}a_{relative}T_{delay}^2\right)+\left(\frac{V_{relative}^2}{2\cdot a_{control}}\right) \quad (11)$$

6. CONCLUSION

In this paper, we have proposed a stereo vision-based obstacle detection and distance measurement method. By introducing traveling lane-based ROI establishment, peak detection by threshold-line, and edge feature correlation-based verification, we were able to overcome the problems of the existing disparity histogram-based method and extend its applications to highway collision warning/avoidance. Inevitable errors caused by simple stereo matching were ignored because the disparity histogram represented a kind of statistical method. By accepting all probable candidates and then verifying the detection principle strictly again, the false detection rate decreased and the true detection rate increased. However, this benefit required high costs, i.e. additional computation load.

Although we were able to verify the improvement of disparity histogram-based obstacle detection by implementing a basic lane detection method, practical applications require an advanced lane detection method. Particularly, a stereo vision-based lane detection method may improve the stability of the detection performance irrespective of the tilt angle change. Furthermore, commercialization requires the VHDL implementation of a core algorithm such as stereo matching, edge-feature correlation, and disparity histogram generation. Possible future work may include obstacle classification and recognition of obstacles, e.g. vehicles, pedestrians, or standing poles.

REFERENCES

Batavia, P. H., Pomerleau, D. A. and Thorpe, C. E. (1997). Overtaking vehicle detection using implicit optical flow. *IEEE Conf. Intelligent Transportation System*, 729–734.

Bertozzi, M. and Broggi, A. (1998). GOLD: A parallel real-time stereo vision system for generic obstacle and lane detection. *IEEE Trans. Image Processing* **7, 1**, 62–81.

Blum, J. and Eskandarian, A. (2002). Software architecture for adaptive collision avoidance systems. *Int. J. Automotive Technology* **3, 2**, 79–88.

Bouguet, J. Y. (2006). *Camera Calibration Tool Box for Matlab*. http://www.vision.caltech.edu/bouguetj/calib_doc/.

Brown, M. Z., Burschka, D. and Hager, G. D. (2003). Advances in computational stereo. *IEEE Trans. Pattern Analysis and Machine Intelligence* **25, 8**, 993–1008.

Chellappa, R., Qian, G. and Zheng, Q. (2004). Vehicle detection and tracking using acoustic and video sensors. *IEEE Int. Conf. Acoustics, Speech, and Signal Processing*, 793–796.

Franke, U. and Kutzbch, I. (1996). Fast stereo based object detection for stop & go traffic. *IEEE Intelligent Vehicle Symp.*, 339–344.

Giachetti, A., Campani, M. and Torre, V. (1998). The use of optical flow for road navigation. *IEEE Trans. Robotics and Automation* **14, 1**, 34–48.

Hancock, J., Hoffman, E., Sullivan, R., Ingimarson, D., Langer, D. and Hebert, M. (1997). High-performance laser ranger scanner. *SPIE Int'l Conf. Intelligent Trans-*

portation Systems.

Herbert, M. (2000). Active and passive range sensing for robotics. *Proc. IEEE Int. Conf. Robotics and Automation*, **1**, 102–110.

Ito, T., Yamada, K. and Nishioka, K. (1995). Understanding driving situations using a network model. *Intelligent Vehicles '95 Symp.*, 48–53.

Jung, H. G., Kim, D. S., Yoon, P. J. and Kim, J. H. (2006). 3D Vision system for the recognition of free parking site location. *Int. J. Automotive Technology* **7, 3**, 361–367.

Kastrinaki, V., Zervakis, M. and Kalaitzakis, K. (2003). A survey of video processing techniques for traffic applications. *Image and Vision Computing*, **21**, 359–381.

McCall, J. C. and Trivedi, M. M. (2006). Video-based lane estimation and tracking for driver assistance: Survey, system and evaluation. *IEEE Trans. Intelligent Transportation Systems* **7, 1**, 20–37.

Park, S., Kim, T., Kang, S. and Heon, K. (2003). A novel signal processing technique for vehicle detection radar. *2003 IEEE MTT-S Int. Microwave Symp. Digest*, 607–610.

Scharstein, D. and Szeliski, R. (2002). A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. *Int. J. Computer Vision*, **47(1/2/3)**, 7–42.

Sun, Z., Bebis, G. and Miller, R. (2005). On-road vehicle detection using evolutionary gabor filter optimization. *IEEE Trans. Intelligent Transportation Systems* **6, 2**, 125–137.

Sun, Z., Bebis, G. and Miller, R. (2006). On-road vehicle detection: A Review. *IEEE Trans. Pattern Recognition and Machine Intelligence* **28, 5**, 1–18.

Wang, C., Thorpe, C. and Suppe, A. (2003). Ladar-based detection and tracking of moving objects from a ground vehicle at high speeds. *Proc. IEEE Intelligent Vehicle Symp.*, 416–421.