A Vision Sensor System for Intelligent Vehicles

Bing-Fei Wu and Meng-Liang Chung*

Institute of Electrical Control Engineering, National Chiao Tung University

ABSTRACT

This paper proposes a vision sensor system to facilitate safer driving. The proposed system utilizes two low-cost compact Complementary Metal-Oxide Semiconductor (CMOS) cameras as vision sensors and an effective real-time detection algorithm to detect front vehicles and measure distances. Further, a custom-made binocular is used to capture pairs of images via a low-cost grabber card. This low-cost vision sensor is an asynchronous vision system because the left and right images have a slight time difference. However, the proposed detection algorithm, which comprises four modules—namely, image preprocessing, vehicle detection, detected vehicle tracking, and distance measurement—is unrestrained by the limitations of the epipolar constraints for synchronous vision systems, and also enables the system to overcome the asynchronous detection problem affecting conventional low-cost vision systems. The results of long-term performance tests conducted on highways and urban and country roads confirm that the proposed system can successfully detect the distance of the front vehicle with detection rates of up to 99%, and detect the distances of front vehicles under various weather and illumination conditions with detection rates of up to 93%.

Keywords: vision sensor, intelligent vehicles, stereo vision system, disparity.

應用於智慧型車輛的視覺感測系統

吳炳飛 鐘孟良*

國立交通大學電控工程研究所

摘　要

本論文提出一種視覺感測系統，以增進安全的駕駛。所提出的系統利用兩個低成本的小型互補式金屬氧化物半導體（CMOS）相機作為視覺感測器，運用有效的即時偵測演算法來偵測前方車輛和測量距離。此外，客製化的雙鏡頭是經由擷取卡來依序擷取成對的影像，因為左和右影像具有輕微的時間差，所以這種低成本視覺感測器屬於非同步的視覺系統。本論文提出的偵測演算法包含四個模組，分別為影像預處理、車輛偵測、追踪已偵測的車輛及距離測量。利用所提出的偵測演算法解決低價的非同步拍攝影像視覺偵測的問題，使得立體視覺的偵測可以不受到同步拍攝系統的限制。針對高速公路、城市和鄉村道路進行長期性能測試的結果，證實該系統成功地偵測前方車輛距離高達 99%的偵測率，在各種天氣狀況下和照明條件下，偵測前方車輛距離的偵測率高達 93%。

關鍵詞：視覺感測，智能車輛，立體視覺系統，視差

文稿收件日期 104.10.16;文稿修正後接受日期 105.01.08; *通訊作者
Manuscript received October 16, 2015; revised Januery 08, 2016; * Corresponding author

I. INTRODUCTION

Safe driving is a worldwide objective that is still proving difficult to attain. In their efforts toward this goal, car manufactures have actively developed various types of onboard driving assistance systems to ensure the safety of drivers, passengers, and pedestrians. These systems are generally divided into passive and active safety systems. Passive safety systems, which include devices such as seatbelts, air bags, and anti-lock braking systems, are widely used in vehicular technology. Combining these systems with active (pre-crash) safety sensors can enhance passenger safety by preventing collisions. Forward collision warning systems (FCWS) are onboard electronic systems that monitor the roadway in front of a vehicle and warn the driver when a potential collision risk exists. Commercially available FCWS use a combination of different sensors such as charge coupled device (CCD) cameras and millimeter-wave radar to detect the distance, azimuth, and relative speed between the host vehicle and the vehicle or obstacle ahead.

The devices currently available for detecting obstacles include millimeter-wave radar, laser radar, and vision sensors. Millimeter-wave radar [1] is essential for the development of radar detection systems to detect and track the front vehicle. Although these systems have been used for adaptive cruise control of vehicles by continuously tracking the front vehicle [2], they require knowledge of the trajectory of the road in front of the vehicle to identify whether the targets ahead are in the same lane. Furthermore, the cost of millimeter-wave radar systems is higher than that of other sensors. Laser radar has also been used as a sensor for distance measurement and target position designation [3]. However, although laser radar is less expensive than millimeter-wave radar, it has the same limitations as other radar systems; specifically, it is unsuitable for classifying obstacles and recognizing lane markings.

Vision sensors can provide high spatial resolution that facilitates extraction of the edges, features, and intensities of objects. Vision-based systems can not only help to detect and track roads, but also to avoid hitting obstacles and pedestrians [4]. They have also been proved suitable for recognizing and tracking static/moving obstacles. Furthermore, the image provided by a vision sensor is also suitable for detecting lane markings. Because vision sensors can cover a large sensing area and can also detect and recognize obstacles and lane markings, vision-based detection techniques have become a very active research area in the field of intelligent vehicles. Consequently, some obstacle detection systems have incorporated vision sensors combined with an active sensor, such as millimeter-wave radar and laser scanners, to increase their detection performance.

Existing vision sensor vehicle detection systems can be categorized as either monocular or stereo vision. Broggi et al. [5] utilized monocular vision to calculate the symmetry map of an image by first searching for the position of the symmetry axis, and then marking the bounding box according to the aspect ratio of the vehicle using the grey-level symmetry of the vehicle and the contrast between the vehicle and the road. Several researchers [6]–[9] have also proposed methods for recognizing the leading vehicles and pedestrians using monocular vision systems. However, vibrations due to the road surface and/or the vehicle suspension system can cause serious distortion in the distance measurements because these approaches rely on a monocular system mounted inside the passenger compartment. Furthermore, distance measurement is very sensitive to depression angles and the altitude parameters of a monocular system. A monocular system is simply not sufficiently accurate for estimating the distance to leading vehicles in real-world applications because variable vehicle loads and passengers will cause the required camera setup parameters to change.

Because stereo vision systems have been used for adaptive cruise control of vehicles through distance measurement, they can be developed as effective measures for enhancing the safe and smooth operation of an automobile in traffic. The underlying operational concept of a stereo vision system is that it fuses images recorded by two or more cameras and exploits the differences between them to provide an indication of the distance [10].

Stereo vision systems utilize two processes: fusion of the features observed by two (or more) cameras and reconstruction of their three-dimensional (3D) information. The well-known epipolar constraint [11] is typically used to describe the geometric relationships between stereo images. Bensrhair et al. [12] and Toulminet et al. [13] presented a stereo vision-based algorithm for vehicle detection and distance computation. In the proposed algorithm, a stereo vision-based system is initially used to recover the most relevant 3D features in a scene. Then, it exploits the vertical symmetry characteristics of the vehicle when it is framed from the rear; subsequently, a symmetry operator investigates the previously computed

3D features, grey level image, and the images obtained by analyzing the horizontal and vertical edges. Ito and Yamada [14] presented the preceding vehicle and road lanes recognition methods for a rear-end collision avoidance system (RCAS) that uses an edge histogram method based on the model based vision concept and edge histogram method that can detect line elements of objects stably at low computation cost.

However, the cost and complicated setup for conventional synchronous binocular systems make them less desirable to consumers and inconvenient for operations. Consequently, this paper proposes a front-vehicle detection system comprising a custom-made binocular and an effective real-time front-vehicle detection algorithm. Because the proposed binocular platform uses two low-cost compact CMOS cameras with asynchronous exposure times, the pair of images from the proposed binocular platform does not meet the epipolar constraints [11]. This problem causes the image pair matching method to be more time-consuming than methods used by synchronous systems. Consequently, this paper also proposes a real-time stereo vision vehicle detection algorithm that overcomes the computational load and image matching problems to provide better performance with an asynchronous stereo camera system.

Ⅱ. MATERIALS and METHOD

2.1 Stereo vision from asynchronous cameras

The proposed binocular platform has two coplanar cameras with the same intrinsic parameters, and a pitch angle, θ, to the ground plane. Fig. 1 illustrates the differences between synchronous and asynchronous binocular systems.

In the coordinate system shown in Fig. 1, the origin of the world-coordinate system (X_w, Y_w, Z_w) is superimposed on the center of the origins of the two camera-coordinate systems (X_l, Y_l, Z_l) and (X_r, Y_r, Z_r). In the X-Y camera-coordinate system, the coordinate of a point in the two-dimensional image plane is given by (u, v). The coordinate of the projection of the optical center is thus denoted by (u0, v0), and is assumed to be at the center of the image. Transformation from the world-coordinate system to the camera-coordinate system is achieved by combining a translation vector and a rotation matrix. T_l and T_r denote the translation vectors of the left and right images, respectively, and R denotes the rotation matrix. Assuming that the left image is captured earlier than the right image, in homogeneous coordinates, the transformation matrices, D_l and D_r, of the left and right images of the asynchronous cameras are given by,

$$D_l = [R \mid T_l] = \begin{bmatrix} 1 & 0 & 0 & +\frac{B}{2} \\ 0 & \cos\theta & -\sin\theta & 0 \\ 0 & \sin\theta & \cos\theta & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \text{ and}$$

$$D_r = [R \mid T_r] = \begin{bmatrix} 1 & 0 & 0 & -\frac{B}{2}+\Delta x \\ 0 & \cos\theta & -\sin\theta & \Delta y \\ 0 & \sin\theta & \cos\theta & \Delta z \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (1)$$

where Δx and Δy are, respectively, the horizontal and vertical displacements caused by the road surface, the vehicle's suspension, and the asynchronous exposure time of the cameras; Δz is the relative distance between the host vehicle and the front vehicle on the right image; and B is the length of the baseline. Because the horizontal displacement is very small when a vehicle drives forward on a freeway or an urban road, Δx can be neglected.

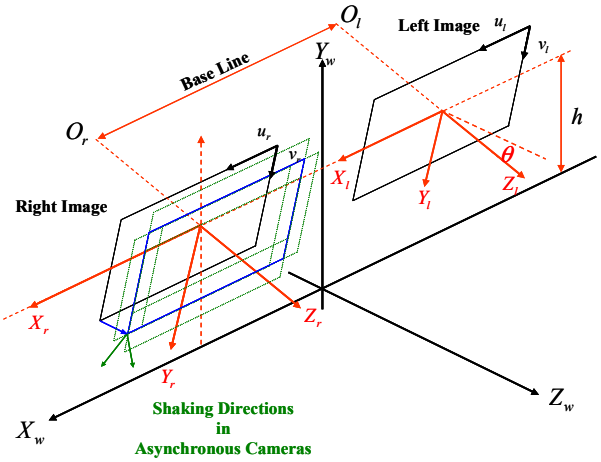


Fig.1. Differences between synchronous and asynchronous binocular systems.

The disparity value Δu is affected by Δz, which is the relative distance error between the host vehicle and the front vehicle in the right image. If the relative speed between the host

vehicle and the front vehicle is 100 km/h and the frame rate of the camera is 30 fps, then the maximum relative distance error is
The relative speed between the host vehicle and the front vehicle is approximately 10–30 km/h on the freeway; therefore, the effect of the relative distance error, Δz, can be neglected.

The correspondence matching algorithm for the asynchronous stereo-pair images appears to be more complex and time-consuming than the algorithm for the synchronous stereo-pair images. Thus, the focus of our proposed algorithm is on the correspondence matching problems caused by the asynchronous stereo vision systems.

Following application of the matching algorithm, the corresponding pixels can be matched by the left and right images. The disparity value of the corresponding pixels can be used to estimate the distance value described by Fig. 2, which has a simple expression for the coplanar cameras:

$$B = dx_L + dx_R = \frac{Z}{f}(dx_l + dx_r) = \frac{Z}{f}dx$$
$$Z = \frac{B \times f}{dx} = \frac{B \times f}{\Delta u \times W_{resolution}} \quad (2)$$

where B is the baseline; dx_L and dx_R are the horizontal distances between the object and the center axes of the left and right cameras, respectively; dx_l and dx_r are the horizontal distances between the corresponding pixels and the centers of the left and right images, respectively; Z is the estimated distance; f is the focal length of the cameras; Δu is the disparity value; and $W_{resolution}$ is the resolution of the pixel width (m/pixel) of the camera.

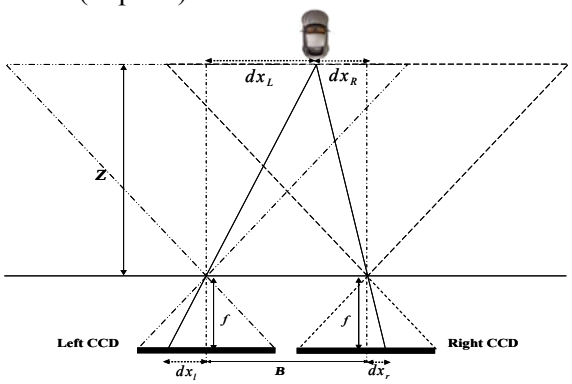


Fig. 2. Stereo geometry of a binocular system.

Parameter α, α = (B×f) / $W_{resolution}$, is defined as an intrinsic coefficient of the camera. Eq. (3) can be simplified to

$$Z = \frac{\alpha}{\Delta u} \quad (3)$$

2.2 The stereo vision vehicle detection algorithm

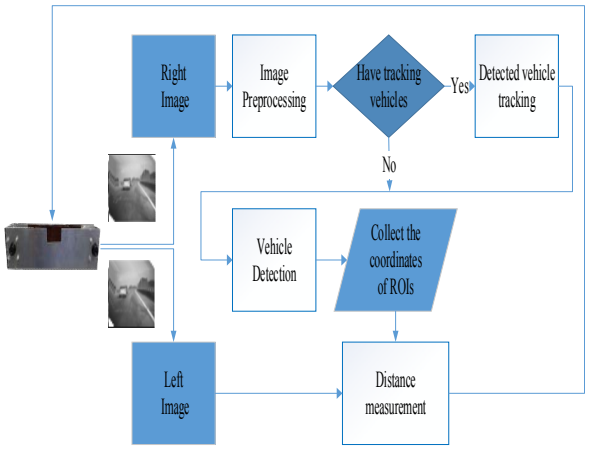


Fig. 3. Flow diagram for the proposed algorithm.

The flow diagram for the proposed algorithm is depicted in Fig. 3. The algorithm comprises four parts: image preprocessing, vehicle detection, detected vehicle tracking, and distance measurement. In the image preprocessing block, downsample, smooth, and Sobel filters are combined into one composite filter. The right image passes through this composite filter for detection of the horizontal edges. When front vehicles are detected in the previous frame, the detected vehicles are tracked by the tracking algorithm in the current frame. After the tracking algorithm, the vehicle detection algorithm detects the front vehicles from the remaining space, excluding the location of the tracked vehicle, of the current image. The regions of the candidates of front vehicles are processed by the distance measurement to detect the disparities in each corresponding point in the image pair. The relative distance and width of each candidate region are used to check the reasonability of a vehicle. The invalid candidates are deleted and the other candidates are retained as the detected front vehicles in this frame. Consequently, the total front vehicles include the tracking vehicles of the tracking phase and the detected vehicles of the detection phase. The details of the four parts are described in the following subsections.

2.3 Image preprocessing

Because the edge features of an image can retain the features of an object after the image is downsized, the right input image is downsized by a factor δ in the x and y directions to reduce computational complexity. In order to reduce

noise interference, the proposed algorithm uses a 3 × 3 average filter to smooth the image following the downsampling process. Then, an edge detection filter, $[1, 0, -1]^T$ and [1, 0, -1], is used to detect the horizontal and vertical edges, respectively. In this paper, the downsampling, smoothing, and edge detection processes are combined into a single image preprocessing process.

Assume that f(x, y) is the original right image, where x and y are the horizontal and vertical coordinates, and that the frame size is M×N pixels. The downsampling and smoothing processes of f(x, y) are combined into a single process, DS(x, y), as shown in Eq. (4):

$$DS(x,y)=\frac{1}{9}\sum_{j=-1}^{1}\sum_{i=-1}^{1} f[\delta x+i\cdot\delta,\delta y+j\cdot\delta] \quad (4)$$

where x=1,2,…,$\left\lfloor\frac{M}{2}-1\right\rfloor$, y=1,2,…,$\left\lfloor\frac{N}{2}-1\right\rfloor$, and δ is the downsampling factor.

After the image is downsized and smoothed, the edge detection filter is used to detect the horizontal and vertical edges. DSEH(x, y) and DSEV(x, y) define the operators for the combined downsampling, smoothing, and edge detection processes in order to detect the horizontal and vertical edges, respectively. Their corresponding equations are Eqs. (5) and (6):

$$DSE_H(x,y)=\left|DS(x,y-\delta)-DS(x,y+\delta)\right| \quad (5)$$

$$DSE_V(x,y)=\left|DS(x-\delta,y)-DS(x+\delta,y)\right| \quad (6)$$

2.3 Vehicle detection

The vehicle detection system searches the regions of the front vehicles in the current image according to their contours. Because the front vehicles are rigid bodies, each vehicle has at least two horizontal and two vertical boundary lines. Further, because a two-dimensional image is a mapping from the 3D real world, the parallel horizontal and vertical lines of the vehicle are located on a vertical plane that is parallel to the image plane. Consequently, the parallel horizontal lines are used to search the regions of interest (ROIs) of the front vehicle. Then, the refining process improves the proper region of each front vehicle by the vertical and horizontal edges within the ROI.

First, the horizontal edges are detected by the DSEH filter, as shown in Eq. (5). Then, the connected-component labeling [15] merges the horizontal edges into different horizontal line segments. Because the distances from the horizontal lines to the bottom of the image frame represent the proximity of the lines to the host vehicle, the processing priority of the horizontal segments is arranged from bottom to top in the image. According to the priority, each horizontal segment is used to search the corresponding horizontal segments of a vehicle. The coordinates of the left and right extreme points of each segment can be acquired after the processing of connected-component labeling. Assume that the coordinates of the left and right extreme points of the seed segment that has the highest priority are denoted by (x_{ls}, y_{ls}) and (x_{rs}, y_{rs}), respectively. Furthermore, also assume that the coordinates of the left and right extreme points of the remaining horizontal segments are denoted (x_{l_i}, y_{l_i}) and (x_{r_i}, y_{r_i}), where i is the number of horizontal segments remaining. Then, the search condition for the corresponding horizontal segments of a vehicle is defined by Eq. (7):

$$\frac{1}{2}\left|x_{ls}-x_{rs}\right|<\left|x_{l\_i}-x_{r\_i}\right|<\frac{3}{2}\left|x_{ls}-x_{rs}\right| \text{ and}$$

$$\left|y_{ls}-y_{l\_i}\right|<\frac{3}{2}\left|x_{ls}-x_{rs}\right| \quad (7)$$

The segments that satisfy Eq. (7) are clustered as the horizontal features of the same vehicle, and a bounding box is chosen to contain the horizontal features as the ROI. The horizontal segments in the ROI are marked and the rest of the horizontal segments continue to search their ROIs.

Because the shadow, background, or vehicle occlusion of the front vehicle may expand the width of the ROI, the region of the ROI should be refined by the vertical and horizontal edges within the ROI. The refining process detects the vertical edges of the ROI via the DSEV filter, Eq. (6), and selects the two largest values from the vertical projection of the vertical edges as the width of the refined region. The horizontal projection of the horizontal edges detected by the DSEH filter, Eq. (5), is detected within the width of the refined region. Assume that the width of the refined region is η. Each horizontal length of the horizontal projection larger than 0.6η is marked. The entire region with marked horizontal lines is considered to be the refined ROI. A sample experimental result for the search module is shown in Fig. 4.

Fig. 4(a) is the right image of the image pairs, whereas Fig. 4(b) is the horizontal segments after vehicle detection. One of the ROIs is illustrated in Figs. 4(c) and (d) to show the refining processing. In Figs. 4(c) and (d), the vertical and horizontal edges are projected onto the vertical and horizontal histograms to segment the refined region, respectively.

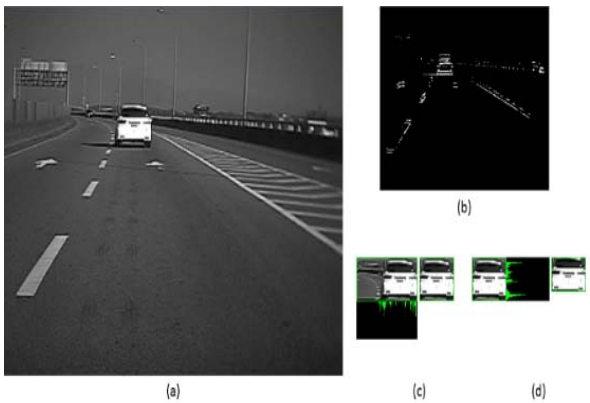


Fig. 4. Sample experimental result for the search module: (a) Original image. (b) Vehicle detection results. (c) Vertical edge projection to refine the ROI. (d) Refined ROI after horizontal edge projection.

2.4 Detected vehicle tracking

The tracking phase tracks the front vehicles detected from the previous right image. Because the frame rate of the cameras is 30 fps, the position of the front vehicles in the current frame is very close to the positions of the detected vehicles in the previous frame. In the current frame, the tracking module searches the locations of vehicles around the refined ROIs of the previous frame.

In order to speed up the processing, the search space for each detected vehicle is enlarged based on the refined ROI of the previous frame. Assume that the left-up and right-down coordinates of the refined ROI of the previous frame are (x_{lu}, y_{lu}) and (x_{rd}, y_{rd}), respectively. Thus, the left-up and right-down coordinates of the search area of the current frame are $(x_{lu}-\Delta, y_{lu}-\Delta)$ and $(x_{rd}+\Delta, y_{rd}+\Delta)$, respectively. In the tracking phase, the vertical and horizontal edges are detected and used to locate the region of the tracking vehicle. Fig. 5 illustrates the processing of the tracking module.

Fig. 5 illustrates the processes employed in the tracking phase. Fig. 5(a) shows the enlarged region of the detected vehicle in the previous frame that is used to track the detected vehicle in the current frame. The extended region for the nth frame is shown in Fig. 5(a). In this paper, parameter Δ is set to 30 pixels. The vertical edge projection processed by Eq. (6) locates the two maximum peaks to cut the tracking vehicle shown in Fig. 5(b). After the vertical cut, the horizontal edge projection is processed by Eq. (5). Because the width of the tracking vehicle is $|x_{rd}-x_{lu}|$, the peak values of the horizontal edge projection greater than $0.6|x_{rd}-x_{lu}|$ are marked as valid horizontal lines. These lines can be used to locate the top and bottom sides of the tracking vehicle, as shown in Fig. 5(c). After tracking the front vehicles detected from the previous image pair, the other region of the image frame is processed by the search module to detect vehicles entering the frame.

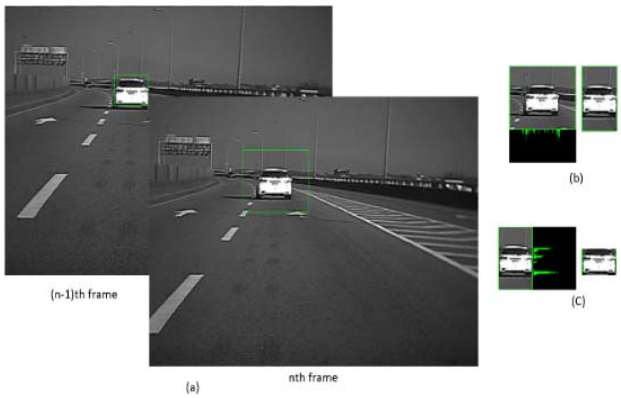


Fig. 5. Tracking the detected vehicles: (a) Enlarge the region of the detected vehicle at the (n-1)th frame in the nth frame. (b) Use the vertical edge projection to locate the left and right sides of the tracking vehicle. (c) Use the horizontal edge projection to locate the top and bottom sides of the tracking vehicle.

2.4 Distance measurement

The ROIs obtained by the vehicle detection and tracking phases situate the positions of the obstacle-like objects in the right image. Because the coordinates of the ROIs are located in the downsampling image, the coordinates have to multiply the factor δ back to the original resolution. Therefore, the coordinates of the edge pixels within each ROI are multiplied by δ, and each edge pixel is enlarged to $\delta \times \delta$ pixels in the original resolution. Then, the distance measurement algorithm calculates the disparities of the edge pixels within each ROI in the original resolution.

Because the pixel in the right image and its corresponding pixel in the left image are not in the same horizontal line for the asynchronous stereo vision system, the effect of vertical shift between the image pairs should be considered. The proposed algorithm calculates the disparity of each edge pixel using the full search block-matching algorithm. Assume that $f_R(x, y)$ and $f_L(x, y)$ represent the right and left images, B_w and B_h denote the width and height of the block size, and S_W and S_H denote the width and height of the search area.

The width of each ROI is a very useful parameter in the deletion of the error detections. Because the widths of different vehicles vary

from 1.7 m to 2.5 m, the reference width of a vehicle is set as 2 m in this paper. The reference width is used to estimate the vehicle width in images under different distances. Fig. 6 is the mapping of the reference width from 5 m to 50 m. In Fig. 6, the x-axis is the relative distance between the front vehicle and the binocular system and the y-axis is the pixel width of 2 m in the images. When the relative distance is calculated, the reference width of the vehicle can be obtained, as shown in Fig. 6. Assume that the width of the ROI is W_{ROI} and the reference width mapped by Fig. 6 is W_{ref}. When the width of the ROI satisfies Eq. (8), the ROI is deleted as an error. After the matching, the coordinates of the ROIs are used to track the vehicles in the next frame.

$$W_{ref} + \frac{W_{ref}}{20} < W_{ROI} \quad \text{or}$$

$$W_{ROI} < W_{ref} - \frac{W_{ref}}{20} \qquad (8)$$

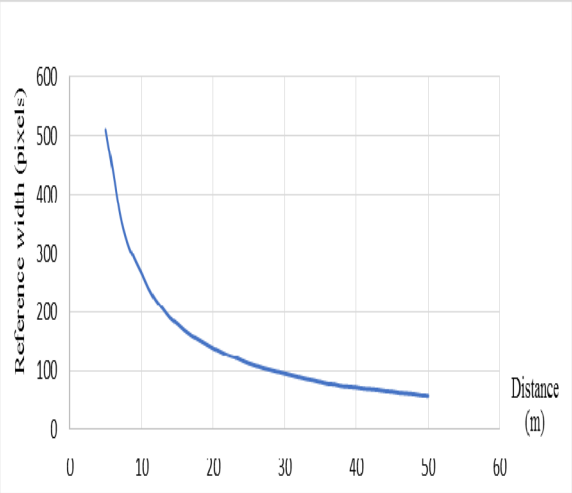


Fig. 6. Reference width under various relative distances.

Ⅲ. RESULTS and DISCUSSION

3.1 System overview

The hardware used to implement the proposed system comprised four components: custom-made binocular vision sensors, an image capture card, a power converter, and an industrial computer. Fig. 7 shows the connections between components. The proposed binocular system is less obtrusive and much easier to install in the back of the rear-view mirror by the driver. In other words, the setup does not interfere with safe driving by blocking the view of the driver. The industrial computer used for the experiments had an Intel Core i7-620LE 2.0 GHz CPU. Images were obtained using a PCI-RTV24 capture card. The focal length of the CMOS camera was 12 mm and the image size was 640 × 480 pixels. The program used was implemented in Microsoft Visual C++. For the pair of images, the average processing time of the detection algorithm was approximately 9×10^{-3}s.

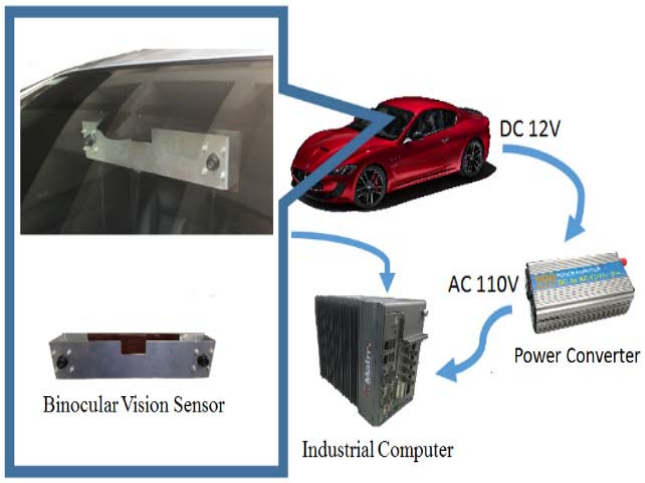


Fig. 7. System overview.

3.2 System reliability analysis on highways

To provide a quantitative assessment of the proposed system to detect the front vehicle, it was implemented on a vehicle and underwent long-term testing on highways with various sources of interference. The image frames captured by the binocular detection system under varying velocities and on various test road conditions, such as tunnel, overpass, and curve, are summarized in Table 1.

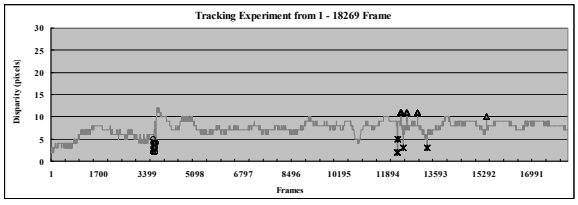
Table 1. Various environmental parameters from long-term testing on highways.

Route	A–B	B (Tunnel)	B–C	C (Curve)
Frames	**1–2341**	**2342–3734**	**3734–10013**	**10014–11498**
Speed	**70–100 km/h**	**60–80 km/h**	**70–100 km/h**	**40–60 km/h**

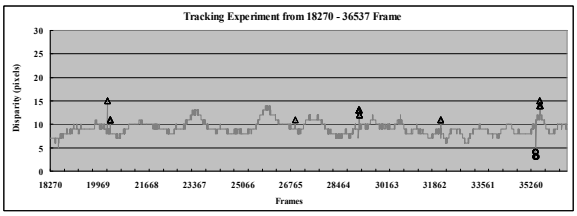
Route	C−D (Overpasses)	D (Overpass)	D−E	E−F (Overpasses)
Frames	**11499−35146**	**35147−35476**	**35477−35537**	**35538−61465**
Speed	**70−100 km/h**	**60−80 km/h**	**70−100 km/h**	**70−100 km/h**
Route	F (Curve)	F−G	G (Tunnel)	G−A
Frames	**61466−62302**	**62303−69426**	**69427−70914**	**70915−74631**
Speed	**40−60 km/h**	**70−100 km/h**	**60−80 km/h**	**70−100 km/h**

The detection system was evaluated on highways, and regulation experiments were conducted to determine the reliability of the proposed system over the long term. The parameters of the proposed system were not adjusted during the test periods on the highways.

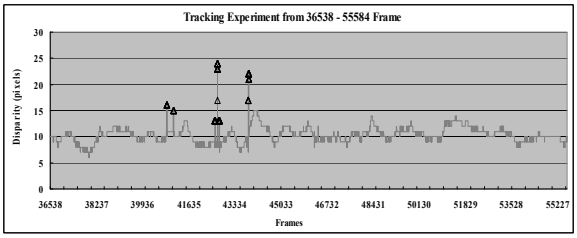
(a)



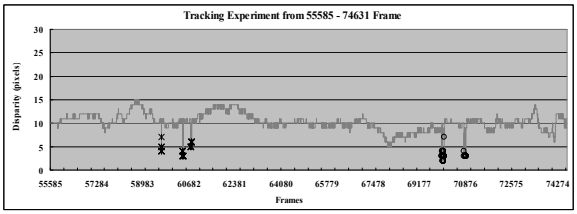
(b)



(c)



(d)



(e)

Fig. 8. (a) Testing route in Taiwan. (b)-(c) Experimental results on route A-B-C-D-E. (d)-(e) Experimental results on route E-F-G-A.

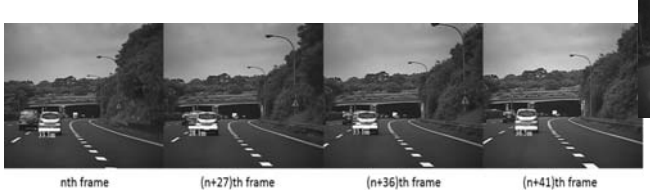
Field tests were conducted in which we attempted to detect front vehicles on Highway No. 2 and Highway No. 3 in Taiwan, as shown in Fig. 8(a). The maximum road curvature is 300−1 and 500−1 m−1 for Highways No. 2 and No. 3, respectively, and the permitted speed ranges from 60 to 110 km/h. Locations B and G are tunnels that have an interval of open space between them that is approximately 100 m long. Locations C and F are slip-roads connecting the two highways; the maximum road curvature is 200−1m−1 and the legal velocity ranges from 40 to 60 km/h. After slip-road C, there are three overpasses, each approximately 15 m wide, at the beginning of C-D and an interval between the two overpasses approximately 100 m in length. In other words, the same overpasses are at the end of E-F before slip-road F. The total length of the field test route was approximately 50 km and the testing time was approximately 43 min with 74,631 pairs of images.

In the field test, the driver maintained various distances while following the same vehicle. On the highways, the proposed detection system successfully detected and

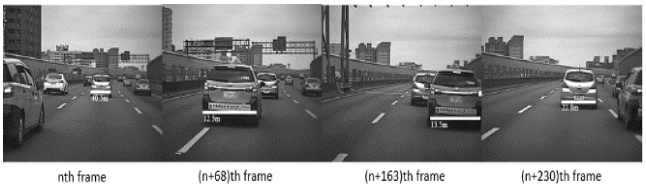
tracked the front vehicle over the entire test route. The overall detection rate was approximately 99.7% (74,456/74,631). There were 175 error frames in the test. Figs. 9(b)–(e) are the detection results for the field test (74,631 pairs of images). The detection error is defined as a front-vehicle detection disparity greater than 2 pixels. Three possible causes for the detection errors in the test are outlined below:

a) False detection: False detection denotes the disparity/matching error. We use "○" as the symbol to mark the false detection. There are 73 such false detection frames in the long-term testing results.
b) Overpassing: Overpass errors occur when a car overpasses from the back of the host vehicle. When the car is too close to the front vehicle, there are some detection errors in locating the front vehicle. We use "Δ" as the symbol to mark the overpassing errors. There are 65 such false detection frames in the long-term testing results.
c) Feature loss: The edges of the front vehicle temporarily disappear when the vehicle passes through the overpass. When this occurs, the detection algorithm may detect other vehicles in front of it to replace the leading vehicle until the leading vehicle reappears. We use "x" as the symbol to mark these errors. There are 37 such false detection frames in the long-term testing results.

Figs. 8(b) and (c) show the experimental results for route A-B-C-D-E; and Figs. 9(d) and (e) are the experimental results for route E-F-G-A. It is interesting to note that the accuracy of the detection system was not seriously affected when the leading vehicle was on the curve road (C and F), in the tunnels (B and G), or on the overpass (D). Most cases of mistaken detections arose when there were sudden changes in illumination causing CCD overexposure, such as in Fig. 9(a) in the tunnels, frames 2342−3734, frames 35147−35476, and frames 69427−70914. Fortunately, these detection errors were non-successive and transient. A feature loss error is shown in Fig. 9(b) and the overpass errors are shown in Fig. 9(c).



(a)



(b)

(c)

Fig. 9. Detection errors in the field test: (a) False detection. (b) Overpassing. (c) Feature loss.

3.2 Experimental Results Obtained under Various Weather/Illumination Conditions

In our evaluation of the performance of the detection system, tests were also conducted on urban and country roads under various weather and lighting conditions. For the tests conducted on the urban and country roads, the detection rate is defined as the percentage of correct detections in the detection zone. This zone is defined as the zone ranging from 5 m to 50 m in front of the binocular system. A detection error arises when the detected disparity of the front vehicle is greater than 2 pixels. In the tests, the host vehicle maintained a speed of 30–50 km/h to test the proposed system on the urban and country roads. All of the front vehicles in the road tests were selected randomly without prior planning. Some of the experimental results under various weather/illumination conditions are provided below. The detected distance is shown at the bottom of each detected vehicle and is measured in meters. There are a total of 37,880 pairs of test images. The average detection rate was approximately 93.2% under the various weather/illumination conditions. Sample detection results are shown in Fig. 10.

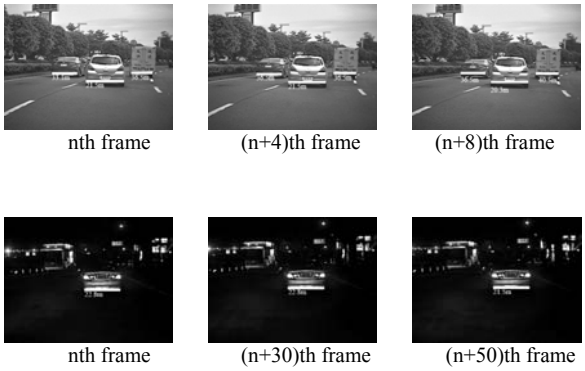


Fig. 10. Detection results under various weather/illumination conditions.

Ⅳ. CONCLUSIONS

This paper proposed a detection system comprising a custom-made binocular and an effective real-time front-vehicle detection algorithm. The proposed detection system is able to detect vehicles under various weather and illumination conditions. The results of long-term tests conducted indicate that the proposed system has a detection rate of approximately 99.7% and 93.2% under normal conditions on highways and under various illumination conditions, respectively. The average processing time of the detection algorithm is approximately 9 ms. Thus, the proposed system can be considered a real-time application. Its low price, small size, and easy implementation makes the proposed asynchronous stereo vision system suitable for vehicles as a vision sensor to detect front vehicles. Moreover, unlike synchronous stereo vision systems, which have to satisfy epipolar constraints to detect feature points, the proposed detection algorithm is unrestrained. The experimental results obtained from various tests indicate that the proposed system can successfully detect front vehicles and estimate their distances under various weather and illumination conditions.

REFERENCES

[1] Nashashibi, A., Sarabandi, K., De, R. R., Frantzis, P., and Ulaby, F., "An ultrafast wide-band millimeter-wave (MMW) polarimetric radar for remote sensing applications," IEEE Trans Geosci. Remote Sens., Vol. 40, No. 8, pp. 1777-1786, 2002.

[2] Shiu, H. T., Hall, P. S., Hoare, E. G., and Clarke, N. J., "Advance path measurement for automotive radar applications," IEEE Trans. Intell. Transp. Syst., Vol. 7, No. 3, pp. 273-281, 2006.

[3] Jung, H. G., Cho, Y. H., Yoon, P. J., and Kim, J., "Scanning laser-radar-based target position designation for parking aid system," IEEE Trans Intell. Transp. Syst., Vol. 9 No. 3, pp. 406-424, 2008.

[4] Trivedi, M. M., Gandhi, T., and McCall, J., "Looking-in and looking-out of a vehicle Computer-vision-based enhanced vehicle safety," IEEE Trans. Consum. Electron., Vol. 8, No. 1, pp. 108 - 120, 2007.

[5] Broggi, A., Bertozzi, M., Fascioli, A., Bianco, C. G. L., and Piazzi, A., "Visual perception of obstacles and vehicles for platooning," IEEE Trans Intell. Transp. Syst., Vol. 1, No. 3, pp. 164-172, 2000.

[6] Suzuki, T., and Kanade, T., "A Measurement of vehicle motion and orientation using optical flow," Proc. IEEE ITSC, Tokyo, pp. 346-349, 1999.

[7] Stein, G. P., Mano, O., and Shashua, A., "A robust method for computing vehicle ego-motion," Proc. IEEE IV, Dearborn, pp. 362-368-397, 2000.

[8] Srinivasa, N., "Vision-based vehicle detection and tracking method for forward collision warning in automobiles," Proc. IEEE IV, New York, pp. 626-631, 2002.

[9] Shashua, A., Gdalyahu, Y., and Hayun, G., "Pedestrian detection for driving assistance systems: Single-frame classification and system level performance," Proc. IEEE IV, Parma, pp.1-6, 2004.

[10] Forsyth, D. A., "Computer Vision: A Modern Approach," New Jersey, Prentice Hall, 2000.

[11] Ma, Y., Soatto, S., Kosecka, J., Sastry, S. S., "An invitation to 3-D vision," Springer, 2000.

[12] Bensrhair, A., Bertozzi, M., Broggi, A., Fascioli, A., Mousset, S., and Toulminet, G., "Stereo vision-based feature extraction for vehicle detection," Proc. IEEE IV, New York, pp. 465-470, 2002.

[13] Toulminet, G., Bertozzi, M., Mousset, S., Bensrhair, A., and Broggi, A., "Vehicle detection by means of stereo vision-based obstacles features extraction and monocular pattern analysis," IEEE Trans. on image processing., Vol. 15 No. 8, pp. 2364-2375, 2006.

[14] Ito, T., and Yamada, K., "Preceding vehicle and road lanes recognition

methods for RCAS using vision system," IEEE IV, Paris, pp. 85-90, 1994.

[15] Suzuki, K., Horiba, I., and Sugie, N., "Fast connected-component labeling based on sequential local operations in the course of forward raster scan followed by backward raster scan.," Pattern Recognition, pp. 434-437, 2000.