2010-01-0713

Design and Implementation of an Image-based Driver Attention Warning System on Low-cost DSP Platform

Yu-Sheng Liao, Ming-Cong Weng and Chia-Tseng Chen
Automotive Research & Testing Center (ARTC)

ABSTRACT

This paper proposes an image-based vehicle safety system, which has the ability to monitor the driver's status during their driving in all environment conditions. The image process technique, used to detect the driver's face by AdaBoost based algorithm, is suitable to be applied under the internal environments of cabin. By analyzing the position of driver's face, once the driver's face disappeared in a reasonable area, the system will regard it as inattentive driving. When the acts of inattentive driving occurred, the system will generate warning signals through a buzzer or other devices to draw the driver's attention to notice on his driving. This system has been implemented on low-cost DSP platform and it can be installed efficiently on varied kind of vehicles. Under day and night driving conditions with series of specific viewing angles (from ahead, left/right-turn and turn-back views, which are corresponded to the angles in 0°~±90° and 180°), the averaged correct warning rate is 92.1%. In addition, under the circumstance of attentive driving, the correct warning rate is 96.08%.

INTRODUCTION

According to the statistical [1, 2] from U.S. National Highway Traffic Safety Administration (NHTSA) shows that inattentive driving is the major factor leads traffic accidents. Behaviors, such as operating hand-held devices, napping and picking objects when driver's driving, may shift the driver's attention on driving and raises the collision rate more than 30%. To avoid this problem, there are relevant solutions which have been proposed. A system is lane departure warning system (LDW) based on vision technology. This system mounts a video camera on the forepart of vehicle to look ahead, and outputs the lane departure rate. When the vehicle diverges from it's expect path, the LDW system raises alarms in the forms of LED light or buzzer etc, al. LDW notices the driver to concentrate on their driving adequately.

Besides, another solution is DD850/Copilot [3] based on vision technology to read the driver's eyelid behaviors to distinguish the driver attentiveness. This solution uses invisible infrared LEDs as an illuminator to cause the bright/dark pupil effect. Then, the pupil is located precisely through a serious of image processing procedure. Moreover, a measurement named PERCLOS (PERCentage of eyelid CLOSure) is applied to determine the level of driver's drowsiness in this solution. This system currently only works well under the low lighting conditions, such as night-time. Another solution has been promoted as a marketed product, FaceLAB [4], which relies on vision technology as well, but two-dimensional (2-D) template searching and three-dimensional (3-D) stereo matching is applied for its facial features extraction task. Due to its computational complexity, FaceLAB has only been deployed on PC-based and not been popular due to its cost.

This study proposes a low-cost vision-based solution to determine the driver's head pose and watch their focuses on their driving in all driving conditions (high-/low- ambient lighting, pose and styling et, al.). Figure 1 shows the framework of this system. In the first, image capture unit grabs a serious of images from the cabin through a video camera. These grabbed images input to a image process unit, which consists of AdaBoost algorithm for driver face detection. In the third, driver status discrimination unit contains a logical rule to determine the driver's status in the end. In the alarm unit, this system generates a warning signal to draw driver's attention back to the traffic, if the driver loses his attentive in driving. Now, this solution has been implemented on a low-cost DSP platform and verified in real application under day and night driving conditions.

Figure 1. System framework

In this paper, section 2 introduces the hardware and software framework and Adaboost face detection technique. Section 3 shows the experimental result and discussions. The conclusion presents in section 4.

2. METHOD

2.1 Framework of Hardware System

In this system, the video input is one mono video (Black and White) camera with 320x240 (QVGA) resolution and with capturing rate in 30 frames/sec., as shown in Figure 2(A). Furthermore, two IR-illuminators and a DSP module are essential, as shown in Figure 2(B) and Figure 2(C). IR-illuminators provide invisible ambient lighting source to ensure the camera could sense the driver's face in a low ambient lighting condition. The DSP module shown in Figure 2(C), process the image data from a video camera. Inside the DSP, a face detection algorithm and driver status estimator has been deployed. Behind the face detection and status estimator , once the driver in the status of inattentive driving, the system raises alarm signals through a buzzer or other devices to draw the driver's attention to notice on their driving.

Figure 2(D) shows the installation of hardware. One mono video camera mounted on the dashboard, two IR-illuminators set on the sides of upper-right and upper-left. The mono video camera mounts with a narrow band-pass filter, which filters out the visible light. Therefore, only the IR wavelength will be received by the mono video camera. When the system lights up the invisible IR-illuminators in the dark environment, it ensures that the camera will received the reflected invisible IR light from the driver's face, and does not disturb the driver himself. The optical design reduces the effects from the vehicle light and other artificial lights in most cases. The system works appropriately both in the dark and bright. Due to the sensitivity of the eye from the IR emission, the power has been followed the recommendation from the International Electrotechnical Commission (IEC) 825-1 and the European Committee for Electrotechnical Standardization (CENELEC) 60825-1 [5].

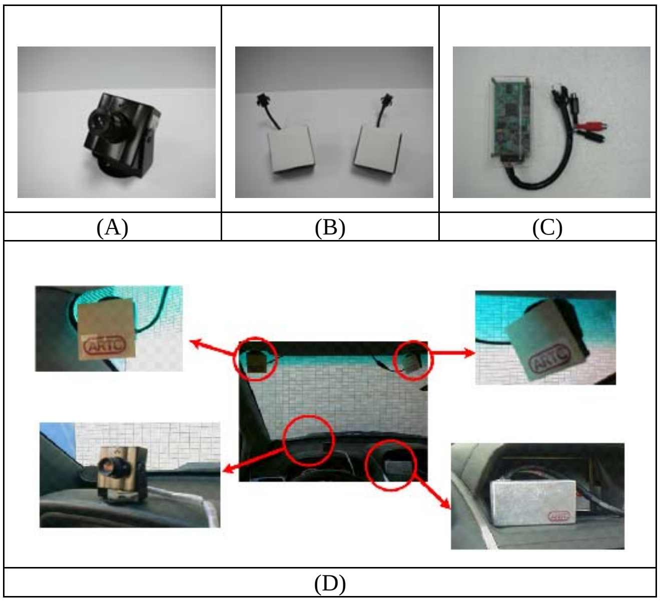

Figure 2. System hardware (A) Video camera (B) IR-illuminator (C) DSP module (D) System installation

2.2 Software Framework

The software framework is shown in Figure 3. After a series of images grabbed by the video camera, the first step is image pre-processing. In the face detection stage, AdaBoost based method [6] and cascade of strong classifiers are adopted. In order to recognize the driver's status, the result from AdaBoost face detection is taken in one single image. To analyze the driving behavior, the driver status estimation uses the continuous results from face detection unit,

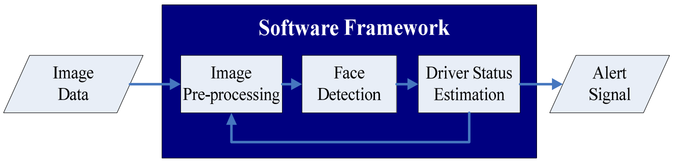

Figure 3. Software framework

2.2.1 Image Pre-processing

Inside the pre-processing procedure, some essential image processing methods, e.g., ROI (Region of Interest) selection, nearest-neighbor interpolation, integral image transformation, and Haar-like features estimation are adopted for Adaboost face detection. In this system, due to the camera FOV and its installation in the cabin, the driver's face must be located in the center when they are driving in a normal condition. Thus, the system assumes that once the driver's face removed from the normal position it should regard as inattentive driving. To make this system becomes efficiently, a ROI (Region of Interest) is selected on each original image as input to reduce computational cost. The nearest-neighbor interpolation [7, 8] is used to rescale the size of input images.

To calculate the face feature parameters for the Adaboost face detection unit, the Haar-like features are used to describe light and shade characteristics of a human face, as shown in Figure 4. In order to calculate these features efficiently, a new image representation called integral image that allows for fast feature evaluation. Once the integral image is established, any one of these Haar-like features can be computed at any scale or location in constant time.

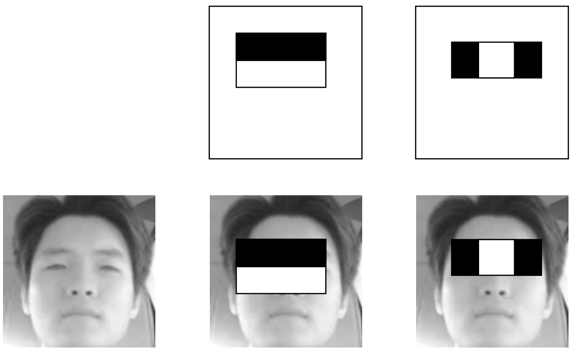

Figure 4. Haar-like features are used to describe light and shade characteristics of a human face

2.2.2 Driver's Face Detection

In the face detection stage, an AdaBoost based face detection model is developed through a training procedure. It generates a facial model to discriminate faces or non-faces on its input image. There are general facial databases can be applied for facial model training, like FERET [9] , AT&T [10] , MIT [11] and UMIST [12] etc. However, they might not suitable for driver monitoring application, due to their illumination environment and computation cost, etc. Under these considerations, training samples are collected under all in-cabin environments by Automotive Research & Testing Center (ARTC). In order to overcome the interference from varied ambient light, all the images are converted to gray level and be normalized before the training procedure. Refer to [6] for the details about the training procedure.

The training result is a set of parameters, which can be used to construct a strong classifier. A strong classifier $C(x)$ consists of a set of weak classifiers $h_j(x)$. The number of weak classifier depends on iteration time T. For each weak classifier, it contains a feature $f_j(x)$, a threshold θ_j and a weight α_t. The form of weak classifier is shown as Equation (1).

$$h_j(x) = \begin{cases} 1 & if \quad f_j(x) > \theta_j \\ -1 & otherwise \end{cases} \tag{1}$$

Here x is a sub-window of an image that it has the same size with detector. For each feature, the weak learner determines the optimal threshold classification function, such that the minimum number of examples is misclassified.

After T times iteration, T weak classifiers ($h_1(x), h_2(x), \ldots, h_T(x)$) will be gained. These weak classifiers can form a strong classifier by linear combination, as shown in Equation (2). In practice, the output of a strong classifier could be regarded as the confidence of the face region in a score. Thus, once the score is higher than a threshold value, the scanning block will be regarded as a candidate region of face. In the end of scan, a region with the maximum scorn will be chosen as a face in an image.

$$C(x) = \begin{cases} 1 & \sum_{t=1}^{T} \alpha_t h_t(x) \geq 0 \\ -1 & otherwise \end{cases} \tag{2}$$

To enhance the classification accuracy, the cascaded classifier introduced by Viola and Jones [6] is adopted, as illustrated in the Figure 5. In the cascaded structure, each strong classifier rejects numbers of sub-windows. A positive result from the first strong classifier triggers the evaluation of a second strong classifier. A negative outcome leads to immediate rejection and it won't be evaluated at next stage. A positive result from the second strong classifier triggers a third strong classifier, and so on.

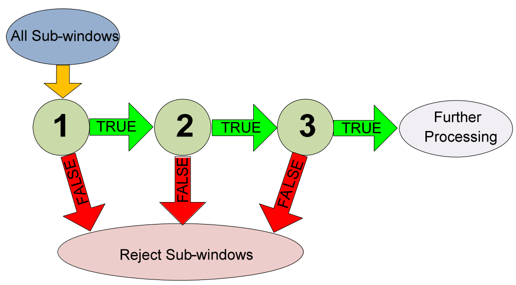

Figure 5. Concept of cascaded classifier

The structure of the cascade reflects the fact that within any single image a majority of sub-windows are negative. As such, the cascade attempts to reject as many negatives as possible at the earliest stage. Only a few numbers of sub-windows are filtered out to be evaluated at last stage.

2.2.3 Driver Status Estimation

Due to the driving behavior is a continuous motion, thus it is hard to analyze these behaviors from a single time slot. In general, people tend to take a series of time slot to observe the driver's behavior instead.

In this study, this system take 2-seconds result from AdaBoost as time slots and it will be recorded and updated continually. Within these 2 seconds, once the driver's face positions was not appeared in a reasonable or fully disappeared in whole detection area, the system will raise an alarm. The rule of estimation is shown in Figure 6.

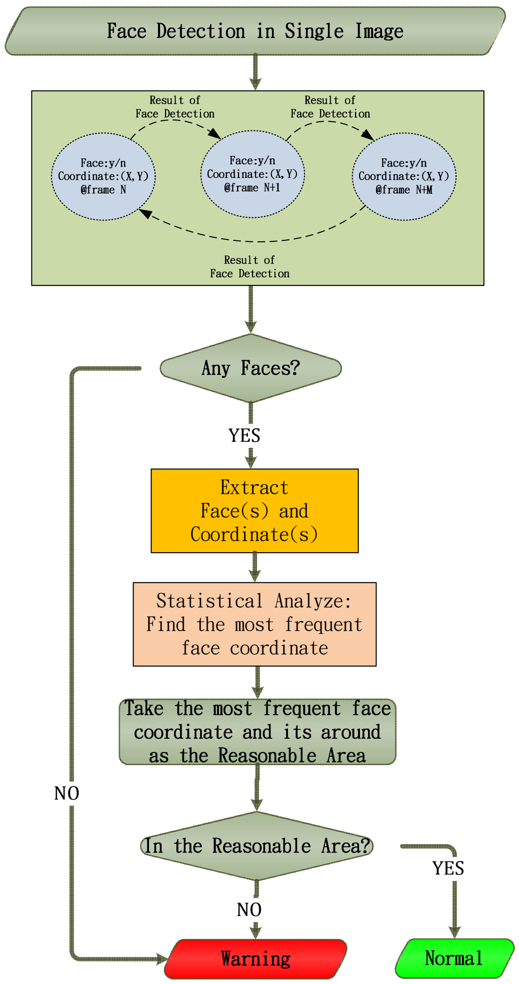

Figure 6. Logic rule of driver's status estimation

To define the reasonable area, the system takes the faces information from a serious of pervious frames. Once a face has been detected in an image frame, its coordinate will be stored for the next 2 seconds and taken into statistical analysis. Thus, there is a queue (sliding time slot), which contains the results from the past 2 second's face detection. When the queue is full, a statistical analysis is proposed. This process finds the most frequent facial coordinate and takes its' around as a reasonable area. This area updates dynamically when a new face been detected. For the following, the system compares the coordinate of the latest detected face with the reasonable area. Once these areas are not overlapped, it takes the newest detected face as inattentive status and checks if this status lasts for 2 seconds or not. If the inattentive status keeps over 2 seconds, the system raises an alarm to the driver. For another case, on the condition of a driver's face not appears in the whole image and keeps over 2 seconds, it raises an alarm as well. If and only if one of these conditions happened, the system raises an alarm. This process flow adapts to driver's pose changing and is robust to variant environment, which is suitable for the usage of driver monitoring field.

3. EXPERIMENTAL RESULT AND DISCUSSION

For each driver with different driving habit, it's hard to well define inattentive driving. Thus, in order to estimate the system performance objectively, a set of viewing angles (forward looking, down, dashboard, left/right reflectors, left/right center pillars and back) which are simulated by driver have been tested. Among these directions, possible circumstances of inattentive driving are included. Figure 7 shows information of this system which is displayed on screen.

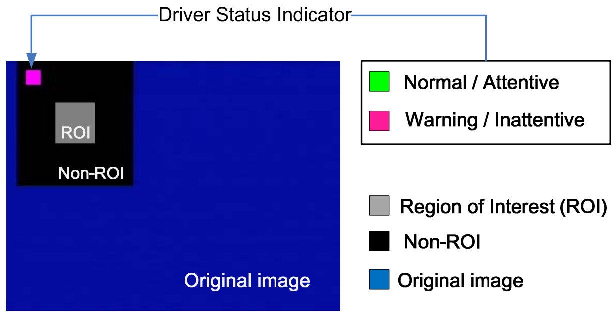

Figure 7. Detection interface displays on screen

The following describes the judgment rules of face detection as correct or false identification cases and the analysis results of the above driving conditions.

3.1 Correct Identification Case

As shown in Figure 8, there are only two negative (attentive driving) scenarios in our system: look forward and dashboard. Both of these head poses are the most frequent in normal driving. Thus, if and only if the driver face on to the camera, the scenarios regard as attentive driving and the results are shown below.

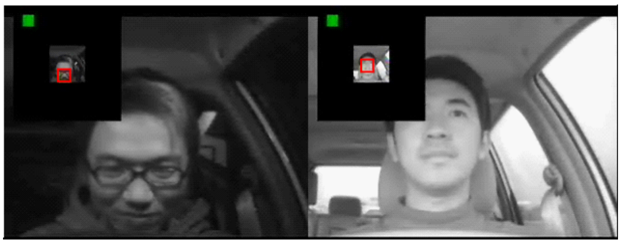

Figure 8. Snapshot of the cases regard as Attentive (True Negative)

As shown in Figure 9, the others are positive (inattentive driving) scenarios in our system and the alarm would be raised. There are 8 scenarios, including (A) Look Back, (B) Left Center Pillar, (C) Right Reflector, (D) Look Down, (E) Left Reflector, (F) Rear-view Mirror, (G) Right Center Pillar and (H) Control Panel. In the cases of (A) to (H), the head poses or the fields of views are removed from the road on the front. Thus, if any of these situations keeps for a period of time (ex. 2 seconds), the alarm would be raised in time.

Figure 9. Snapshot of the cases regard as Inattentive (True Positive) (A) Look Back, (B) Left Center Pillar, (C) Right Reflector, (D) Look Down, (E) Left Reflector, (F) Rear-view Mirror, (G) Right Center Pillar and (H) Control Panel

3.2 False Identification Case

The cases shown below are respected to false identification cases (false negative). Most of them are referred to the robustness of AdaBoost face detector. No matter day or night, when the driver's head moved away from the straight front, our AdaBoost face detector still remained its robustness in some ways. The results are shown in the following figures from 10 to 12. Theses cases are reduced by this study effective compared to the traditional Adaboost face detection.

Figure 10. Snapshot of the cases in miss warning / false negative (Look Down)

Figure 11. Snapshot of the cases in miss warning / false negative (Rear-view Mirror)

Figure 12. Snapshot of the cases in miss warning / false negative (Left Reflector)

3.3 System Performance

The system is verified under all driving conditions (day/night, rain/shine) with series of specific viewing angles (from ahead, left/right-turn and tune-back views, which are corresponded to the angles in 0°~±90° and 180°). Experimental result shows that the averaged correct warning rate can reach to 92.1%.

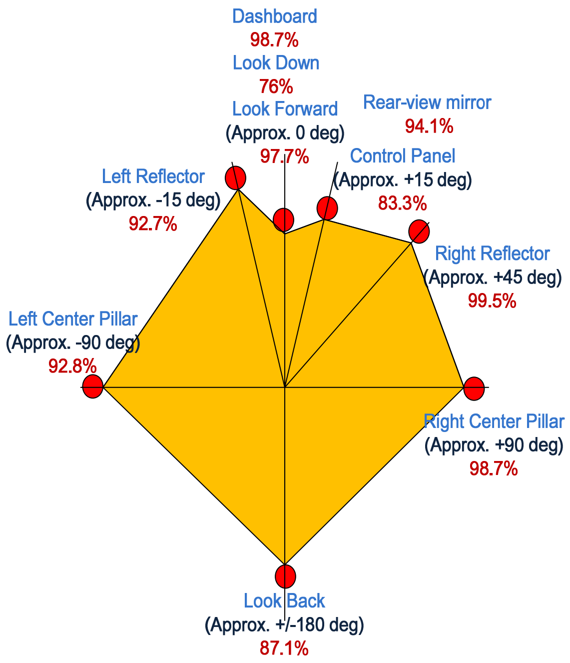

Figure 13. Warning accuracy of each direction

In addition, the system is tested under circumstance of attentive driving (Driver drives on a straight lane and the viewing angle less than 15°). The drivers including both males and females are without any limitations of their hair styles and glasses affection. The results are listed in Table 1. The mean accuracy of all testing condition is 96.08 %. Therefore, the Driver Attention Warning System based on the low cost DSP platform achieves good performance to draw the driver attention on real road environment testing.

Table 1. True Negative Rate in Attentive Driving

Attentive/ Normal Driving	True Negative Frames	Total Frames	True Negative Rate(%)	Averaged(%)
Day	203301	220289	92.29%	96.08%
Night	220099	220384	99.87%	
Short Hair/ Male	331863	346715	95.72%	
Long Hair/ Female	91537	93958	97.42%	
Naked Eye	46510	48713	95.48%	
With Glasses	15070	391960	96.16%	

4. CONCLUSION

The driver attention warning system based on AdaBoost face detection and statistical estimation technique is developed in this study. This system operates both in the day and night time, and under all weather conditions. Besides, this system has been deployed on low-cost DSP platform and been installed efficiently on various kind of vehicles, e.g., truck, van, and passenger car. By validating a series of specific viewing angles, the mean of correct warning rate is 92.1 %, and under a condition of attentive driving, the correct warning rate is 96.08%. The experimental results show that it has a ability to raise an alarm and draw back the driver's attention adequately.

ACKNOWLEDGMENTS

This research was supported by Ministry of Economic Affairs – Department of Industrial Technology, Taipei, Taiwan.

REFERENCES

1. The 100-Car Naturalistic Driving Study Phase II – Results of the 100-Car Field Experiment, National Highway Traffic Safety Administration (NHTSA), April 2006.

2. The Impact of Driver Inattention on Near-Crash/Crash Risk: An Analysis of 100-Car Naturalistic Driving Study Data, National Highway Traffic Safety Administration (NHTSA), April 2006.

3. Grace, "Drowsy driver monitor and warning system," in Proc. Int. Driving Symp. Human Factors Driver Assessment, Training and Vehicle Design, Aspen, CO, Aug. 2001, pp. 64–69.

4. Seeing Machines. (2004, Aug.). Facelab Transport. [Online]. Available:http://www.seeingmachines.com/transport.htm

5. Application Note 1118: Compliance of Infrared Communication Products to IEC 825-1 and CENELEC EN 60825-1, Palo Alto, CA: Agilent Technologies, Inc., 1999.

6. Viola, P., and M. Jones., "Robust Real-time Object Detection," in Proceedings of IEEE Workshop on Statistical and Theories of Computer Vision, 2001.

7. Jensen, K., and D. Anastassiou, "Subpixel edge localization and the interpolation of still images," IEEE Trans. On Image Processing, Vol. 4, pp. 285-295, Mar. 1995.

8. Sonka, M., V. Hlavac, and R. Boyle, Image Porcessing, Analysis, and Machine Vision, PWS Pubishing, pp. 67-68, 1999.

9. Phillips, P. J., H. Moon, P. J. Rauss, and S. Rizvi, "The FERET evaluation methodology for face recognition algorithms", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 22, No. 10, October 2000.

10. ORL face database, website: http://www.cam-orl.co.uk/facedatabase.html, AT&T Laboratories Cambridge.

11. Turk, M., and A. Pentland, “Eigenfaces for recognition,” J. Cognitive Neuroscience, vol. 3, no. 1, pp. 71–86, 1991.

12. Graham, D. B., and N. M. Allinson, “Characterizing virtual eigensignatures for general purpose face recognition”, in Face Recognition: From Theory to Applications, NATO ASI Series F, Computer and Systems Sciences, H. Wechsler, P. J. Phillips, V. Bruce, F. Fogelman-Soulie, and T. S. Huang, Eds., vol. 163, pp. 446–456. 1998.

CONTACT

Yu-Sheng Liao

Automotive Research & Testing Center (ARTC)

R & D Division

Taiwan

e-mail: ysliao@artc.org.tw