DRIVER'S DROWSINESS DETECTION BASED ON VISUAL INFORMATION

Marco Javier Flores, José María Armingol and Arturo de la Escalera
Intelligent System Laboratory, Universidad Carlos III de Madrid, Leganés 28911, Madrid, Spain
{mjflores, armigol, escalera}@ing.uc3m.es

Keywords: Drowsiness, driver assistance system, object detection, support vector machine, intelligent transportation technology.

Abstract: In this paper, a new Driver Assistance System (DAS) for automatic driver's drowsiness detection based on visual information and image processing is presented. This algorithm works on several stages using Viola and Jones (VJ) object detector, expectation maximization algorithm, the Condensation algorithm and support vector machine to compute a drowsiness index. The goal of the system is to help in the reduction of traffic accidents caused by human errors. Examples of different driver's images taken over a real vehicle are shown to validate the algorithm.

1 INTRODUCTION

Active Security, whose objective is to endow vehicles with intelligent systems that predicts and avoids accidents, has acquired a growing interest and it has become one of the most important research fields in the transport security. Indeed, DAS objective is to contribute in traffic accident reduction by using new technologies; this is, increasing the vehicles security, and at the same time, decreasing the danger situations that may be generated during driving process.

Current research is interested in the study of driver's state behavior; in this ambitious research, it has taken relevance the driver's drowsiness study, also denominated fatigue and related closely with distraction. Drowsiness is presented in stress and fatigue situations in an unexpected and inopportune way. The dream sensation generates the decrease vigilance level state, and this factor produces danger situations and increases the probability of causing some accident. Drowsiness may also be produced by dream's illnesses, certain type of medications, and even, bored situations, such as driving for a long time. It has been estimated that drowsiness produces among 10% and 20% of traffic accidents with dead drivers (Tian and Qin, 2005) and hurt drivers (Dong and Wu, 2005). Whereas trucking industry produces 57% of fatal truck accidents for this fatality (Ji and Yang, 2002; Bergasa et al., 2004). Fletcher (Fletcher et al., 2003) goes further on and has mentioned that 30% of total traffic accidents have been produced by drowsiness. For these reasons, it is important to design systems that allow monitoring the drivers and measuring their level of attention during whole driving process. Fortunately, people in drowsiness produce several typical visual cues that are detected on the human face: yawn frequency, eye-blinking frequency, eye-gaze movement, head movement and facial expressions. Taking advantage of these visual characteristics; computer vision is the feasible and appropriate technology to treat this problem.

The organization of the paper is as follows. Section 2 presents an extended state of the art. Section 3 introduces the proposed method for face location and eye detection in detail. Finally, in section 4 results and conclusions are shown.

2 PREVIOUS WORK

Ji and Yang (2002) has presented a detection drowsiness system based on infrared light illumination and stereo vision. This system localizes the eye position using image differences based on the bright pupil effect. Afterwards, this system computes the blind eyelid frequency and eye gaze to build two drowsiness indices: PERCLOS and AECS. Bergasa and his colleagues (Bergasa et al., 2004) has developed a non-intrusive system that also uses infrared light illumination, this system computes driver vigilance level using a finite state automata with six eye states that computes several indices, among them, PERCLOS; on the other hand, the system is able to detect inattention through face pose. Horng et al. (2004) has shown a system that uses a skin color model over HSI space for face detection, edge information for eye localization and dynamical template matching for eye tracking. Using color information of eyeballs, it identifies the eye state and computes the driver's state. Brandt et al. (2004) has shown a system that monitors the driver fatigue and inattention. For this task, he has used VJ method to detect the driver's face. Using the optical flow algorithm over eyes and head this system is able to compute the driver state. Tian and Qin (2005) have built a system for verifying the driver's eye state. Their system uses Cb and Cr components of the YCbCr color space; with vertical projection function this system localizes the face region and with horizontal projection function it localizes the eye region. Once the eyes are localized the system computes eye state using a complexity function. Dong and Wu (2005) have presented a system for driver fatigue detection, which uses a skin color model based on bivariate Normal distribution and Cb and Cr components of the YCbCr color space. After localizing the eyes, it computes the fatigue index utilizing the eyelid distance to classify between open eyes and closed eyes.

(a) (b)

Figure 1: (a) IvvI vehicle, (b) Driver's camera.

and near real-time with 90% of correct detection. A further important aspect of this method is its robustness under changing light conditions. However, in spite of the above-mentioned, its principal disadvantage is that it can not extrapolate and does not work appropriately when the face is not in front of the camera axis. Such would be the case when the driver moves his/her head; however, this shortcoming will be analyzed later on.

(a) (b) (c) (d)

Figure 2: Face and eye detection on different drivers.

3 PROPOSED SYSTEM

In this paper, a system to detect the driver's drowsiness is presented; it works on grayscale images taken with the camera inside the IvvI (Intelligent Vehicle based on Visual Information) vehicle, Figure 1 (b). IvvI is an experimental platform used to develop driver assistance systems in real driver conditions. IvvI is a Renault Twingo vehicle, Figure 1 (a), equipped with a processing system which processes the information comes from the cameras. This system consists of several parts that will be described throughout this section.

3.1 Face Detection

To localize the face, this system uses VJ object detector which is a machine learning approach for visual object detection. It uses three important aspects to make an efficient object detector based on the integral image, AdaBoost technique and cascade classifier (Viola and Jones, 2001). Each one of these elements is important to process the images efficiently

Continuing with the algorithm description, when driver's face is detected, it is enclosed with a rectangle R which is addresses by left-top corner coordinates $P_0 = (x_0, y_0)$ and right-bottom corner coordinates $P_1 = (x_1, y_1)$, as can be observed in Figure 2 (a), (c). Indeed, rectangle size comes from experimental analysis developed on the face database that has been created for this task.

3.2 Face Tracking

The principal problem of VJ method is that it is only able to localize the human face when it is in frontal position at camera. This drawback induces to have an unreliable system to driver's analysis during all driving process. Much effort has been put to correct

this problem; so, using a dual active contour (Gun and Nixon, 1994; Dokladal et al., 2004) is able to solve this disadvantage and to track the face in the driving process appropriately.
This face tracker needs to be initialized to extract an approximation of the head boundary. This is done once the face has been located through the rectangle R of the previous section, the system automatically generates an internal and external ring around the face based on gradient information, continuing with the calculation of the energy of the two active contours, after that, it corrects the position that corresponds to the face contour model.

3.3 Eye Detection

Once the face has been located through the rectangle R in previous section, using the face anthropometric properties (Gejgus and Sparka, 2003) which come from face database analysis, two rectangles containing the eyes are obtained. Preliminary, we use R_L for left eye rectangle and R_R for right eye rectangle. The rectangles coordinates are presented in Table 1 and Figure 3 (a) shows some examples.

Table 1: Preliminary rectangles that contain the eyes.

	Left eye
Left top corner	$(u_{0L}, v_{0L}) = (x_0 + w/6, y_0 + h/4)$
Right bottom corner	$(u_{1L}, v_{1L}) = (x_0 + w/2, y_0 + h/2)$
	Right eye
Left top corner	$(u_{0R}, v_{0R}) = (x_0 + w/2, y_0 + h/4)$
Right bottom corner	$(u_{1R}, v_{1R}) = (x_1 - w/6, y_1 - h/2)$

where $w = x_1 - x_0$ and $h = y_1 - y_0$.
After the previous step; the exact position of each eye will be localized, incorporating information from grey-level pixels through the following algorithm:

- Generate the image J by means of the following equation:

$$J(x,y) = \frac{I(x,y) - m}{\sigma} \quad (1)$$

  where m and σ are the mean and the standard deviation, respectively. They are computed over the eye rectangles described in Table 1, and $I(x,y)$ is the pixel value in the position (x,y).
- Generate the image K using the equation:

$$K(x,y) = \begin{cases} J(x,y) - 256\delta_1 & if \ \ J(x,y) \geq 0 \\ 256\delta_2 + J(x,y) & if \ \ J(x,y) < 0 \end{cases} \quad (2)$$

  where $\delta_1 = \max(0, ceil(J(x,y)/256) - 1)$, $\delta_2 = \max(1, ceil(|J(x,y)|/256))$ and $ceil(x)$ is the function that returns the smallest integer larger than x.
- Obtain the binary image , B , from image K through the equation (3), namely,

$$B(x,y) = \begin{cases} 255 & if \quad K(x,y) \geq \kappa \\ 0 & other \quad case \end{cases} \quad (3)$$

  where κ is computed by Ostu's method (Otsu, 1979), Figure 3 (b).
- Compute the gradient image, G, using the Sobel horizontal and vertical edge operator followed by an image contrast enhancement (Jafar and Ying, 2007), Figure 3 (d).
- Compute the logarithm image, L, with the objective to enhance the iris pixels that are the central part of the eye (Wu et al., 2004), Figure 3 (e).

All previous information produces a random sample that comes from a distribution function that it has an elliptic shape; i.e., the pixels coming from each eye through the images B, G and L can be viewed as a realization of a random variable. Having specified all the data describing the model, to obtain the parameters of this function the expectation maximization algorithm (EM) has been used. Special attention has received the ellipse center, because, it allows to obtain the exact position of the eye center. The ellipse axes determine the width and height of the eyes. The result is shown in Figure 3 (c), (f), while in Figure 2 (b), (d) the eye position generated for this procedure is depicted. The expectation maximization algorithm computes the mean, variance and the correlation of X and Y coordinates that belong to the eye. The initial parameters to run EM are obtained from a regression model adjusted with the last square method. These parameters will be used in the eye state analysis below.

3.4 Eye Tracking

There are a number of reasons for tracking. One is the VJ's problems mentioned above. Another is the necessity to track the eyes continuously from frame to frame. A third reason is to satisfy the real-time conditions reducing the eye search space. For this task; the Condensation algorithm that was proposed

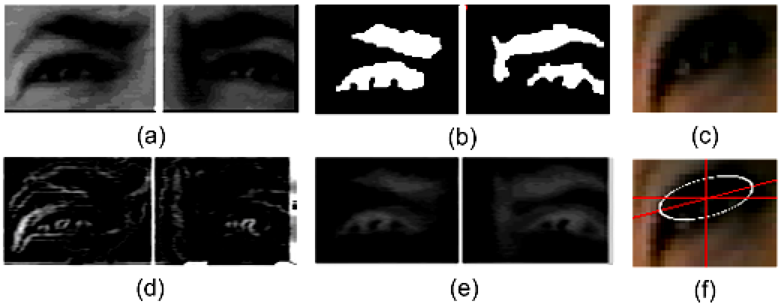

Figure 3: Eye location through R_L and R_R and Expectation Maximization algorithm over a spatial distribution of the eye pixels: (a) grayscale image, (b) binary image B, (d) gradient image G, (e) logarithm image L, (c) right eye image, and (f) ellipse parameters: center position, axes and inclination angle.

by Isard and Blake (1998) for tracking active contours using a stochastic approach has been used. The Condensation algorithm combines factored sampling with a dynamical model that is governed through the state equation:

$$X_t = f(X_{t-1}, \xi_{t-1}) \quad (4)$$

where X_t is the state at instant t, $f(\cdot)$ is an nonlinear equation and depends on a previous state plus a white noise. The goal is to estimate the state vector X_t with the help of systems observation which are realization of the stochastic process Z_t governed by the measurement equation:

$$Z_t = h(X_t, \eta_t) \quad (5)$$

where Z_t is the measure system at time t, $h(\cdot)$ is another nonlinear equation that links the present state plus a white noise. The processes ξ_t and η_t are each one white noise and independent among them. It must be pointed out that X_t is an unobservable underlying stochastic process, for this problem, it is eye position over the image and its velocity:

$$X_t = (x, y, \dot{x}, \dot{y})^T \quad (6)$$

The Condensation is initialized when the eyes are localized with the method of previously explained. Table 1 shows the eye tracking results that has been developed in two sequences of images.

3.5 Eye State Detection

To identify drowsiness through eye analysis is necessary to know its state (open or closed) and develop an analysis over the time. The classification among the open and closed state is complex due to the changing shape of the eye, among other factors, changing position and face rotating, twinkling and illumination variations. All this makes difficult to use only color cues to analyze eye in a reliable manner. For the problems that have been exposed a supervised classification method has been used for this challenging task, in this case, support vector machine (SVM) classification (Cristianini and Shawe-Taylor, 2000; Chang and Lin, 2001) which is rooted in statistical learning theory. SVM uses a training set, $S = \{(x_i, y_i) : i = 1, \cdots, m\}$, where x_i is the characteristic vector in R^n, $y_i \in \{1,2\}$ represents the class, in this case 1 for open eyes and 2 for closed eyes, and m is the number of elements of S. To do this work a training set has been built that consists of images of open eyes and images of closed eyes. The images come from diverse sources, under several illumination conditions and different races. A further important aspect of this eye database is that contains images of different eye colors, i.e., blue, black, green. Previous to SVM training, it is indispensable to process each image that consists on histogram equalization, filter with the median filter, followed by the sharpen filter and to normalize in the $[0,1]$ interval. The median filter is used to reduce the image noise, whereas the sharpen filter is used to enhance the borders. The main objective of training SVM is to find the best parameters and the best kernel that minimizes the optimization problem (Chang and Lin, 2001), so, after several training experiments of the SVM algorithm, it has decided to use the RBF kernel, i.e., $K(x_i, x_j)$ is $\exp(-\gamma \|x_i - x_j\|^2)$, $C = 35$ and $\gamma = 0.0128$; these parameters reach high training classification rate that is about 94%.

Table 2: Result of eye tracking and eye state analysis.

	Total frames	Eye tracking		Eye state analysis		
		Tracking failure	Correct rate (%)	Eyes Open	Eyes Closed	Correct rate (%)
Video1	960	20	97.91	690/700	258/260	98.90
Video2	900	30	96.60	520/560	339/340	96.27

Figure 4: Different stage of the proposed algorithm on several instants of time and driving conditions.

3.6 Drowsiness Index

The eye-blinking frequency is an indicator that allows to measure driver's drowsiness (fatigue) level. As in the works of Horng et. al. (2004) and Dong and Wu (2005), if five consecutive frames or during 0.25 seconds are identified as eye-closed the system is able to issue an alarm cue. Table 2 also presents the result of eye state analysis over two sequences of images.

4 CONCLUSIONS

A non-intrusive driver's drowsiness system based on computer vision has been presented in this paper. This system uses visual information to analyze and to monitor driver's eye state at near real-time and real-driving conditions, i.e., external illuminations interference, vibrations, changing background and facial orientations changing. Experiments were carried out in the IvvI vehicle with different drivers. This guarantees and confirms that these experiments have proven robustness and efficiency in real traffic scenes. Another drowsiness indexes will be implemented as future works and they will be compared. Figure 4 shows an example that validates this system.

REFERENCES

Viola P. and Jones M., 2001: Rapid Object Detection using a Boosted Cascade of Simple Features. *Accepted Conference on Computer Vision and Pattern Recognition*.

Horng W., Chen C. and Chang Y., 2004: Driver Fatigue Detection Based on Eye Tracking and Dynamic Template Matching. *Proceedings of the IEEE International Conference on Networking, Sensing & Control*.

Tian Z. and Qin. H., 2005: Real-time Driver's Eye State Detection. *IEEE International Conference on Vehicular Electronics and Safety, Pg. 285-289.*

Ji Q. and Yang. X., 2002: Real-Time Eye, Gaze, and Face Pose Tracking for Monitoring Driver Vigilance. *Real Time Imaging, Nr. 8, Pg. 357-377, Elsevier Science Ltd.*

Bergasa L., Nuevo J., Sotelo M. and Vazquez M., 2004: Real Time System for Monitoring Driver Vigilance. *IEEE Intelligent Vehicles Symposium.*

Isard M. and Blake A., 1998: Condensation: conditional density propagation for visual tracking. *International Journal on Computer Vision, 29(1), pp. 5-28.*

Cristianini N. and Shawe-Taylor J., 2000: An introduction to Support Vector Machines and other kernel-based learning methods. *Cambridge University Press.*

Chang C. and Lin C., 2001: LIBSVM: a library for support vector machine, URL: www.csie.ntu.edu.tw /~cjlin/libsvm

Otsu N., 1979: A threshold selection method from gray-level histograms. *IEEE Trans. Systems, Man and Cybernetics, Vol. 9, pp. 62-66.*

Gejgus P. and Sparka M., 2003: Face Tracking in Color Video Sequences. *The Association for Computing Machinery Inc.*

Brandt T., Stemmer R., Mertsching B., and Rakotomirainy A., 2004: Affordable Visual Driver Monitoring System for Fatigue and Monotony. *IEEE International Conference on Systems, Man and Cybernetics. Vol. 7, pp. 6451-6456*.

Fletcher L., Petersson L. and Zelinsky A., 2003: Driver Assistance Systems based on Vision In and Out of Vehicles. *IEEE Proceedings of Intelligent Vehicles Symposium, pp. 322-327.*

Wu Y., Liu H. and Zha H., 2004: A New Method of Detecting Human Eyelids Based on Deformable Templates. *IEEE International Conference on Systems, Man and Cybernetics, pp. 604-609*.

Jafar I., Ying H., 2007: A new method for Image Contrast Enhancement Based on Automatic Specification of Local Histograms. *IJCSNS International Journal of Computer Science and Network Security, Vol.7 No.7, July*.

Dong W. Wu X., 2005: Driver Fatigue Detection Based on the Distance of Eyelid. *IEEE Int. Workshop VLSI Design & Video Tech. Suzhou-China*.

Gunn S. R. and Nixon M. S., 1994: A Dual Active Contour for Head Boundary Extraction. *IEEE Colloquium on Image Processing for Biometric Measurement, pp. 6/1 - 6/4, London.*

Dokladal P., Enficiaud R. and Dejnozkova E., 2004: Contour-Based Object Tracking with Gradient-Based Contour Attraction Field. *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP'04), vol. 3, pp. iii-17-20*.