2010 Fourth IEEE International Conference on Self-Adaptive and Self-Organizing Systems Workshop

Driver drowsiness identification by means of passive techniques for eye detection and tracking

Andrea Cristiani, Marco Porta, Davide Gandini, Gian Mario Bertolotti
Dip. Informatica e Sistemistica
Università di Pavia
Pavia, Italy
e-mail: andrea.cristiani@unipv.it

Nikola Serbedzija
Fraunhofer FIRST
Berlin, Germany
e-mail: nikola@first.fraunhofer.de

Abstract—**The aim of this paper is to describe a system whose final goal is to detect if a driver is drowsy, in order to prevent potentially danger situations. The system is based on the processing of the driver's face image, acquired by a webcam installed on the dashboard of the car. After a brief introduction explaining the connection of the present work to the European project REFLECT, the relashonship between drowsiness condition and fatal car crashes is dicussed. Then, an overview of the most used techiques for face and eye detection is given, and the developed algorithm is described in detail. Finally, preliminary results of in-laboratory and in-car tests are presented and commented.**

Keywords-eye blink detection; drowsiness; driving experience; face tracking

I. INTRODUCTION

This paper describes a prototype of webcamera-based system aimed at detecting and tracking the eye of a driver. The final goal is to extract parameters (e.g. eye blinks durations) which are useful to identify drowsiness and alert the pilot in order to prevent potentially danger situations.

The presented system is part of the work carried on within Workpackage 5 of the REFLECT project [1], which is funded by the EU Seventh Framework Programme, ICT Thematic Priority, Challenge 8: Future and Emerging Technologies, Objective ICT-2007.8.2: Pervasive adaptation. Actually, the main objective of WP5 is the development of a system which is intended to identify and, if possible, to enhance the physical state of end users. The core of this "Physical Loop" is the "seat adaptation system" (not presented here), which aims at improving sitting comfort by detecting subject's postures and acting on the seat accordingly. A vehicle environment is the preferred scenario where the system will be implemented. In particular, a top level car (F149 California by Ferrari, Italy) will be used as a vehicle-demonstrator of the overall project.

As the Physical Loop deals with physical conditions, it is reasonable, and important, to take also into account possible dangerous postures and, more generally, physical states (e.g. drowsiness), especially if the subject is driving a vehicle.

Many studies and official statistics highlight how the high level of fatigue can decrease the driver's safety and cause fatal or serious crashes. This is confirmed by studies performed by U.S.A. National Highway Traffic Safety Administration, according to which about 100000 accidents are caused by drivers falling asleep. Also European studies showed that drowsiness is the cause of one-third of fatal accidents [2-4]. Often, these crashes involve only one vehicle, during long trips, and they occur outside town: this could depend on the fact that monotony environments could quickly decrease the driver's attention and therefore increase the risk of falling asleep.

Driver's attention can also decrease due to events that occur inside the cockpit, such as the sudden ring of a cell phone, an object falling under the seat, and so forth [5,6].

Looking at these statistics, it seems quite clear how serious the hazard for the driver is when he is driving with a diminished level of vigilance.

It appears obvious that it would be very useful, in terms of driver's safety, to have an active system inside the cockpit able to monitor the vigilance level of the driver and, if needed, to give him or her an alarm signal. In the last years many systems have been developed to accomplish this task, based on different kinds of techniques, but many of these techniques are intrusive, since they involve the use of electrodes or sensors placed on driver's body. In order to avoid these limitations, systems have been developed which use image analysis to perform measures of frequency and total duration of prolonged blinks. Numerous results demonstrate that the measurement of eye blink parameters provides reliable information about drowsiness. [7,8]

II. BACKGROUND

To check the potential sleepiness state of the driver it is necessary to analyze his or her eyes, to assess their openness degree.

In most cases, eye detection is preceded by a face identification stage, in which the acquired frames are examined to find the user's face.

Main passive methods for face detection include color segmentation, image differencing and machine learning.

Color segmentation looks for skin-like areas in the image, but is likely to fail in presence of complex

978-0-7695-4229-4/10 $26.00 © 2010 IEEE
DOI 10.1109/SASOW.2010.30

backgrounds. Similarly, image differencing methods, which try to locate the face on the basis of head movements, are hardly applicable with dynamic backgrounds or when there are frequent changes in illumination. Machine learning algorithms are usually a better choice, due to both their robustness and precision in different background and illumination conditions. Among these methods, the Viola-Jones technique [9] (partially relying on the Haar basis functions [10]) is certainly one of the most exploited. An implementation of this face recognition algorithm is available under open source license from the OpenCV project [11].

Once the face region has been localized, it is necessary to identify eye areas within it.

Like for face detection, the difference between consecutive frames in a video stream can be used to find the eyes through their movements, mainly blinks (e.g. [12]). While simple, this technique cannot however produce reliable results in our context, because correct eye detection may be misled by large head shifts or sudden changes in lighting. Anthropometric methods (e.g. [13]), which exploit "rules" about positions and proportions of facial traits to infer the most probable eye area in the face region, can be of some help as well (although they usually tolerate only small head rotations). The eyes can be also detected through wavelet transforms, for instance Gabor wavelets [14]. Haar-like features, an extension of Haar wavelets, are often used to build classifiers able to codify eye orientation, shape and size [9,15].

Edge detection is another possible approach to eye recognition in digital videos (e.g. [16]). However, the potential noise introduced by eye elements such as folds in the skin, eyelashes and wrinkles, along with the need for fine tuning of some thresholds, make this technique little suitable for our application context.

A further method for eye detection considers projections of intensity distributions within face images. A vertical or horizontal integral projection is calculated by summing the intensities of pixels on each column or row of the image (usually after an edge detection process). Since eye regions correspond to areas of high spatial frequency, the peak positions in the projections identify the eye zones in the face image [17].

Deformable parametric models are another category of eye detection techniques. Essentially, a model is used to describe the eye and its features, allowing to discriminate among different candidates. Examples of simple models include shapes such as circles, ellipses and parabolas. While generally computationally challenging, the main advantage of deformable models is that they can adapt to the specific size and position of the eye. Among the main techniques pertaining to the category, it is worth citing the following.

1. Snakes (also called Active Contours) [18]: use deformable splines to model shapes.
2. Active Shape Models (ASM) [19]: can be considered an evolution of snakes and exploit a priori information (obtained from sample images) about the shape to be searched for.
3. Active Appearance Models (AAM): consider both the geometric shape and the appearance (texture) of the elements to be recognized. These models are built through statistical analyses of sets of samples, and can be also applied in real-time (the Inverse Compositional Algorithm [20], or ICA, is one of the most effective methods to this purpose). Statistical models are usually created using the Principal Component Analysis (PCA) technique, a powerful tool for data examination and compression. In general, the geometrical shape of an object is represented by the coordinates of landmark points placed on it, while the texture is typically given by the intensity of pixels within the shape.

III. ALGORITHM DESCRIPTION

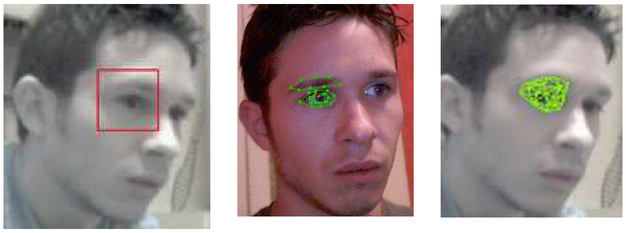
Initially, the driver's face is identified within the acquired frames by means of the Viola-Jones algorithm [9]. False positives are avoided by selecting, among possible multiple face candidates, the one whose size and position match the typical location and dimension of the driver's head.

Once the rectangular area containing the driver's face has been identified, it becomes the region of interest in which to look for the eyes. Because of the position of the webcam (which is not exactly in front of the driver but placed at his or her right, as shown in Fig. 1), and to reduce the algorithm computational load, only the right eye is tracked. Like for the face, the eye area is initially broadly identified using the Viola-Jones technique.

Figure 1. Webcam position inside the cockpit.

The result of this process is a rectangular region enclosing the eye, as shown in Fig. 2a. Human morphology is also exploited, as the eyes are usually located, vertically, in a region occupying one third of the face and placed slightly below the upper limit of the head.

To precisely identify the eye and its features within the found rectangular region, in a first implementation of the project we employed an AAM eye detection method. After creating a suitable training set of face images to build the model, each picture was manually annotated by placing proper points (landmarks) on the eye and eyebrow (Fig. 2b). The ICA algorithm was then used to build the best-fitting active appearance model (Fig. 2c), after an initialization of the model parameters according to the position and size of the just found eye region.



(a) (b) (c)

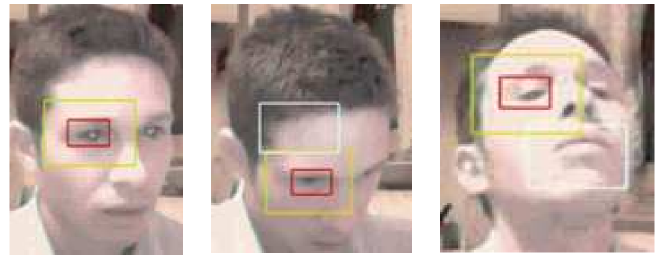
Figure 2. (a) Identification of the eye region through the Viola-Jones technique; (b) Manual annotation of eye and eyebrow through landmark points (AAM technique); (c) Eye tracking result using the ICA method

While effective, however, such an implementation turned out to be little efficient for our purposes - on average, only seven frames per second could be processed, a rate which is unacceptable for reliable eye-blink detection. We therefore opted for a simpler and faster (but still effective) algorithm, which is practically a combination of the Viola-Jones technique and a template matching approach (a normalized squared difference matching method) . To reduce the computational load, the algorithm is applied to an area A centered on the eye position found in the previous frame and double-size with respect to the eye rectangle (Fig. 3a). The algorithm can be briefly summarized as follows:

1. The eye is precisely searched within A using the Viola-Jones technique;
2. If the eye is found, a copy of it is saved as a template;
3. If the eye is not found, the last valid template saved is used in a template-matching procedure;
4. After a certain number of consecutive template-matching procedures (i.e. successive failures of the Viola-Jones method), a re-initialization process (face detection, etc.) is carried out, to avoid a potential general failure of the tracking due to a sequence of wrong matchings.

The peculiarity of this algorithm is that it exploits the good qualities of both the Viola-Jones and template-matching techniques, to reduce the error rate: when, for any reason, the Viola-Jones technique fails, the tracking continues with the template-matching method, which can rely on a constantly updated template.

The analysis of the eye status should be carried out only if the driver's head is about still and directed forward, since a moving or turned head are usually, as such, signs of a wakefulness state. The eye test is thus performed only if the face is detected in a "rest" position. To identify such posture, the average eye location over the last n frames is calculated. This position becomes the center of an area used to test both the eye status and the inclination of the head: the vertical position of the eye within such area indicates how much the driver's head is vertically tilted, thus allowing the identification of potentially danger situations (Fig. 3b and Fig. 3c).



(a) (b) (c)

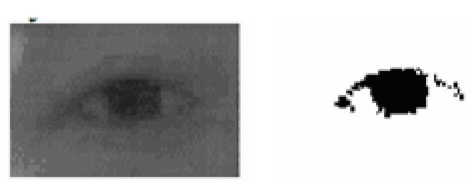
Figure 3. (a) Inner rectangle: eye area (found in the previous frame). Outer rectangle: eye search area in the current frame; (b) Inclination level of the driver's head: downward tilting; (c) Inclination level of the driver's head: upward tilting

If the eye is within the rest area, the blink test can be carried out. To this purpose, the eye image is first converted to grayscale and then thresholded to obtain a binary picture (Fig. 4).

Since the binary picture (especially with infrared lighting) is often characterized by reflections on the pupil, which may compromise the study of the openness degree, a further processing is necessary. After an edge detection procedure, the closed boundary with maximum area is selected and filled, in order to eliminate the gaps generated by pupil reflections. Through this process, also the noise due to shadows and/or portions of eyebrows can be usually reduced.

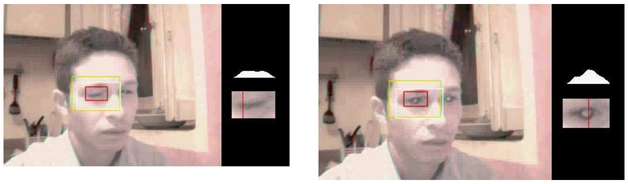
Subsequently, the vertical projection histogram of the binary image is obtained. To remove micro peaks generated by eyelash edges, a histogram smoothing is performed by calculating, for each column, the average value over the previous ten. The peak in the obtained histogram corresponds to the position of the pupil. The eye openness level can then be assessed by calculating the average value over all the columns and comparing it with that relative to the "fully-open eye" case (Fig. 5). Moreover, the peak position provides also a broad indication about the driver's horizontal gaze direction.

When the eye remains totally or partially closed for more than a defined timeout, the driver may have fallen asleep, and proper actions must be taken.



(a) (b)

Figure 4. (a) Gray-level eye image; (b) Binary eye image



(a) (b)

Figure 5. (a) Eye blink; (b) Open eye

IV. EXPERIMENTS

To test the performance of the developed eye detection and tracking algorithm, we carried out several tests in different environments (even if the natural setting is, of course, a car interior). In particular, the following three cases have been considered: (1) the Ferrari California car; (2) various cars; and a room (university laboratory) illuminated with fluorescent light.

For each one of the above reported contexts, the experiments have been performed in different light conditions, namely (1) daylight, (2) little lighting or zero lux, and (3) changing lighting.

Two USB web-cameras have been used, depending on light conditions: a Hercules Classic Silver (by Hercules) in condition (1), and a X-33850 Xtreme Night Vision Camera (by Xtreme) in conditions (2) and (3).

The first device is a traditional web camera, the second one is provided with 6 infra red LEDs and an infra red sensor, then it can acquire images even in low light conditions.

Maximum frame rate and resolution are 30 frame per second and 640 x 480 for both devices.

Each camera has been connected, through a USB port, to a Vostro 1310 notebook (by Dell), with the following main characteristics:

- Processor: Intel Core 2 Duo T8100 (2.1GHz)
- Memory: 2GB - 2 DIMM (DDR2-667)
- HDD: 160GB 5400RPM HDD
- Graphics: 128MB NVIDIA GeForce 8400M GS
- Operative System: Microsoft XP

Five sets of data have been collected, related to the correct recognition of slow/fast blinks and accurate tracking of slow/fast eye movements.

The following main kinds of errors were detected:

- tracking miss: when, in successive frames, the tracking does not occur correctly;
- detection miss: when the user's face or eye are not detected because of shadows or occlusions;
- blink miss: when the eye blink is not detected (especially with rapid blinks)
- blink fail: when a blink is erroneously detected (usually because of particular head positions)

Unfortunately, a limited number of tests could be carried out on the Ferrari (30 video clips lasting 20 seconds each). More experiments could instead be performed in the other two settings (100 video clips, lasting 20 seconds each). Each test was performed by the same subject, a 25 years old male.

Tables 1 to 4 in the following summarize success percentages for the recognition of slow and fast blinks and eye movements, while Table 5 reports fail percentages for blink detection. During the experimental sessions in settings 1 and 2, the tester was asked to "behave normally", driving the car. In setting 3, this behaviour could of course only be simulated (but also prolonged blinks could be tested, all correctly recognized).

Table 1: fast blinks

Lighting	*Ferrari*	*Various cars*	*Room*
Daylight	97.3	95.4	98.6
Little/no lighting	--	97.5	98.9
Changing light	95.9	94.5	--

Table 2: slow blinks

Lighting	*Ferrari*	*Various cars*	*Room*
Daylight	98.3	97.8	99.1
Little/no lighting	--	97.2	98.8
Changing light	97.4	96.4	--

Table 3: slow eye movements

Lighting	*Ferrari*	*Various cars*	*Room*
Daylight	99.0	98.6	99.7
Little/no lighting	--	96.4	97.8
Changing light	98.4	97.1	--

Table 4: fast eye movements

Lighting	*Ferrari*	*Various cars*	*Room*
Daylight	96.0	95.9	96.1
Little/no lighting	--	94.2	94.6
Changing light	94.8	95.1	--

Table 5: blink fail

Lighting	*Ferrari*	*Various cars*	*Room*
Daylight	1.3	1.4	0.9
Little/no lighting	--	1.7	1.3
Changing light	1.6	2.1	--

V. CONCLUSIONS

Although only one tester was involved in this preliminary experimental phase, the results obtained are quite comforting; we are currently planning further investigations with additional testers. We are also considering to improve system performance in low light condition, by adding an array of infrared LEDs to the night vision camera.

The proposed algorithm seems to be very promising in detecting the eyeblinks of a driver. Thanks to the combination of the Viola-Jones and template matching techniques - the main peculiarity of our system - the algorithm can robustly deal with different driving conditions. Our next efforts will be directed towards finding clear correlations between eyeblink data (such as duration and speed) and the vigilance state of the driver.

ACKNOWLEDGMENT

This work has been partly supported by the EU Seventh Framework Programme project REFLECT (Pr. No.: FP7-215893).

REFERENCES

[1] REFLECT project website. Available: http://reflect.pst.ifi.lmu.de/.

[2] F. Streff, and H. Spradlin, "Driver distraction, aggression, and fatigue: a synthesis of the literature and guidelines for Michigan planning," Ann Arbor, vol. 1001, p. 48109.

[3] A. C. (IST 2000-28062), "Awake-system for effective assessment of driver vigilance and warning according to traffic risk estimation," Sept. 2001-2004.

[4] L. Bergasa, J. Nuevo, M. Sotelo, R. Barea, and M. Lopez, "Real-time system for monitoring driver vigilance," IEEE Transactions on Intelligent Transportation Systems, vol. 7, no. 1, pp. 63–77, 2006.

[5] J. Stutts, J. Feaganes, D. Reinfurt, E. Rodgman, C. Hamlett, K. Gish, and L. Staplin, "Driver's exposure to distractions in their natural driving environment," Accident Analysis and Prevention, vol. 37, no. 6, pp. 1093–1101, 2005.

[6] T. Sheridan, "Driver distraction from a control theory perspective," Human factors, vol. 46, no. 4, p. 587, 2004.

[7] M. Ingre, T. ÅKerstedt, B. Peters, A. Anund, and G. Kecklund, "Subjective sleepiness, simulated driving performance and blink duration: examining individual differences," Journal of sleep research, vol. 15, no. 1, pp. 47–54, 2006.

[8] P. Caffier, U. Erdmann, and P. Ullsperger, "Experimental evaluation of eye-blink parameters as a drowsiness measure," European Journal of Applied Physiology, vol. 89, no. 3, pp. 319–325, 2003.

[9] P. Viola, and M. J. Jones, "Rapid Object Detection using a Boosted Cascade of Simple Features," Proc. IEEE CVPR, 2001.

[10] C. Papageorgiou, M. Oren, and T. Poggio, "A general framework for object detection," International Conference on Computer Vision, 1998.

[11] Open Computer Vision Library. Available: http://sourceforge.net/projects/opencvlibrary/.

[12] K. Grauman, M. Betke, J. Gips, and G. R. Bradski, "Communication via eye blinks: Detection and duration analysis in real time," Proc. 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2001.

[13] M. Zobel, A. Gebhard, D. Paulus, J. Denzler, and H. Niemann, "Robust facial feature localization by coupled features," Proc. 4th IEEE International Conference on Automatic Face and Gesture Recognition, 2000.

[14] S. A. Sirohey, and A. Rosenfeld, "Eye detection in a face image using linear and nonlinear filters," Pattern Recognition, Vol. 34, No. 7, 2001.

[15] R. Lienhart, and J. Maydt, "An extended set of Haar-like features for rapid object detection," Proc. International Conference on Image Processing, 2002.

[16] J. G. Wang, E. Sung, and R. Venkateswarlu, "Eye gaze estimation from a single image of one eye," Proc. 9th IEEE International Conference on Computer Vision, 2003.

[17] M. Eriksson, and N. P. Papanikolopoulos, "Driver fatigue: a vision-based approach to automatic diagnosis," Transportation Research, Part C: Emerging Technologies, Vol. 9, No. 6, 2001.

[18] M. Kass, and D. Terzopoulos, "Snakes: Active contour models," Proc. 1st International Conference on Computer Vision, 1987.

[19] T. Cootes, and C. Taylor, "Active shape models - smart snakes," Proc. British Machine Vision Conference, 1992.

[20] I. Matthews, and S. Baker, "Active appearance models revisited," International Journal of Computer Vision, 60(2), 2004.