2004 IEEE Intelligent Vehicles Symposium
University of Parma
Parma, Italy • June 14-17, 2004

Real-Time System for Monitoring Driver Vigilance

Luis M. Bergasa, Jesús Nuevo, Miguel A. Sotelo

Departamento de Electrónica
Universidad de Alcalá
Alcalá de Henares
28805 Madrid, Spain
email: bergasa@depeca.uah.es

Manuel Vázquez

Dpto. Sistemas Electrónicos y de Control
Universidad Politécnica de Madrid
28031 Madrid, Spain
email:mvazquez@sec.upm.es

Abstract

In this paper we present a non-intrusive prototype computer vision system for real-time monitoring driver's vigilance. It is based on a hardware system, for real time acquisition of driver's images using an active IR illuminator, and their software implementation for monitoring some visual behaviours that characterize a driver's level of vigilance. These are the eyelid movements and the pose face. The system has been tested with different sequences recorded on night and day driving conditions in a motorway and with different users. We show some experimental results and some conclusions about the performance of the system.

1. Introduction

The increasing number of traffic accidents in Europe due to a diminished driver's vigilance level has become a serious problem for society. Statistics show that between 10% and 20% of all the traffic accidents are due to drivers with a diminished vigilance level. In the trucking industry, about a 60% of fatal truck accidents are caused to driver fatigue [1]. For this reason, developing systems for monitoring a driver's level of vigilance and alerting the driver, when he is not paying adequate attention to the road, is essential to prevent accidents.

In the last few years many researchers have been working on the development of these kinds of systems using different techniques. The most accurate techniques are based on physiological measures like brain waves, heart rate, pulse rate, respiration, etc. These techniques are intrusive, since they need to attach some electrodes on the drivers, causing annoyance to them. Some representative projects in this line are the MIT Smart Car [2], and the ASV (Advanced Safety Vehicle) project performed by Toyota, Nissan and Honda [3]. Others techniques monitor eyes and gaze movement using a helmet or special contact lens. These, though less intrusive, are still not acceptable in the practice [4].

A driver's state of vigilance can also be characterized by indirect behaviours of the vehicle like lateral position, steering wheel movements, time to line crossing, etc. Although these techniques are not intrusive they are subjected to several limitations as the vehicle type, driver experience, geometric characteristics and state of the road, etc. On the other hand, this takes too much time to analyse user behaviours and thereby it doesn't work with the known as micro-sleeps [5]. In this line we can find an important Spanish project called TCD (Tech Co Driver) [6] and the Mitsubishi advanced safety vehicle system [3].

People in fatigue show some visual behaviours easily observable from changes in their facial features like eyes, head and face. Computer vision can be a natural and non-intrusive technique to monitor driver's vigilance. Many researches have been reported in the literature on developing image-based driver alertness using this technique. Several of them have worked on head tracking. In [7], [8] two methods are presented based on 3D stereo matching. In [9] we can find a system based on templates matching with one camera. All these systems rely of manual initialization. A successful head/eye monitoring and tracking of drivers to detect drowsiness by use of one camera and based on colour predicates is presented in [10]. The described systems are based on passive vision techniques and their functioning can be problematical in poor and very bright light conditions. Moreover, they don't work at nights, where the monitoring is more important.

In order to work at nights some researches use active illumination based on infrared LED [11], [12]. Almost all the active systems that have been reported in the literature have been tested in laboratories but not in real vehicles moving. A moving vehicle presents new challenges like variable lighting, changing background and vibrations

0-7803-8310-9/04/$20.00 © 2004 IEEE

that must be haven in mind in real systems. In [13] an industrial prototype called *Copilot* is presented. This system has been tested with truck's drivers. It uses a simple subtraction process for finding the eyes and it only calculates PERCLOS (percent eye closure), in order to measure driver's drowsiness.

Systems relying on a single visual cue may encounter difficulty when the required visual features cannot be acquired accurately or reliably, as it occurs in real conditions. Then, a single visual cue may not always indicative of one's mental conditions [12]. The use of multiple visual cues reduces the uncertainty and the ambiguity present in the information from a single source. In this line, currently, it is been developing an ambitious European project called AWAKE [1]. At the moment only some partial results have been presented.

This paper describes a real-time prototype computer vision system for monitoring driver vigilance using infrared illumination. As we have justify before, we considerer that using active illumination is the best option because our goal is to monitor the driver 24 hours on real conditions (vehicle moving) and in a very robust and accurate way. As a difference respecting to other previous systems we propose monitor several visual behaviours that typically characterize a person's level of alertness while driving. Moreover, we have tested our system in a car moving in a motorway and with different users.

The rest of the paper is structured as follows. In section 2 we present the general system architecture, explaining its main parts. Experimental results are presented in section 3. Finally, in section 4 we show the conclusions and the future works.

2. System Architecture

The general architecture of our system is shown in figure 1. It consists of four major parts: 1) Image acquisition, 2) Pupil detection and tracking, 3) Visual behaviours and 4) Driver vigilance. As image acquisition system we use a CCD micro-camera, a commercial frame-grabber and an IR illuminator controlled by the acquisition system. The pupil detection and tracking stage stars with pupil detection based on the bright pupil effect, similar to the red eye effect in photography. Then, we use two Kalman filters in order to track the pupils robustly in real-time. In the visual behaviours stage we use some visual cues as pupil movements and face pose in order to detect some visual behaviours easily observable in people in fatigue as: slow eyelid movement, smaller degree of eye openness, frequent nodding, blink frequency, face pose, etc. Finally, in the driver vigilance stage we fusion all the individual behaviours obtained in the before stage using a Fuzzy system. This way we can conclude if the driver is fatigued and active a warning.

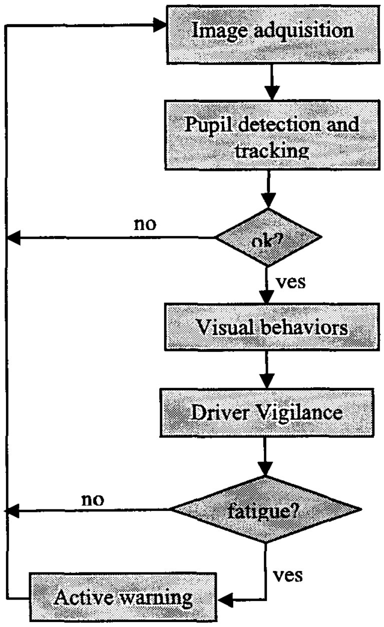


***Figure 1**. General architecture*

2.1. Image Acquisition System

The purpose of this stage is to acquire the image of the driver face in real-time. The acquired images should be relatively invariant to light conditions (day and night) and should facilitate the eye detection and tracking. The use of infrared IR illuminator serves these goals. First, it minimizes the impact of changes in the ambient light, and second, it produces the bright pupil effect, which constitutes the foundation of our detection and tracking system. A bright pupil can be obtained if the eyes are illuminated with an IR illuminator beaming light along the camera optical axis. At the IR wavelength, pupils reflects almost all IR light they receive along the path back to the camera, a bright pupil effect will be produced in the image. If illuminated off the camera optical axis, the pupils appear dark since the reflected light will not enter the camera lens. An example of the bright/dark pupil effect can be seen in figure 3. This pupil effect is clear with and without glasses, contact lenses and it even works to certain degree with sunglasses.

Figure 2 shows the image acquisition system configuration. It is composed by a miniature CCD camera sensitive to the near-infrared and located on the dashboard of the vehicle. This camera focuses on the driver's head for detecting the multiple visual cues. The IR illuminator is composed by two sets of IR LEDs distributed symmetrically along two rings as shown in figure 2. An embedded PC with a low cost frame-grabber is used for video signal acquisition and signal processing.

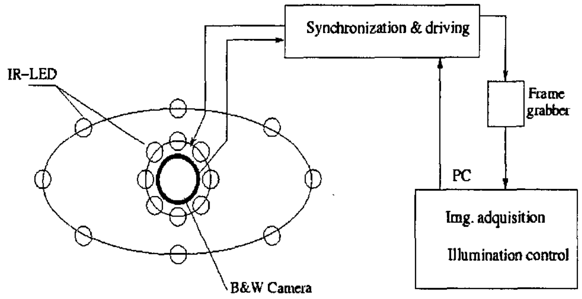


***Figure 2**. Block diagram of the prototype*

The size of the rings has been calculated empirically in order to obtain a dark pupil image if the outer ring is turned on and a bright pupil image if the inner ring is turned on. The inner ring configuration obtains the bright pupil effect because the centre of the ring coincides with the camera optical axis, actuating as it was only a LED locate along the optical axis of the lens. For minimizing interference from light sources beyond the IR light emitted by our LEDs, a narrow bandpass filter has been attached between the CCD and the lens.

2.2. Pupil Detection and Tracking

This stage starts with pupil detection. We have designed a specific hardware that detects from each interlaced image frame (camera output) the even and odd field signal starts, which are then used to alternately turn the outer and inner IR rings on to produce the dark and bright pupil image fields. Each frame is separated into two images fields (even and odd), representing the bright and dark pupil images separately. The even image field is then digitally subtracted from the odd image field to produce the difference image. In this image pupils appear as the brightest parts in the image as can be seen in figure 3. This system minimizes the ambient light illumination influence because this is subtracted in the generation of the difference images.

Pupils are detected on the difference image, by searching the entire image lo locate two bright blobs that satisfy certain size, position and distance constraints. The image is binarized, using a priori threshold for detecting the brighter blobs in the image. The blobs that are out of some size constraints are removed and for the rest we take all the possible pairs between two blobs and we applying some constraints respecting their size and distance in order to detect the most probably pair of be pupils. The centroids of the blob are returned as the position of the detected pupils.

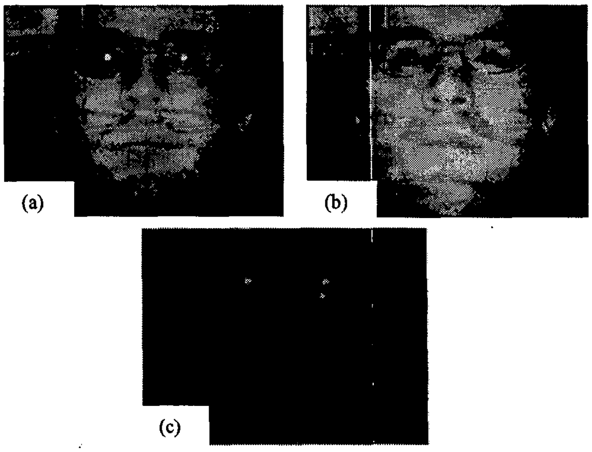
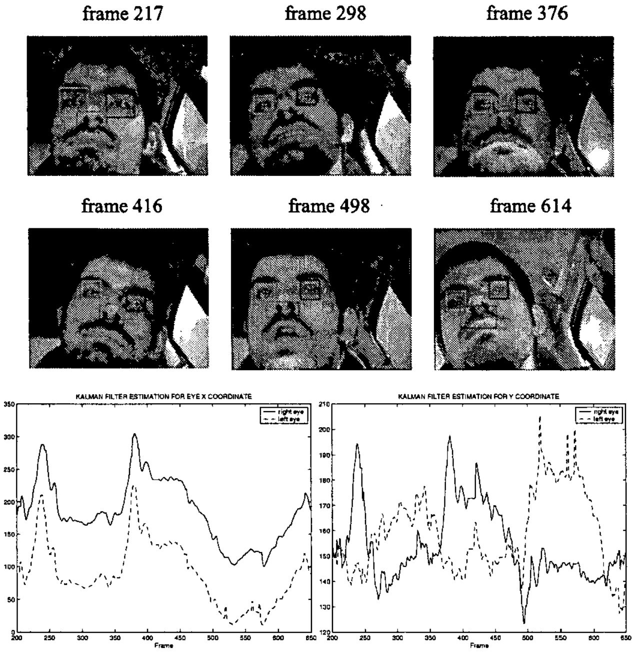


***Figure 3**.(a) Image obtained with internal IR, (b) Image obtained with external IR, (c) Difference Image*

To continuously monitor the driver it is important to track his pupils from frame to frame. This can be done efficiently by using two Kalman filters, one for each pupil, in order to predict pupil positions in the image. We have used a pupil tracker based on [12] but we have tested it with images obtained from a car moving in a motorway. Then, in our case the search window is determined automatically based on pupil size and location error. The state vector of the filter is represented as $\mathbf{x}_t=(\mathbf{c}_t, \mathbf{r}_t, \mathbf{u}_t, \mathbf{v}_t)^t$, where $(\mathbf{c}_t,\mathbf{r}_t)$ indicate the pupil pixel position (its centroid) at time t and $(\mathbf{u}_t, \mathbf{v}_t)$ be its velocity at time t in c and r directions respectively.

Figure 4 shows a sequence of the pupil tracker working. Rectangles on the images indicate the search window of the filter and crosses indicate the locations of the detected pupils. (a) and (b) draw estimation of the pupil positions for the sequence under test. The tracker is found to be rather robust under different users without glasses, light conditions, face orientations and distances in a car. Performance of the tracker gets worse when users wear eyeglasses because different bright blobs appear in the image due to IR reflections in the glasses. It automatically finds and tracks the pupils and can recover from tracking-failures. The system runs at a frame-rate of 25 frames/s.



(a) (b)

Figure 4.Tracking results for a sequence

2.3. Visual behaviours

Eyelid movements and pose face are some of the visual behaviours that reflect a person's level of fatigue. There are several ocular measures to characterize eyelid movements such as eye closure duration, blink frequency, eye closure/opening speed, and the recently developed parameter PERCLOS [14]. This last measure percentage of eye closed over time, excluding the time spent on normal closure. It has been found to be the most valid ocular parameter for characterizing driver fatigue [3]. Nevertheless, we have calculated all of them in order to be able to evaluate its performance.

To obtain the ocular measures we continuously track the subject's pupils and fit two ellipses, to each of them, using a modification of the LIN algorithm [15]. In order to detect the difference between a blink and an error in the tracking of the pupils, we use a Finite State Automata (FSM) as we depict in figure 5. Apart of the *init_state*, five states have been defined: *tracking_ok, closing, closed, opening and tracking_lost.* Transitions between states are achieved from frame to frame as a function of the width-height ratio of the pupils.

The system starts at the *init_state*. When the pupils are detected, the FSM pass to the *tracking_ok* state indicating that the pupil's tracking is working correctly. Being in this state, if the pupils are not detected in a frame, a transition to the *tracking_lost* state is produced. The FSM stays in this state until that the pupils are correctly detected again. In this moment, the FSM pass to the *tracking_ok* state. If the width-height ratio of the pupil increases above a threshold, being in *tracking_ok* state, a closing eye action is detected and the FSM changes to the *closing_state.* Because the width-height ratio may increase due to other reasons, such as segmentation noise, it is possible to return to the *tracking_ok* state if the ratio does not constantly increase.

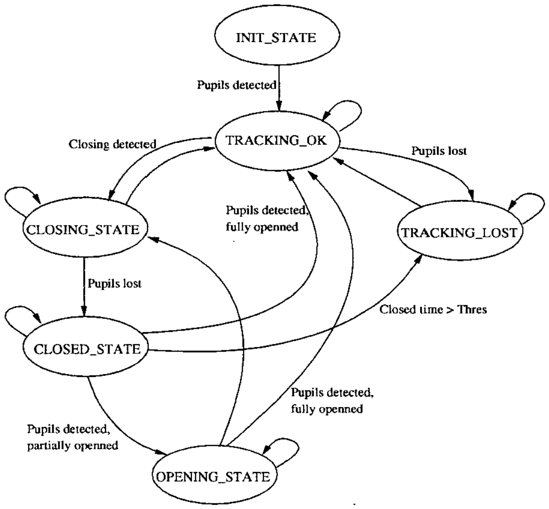


Figure 5.Finite State Automata for ocular measures

The pupils loss being in *closing_state* provokes the transition of the FSM to *closed_state,* which means that the eyes are closed. A new detection of the pupils from the *closed_state* produces a change to *opening_state* or *tracking_ok state*, depending on the degree of openness of the eyelid. Being in the *closed_state,* a transition to the *tracking_lost* state is produced if the closed time is over a threshold. That means that the tracker is not working correctly. A transition from *opening* to *closing* is possible if the width-height ratio decreases again.

Ocular parameters that characterize eyelid movements have been calculated as function of the FSM. PERCLOS is calculated from all the states, except from the *tracking_lost* state, analysing the pupil height. We consider that an eye closure occurs when the pupil height is less than the 30% of its nominal size. We compute the PERCLOS measure based on the percentage of eye closure, in 30 s. Eye closure duration measure is the time that the system is at the *closed_state* evaluated in 30 s. Eye closure/opening speed measures are calculated in the same way but for the *closing_state* and *opening_state* respectively. Blink frequency is the number of blinks detected in 30 s. A blink action will be detected as a transition among the following states: closing, closed and opening.

On the other hand, face pose can detect fatigue and inattentive drivers. The nominal face orientation while driving is frontal. If the driver's face orientation is in other directions for an extended period of time or occurs frequently (in case of various head tilts), this is due to fatigue or inattention. In our application, the precise degree of face orientation is not necessary. What we are interested in is to detect if the driver's head deviates from its nominal position and orientation for an extended time or too frequently (nodding detection).

We have used a model-based approach for recovering the face pose by establishing the relationship between 3D face model and its two-dimensional (2D) projections. The inputs of the model are the pupils and the nostrils positions. Then, we have developed a new tracker for the nostrils positions in the same way as the used for the pupils tracking. Nostrils appear in the images as dark pixels surrounded by not so dark pixels (skin), and they are easily detectable.

As function of the calculated angles from the model and using the speed data of the pupil's movements from the Kalman filters, we quantized the face direction in nine areas: frontal, left, right, up, down, upper left, upper right, lower left and lower right. Nodding measure indicates the number of head tilts detected in the last 2 minutes. These actions are calculated analysing the temporal evolution of the face model and the vertical speed data of the pupils.

2.4. Driver Vigilance Computation

The fatigue visual behaviours, obtained in the previous section (eyelid movements and pose face) are subsequently combined to form a fatigue parameter that can robustly and accurately characterize one's vigilance level. This information fusion is obtained using a Fuzzy system. We have chosen this technique for its well known linguistic concept modelling ability. The fuzzy rule expression is close to expert natural language. On the other hand, as they are universal approximators, fuzzy inference systems can be used for knowledge induction processes. The objective of the fuzzy system is to provide a driver's vigilance level (DVL) from the fusion of several ocular and face pose measures and using expert knowledge. This knowledge has been extracted from the data analysis of the parameters in some simulated tests with different users.

Various studies have shown the possibility to detect fatigue by means of ocular measures. Taking in mind the study from US Department of Transportation [3] and our experience, we have concluded that the best set of input variables for the fuzzy system are: PERCLOS, eye closure duration, blink frequency, nodding frequency and face direction. This last variable evaluates the number of frames where the face direction is not frontal in the last 30 s. All the fuzzy inputs have been normalized and different linguistic terms and its corresponding fuzzy sets have been distributed in each of them. For the output variable we have distributed some fuzzy sets in vigilance level range between 0 and 100%. Based on the variables mentioned before, we have defined some rules for correct vigilance monitoring. If this level is higher than a predefined threshold, an acoustic warning is activated.

3. Experimental Results

The system is currently running on PC Pentium IV (1,8 Ghz) in real time (25 frames/s) with a resolution of 400x320 pixels. For testing its performance ten sequences, simulating some drowsiness behaviours, were done. These were achieved following the physiological rules explained in [3] to identify drowsiness in drivers. Test sequences were recorded from a car in a motorway using different users without glasses and with different light conditions. These images have been used as inputs of our algorithms, obtaining some quite robust, reliable and accurate results.

Figure 6 plot DVL obtained from the fuzzy system for a test sequence of 6 minutes. As can be seen, until the frame 1000 (40 s) DVL is below 45 %, which represents the alert state. DVL is overs 50%, between frames 1000 and 1900 (76 s), due to the blink frequency increase. Then, between frames 1900 and 5500 (220 s) DVL decreases because the blink frequency returns to its normal rate. Beyond the frame 5500 the DVL rises due to the PERCLOS enhance. Between the frame 7000 (280 s) and 8000 (320 s) the duration of the eye closure rise provoking a higher enhance of the DVL. In the frame 7900 (316 s) a nodding is produced. This drowsiness behaviour provokes that the DVL is over the 60%, showing a fatigue state. After the nodding, the user temporally awakes provoking the DVL decrease. Our experiments show that, if a driver is alert while driving, the DVL should be less than 50%. A DVL over 70% indicates a clear drowsiness state.

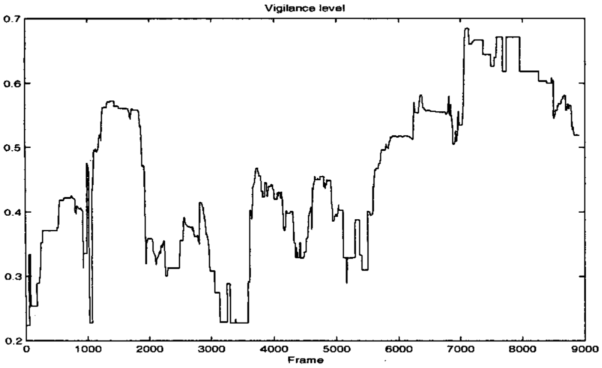


***Figure 6.**Parameter over a sequence of 6 min*

General performance of the measured variables is presented in table 1. Performance was measured by comparing system performance to the observer measures that were in the recorder video sequences.

Parameters	Percentage detected
PERCLOS	90
Eye closure duration	85
Blinks frequency	80
Nodding frequency	70
Face direction	85

***Table 1.** Accuracy of the results*

As we can see, the percentage detected for all the variables is quite good taking in mind that the sequences were achieved in real situations. The system performance decreases when drivers wear eyeglasses.

4. Conclusions and future works

We have developed a non-intrusive prototype computer vision system for real-time monitoring driver's vigilance. It is based on a hardware system, for real time acquisition of driver's images using an active IR illuminator, and their software implementation for real time pupil tracking, ocular measures and face pose estimation. Finally, driver's vigilance level is determined from the fusion of the measured parameters into a fuzzy system.

In the future we have the intention to test the system with more users, in order to generalize drowsiness behaviours, and to improve it for users with eyeglasses.

Acknowledgments

This work has been supported by grants FOM2002-002 from the Spanish Ministry of Public Works and by grants DPI2002-02193 (SIRAPEM Project) from the Spanish Ministry of Science and Technology (MCyT)

References

[1] AWAKE Consortium (IST 2000-28062). sept 2001– sept. 2004. http://www.awake-eu.org

[2] J. Healey and R. Picard, "SmartCar: Detecting Driver Stress". *15th International Conference on Pattern Recognition*, Volumen 4, pp.218-221. Barcelona, 2000.

[3] Albert Kircher, Marcus Uddman, Jesper Sandin, "Vehicle control and drowsiness", *Swedish National Road and Transport Research Institute*. 2002.

[4] Anon, "Perclos and eyetracking: Challenge and Opportunity". *Technical Report*, Applied Science Laboratories, Bedford, MA. 1999.

[5] Ueno, H., Kaneda, M. & Tsukino, M. "Development of drowsiness detection system". *Proceedings of Vehicle Navigation and Information Systems Conference*, Yokohama, Japan, pp. 15–20, 1994.

[6] Jesús Soria, "Llega el piloto tecnológico", *Revista Tráfico, Dirección General de Tráfico. Ministerio de Interior.* pp: 20-21, Marzo-Abril 2002.

[7] Y. Matsumoto and A. Zelinsky "An algorithm for real-time stereo vision implementation of head pose and gaze direction measurements". *Procs. IEEE 4th Int. Conf. Face and Gesture Recognition*, pp: 499-505, March 2000.

[8] T. Victor and A. Zelinsky "Automating the measurement of driver visual behaviours using passive stereo vision". *Procs. Int. Conf. Series Vision in Vehicles VIV9*, Brisbane, Australia, August 2001.

[9] Boverie, S., Leqellec, J.M. & Hirl, A. "Intelligent systems for video monitoring of vehicle cockpit". *International Congress and Exposition ITS. Advanced Controls and Vehicle Navigation Systems*, pp. 1-5. 1998

[10] P. Smith, M. Shah and N. Da Vitoria Lobo. "Determining Driver Visual Attention with One Camera". *IEEE Transactions on Intelligent Transportation Systems*, Vol. 4, nº 4, pp: 205-218, December 2003.

[11] W. Shih and Liu, "A calibration-free gaze tracking technique". *in Proc. 15th Con Patterns Recognition*, vol. 4, Barcelona, Spain, pp:201-204. 2000.

[12] Qiang Ji and Xiaojie Yang, "Real-time Eye, Gaze and Face Pose Tracking for Monitoring Driver Vigilance", *Real-Time Imaging 8*, pp:357-377, 2002.

[13] R. Grace, "Drowsy Driver Monitor and Warning System", *Int. Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design*, August, 2001.

[14] Dinges D. "F. PERCLOS: A valid psychophysiological measure of alertness as assesed by psychomotor vigilance", *Federal Highway Administration. Office of motor carriers.* FWHA-MCRT-98-006. 1998.

[15] Andrew W. Fitzgibbon , Robert B. Fisher, "A buyer's guide to conic fitting", *Proceedings of the 6th British conference on Machine vision (Vol. 2)*, p.513-522, Birmingham, United Kingdom, 1995.