IEEE ISIE 2005, June 20-23, 2005, Dubrovnik, Croatia

Real-Time System for Monitoring Driver Vigilance

Luis M. Bergasa*, Jesus Nuevo*
*Departamento de Electronica
Escuela Politecnica, Univ. Alcala
Alcala de Henares, 28885 Madrid
Email: bergasa@depeca.uah.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 behaviors 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.**

I. INTRODUCTION

The increasing number of traffic accidents due to a diminished driver's vigilance level has become a serious problem for society. In Europe, statistics show that between 10% and 20% of all the traffic accidents are due to drivers with a diminished vigilance level caused by fatigue. In the trucking industry, about a 60% of fatal truck accidents are caused to driver fatigue. It is the main cause for heavy truck crashes [1].

According to the U.S.A. National Highway Traffic Safety Administration (NHTSA), falling asleep while driving is responsible for at lest 100.000 automobile crashes annually. An annual average of roughly 40,000 nonfatal injuries and 1,550 fatalities results from this crashes [2]. These crashes happen between the hours of midnight and 6 am, involve a single vehicle and a sober driver travelling alone, with the car leaving the roadway without any attempt to avoid the crash. These figures underestimate the true level of the involvement of drowsiness because they do not include crashes involving daytime hours, multiple vehicles, alcohol, passengers or evasive manoeuvres. These statistics also do not deal with crashes causes by driver distraction, which is believed to be a larger problem. As car markers add intelligent vehicle applications to satisfy consumers ever increasing demand for a wired, connected world, the level of cognitive stress on drivers is being incrementally raised. That is, the more assistant systems for comfort, navigation or communication thereby causing inattention from the most basic task at hand, driving the vehicle.

With this background, developing systems for monitoring a driver's level of vigilance and alerting the driver, when he is not paying adequate attention to the road due to fatigue or distractions, is essential to prevent accidents.

The rest of the paper is structured as follows. In section II we present a review of the main previous work in this line. Section III presents the general system architecture, explaining its main parts. Experimental results are presented in section IV. Finally, in section V we show the conclusions and the future works.

II. PREVIOUS WORK

In the last few years many researchers have been working on the development of safety 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. A representative project in this line is MIT Smart Car [3], where several sensors (electrocardiogram, electromyogram, respiration and skin conductance) embedded in a car and visual information for sensors confirmation are used. In ASV (Advanced Safety Vehicle) project, held by Toyota [4], the driver must wear a wristband in order to measure his heart rate. Others techniques monitor eyes and gaze movements using a helmet or special contact lens [5]. These techniques, though less intrusive, are still not acceptable in practice.

A driver's state of vigilance can also be characterized by indirect vehicle behaviors like lateral position, steering wheel movements, and time to line crossing. Although these techniques are not intrusive they are subject to several limitations such as vehicle type, driver experience, geometric characteristics, state of the road, etc. On the other hand, these procedures require a considerable amount of time to analyze user behaviors and thereby they do not work with the so-called micro-sleeps: when a drowsy driver falls asleep for some seconds on a very straight road section without changing the lateral position of the vehicle [6]. In this line we can find different experimental prototypes, but none of them has been commercialized so far. Among them there is an important Spanish

0-7803-8738-4/05/$20.00 ©2005 IEEE

system called TCD (Tech Co Driver) [7] based on steering wheel and lateral position sensors. Toyota [4] uses steering wheel sensors and pulse sensor to record the heart rate, as explained above. Mitsubishi has reported the use of steering wheel sensors and measures of vehicle behavior (such as lateral position of the car) to detect driver drowsiness in their advanced safety vehicle system [4]. Daimler Chrysler has developed a system based on vehicle speed, steering angle and vehicle position relative to road delimitation (recorded by a camera) to detect if the vehicle is about to leave the road [8].

People in fatigue show some visual behaviors easily observable from changes in their facial features like eyes, head, and face. Typical visual characteristics observable from the images of a person with reduced alertness level include longer blink duration, slow eyelid movement, smaller degree of eye opening (or even closed), frequent nodding, yawning, gaze (narrowness in the line of sight), sluggish facial expression, and drooping posture. Computer vision can be a natural and non-intrusive technique for extracting visual characteristics that typically characterize a driver's vigilance from the images taken by a camera placed in front of the user. Many researches have been reported in the literature on developing image-based driver alertness using computer vision techniques. Some of them are primarily focused on head and eye tracking techniques using two cameras [9] [10]. In [11] a system called FaceLAB developed by a company called Seeing Machines is presented. All systems explained above rely on manual initialization of feature points. The systems appear to be robust but the manual initialization is a limitation, although it makes trivial the whole problem of tracking and pose estimation.

In [12] we can find a 2D pupil monocular tracking system based on the differences in color and reflectivity between the pupil and iris. The system monitors driving vigilance by studying the eyelid movement. Another successful head/eye monitoring and tracking of drivers system to detect drowsiness using of one camera, and based on color predicates, is presented in [13]. This system is based on passive vision techniques and its functioning can be problematical in poor or very bright lighting conditions. Moreover, it does not work at nights, when the monitoring is more important.

In order to work at nights some researches use active illumination based on infrared LED. In [14] a system using 3D vision techniques to estimate and track the 3D line of sight of a person using multiple cameras is proposed. In [15] a system with active infrared LED illumination and a camera is implemented. Almost all the active systems reported in the literature have been tested in simulated environments but not in real moving vehicles. A moving vehicle presents new challenges like variable lighting, changing background and vibrations that must be taken into account in real systems. In [16] an industrial prototype called *Copilot* is presented. This system uses infrared LED illumination to find the eyes and it has been tested with truck's drivers in real environments. It uses a simple subtraction process for finding the eyes and it only calculates a validated parameter called PERCLOS (percent eye closure), in order to measure driver's drowsiness. This system currently works under low light conditions.

Systems relying on a single visual cue may encounter difficulties when the required visual features cannot be acquired accurately or reliably, as happens in real conditions. Then, a single visual cue may not always be indicative of the overall mental condition [15]. The use of multiple visual cues reduces the uncertainty and the ambiguity present in the information from a single source. The most recent researches in this line use this hypothesis. Currently, the ambitious European project AWAKE [1] is under development. A multi-sensor approach is proposed in this project adapted to the driver, the vehicle, and the environment in an integrated way. The system is under exhaustive pilot testing for determining its functional performance and the user acceptance of the application [17]. At the moment only some partial results have been presented.

This paper describes a real-time prototype system based in computer vision for monitoring driver vigilance using active infrared illumination and a single camera placed on the car dashboard. We have employed this technique because our goal is to monitor a driver on real conditions (vehicle moving) and in a very robust and accurate way, mainly at nights (when the probability to crash due to drowsiness is the highest). The proposed system does not need manual initialization and monitors several visual behaviors that typically characterize a person's level of alertness while driving. In a different fashion than other previous works, we have fused different visual cues from one camera using a fuzzy classifier instead of different cues from different sensors. We have analyzed the different visual behaviors that characterize a drowsy or inattentive driver and we have studied the best fusion for optimal detection. Moreover, we have tested our system during several hours in a car moving in a motorway and with different users.

III. SYSTEM ARCHITECTURE

The general architecture of our system is shown in figure 1. It consists of four major modules: 1) Image acquisition, 2) Pupil detection and tracking, 3) Visual behaviors and 4) Driver vigilance. Image acquisition is based on a low-cost CCD micro-camera sensitive to near-IR. The pupil detection and tracking stage is responsible for segmentation and image processing. Pupil detection is simplified by 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 following stage we calculate some parameters from the images in order to detect some visual behaviors easily observable in people in fatigue, such as slow eyelid movement and face pose, to detect some visual behaviours easily observable in people in fatigue. Finally, in the driver vigilance evaluation stage we fusion all the individual parameters obtained in the previous stage using a fuzzy system, yielding the driver inattentive level. An alarm is activated if this level is over a certain threshold.

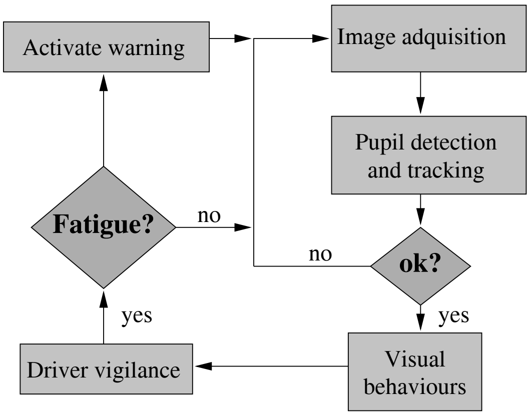


Fig. 1. General architecture

A. Image Acquisition System

The purpose of this stage is to acquire images of the driver's face. In order to make acquired images as invariant as possible to light conditions and ease eye detection and tracking, we use active illumination in the Near-IR band. The illuminator is based on two rings of IR-LEDS, as can be seen on figure III-A. The inner ring beams light along the camera optical axis, producing the "bright pupil" effect as light is reflected on the driver's retina. The outer ring provides ambient illumination that is used for contrast enhancing. Those LEDs do not produce the bright pupil effect, and only a glint can be observed on the eyes. An example of those effects can be seen in figure 3. This pupil effect is clear with and without glasses or contact lenses and it even works to certain degree with sunglasses.

Image acquiring from the camera and LED excitation is synchronized. LED rings light driver's face alternatively, one in each of the image fields, providing different light conditions for almost the same image, once the fields are de- interlaced on the following stages. The camera includes a narrow bandpass filter, centered at LED wavelenght, to reduce interference from other light sources, and improve control over the illumination of the resulting images.

Figure III-A shows some pictures of the image acquisition system. It is composed by a miniature CCD camera sensitive to Near-IR and located on the dashboard of the vehicle, in the centre of the IR-LED rings. This camera focuses on the driver's head for detecting the multiple visual cues. An embedded PC with a low cost frame-grabber is used for video signal acquisition and signal processing.

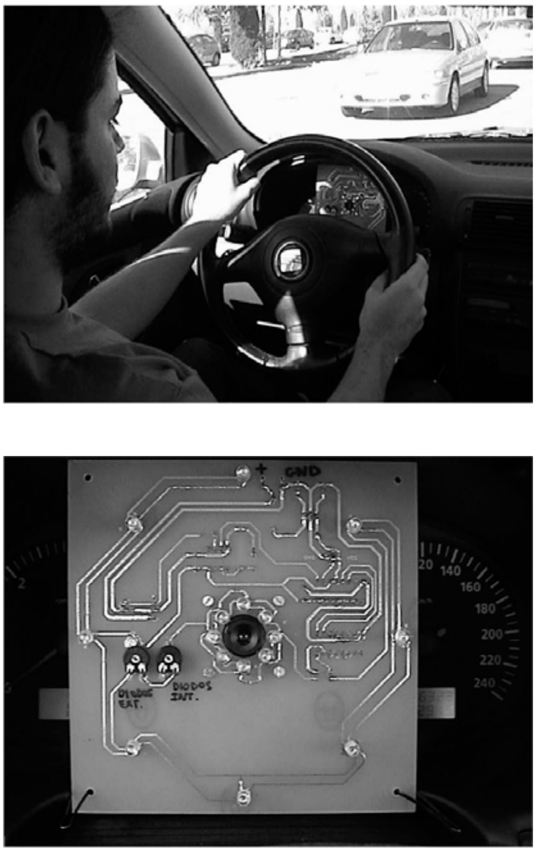
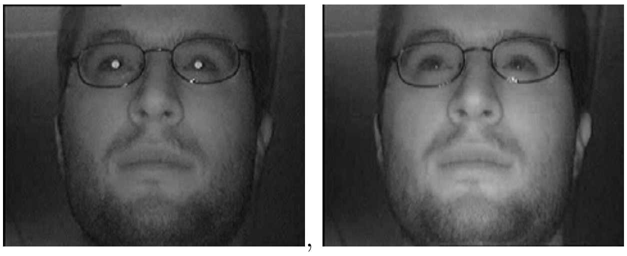


Fig. 2. Pictures of the prototype

B. Pupil detection and tracking

This stage starts with pupil detection. As mentioned above, each frame obtained from the frame-grabber is de-interlaced in even and odd fields. 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 ambient light illumination influence because this is subtracted in the generation of the difference images.

This procedure yields images with high contrast where the pupils are easily found. It can be observed that the glint produced by the outer ring usually falls close to the pupil, and with the same grey level as the bright pupil. Then, the shape of the pupil blob in the difference image is not the same as the original. Modelling of the pupil blobs is done by fitting an ellipse [18]. However, as the algorithm used does not perform well in this situation, the system only uses subtracted images during initialization, and when light conditions are poor. On other cases, only the field obtained with the inner ring on is processed, increasing accuracy and reducing computation time.



(a) Image obtained with inner IR ring

(b) Image obtained with outer IR ring

(c) Difference Image

Fig. 3. Fields captured and subtraction

Pupils are detected on the resulting image, by searching the entire image to locate two bright blobs that satisfy certain constraints. The image is binaryzed, using an adaptive threshold, for detecting the brightest blobs in the image. A connected components analysis is then applied to the binaryzed difference image to identify binary blobs that satisfy certain size and shape constraints. The blobs that are out of some size constraints are removed, and for the others an ellipse model is fit to each one. Depending on their size, intensity, position and distance, best candidates are selected, and all the possible pairs between them are evaluated. The pair with the highest qualification is chosen, and its parameters are used to update the statistics that set thresholds and margins in all processing stages. Those statistics include size, grey level, position and apparent distance and angle between pupils, calculated over a time window of 2 seconds. The thresholds also modify their values if the pupils are not found, increasing the margins to make more candidates available to the system. To continuously monitor the driver it is useful 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 [15] but we have tested it with images obtained from a car moving in a motorway. In our case, the search window is detemined automatically based on pupil size, velocity and location error. The state vector of the filter is represented as $\mathbf{x}_t = (\mathbf{c}_t, \mathbf{r}_t, \mathbf{u}_t, \mathbf{v}_t)$, where $(\mathbf{c}_t, \mathbf{r}_t)$ indicates the pupil pixel position (its centroid) at time $\mathbf{t}$ and $(\mathbf{u}_t, \mathbf{v}_t)$ be its velocity at time t in $\mathbf{c}$ and $\mathbf{r}$ directions respectively.

C. Visual behaviors

On this section we show how our system recognizes and process a series of visual behaviours that are found to characterize a person's level of fatigue. The recently developed PERCLOS [19] parameter has been found to be the most valid ocular parameter for characterize driver's fatigue. This measures reflects the percentage of time the eye is closed, excluding the time spent on normal closure, such as a regular blink. There are other several interesting measures to characterize eyelid movements, such as eye closure duration, blink frequency and eye closure/opening speed. All of them have been calculated to make the system more robust. Explanation of the developed methods to detect these measures has been presented by the authors in [20]. In this paper we will focus our attention in another interesting eyelid movement parameter, known as "fixed gaze", and a measure of the face pose as is the nodding.

One of the behaviours that appear on drowsy driver is that of loosing the focus of the gaze, not giving attention to any of the elements of the traffic. This lost of concentration usually takes place before other sleepy behaviours, such as nodding. In order to develop a method to measure this behaviour in a simple way, we have found that the speed data from the Kalman filters used to track the pupils can be used.

A driver in good condition moves his eyes frequently, focusing to the changing traffic conditions, specially if the road is busy. This has a clear reflection on the speed data of the Kalman filters, as can be seen on Figure 4, where "fixed gaze" behaviour is present from second 130 to 250.

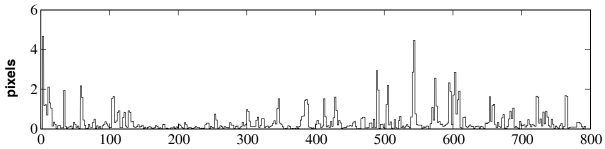


Fig. 4. Measures of the "fixed gaze" parameter

Our system monitors the speed on the $\mathbf{x}$ coordinate, and calculates the mean and variance of the speed in 2-second time windows. If both statistics are below an a priori threshold, the window is marked as "positive". The mean of the last 30 seconds is passed to the next processing stage. Coordinate $\mathbf{y}$ is not used, as difference between drowsy and awake drivers is not clear.

The results obtained with this parameter have been encouraging. Although using the same a priori threshold for different drivers and situations, detection was always correct. Even more remarkable was the absence of false positives.

Nodding detection is also based on data from the Kalman filters. We have observed that noddings follow a similar pattern in many occasions. The driver tilts his head while closing his eyes, to open his eyes and rise the head again quickly after that. Those movements produce a pattern both in positions and speeds of the pupils, as well as the moments when the eyes are closed and opened. We have found that above 80% of the noddings in our video database fit in that pattern and can be easily detected. This percent of accuracy can be considered high for such a simple method.

Two of the parameter used in our system, blinking frequency and closed time duration, rely completely on the accuracy of the detection of blinks. The greatest challenge is to differentiate between a blink and an error in the tracking of the pupils, as in both cases the tracking is lost. To do that, we use a Finite State Automata (FSM) as we depict in figure 5. Apart from 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 initializes and remain in *init_state* until the pupils are detected. The FSM passes 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 the pupils are correctly detected again. In this moment, the FSM passes to the *tracking_ok* state. If the width-height ratio of the pupil increases above a threshold (20% of the nominal ratio), 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.

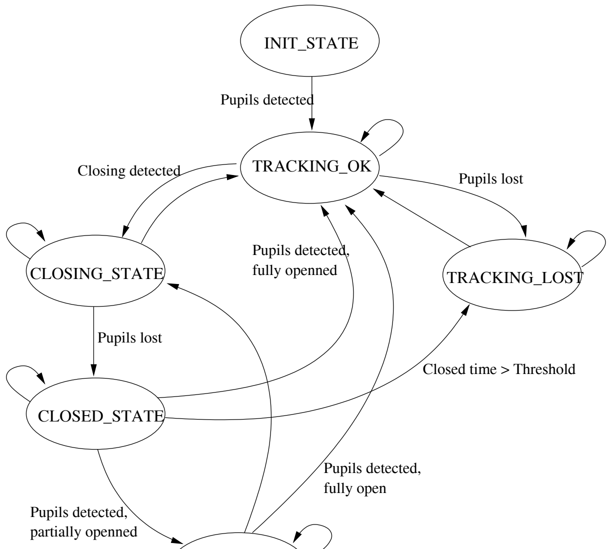


Fig. 5. Finite State Machine for ocular measures

Pupil loss while in *closing_state* provokes a transition of the FSM to *closed_state*. 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 opening of the eyelid. Being in *closed_state*, a transition to the *tracking_lost* state is produced if the closed time goes over a threshold. A transition from *opening_state* to *closing_state* is possible if the width-height ratio increases again. A blink will be only counted and processed if the transition between *closing*, *closed* and *openning* states completes.

Ocular parameters that characterize eyelid movements have been calculated as a function of the FSM. PERCLOS is calculated from all the states, except from the *tracking_lost* state, analyzing the pupil width-height ratio. We consider that an eye closure occurs when the pupil ratio is above the 80% of its nominal size. We compute this parameter by measuring the percentage of eye closure in a 30 seconds window. 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, for the *closing_state* and *opening_state* respectively. Blink frequency is computed as the number of blinks detected in 30 s.

D. Driver vigilance computation

The fatigue visual behaviours, obtained in the previous stage, are subsequently combined to form a fatigue parameter that can robustly and accurately characterize driver's vigilance level. This information fusion is obtained using a Fuzzy Inference System. The objective of the fuzzy system is to provide a driver's inattentiveness level (DIL) 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. Generation of the rules was done with the inference tool KBCT (Knowledge Base Configuration Tool) [21] developed by a research group of the UPM (Polytechnic University of Madrid). The FIS (Fuzzy Inference System) contains 94 rules, 86 inferred directly with KBCT and 8 included manually by the developers.

The inputs to the FIS are: PERCLOS, eye closure duration, blink frequency, nodding frequency, nodding magnitude and the "fixed gaze" parameter. As the system is adaptive to the user, the values of the inputs to the FIS are approximately the same for all users. The fuzzy sets where distributed by the inference tool. For the output variable, the fuzzy set was chosen manually. The inattentiveness level range is between 0 and 100%, with a normal value up to 50%. When its value is between 50% and 80%, the driver the fatigue is medium, but it the DIL is over 80% the driver is considered to be fatigued, and an alarm is activated.

IV. RESULTS

The system is currently running on PC Pentium IV (1,8 Ghz) in real time (50 fields-25 frames/s) with a resolution of 400x320 pixels. For testing its performance ten sequences, simulating some drowsiness behaviours, were recorded. These were achieved following the physiological rules explained in [4] 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 plots DIL obtained from the fuzzy system for a test sequence of 6 minutes. As can be seen, until the second 90, DIL is 50 %. After that, the driver was told to simulate drowsiness twice times during the recording, showing different behaviours, so the performance of all the parameters could be evaluated. General performance of the measured variables is presented in table I. Performance was measured by comparing system performance to the observer measures that were in the recorder video sequences.

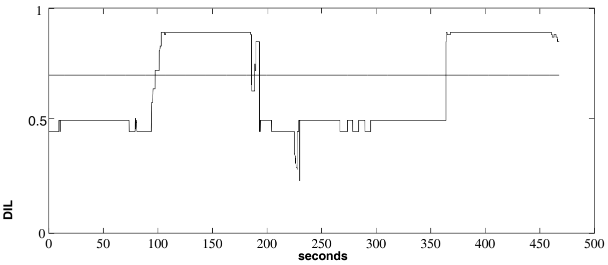


Fig. 6. DIL over a sequence of 6 minutes

TABLE I

Parameter performance

Parameters	**Total correct percentage**
PERCLOS	93.12%
Eye closure duration	84.37%
Blink freq.	80%
Nodding freq.	72.5%
Face position	87.5%
Fixed gaze	95.62%
DIL	98.12%

V. 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 wearing glasses.

Acknowledgement

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

[2] D. Royal, "Volume I - Findings; National Survey on Distracted and Driving Attitudes and Behaviours, 2002," The Gallup Organization, Tech. Rep., Mar. 2003.

[3] J.Healey and R. Picard, "Smartcar: Detecting driver stress," in *15th International Conference on Pattern Recognition*, vol. 4, 2000, pp. 218–221.

[4] A. Kircher, M. Uddman, and J. Sandin, "Vehicle control and drowsiness," Swedish National Road and Transport Research Institute, Tech. Rep. VTI-922A, 2002.

[5] Anon, "Perclos and eyetracking: Challenge and opportunity," Applied Science Laboratories, Bedford, MA., Tech. Rep., 1999.

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

[7] J. Soria, "Llega el piloto tecnológico," pp. 20–21, Mar./Apr. 2002.

[8] DaimerChryslerAG. (2001, June) The electronic drawbar. [Online]. Available: http://www.daimlerchrysler.com

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

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

[11] Seeing Machines. (2004, Aug.) Face-lab transport. [Online]. Available: http://www.seeingmachines.com/transport.htm

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

[13] P. Smith, M. Shah, and N. D. V. Lobo, "Determining driver visual attention with one camera," *IEEE Trans. Intell. Transport. Syst.*, vol. 4, no. 4, pp. 205–218, dec 2003.

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

[15] Q. Ji and X. Yang, "Real-time eye, gaze and face pose tracking for monitoring driver vigilance," *Real-Time Imaging*, vol. 8, pp. 357–377, Oct 2002.

[16] R. Grace, "Drowsy driver monitor and warning system," in *Int. Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design*, Aug 2001.

[17] A. Polychronopoulos, A. Amditis, and E. Bekiaris, "Information data flow in awake multi-sensor driver monitoring system," in *IEEE Intelligent Vehicles Symposium*, Parma, Italy, Jun 2004, pp. 902–906.

[18] A. W. Fitzgibbon and R. B. Fisher, "A buyer's guide to conic fitting," in *Proc. of the 6th British conference on Machine vision*, vol. 2, Birmingham, United Kingdom, 1995, pp. 513–522.

[19] D. Dinges, "F. perclos: A valid psychophysiological measure of alertness as assesed by psychomotor vigilance," Federal Highway Administration. Office of motor carriers, Tech. Rep. MCRT-98-006, 1998.

[20] L. Bergasa, R. Barea, E. Lopez, and M. Escudero, "Facial features tracking applied to drivers drowsiness detection," in *21st IASTED Int. Multi-Conference APPLIED INFORMATICS*, Innsbruck, Austria, Feb. 2003, pp. 231–235.

[21] J. M. A. Moral, S. Guillaume, and L. Magdalena. (2003) KBCT, knowledge base control tool. [Online]. Available: http://www.mat.upm.es/projects/advocate/en/index.htm