Gaze Fixation System for the Evaluation of Driver Distractions Induced by IVIS

Pedro Jiménez, Luis M. Bergasa, Jesús Nuevo, Noelia Hernández, and Ivan G. Daza

***Abstract*—We present a method to monitor driver distraction based on a stereo camera to estimate the face pose and gaze of a driver in real time. A coarse eye direction is composed of face pose estimation to obtain the gaze and driver's fixation area in the scene, which is a parameter that gives much information about the distraction pattern of the driver. The system does not require any subject-specific calibration; it is robust to fast and wide head rotations and works under low-lighting conditions. The system provides some consistent statistics, which help psychologists to assess the driver distraction patterns under influence of different *in-vehicle information systems* (IVISs). These statistics are objective, as the drivers are not required to report their own distraction states. The proposed gaze fixation system has been tested on a set of challenging driving experiments directed by a team of psychologists in a naturalistic driving simulator. This simulator mimics conditions present in real driving, including weather changes, maneuvering, and distractions due to IVISs. Professional drivers participated in the tests.**

***Index Terms*—Distraction monitoring, driver, gaze fixation, inattention, in-vehicle information systems (IVISs), naturalistic simulator, percent road center (PRC).**

I. Introduction

DRIVING inattention is a major factor in traffic accidents. In the EU-27, 38 900 people died in 2008 in traffic accidents, and 34 500 people lost their lives in 2009 [1]. That year, over 1.25 million accidents took place, and more than 1.5 million people were injured [2]. On the other hand, the *National Highway Traffic Safety Administration* (NHTSA) estimates that approximately 25% of police-reported crashes involve some form of driving inattention, including fatigue and distraction [3]. Driving distraction is more diverse and implies a more risky factor than fatigue, and it is present in more than half of inattention-involved crashes, resulting in as many as 5000 fatalities and $40 billion in damages each year [4].

Manuscript received May 12, 2011; revised September 28, 2011 and January 10, 2012; accepted January 20, 2012. Date of publication March 6, 2012; date of current version August 28, 2012. This work was supported in part by the Ministerio de Ciencia e Innovación under Project DRIVER-ALERT (TRA2008-03600), Project ADD-Gaze (TRA2011-29001-C04-01), and Project CABINTEC (PSE-370000-2009-12), and in part by the Comunidad de Madrid under Project Robocity 2030 (S2009/DPI-1559). The Associate Editor for this paper was S. S. Nedevschi.

P. Jiménez is with Gamesa, 48170 Zamudio, Spain.

L. M. Bergasa and N. Hernández are with the Department of Electronics, University of Alcalá, 28805 Madrid, Spain.

J. Nuevo is with the Commonwealth Scientific and Industrial Research Organisation ICT Centre Computer Vision Group, Brisbane, QLD 4069, Australia.

I. G. Daza is with the Computer Engineering Department, University of Alcalá, 28805, Madrid, Spain.

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TITS.2012.2187517

Driving distraction is defined by the *American Automobile Association Foundation for Traffic Safety* as occurring "when a driver is delayed in the recognition of information needed to safely accomplish the driving task because some event, activity, object, or person within or outside the vehicle compelled or tended to induce the driver's shifting attention away from the driving task" [5]. Thirteen types of potentially distracting activities are listed [4]. Since the distracting activities take many forms, NHTSA classifies distraction into four categories from the view of the driver's functionality [3]: 1) visual distraction; 2) cognitive distraction; 3) auditory distraction (e.g., responding to a ringing cell phone); and 4) biomechanical distraction (e.g., manually adjusting the radio volume). Many distracting activities can involve more than one of these components (e.g., talking to a phone while driving creates a biomechanical, auditory, and cognitive distraction). Increasing use of *in-vehicle information systems* (IVISs) such as cell phones, Global Positioning System (GPS), DVD players, and other onboard devices has exacerbated the problem by introducing additional sources of distraction [6]. Enabling drivers to benefit from IVISs without diminishing safety is an important challenge.

One promising strategy to mitigate the effects of distraction involves monitoring and classifying the driver state and then using this classification to adapt the IVIS. Driver inattention monitoring has been an active research field for decades, mainly focusing on fatigue, and various methods have been proposed. Some auto companies have already installed some fatigue-monitoring systems in their high-end vehicles. However, there is still a great need to develop a more reliable and fully functional system using cost-efficient methods for a real driving context.

To date, realistic studies that provide information on the impact of distracting activities have been developed as small-scale studies. An effort is needed to study distraction under naturalistic situations. Simulation is an optimal method of experimentation to acquire knowledge of the driver's behavior, which is close to a real scenario but without the safety risks of having inattentive drivers in an open road. The simulation methodologies applied in Europe to the road transport sector research are demonstrating their profitability and efficiency [7]. The main objective through the simulation is to immerse the driver in his normal work environment. To do this, a fully equipped cockpit is required to perform the driving task. Previous work scoping the prediction of driver behavior mostly relies on lane position and vehicle sensors. In addition, a video signal of the driver is frequently used, but manual annotation or very simple head position estimation is used [8].

In this paper, an automatic gaze fixation system for the evaluation of IVIS-induced distraction is presented. Driver's gaze fixation is estimated using a nonintrusive vision-based approach. The system has been tested with professional drivers in a naturalistic simulator running tests directed by a team of psychologists.

We show results of the performance of our system and consistent statistics to infer the distraction behavior of the drivers during these exercises.

The remainder of this paper is organized as follows: In Section II, a review of the main approaches of the state of the art is presented. Sections III and IV present our 3-D face pose estimation and eye direction estimation proposals. Evaluation of our gaze fixation system performance and an analysis of the statistics proposed to study distractions are addressed in Section V. This paper closes with conclusions and future work presented in Section VI.

II. Driver-Distraction-Monitoring Approaches

In the literature, there are three main groups of works according to the measurements they used to detect distractions: 1) biological signals; 2) driving signals; and 3) driver images.

Biological signals include electroencephalogram, electrocardiogram, etc. These signals are collected through electrodes in contact with the skin of the human body, and consequently, they are intrusive systems [9]. Only few works focusing in cognitive distractions have been reported in the literature using this approach [10]. Most of them have only been tested in operational environments.

Vehicle signals reflect driver's action, so driver's state can be characterized in an indirect way. Force on pedals, vehicle velocity changes, steering wheel motion, lateral position, or lane changes are normally used in this category. The advantage of these approaches is that signal acquisition is easier. This is the reason the few commercial systems existing nowadays use this technique [11], [12]. However, they are subject to several limitations such as vehicle type, driver experience, or road geometric characteristics. Moreover, some results showed that the accuracy using this approach varied between individuals [13].

Approaches based on image processing are effective because the occurrence of distraction is reflected on the driver's face appearance and head and eye activity. Different kinds of cameras and analysis algorithms have been employed in this approach. We group them according to the cameras they adopted, including visible spectrum monochrome cameras, infrared (IR) cameras, or stereo cameras.

A. Methods Based on Visible Spectrum Camera

The simplest and most affordable hardware setup is a visible spectrum image acquisition system, at the cost of requiring more complicated algorithms to compensate for the lack of data when compared to infrared or stereo systems. Rongben *et al.* [14] used skin color to segment the face. This method needs initialization and is not robust to lighting conditions or user race. Sun *et al.* [15] detected the face using adaptive boosting, locating the eyes using a template matching method, and estimated gaze combining Hough transform and gradient direction. Eye activities contain not only fatigue information but also distraction information.

A commercial eyetracker is also employed in some research works. Blaschke *et al.* [16] used an off-the-shelf eyetracker to get head pose and eye gaze signal. They model the visual distraction level as a time dependency of the visual focus, with the assumption that visual distraction increases with time as the driver looks away from the road but nearly instantaneously decreases when the driver refocuses on the road. Based on the pose and eye signals, they propose a two-stage detection method: First, the instantaneous distraction level is estimated, and then, a classifier determines if the current level corresponds to a distracted driver.

B. Methods Based on IR Camera

Many researchers have adopted image acquisition systems based on IR illumination. The use of IR serves three purposes: 1) It minimizes the impact of different ambient lighting conditions. 2) It produces the bright pupil effect. 3) It increases illumination without disturbing the driver. Because of bright pupil effect, the eye can be more easily detected, eliminating the face segmentation and reducing computation times.

Cudalbu *et al.* [17] used a headband with IR reflective markers to estimate the head pose, from which they get a 6 degree-of-freedom head pose with the average error of 0.2°. Together with this headband, they use a simplified 3-D eyeball model to estimate the gaze orientation with an accuracy between 1 and 3°. Jiao and He [18] proposed a Round Template Two Values Matching algorithm to locate the bright pupil, which obtains an accuracy of 96.4% but takes 1.01 s/frame on a PIII 800-MHz computer.

Some commercial products measuring driver's states are already available on the market, such as SmartEye AntiSleep [9] and SeeingMachines DSS [10], [19]. Both of them use two IR illuminators to enhance their robustness to lighting conditions and employ only one camera to give 3-D information. However, they are focusing on detecting fatigue, not distraction, and they are still limited to some well-controlled environments.

C. Methods Based on Stereo Camera

Stereo cameras are also employed to estimate driver state. In [20], two standard web cameras are employed to make a 3-D image acquisition system. They extract the face from the disparity map on the assumption that the driver face has smaller depth than background. After the face region is extracted, they perform embedded hidden Markov model to recognize the forehead, eyes, nose, mouth, and chin, from which the driving fatigue level can be estimated. The commercial products based on 3-D camera technology such as Smart Eye Pro [9] and Seeing Machines faceLAB [21] can provide measurements of head pose, eyebrow, eye, nose, and mouth.

Commercial products are still limited to some well-controlled environments, and more importantly, no technical information about their algorithms and effectiveness has been published.

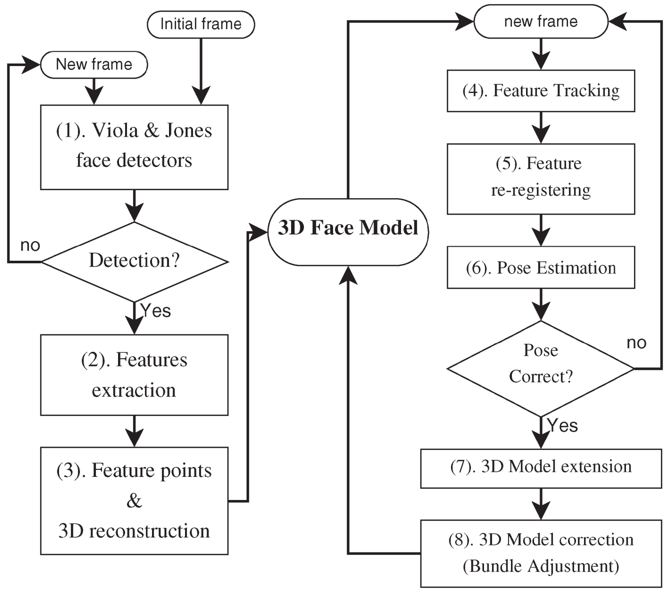


Fig. 1. Architecture of the face pose estimation algorithm.

D. Contributions

Most of the works in the state of the art were designed for visual distractions but only a few for cognitive, auditory, or biomechanical distraction. In this paper, we introduce a novel gaze fixation system that could potentially be used to extrapolate distractions from any kind, provided all the information from the driving exercises and simulator. Our approach comprises a face pose estimation system and a gaze classification system. Both are nonintrusive and designed to work under very low illumination conditions. The face pose estimation works in the full yaw rotation range. In contrast with existing commercial solutions, it does not require any subject-specific calibration and could then be used in commercial vehicles. With only two cameras, the system is robust to fast and wide head rotations. The gaze is classified in several areas in the cabin and the road, which may draw the attention of the driver when a distraction is taking place. We use this system in a naturalistic simulator to study changes in driver behavior due to distractions caused by IVISs. Our system is able to automatically generate objective distraction statistics. It does not require user self-reports, which are subjective, or additional input from experts. The experimental setup was designed by a team of psychologists, and it is described in detail in this paper.

III. Three-Dimensional Face Pose Estimation

In this section, we describe the main characteristics of our face pose estimation method. A detailed description of the face pose estimation can be found in [22]. A diagram of the implemented face pose estimation system is shown on Fig. 1.

A requirement is that the pose estimation must be user independent. The proposed face pose estimation system is based on tracking a set of features, which are automatically detected on the subject's face with a calibrated stereo rig.

The features are selected from high-contrast regions of the face using a Harris interest point detector [23]. Taking both camera views, the features are arranged in the form of a sparse 3-D face model. This user-specific 3-D model provides a prior to the feature tracking that the method performs on each frame: features that drift from their expected positions on the model are discarded as outliers.

The appearance of the features in the face changes when the head rotates, and our method proposes a feature template registering using a novel mixed-view technique: the samples of the features taken by both cameras are used to build a joint appearance model, so that views of the face from one camera can be used to anticipate what the other camera will see as the face rotates. This strategy was inspired by the work of Nuevo *et al.* [24]. From the detected 2-D position of the features, the pose is estimated using Levenberg-Marquardt [25].

A model extension process adds new features to the model when the face rotates, exposing new areas to the camera. This model extension and the shared template registering allow for face pose estimation over the full yaw rotation range from -90° to $+90^\circ$. The online model creation process is subject to errors. We correct the 3-D feature coordinates of the model with bundle adjustment (BA) optimization [26], which is executed at certain key frames after initialization and after the addition of new features.

IV. Eye Direction Estimation

To obtain the gaze fixation areas of the driver, we need to take into account both the face pose and the eye direction. Gaze estimation has been applied in human–computer interaction and in studies of cognition, which also have attracted interests from marketing research. A survey of the techniques can be found in Hansen and Ji [27]. A common step of all these techniques is the need of person-specific calibration, and recent developments focus on reducing the length of this step. Guestrin and Eizenman [28] presented a system based on calibrated light sources that achieved accuracies of 1°, requiring the subjects to look at a single point. Saliency maps of the images shown to the users have been used as a prior to estimate the eye parameters. Sugano *et al.* [29] reported errors of 6°, and more recently, Chen and Qi [30] presented a probabilistic approach with errors below 2°.

The approach that we take is substantially different from these methods. We aim to classify the gaze fixation point in a discrete number of areas of interest, and thus, we are only interested in a rough estimation of the gaze. Compared to the works previously mentioned, our setup is much harder as the head can freely and quickly move in a wide range and the illumination is low. In addition, we cannot do any offline user calibration, and we cannot precompute the saliency maps on the scene. Our only prior is the position of the different areas, which is much less informative than a saliency map because it is not based on human cognition.

We calculate the eye direction $\vec{e}$ with respect to the face model coordinate system. Consequently, gaze $\mathbf{G}$ is computed as the composition of face pose and eye direction, as shown in Fig. 2. It can be expressed as a vector $\vec{g}$ and an origin T_g as

$$\mathbf{G} = \{T_g, \vec{g}\}, \qquad T_g = \vec{T} + R \cdot \vec{e}_{\text{off}}, \quad \vec{g} = R \cdot \vec{e} \qquad (1)$$

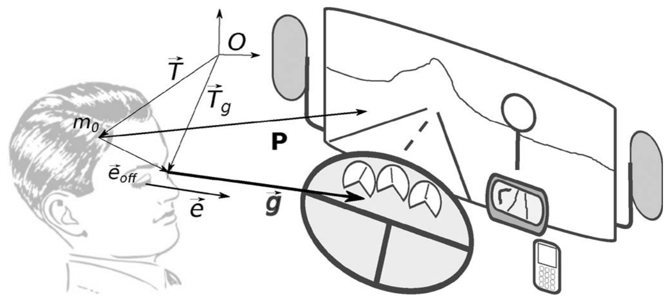


Fig. 2. Difference between gaze and face pose.

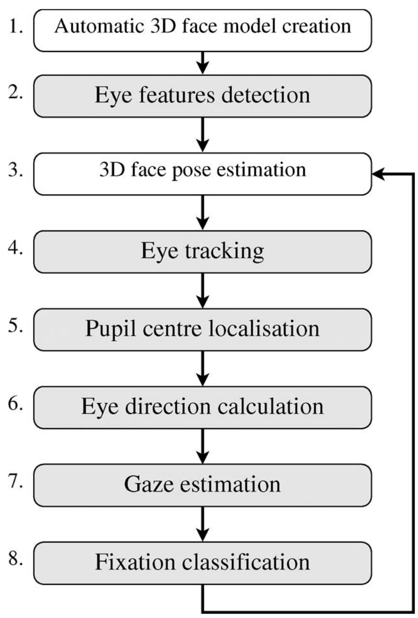


Fig. 3. Main blocks of the gaze estimation algorithm.

where $\mathbf{P} = \{\vec{T}, R\}$ is the face pose (translation to the center of the face and rotation), and $\vec{e}_{\text{off}}$ is the distance between the center of the eyes and the center of the face model $\vec{T}$.

At the gaze estimation step, the face pose is already known because it has been calculated during the previous stage, so $\mathbf{P}$ is given by the face pose translation vector. If the head rotation is wide enough, eyes might not be visible, so the gaze is calculated using only the known face pose.

A. Algorithm Description

The steps of the coarse algorithm for gaze estimation are shown in Fig. 3. It consists of the following steps:

1) Automatic 3-D face model creation: The 3-D face model is first created, as explained in Section III.

2) Initial Eye Features Detection: At the model creation stage, some characteristic features around eyes are detected within the face using the Stacked Timmed Active Shape Models algorithm [31] and added to the face model. Fig. 4 shows these features.

3) Three-Dimensional Face Pose Estimation: In a loop, the face pose is estimated from frame to frame, as described in Section III.

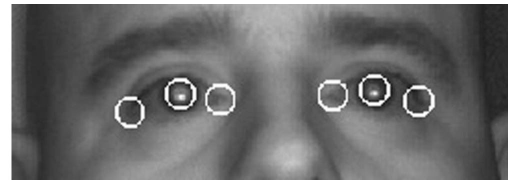


Fig. 4. Initial eye features: Eye corners and pupil.

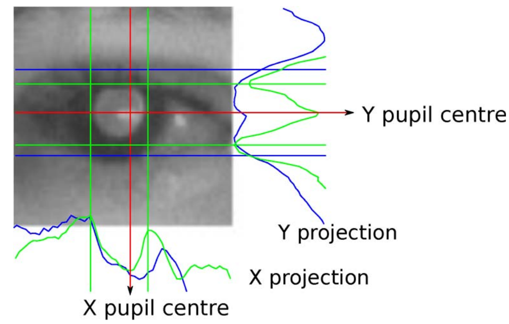


Fig. 5. Samples of the integral projection algorithm.

4) Eye Tracking: The eye features are tracked frame to frame using the same technique as for the rest of the model points. This gives the approximate position of the pupils.

5) Pupil Center Localization: Pupil position is located for each eye using the integral projections algorithm and a Gaussian approximation, obtaining an approximated localization error of 20% of the pupil size (see Fig. 5).

6) Eye Direction Calculation: The eye direction is calculated from the relative displacement of the pupil center with respect to its original 3-D coordinates.

7) Gaze Estimation: Gaze is computed rectifying the face pose estimation with the eye direction estimation, according to (1).

8) Fixation Classification: Fixation is calculated as the intersection point between the gaze vector and the simulator scene. The fixation point is classified among a set of interest areas at which the subject can be looking.

Fig. 6 shows an example of the face pose and gaze estimation monitoring application. At the bottom of the image is shown the fixation point on the scene.

B. Gaze Fixation and Classification

The objective of this classification is to determine the fixation area of the driver and to know where of a set of key areas she/he can be looking at. The 11 different fixation areas defined for this project are shown in Fig. 7.

1) Front: The road itself and traffic ahead. Victor and Joanne, basing on experiments on various simulators and on real traffic, defined this as an area between 16° and 20° in diameter centered on the road [32].

2) Left and Right Signals: Denote the signaling on sides of the road, overtaking cars, crosses, or other objects present in the proximity of the truck. When the driver is looking at any of these points, the fixation is slightly diverted horizontally to the left or to the right.

Fig. 6. Example of face pose estimation and gaze estimation, with fixation point on the scene.

Fig. 7. Set of key fixation areas.

3) Lateral Rear Mirrors: The external rear mirrors located at both sizes of the cabin. Most times, it is not possible to localize the pupils when the driver is looking to them because they are occluded, but it is easily recognizable because he/she needs to largely turn his/her head horizontally.

4) Onboard Computer, GPS, and Hands-Free: Usually, the driver tries to look at these IVISs with very little head movement to not lose attention to the road.

5) Tachograph: This IVIS is located overhead, over the windscreen, and looking at it requires a head movement.

6) Overhead Signaling and Near Road: Looking at these points requires no head movement and very little vertical pupil displacement, so it is difficult to distinguish when the driver is looking there from the front road itself. We do not classify these areas.

The cameras' position inside the cabin is fixed, and the geometric layout of the simulation room, cabin, and projection panels are known. Therefore, the 3-D centroid in the scene of each of the regions previously described can be measured and referenced to the right camera frame system. The fixation area is calculated as the closest area to the gaze fixation point using Mahalanobis distance, which takes into account the different sizes of the areas. Using this algorithm, it is possible to know the area where the driver is looking at or how long the driver's gaze remains fixed on an area. This is very important in studying drivers' reactions to different IVISs to know the potential distraction caused by each of them.

V. Distraction Analysis Using Gaze Estimation

This section presents the tests and results of our nonintrusive approach to monitoring driver's distraction. Fixation in the scene is calculated to infer the driver's distraction state. Different distraction tasks or activities were inferred in a naturalistic simulator, and a study of the incidence of these distracting tasks in the driver's behavior was carried out.

Professional drivers were invited to drive the truck through a few scenarios carefully designed by a team of psychologists from the Safety and Human Factors Investigation and Training Center (ESM) [33], who later examined the generated data to extrapolate behavior. The scenarios were designed and prepared to require a high level of attention from the driver, and some tasks were intentionally programmed during the driving activity to stress the driver to study his/her behavior under such conditions. The simulator cabin was fully equipped with a variety of IVISs. Experiment layout, driver behavior study results, and conclusions are presented.

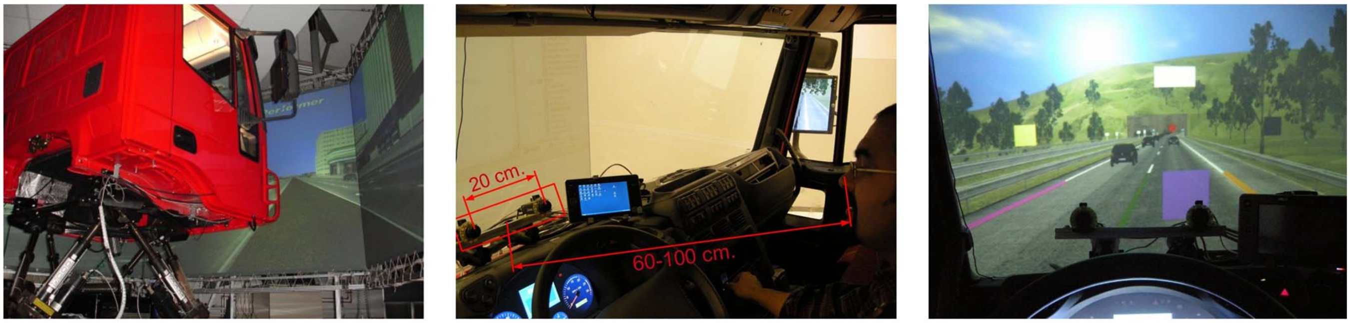
A. Experimental Environment

Different aspects must be considered in the experimental environment: the physical simulator layout, the camera vision system for gaze estimation, the experiment setup, and the subjects.

1) Naturalistic Driving Simulator: The experiments were performed in the research facilities at CEIT (San Sebastián, Spain) [34], in a room with controlled light and sound environment. The given naturalistic simulator *TUTOR* [35], as shown in Fig. 8(a), consists of a real truck cabin motorized to simulate movement and equipped with common IVIS. The cabin is assembled on a movement platform with 6° of freedom on which drivers can feel the vehicle accelerating, braking, its centrifugal force, etc. The devices send information to the host, located at the Instructor Position (*PI*), where the psychologists can control the whole simulator, analyze all the data, and reproduce stored simulations. The main computers are placed in the PI, which are located near the cabin. A dedicated computer processes face pose and gaze estimation using the algorithms presented in this paper and sends this information to the PI, where the psychologists can access to the data.

The visualization system is made of three back-projection panels with a total surface of 22 m^2. The fact that the screens have no marked separation and the geometry of the image system make for a flawless overall impression. Moreover, two computer screens are used as rear mirrors attached to both sides of the cabin.

The cabin is fully equipped and contains a GPS, a hands-free, the onboard computer, and a tachograph. These are some of the key locations that the gaze estimator must differentiate.



(a) (b) (c)

Fig. 8. Naturalistic driving simulator cabin. (a) Motorized simulator cabin and projection panels. (b) Camera position. (c) Driver's view of the road and cameras.

2) Vision System: The stereo rig is located inside the cabin, over the dashboard, between the windscreen and the driving wheel, facing the driver at a distance between 70 and 100 cm. The baseline is 23 cm, and cameras are slightly turned inward to better point to the driver's face, closer to the rig. Fig. 8(b) and (c) shows the camera layout inside the naturalistic simulator cabin. The cameras [36] are 1392 $\times$ 1040 pixels and capture images at of 30 frame/s, with a Firewire 800b interface. The custom-built rig and IR illumination board are attached to the windscreen using suction pads and rest on foam over the dashboard to reduce vibration. The hardware layout can vary without affecting system operation, as long as the stereo rig is calibrated and referenced to the world coordinates, the fixation area locations are approximately known, and the cameras are ahead of the driver but not necessarily in front. Our cameras mount a 9-mm lens, and capture is synchronized with the IR pulsed illumination. The capture system is coded using C++ and GNU libraries and tools, running on a Core 2 Duo processor commanded by Ubuntu. This vision system is an evolution of the monocular one developed by the group for drowsiness detection [37].

3) Experiment Setup: To design the experimental protocol, the team of psychologists built on the following initial hypothesis: "The potential driver distraction due to IVISs is determined by the level of attention demand required by them while driving, decreasing the effectiveness of the primary task: driving."

By analyzing professional driver behavior, the basic and most representative features in the context of this activity are identified [38]. Some scenarios, types of vehicles, incidents, onboard system utilization, and critical situations are selected to infer distraction in drivers. Thus, the professional driver behavior should be generically represented. Taking into consideration this basis, which involves observing and information recording during the driving activity, the next step was to define the simulation exercises.

Experiments were designed with the goal of refuting the initial hypothesis of the research regarding the potential distraction of four different onboard systems, which are commonly used in professional driving. These devices were digital tachograph, GPS, hands-free, and onboard computer. Under these conditions, four scenarios were created: mountain, intercity, urban, and long-distance. Different exercise setup were prepared for each scenario, each containing different tasks, events, weather conditions, and IVIS requirements.

According to Victor *et al.* [32], three different tasks are of special importance to study distraction: visual tasks, auditory tasks, and cognitive tasks. During the experiments, visual tasks require to use the GPS. Auditory ones involve making a call to the hands-free telephone and engaging in a trivial conversation. For the last one, a cognitive task is enforced in one of the exercises by making a phone call in which the driver is asked to describe the route from one point to another on a city he knew. During the exercises, events are inserted proximal to tasks, such as motor, tire or Automatic Breaking System breakdown, sudden braking of the precedent vehicle, broken down vehicles on the road, vehicles running a red light, etc. A summary of the different exercise setup is shown in Table I.

These tests were implemented using 16 different exercises: five based on the intercity scenario, four on the mountain scenario, three on the urban scenario, and the last four on the long-distance scenario. The defined procedure to evaluate these exercises consists of different drivers driving through different scenarios.

The first exercise of each scenario was the *control exercise*, which corresponds to the exercise undertaken by each driver without external perturbations. This exercise provides a reference for the other exercises, in terms of driver gaze fixation patterns on the controls, to which the results from distractions can be compared.

4) Subjects: Twelve professional drivers from different genders, ages, and experiences were selected, having different participants for each test configuration to detect user-dependent behavior variables. Each driver drove in two exercises on each scenario (the control exercise plus another), and duration varies from 20 to 45 min, depending on the scenario.

Previous studies with similar conditions used groups from seven to 30 participants [39], [40]. All subjects were informed of the purpose of the experiment and the security procedures in the simulator facilities.

B. GT Generation

To assess the performance of our system, we obtained the ground-truth (GT) for the face pose and gaze. We obtained the pose GT with a calibration chessboard attached to a helmet that the subjects wore during the experiments. The pose of the chessboard was obtained using camera calibration techniques, with an average error below 1°. The GT data were obtained for six users on sequences that were over 10 min long.

TABLE I
EXERCISE SETUP

Scenario		Exercise	Events	IVIS
A. Inter-city	1	Control exercise	A vehicle running a STOP. Mechanical fault in air filter†‡. Cyclists on road. Sudden speed down of preceding vehicle. Slow vehicle on road.	GPS Hands-free
	2	GPS guidance		
	3	Faulty GPS guidance		
	4	Telephone guidance		
	5	GPS guidance Distorted voice call		
B. Mountain	1	Control exercise	Obstacle on road A vehicle running a STOP. A vehicle stopped on road. Sudden speed down of preceding vehicle. Slow vehicle on road. Tyre blowout‡.	GPS Hands-free Tachograph
	2	GPS guidance		
	3	Telephone guidance		
	4	GPS guidance Faulty voice call Tachograph speed warning		
C. Urban	1	Control exercise	ABS fault† A vehicle running red light. Mechanical fault in air filter†‡. A pedestrian crossing the street. A dog crossing the street.	GPS Hands-free Tachograph Computer
	2	GPS guidance tachograph error		
	3	GPS Guidance Distorted voice assistance call		
D. Long-distance	1	Control exercise	Obstacle on road A vehicle running a STOP. Vehicles stopped on the road. Slow vehicle on road, and a car overtaking a bus downhill.	Hands-free Tachograph Computer
	2	Phone calls On-board computer warnings		
	3	Phone call Cognitive phone call* Tachograph warnings		
	4	Phone call On-board computer data		

†Marked on the on-board computer
‡Truck dynamic model changes
*Phone call with important cognitive charge: The driver is asked to explain a route within a known city.

TABLE II
MEAN FACE POSE ESTIMATION ERROR. THE ERROR IS DIVIDED INTO yaw, $pitch$, AND $roll$, AND EVALUATED IN DIFFERENT RANGES OF THE ABSOLUTE ROTATION ANGLE IN THE GROUND TRUTH α

Rotation	BA?	$\alpha < 15°$	$\alpha < 30°$	$\alpha < 45°$	$\alpha \geq 45^o$
yaw	no	1.92	2.44	6.72	12.83
yaw	**BA**	**0.98**	**1.54**	**3.04**	**8.54**
pitch	no	3.82	7.86	8.59	-
pitch	**BA**	**1.81**	**4.70**	**6.34**	-
roll	no	1.27	2.06	-	-
roll	**BA**	**1.16**	**1.75**	-	-

With respect to gaze, we used a video database of 15 videos that were more than 10 min long each. We have obtained the GT for the gaze fixation classification by manually labeling the driver's fixation area on each moment.

C. Face Pose Experimental Results

First, we analyze the general performance of our face pose estimation system. The mean pose estimation error in the three angle rotations is shown in Table II. The mean error is computed by rotation ranges for each direction. Pitch and roll rotations present a smaller output range than that for yaw since there are no wider head rotations for these directions on the driving exercises. Because the registering technique cannot be applied to pitch variations, the system error is higher in this direction, and it can be observed how it increases for angles $\alpha_{\mathrm{pitch}} > 30°$. Still, the BA slightly improves the results. Evaluation of the error in a wider pitch and roll range is not possible because significant rotations are not natural while driving.

As we can see, our proposal has a very low error due to the BA corrections. The error remains low for the full range $\pm 90°$ of yaw rotations. These results show equal or lower errors than other important works in the literature [41]–[43], despite us testing the system with more challenging scenarios. The low-lighting conditions, along with the fast head movements, make the face appear very blurry when the head is moving side to side. Any other tracking system would lose track of the head under these conditions. On the other hand, we can keep tracking the face due to the extended face model and the mixed-view technique.

The tests used in the different works presented in Table II differ, but a qualitative comparison among them is possible, taking into account the performance results published for each work.

D. Gaze Fixation Experimental Results

Table III presents the gaze fixation classification error. The error has been calculated as the relation between the incorrect classifications and the total number of frames (from the GT) that the driver has gaze fixation to a specific area. A false negative is when we do not detect that the driver is looking at a specific fixation area. A false positive happens when the classifier says that the driver is looking at the fixation area but is actually looking at another area. For reference, we divide the false positives into the different areas where the driver was really looking at. As it can be observed in the table, the classification errors for the Front fixation area, rear mirrors, and onboard computer are below 5%. This are the areas that give

TABLE III
GAZE CLASSIFICATION ERROR PERCENTAGE

Fixation area	**False Negative** (not detected)	**False Positive** (detected while looking to:)						
		Total	Front	Left mirror	Right mirror	Computer	GPS	Hands-free
Front	4.64%	2.33%	-	0.62%	0.41%	0.36%	1.33%	1.02%
Feft mirror	2.79%	1.31%	0.27%	-	0.02%	0.60%	1.20%	0.10%
Right mirror	3.22%	2.14%	0.15%	0.01%	-	0.46%	0.90%	1.30%
Computer	4.96%	4.78%	1.75%	0.36%	0.28%	-	1.30%	0.87%
Gps	11.16%	6.56%	2.86%	0.50%	0.60%	2.40%	-	4.70%
Hands-free	10.78%	15.3%	1.28%	0.60%	0.50%	3.20%	5.20%	-
Tachograph	5.45%	4.70%	2.10%	0.84%	0.34%	1.67%	0.30%	0.20%
Left and Right signals	13.95%	17.30%	6.30%	2.10%	1.40%	1.40%	1.70%	0.90%

TABLE IV
CLASSIFICATION OF REGISTERED INCIDENTS

Distraction	Time	Crashes	Near Crashes	Incidents
No distraction	95%	0%	1%	0%
GPS	2%	42%	47%	40%
On board computer	1%	0%	0%	12%
Tachograph	0.1%	0%	0%	0%
Hands-free	2%	58%	52%	48%

more information on the distraction state of the driver. The left and right signals are the most difficult to detect since they are very proximal to the front fixation area.

E. Statistics Results to Infer Distractions

In this section, we provide some consistent statistics, based on gaze fixations, that are useful to assess driver distraction. The driver's behavior interpretation must be carried out by psychologists and is out of the scope of this paper.

As explained in Section V-A, the drivers were asked to perform demanding tasks while driving, such as answering the phone and following indications, and receive corrections of previous indications to get to a place or to operate the onboard computer or the tachograph. In Table IV, the recorded incidents while using the different IVISs are shown.

Most of the accidents occurred while receiving or executing the indications from the GPS or the hands-free device. As expected, the more cognitive demanding a task is the higher the risk of being involved in a distraction-related incident. Results on Table IV also show that using the tachograph and onboard computer is not very cognitive demanding since it only requires quick looks, similar to a mirror checking. The challenge now is to study driver reactions to different IVISs and find out the optimal IVIS location and the optimal way to deliver indications and warnings to them. Hereafter, we analyze two concrete cases for the most critical IVIS. Using our system, statistics are automatically generated while many other works in the state of the art are based on user self-report, which is subjective.

1) Hands-Free Device: Most incidents were registered while the driver was using the hands-free device. This, together with the fact that 85% of people admitted using a cell phone while driving at least occasionally, makes phone use one of the most dangerous activities while driving. The recorded consequences of using the hands-free device are an increase in the corrections of the in-lane position, short braking, nonconstant speed, and fixed gaze.

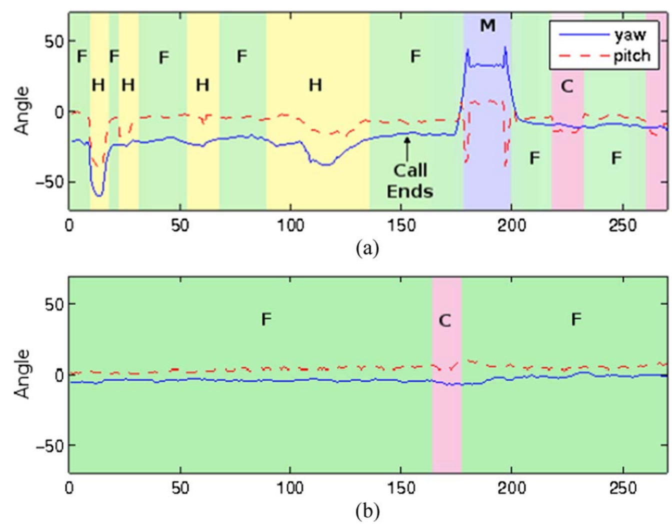


Fig. 9. Driver's gaze focalization while answering the phone. (a) Answering a phone call (seconds). (b) Without external perturbations (seconds).

Fig. 9(a) shows gaze focalization of the driver for a piece of the exercise D2 starting at the moment when a phone call is received. Continuous looks to the hands-free device (H) and no checking to the vehicle speed (C) or rear-view mirrors (M) are detected while answering the phone call. On the other hand, on the control exercise (D1) for the same piece, as shown in Fig. 9(b), the driver remains attentive to the road (F), checking the vehicle speed and the rear-view mirrors from time to time.

2) GPS: The second higher cause of incidents was the use of the GPS device. In our experiment, incidents were caused while setting up the GPS device and while following the indications. When setting up the GPS, the loss of attention to the road was the main cause of incidents. During these losses of attention, the gaze was fixed on the GPS device for long periods of time. However, the indicators received from the GPS also caused an increment in the incident rate, which is mainly due to a decrease in the attention paid to the road information (traffic signs and overhead panels) and a loss of attention to the road with frequent looks to the GPS device.

Fig. 10 shows gaze focalization of the driver for a piece of the exercise B2 versus the control exercise for the same piece. In this case, the driver follows guidance GPS instructions. The lack of attention paid to the road information while using the GPS makes the driver start an overtaking maneuver while approximating the exit that he has been asked to take. The driver is overtaking a very slow moving vehicle when GPS informs him of taking the next exit. To accomplish this instruction, a

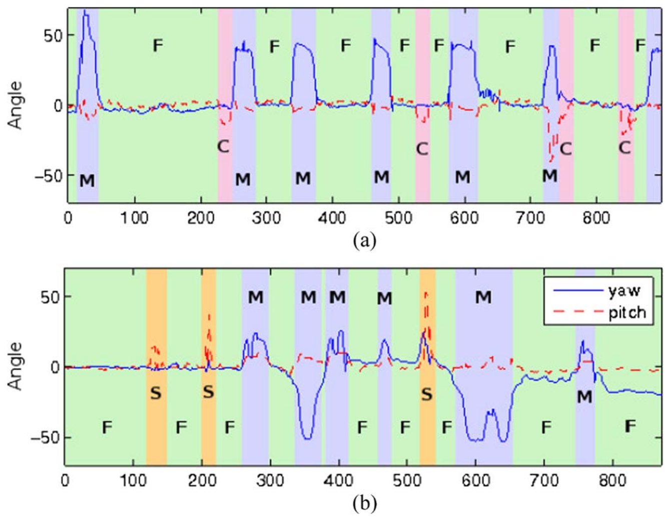


Fig. 10. Driver's gaze focalization previous to overtaking. (a) Following GPS instructions (seconds). (b) Without external perturbations (seconds).

two-lane changing is required in less than 4 s. Instead, when driving without external perturbations, the driver checks the traffic signs to reach his target. The driver is aware that he has to take the next exit in advance and decides to wait behind the slow moving vehicle.

As shown in Fig. 10(a), reiterative checks of the rear-view mirrors (M) and the vehicle speed (C) are performed in the slow-moving vehicle overtaking maneuver. Then, traffic sign checking is not carried out. On the control exercise, the overtaking intentions of the driver are also detected in the mirrors (M) and signals (S) checking, but the driver is aware of the closeness of the highway exit and decides not to start that dangerous maneuver and to wait behind the slow vehicle until taking the exit. Predicting the driver intentions, the indications delivered by the GPS can be modified and can warn the driver about the incoming exit or advise the driver to stay on the right lane. Dangerous situations, as well as annoying indications or warnings, can be avoided by predicting driver intentions.

3) General Statistics: Human factor studies have shown that reaction times are influenced by secondary tasks, such as IVIS usage. Table V shows some statistics obtained after the complete evaluation of the gaze fixation for all the subjects and exercises (where A1, B1, C1, and D1 are control exercises and A[2-5], B[2-4], C[2-3], and D[2-4] are distracting exercises). It compares the reaction times needed before or after an inserted event in the case of the control exercise or when the subject is forced to a distractive driving by requiring a high degree of device utilization. As an example, on exercise D3, a few drivers are unable to avoid hitting an on-road obstacle, whereas they perfectly do it up to 12 s in advance if they were not distracted. Many of them overpass speed limits more often and need more time to notice a mechanical failure. One of the subjects needs more than 2 min to notice that he is driving a fully loaded truck on a mountain road with a flat tire. Being undistracted, he needed a few seconds to notice the same anomaly. The gaze estimation system allows studying what was the subject doing before noticing the anomaly and why he was not aware of that for such a long period. This parameter is important to avoid dangerous situations and to study the arrangements and the ways of delivering information in which the IVIS create a higher level of distraction. Warnings and preventive measures, such as early prebreaking, require a very precise knowledge of not only the driver's state and intentions but the vehicle's surroundings state as well.

To help psychologists understand the distraction pattern of a driver, two more measurements can be taken, apart from reaction time. Traditionally, a glance-based measure has been used. This technique measures the duration of individual fixations on different zones, frequency, number of glances, or total task duration. However, these measurements heavily depend on the task, the driver experience, and other factors. In [32], Victor *et al.* found that the Percent Road Center (PRC) measurement is more stable across users and different experiments. PRC measures how much time is spent looking at the road center area while performing a task. This zone includes the road, signaling, and visual elements proximal to the road. We have analyzed this parameter in our experiments.

Fig. 11 shows the fixation time percents in the different zones of the scene during the execution of a task. This figure was generated after the analysis of the results obtained with the gaze estimation algorithm and tasks schedule during the exercises. The first column shows the average percents for the control experiments (exercises A1, B1, C1, and D1). The second column depicts distractions inferred using the GPS, which are obtained from exercises A2, A3, B2, B4, and C2. The third column depicts distractions inferred by tasks requiring talking by the hands-free phone, on exercises A4, A5, B3, D2, and D4. The last one represents a cognitive task on exercise D3, which induced distraction with a phone call to explain a route.

We have found the PRC to be a good parameter to assess driver distractions. In contrast with other parameters used in state-of-the-art works, this is automatically calculated. Moreover, we also found that the distraction pattern inferred for the different IVISs is different. Unlike most of the works in the state of the art, we analyze visual, cognitive, and auditory distractions. While the GPS shows an important reduction in PRC, phone calls do not reduce the PRC; instead, it slightly increases. On cognitive tasks, we could not infer any important variation of this parameter. However, the time used looking at the signaling and the road proximities are reduced for all tasks. This behavior is clearly observable in Fig. 11, and it is in line with the conclusions presented by [32].

It can be observed how the time that the driver spends looking at the mirrors, signals, and onboard computer is drastically reduced for any of the tasks, in comparison with the control exercises. During phone calls, the driver increases the time in which he is looking at the front but reduces the fixations on the mirrors and signals. This could mean that the driver is not actually paying attention to the road, although more work would be needed to extract a precise conclusion.

VI. Conclusion and Future Works

We have applied a gaze fixation technique, based on face pose and gaze estimation algorithms, to monitor the distraction state of a driver in a naturalistic driving simulator. Our face

TABLE V
DRIVER BEHAVIOR AND REACTION TIME STATISTICS

Scenario	Exercise	Overpass speed limit [#]	Reaction time to an event [seconds]					
			Obstacle‡		Mechanical fault*		Answer a call*	
			max	min	min	max	min	max
Inter-city	A1	0	25	5	(1) 5	9	1.5	5
	A[2-5]	2	15	3	32	81	2	9
Mountain	B1	2	19	5	(2) 0.6	2	1	3
	B[2-4]	6	13	2	24	2min 11s	2	11
Urban	C1	0	16	3	(3) 4	11	2	8
	C[2-3]	0	4	0	4	43	4	miss
Long-Distance	D1	1	34	12	(4) -		1	4
	D[2-4]	4	20	0	-		3	miss

* Reaction time after the event. ‡ Reaction time before the event. Moment at which the subject is aware of the obstacle and takes an action before colliding. (1) Mechanical fault in the air filter. Warning marked on the on-board computer. (2) Tyre blowout. Marked trough audible sound and truck dynamic model changes. (3) ABS fault. Warning marked on the on-board computer. (4) No mechanical faults scheduled.

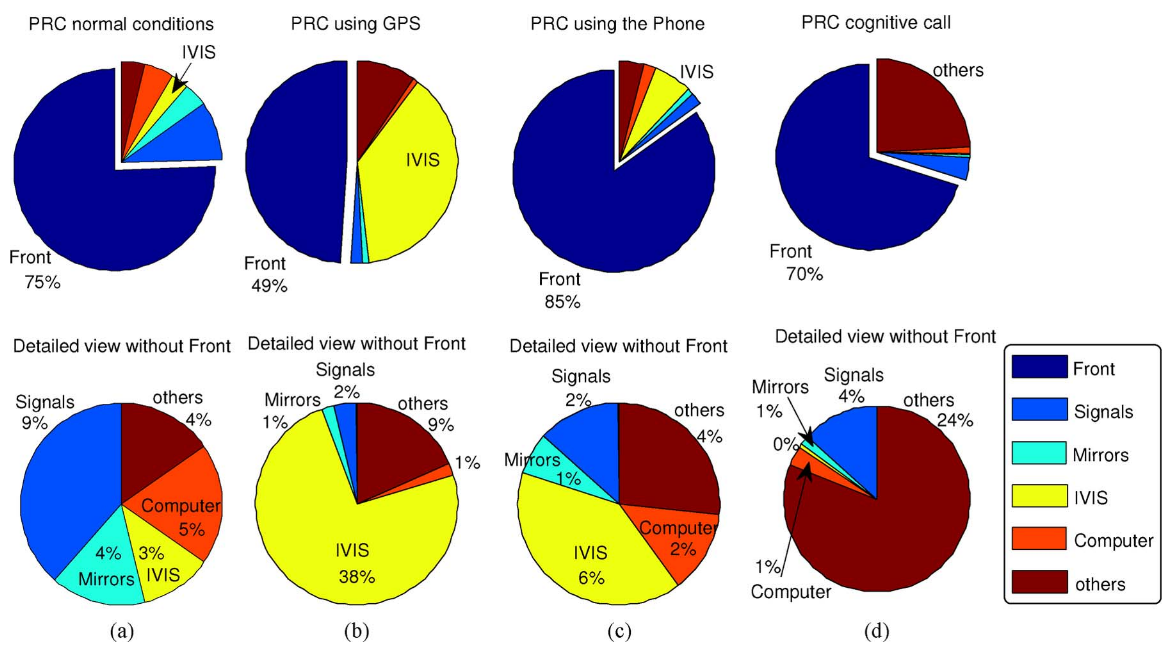


Fig. 11. PRC statistics. (a) Normal driving. (b) Visual tasks. (c) Auditory tasks. (d) Cognitive tasks.

pose estimation approach is based on a sparse 3-D face model obtained from face features. A re-registering algorithm, online model extension, and a BA process for error corrections allow face tracking for the full yaw rotation range $\pm 90^\circ$. The system has been tested under challenging conditions, such as low illumination, variety of drivers, heavy IVIS utilization, and fast and wide head movements. Under these conditions, the system has proven to have good results. It shows a mean error below 1° for rotations in the $\pm 15^\circ$ range and 1.54° in the $\pm 30^\circ$ range, improving the results of other works in the literature.

An eye direction estimation method has been added to the face pose to generate the gaze estimation. The eye direction is based on pupil displacement with respect to their original positions. Gaze fixation information can be used to help psychologists assess the distraction state of the driver. The performance of our gaze fixation system has been analyzed, showing good figures for this application.

Different driving exercises directed by a team of psychologists were recorded in a naturalistic driving simulator, through 12 professional drivers, generating more than 15 hours of driving. Four scenes were created, and distractions induced by different IVIS were inferred and compared to the control exercises without external perturbations. This way, it was possible to compare driver behavior under nondistracted driving conditions versus distracting driving conditions. This has been a challenge data set to evaluate our system.

Reaction times and gaze focalization behavior patterns have been measured to draw conclusions about the capacity of our monitoring system to help psychologists study driver's reactions under different situations while using IVIS. At the time of writing, psychologists are finishing the interpretation of data provided by our system to assess driver distractions over the explained tests.

Further studies about the optimal location of the different IVISs and the way the information is delivered are being designed by psychologists to reduce the distraction New tests with these modified IVIS will be performed on the simulator to evaluate the improvements of the new designs.

ACKNOWLEDGMENT

The authors would like to thank the drivers that participated in the experiments, the psychologists from ESM, and staff from CEIT for their contributions to this paper.

REFERENCES

[1] Y. Mahieu, Highlights of the panorama of transport, Eurostats, Luxembourg, Luxembourg, Tech. Rep. [Online]. Available: http://epp.eurostat.ec.europa.eu/portal/page/portal/product_details/publication?p_product_code=KS-SF-09-042

[2] Annual statistical report, SafetyNet, European Road Safety Observatory, Brussels, Belgium. [Online]. Available: www.erso.eu

[3] T. A. Ranney, E. Mazzai, R. Garrott, and M. J. Goodman, "NHTSA driver distraction research: Past, present, and future," Nat. Hwy. Traffic Safety Admin., Washington, DC, Tech. Rep., 2001.

[4] J. C. Stutts, D. W. Reinfurt, L. Staplin, and E. A. Rodgman, "The role of driver distraction in traffic crashes," AAA Found. Traffic Safety, Washington, DC, Tech. Rep., 2001.

[5] K. Young and M. Regan, *Driver Distraction: A Review of the Literature*. Mawson, NSW, Australia: Australian College of Road Safety, 2007, pp. 379–405.

[6] T. A. Ranney, "Driver distraction: A review of the current state-of-knowledge," U.S. Dept. Transp., Nat. Hwy. Traffic Safety Admin., Washington, DC, Tech. Rep., 2008.

[7] *Integrated System for driver TRaining and Assessment using Interactive education tools and New training curricula for ALL modes of road transport*, 2006–2009, Eur. Project FP6. [Online]. Available: http://cordis.europa.eu/fetch?CALLER=FP6_PROJ&ACTION=D&DOC=3551&CAT=PROJ&QUERY=1170700807981&RCN=79988

[8] J. McCall and M. Trivedi, "Driver behavior and situation aware brake assistance for intelligent vehicles," *Proc. IEEE*, vol. 95, no. 2, pp. 374–387, Feb. 2007.

[9] B. Skinner, H. Nguyen, and D. Liu, "Classification of eeg signals using a genetic-based machine learning classifier," in *Proc. 29th Annu. Int. Conf. IEEE EMBS*, Aug. 2007, pp. 3120–3123.

[10] J. Liu, C. Zhang, and C. Zheng, "EEG-based estimation of mental fatigue by using KPCA-HMM and complexity parameters," *Biomed. Signal Process. Control*, vol. 5, no. 2, pp. 124–130, Apr. 2010.

[11] V. C. Corp., Inside the New Volvo V60—Extra Flexibility Combined With Exclusive Quality, Jul. 2010. [Online]. Available: https://www.media.volvocars.com/global/enhanced/en-gb/Media/Preview.aspx?mediaid=33305

[12] Mercedes-Benz to introduce attention assist into series production in spring 2009, Aug. 2008, Mercedes-Benz. [Online]. Available: http://www.seeingmachines.com/transport.htm

[13] T. Ersal, H. Fuller, O. Tsimhoni, J. Stein, and H. Fathy, "Model-based analysis and classification of driver distraction under secondary tasks," *IEEE Trans. Intell. Transp. Syst.*, vol. 11, no. 3, pp. 692–701, Sep. 2010.

[14] W. Rongben, G. Lie, T. Bingliang, and J. Lisheng, "Monitoring mouth movement for driver fatigue or distraction with one camera," in *Proc. 7th Int. IEEE Conf. Intell. Transp. Syst.*, Oct. 2004, pp. 314–319.

[15] X. H. Sun, L. Xu, and J. Y. Yang, "Driver fatigue alarm based on eye detection and gaze estimation," in *Proc. MIPPR—Automatic Target Recognition Image Analysis Multispectral Image Acquisition*, 2007, p. 678 612.

[16] C. Blaschke, F. Breyer, B. Frber, J. Freyer, and R. Limbacher, "Driver distraction based lane-keeping assistance," *Transp. Res. F, Traffic Psychol. Behav.*, vol. 12, no. 4, pp. 288–299, Jul. 2009.

[17] C. Cudalbu, B. Anastasiu, R. Radu, R. Cruceanu, E. Schmidt, and E. Barth, "Driver monitoring with a single high-speed camera and IR illumination," in *Proc. ISSCS*, Jul. 2005, vol. 1, pp. 219–222.

[18] F. Jiao and G. He, "Real-time eye detection and tracking under various light conditions," *Data Sci. J.*, vol. 6, pp. S636–S640, 2007.

[19] O. Carsten and K. Brookhuis, "Issues arising from the HASTE experiments," *Transp. Res. F, Traffic Psychol. Behav.*, vol. 8, no. 2, pp. 191–196, Mar. 2005.

[20] H. Eren, U. Celik, and M. Poyraz, "Stereo vision and statistical based behavior prediction of driver," in *Proc. IEEE Intell. Veh. Symp.*, Jun. 2007, pp. 657–662.

[21] Seeing Machines, Aug. 2010, faceLAB 5. [Online]. Available: http://www.seeingmachines.com

[22] P. Jiménez, J. Nuevo, L. Bergasa, and M. Sotelo, "Face tracking and pose estimation with automatic three-dimensional model construction," *IET Comput. Vis.*, vol. 3, no. 2, pp. 93–102, Jun. 2009.

[23] C. Harris and M. Stephens, "A combined corner and edge detector," in *Proc. 4th Alvey Vis. Conf.*, 1988, pp. 147–151.

[24] J. Nuevo, L. M. Bergasa, and P. Jiménez, "RSMAT: Robust simultaneous modeling and tracking," *Pattern Recognit. Lett.*, vol. 31, no. 16, pp. 2455–2463, Dec. 2010.

[25] M. Lourakis, levmar: Levenberg–Marquardt nonlinear least squares algorithms in C/C++, 2004. [Online]. Available: http://www.ics.forth.gr/~lourakis/levmar/

[26] B. Triggs, P. McLauchlan, R. Hartley, and A. Fitzgibbon, "Bundle adjustment—A modern synthesis," in *Vision Algorithms: Theory and Practice*, W. Triggs, A. Zisserman, and R. Szeliski, Eds. New York: Springer-Verlag, Sep. 1999, ser. ser. LNCS, pp. 298–375.

[27] D. W. Hansen and Q. Ji, "In the eye of the beholder: A survey of models for eyes and gaze," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 3, pp. 478–500, Mar. 2009.

[28] E. Guestrin and M. Eizenman, "Remote point-of-gaze estimation requiring a single-point calibration for applications with infants," in *Proc. Symp. Eye Track. Res. Appl.*, 2008, pp. 267–274.

[29] Y. Sugano, Y. Matsushita, and Y. Sato, "Calibration-free gaze sensing using saliency maps," in *Proc. IEEE Conf. CVPR*, Los Alamitos, CA, 2010, pp. 2667–2674.

[30] J. Chen and Q. Ji, "Probabilistic gaze estimation without active personal calibration," in *Proc. IEEE Conf. CVPR*, 2011, pp. 609–616.

[31] S. Milborrow and F. Nicolls, "Locating facial features with an extended active shape model," in *Proc. ECCV*, 2008, pp. 504–513.

[32] T. Victor, J. Harbluk, and J. Engström, "Sensitivity of eye-movement measures to in-vehicle task difficulty," *Transp. Res. F, Traffic Psychol. Behav.*, vol. 8, no. 2, pp. 167–190, Mar. 2005.

[33] ESM ESM Safety and Human Factors Investigation and Training, 2011. [Online]. Available: http://www.esm.es/index.php?lang=en

[34] Centre of studies and technical research of gipuzkoa, 2011, CEIT. [Online]. Available: http://www.ceit.es/

[35] Lander Simulation & Training Solutions, S.A., 2010. Tutor. [Online]. Available: http://www.landersimulation.com/eng/solutions/automotive/

[36] Basler scout sca1400-30fm cameras, 2011, Basler Vision Technol. [Online]. Available: http://www.baslerweb.com/beitraege/unterbeitrag_en_36374.html

[37] L. M. Bergasa, J. Nuevo, M. A. Sotelo, R. Barea, and E. López, "Real-time system for monitoring driver vigilance," *IEEE Trans. Intell. Transp. Syst.*, vol. 7, no. 1, pp. 63–77, Mar. 2006.

[38] B. Kay, "How can the instructor improve the long-term education process on the simulator," in *Proc. 10th INSLC*, 1998, pp. 3.1–3.4.

[39] P.-H. Ting, J.-R. Hwang, J.-L. Doong, and M.-C. Jeng, "Driver fatigue and highway driving: A simulator study," *Physiol. Behav.*, vol. 94, no. 3, pp. 448–453, Jun. 2008.

[40] H. B. Lee, J. M. Choi, J. S. Kim, Y. S. Kim, H. J. Baek, M. S. Ryu, R. H. Sohn, and K. S. Park, "Nonintrusive biosignal measurement system in a vehicle," in *Proc. Conf. IEEE Eng. Med. Biol. Soc.*, 2007, pp. 2303–2306.

[41] T. Sheerman-Chase, E.-J. Ong, and R. Bowden, "Online learning of robust facial feature trackers," in *Proc. IEEE 12th ICCV Workshops*, Oct. 2009, pp. 1386–1392.

[42] Y.-T. Lin, C.-M. Huang, Y.-R. Chen, and L.-C. Fu, "Real-time face tracking and pose estimation with partitioned sampling and relevance vector machine," in *Proc. IEEE ICRA*, May 2009, pp. 453–458.

[43] E. Murphy-Chutorian and M. Trivedi, "Hyhope: Hybrid head orientation and position estimation for vision-based driver head tracking," in *Proc. IEEE Intell. Veh. Symp.*, Jun. 2008, pp. 512–517.

Pedro Jiménez received the M.S. degree in telecommunications engineering and the Ph.D. degree from the University of Alcalá, Madrid, Spain, in 2005 and 2011, respectively.

He started as a Researcher with the University of Alcalá in 2004 and, since then, has tackled problems in electronics, field-programmable gate arrays, and computer vision. In the last field, he has worked on automotive and face-tracking applications. He is currently with Gamesa, Zamudio, Spain, working in the field of renewable energies.

Luis M. Bergasa received the M.S. degree in electrical engineering from the Technical University of Madrid, Madrid, Spain, in 1995 and the Ph.D. degree in electrical engineering from the University of Alcalá, Madrid, in 1999.

He is currently a Full Professor with the Department of Electronics, University of Alcalá. He has been Head of the Department of Electronics during the period 2004–2010. He has been the Coordinator of the RobeSafe Research Group since 2010. Since 2009, he has been the Research Director of Vision Safety Technologies Ltd., which is a spin-off company established to commercialize computer vision systems for road infrastructure inspection. He is an Associate Editor of the *Physical Agents Journal*. He is the author of more than 150 refereed papers in journals and international conference proceedings. He is the holder of six national patents and one Patent Cooperation Treaty patent. His research interests include real-time computer vision and its applications, particularly in the field of assistant robotics and intelligent transportation systems.

Dr. Bergasa is a member of the IEEE Robotics and Automation Society Technical Committee on Autonomous Ground Vehicles and Intelligent Transportation Systems. He has received eight Spanish prizes related to the robotics and automotive fields from 2004–2011. He is a habitual reviewer in 12 journals included in the JCR index. He has served on program/organizing committees in more than 15 conferences.

Jesús Nuevo received the Ph.D. degree from the University of Alcalá, Madrid, Spain, in 2009.

He has worked on driver monitoring systems using computer vision since 2004. He is currently an OCE Postdoctoral Fellow with the Commonwealth Scientific and Industrial Research Organisation ICT Centre Computer Vision Group, Brisbane, Australia. His research interests are theoretical and applied computer vision, especially in face tracking, and a nonrigid structure from motion.

Noelia Hernández was born in Valladolid, Spain, in 1984. She received the Bachelor of computer science and the Masters degree in advanced electronic systems (intelligent systems) from the University of Alcalá, Madrid, Spain, in 2007 and 2009, respectively. She is currently working toward the Ph.D. degree with the Department of Electronics, University of Alcalá.

Her research interests include fuzzy control, WiFi localization systems, robotics, image processing, and intelligent vehicles.

Iván G. Daza received the M.S. degree in telecommunications engineering and the Ph.D. degree from the University of Alcalá, Madrid, Spain, in 2004 and 2011, respectively.

He worked on drowsy driver detection systems. His research interests include computer vision, pattern recognition, machine learning, stochastic process optimization, and control theory.