

Vision-Based Distraction Analysis Tested On A Realistic Driving Simulator

N. Hernández, P. Jiménez, L.M. Bergasa
Department of Electronics
University of Alcalá, Madrid (Spain)

{nhernandez,pjimenez,bergasa}@depeca.uah.es

B. Delgado, M. Sevillano
Research and Project Department, ESM
Asturias (Spain)

{beatriz,matiassevillano}@esm.es

Abstract— This paper presents a non intrusive approach to obtain driver's face pose estimation based on stereo gray-level image processing. Face pose estimation is based on an automatic and incremental 3D model creation and its correct tracking. From this information, gaze focalization in the scene is calculated in order to detect driver distraction. Different distraction activities are inferred in a realistic simulator and a study of the incidence of these distracting activities in the driver's behaviour is carried out. Some experimental results and conclusions are presented.

I. INTRODUCTION

Driving inattention is a major factor to highway crashes. The National Highway Traffic Safety Administration (NHTSA) estimates that approximately 25% of police-reported crashes involve some form of driving inattention [1]. The study of AAA FTS (American Automobile Association Foundation for Traffic Safety) showed the driving attention status has five categories: attentive, distraction, cognitive distraction (looking without seeing), fatigue and unknown [2]. In this paper, we will focus on distraction category.

Driving distraction is defined as "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" [3]. Thirteen types of potentially distracting activities are listed in [2]: eating or drinking, outside person, object or event, talking or listening on cellular phone, dialling cellular phone, using in-vehicle-technologies, etc. Since the distracting activities take many forms, NHTSA classifies distraction into 4 categories from the view of the driver's functionality: visual distraction, cognitive distraction, auditory distraction (e.g., responding to a ringing cell phone), and biomechanical distraction (e.g., manually adjusting the radio volume) [1]. 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). As it can be seen, driving distraction is more diverse and commits more risky factor than fatigue and it is present over half of inattention involved crashes, resulting in as many as 5,000 fatalities and \$40 billion in damages every year [2]. Increasing use of in-vehicle information systems (IVISs) such as cell phones, GPS navigation systems, DVDs and satellite radios has exacerbated the problem by introducing additional sources of distraction [4]. Enabling drivers to

benefit from IVIS without diminishing safety is an important challenge.

One promising strategy involves classifying the driver state and then using this classification to adapt the in-vehicle technologies to mitigate the effects of distraction. The literature contains 3 main categories according to the measurement signals they use to detect distractions: biological signals, driving signals and driver images. The biological signal processing approaches directly measure biological signals (EEG, ECG, EOG, EMG, etc) from driver's body and as consequence they are intrusive systems [5], [6]. Only few works, focusing in cognitive distractions, have been reported using this kind of approach. The main reason may be that using biological signal to analyze distraction level is too complicated and no obvious pattern can be found. Vehicle signal reflects driver's action, then measuring it, the 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 [7] [8] [9]. The advantage of these approaches is that the signal is meaningful and its acquisition is quite easy. This is the reason because a few commercial systems existing nowadays use this technique [10] [11] [12]. However, they are subject to several limitations such as vehicle type, driver experience, geometric characteristics, condition of the road, etc. Then, these procedures require a considerable amount of time to analyze user behaviours and therefore, they do not work with the so called micro-sleeps when a drowsy driver falls asleep for a few seconds on a very straight road section without changing the vehicle signals. The image processing based approaches are effective because of the occurrence of distraction are reflected through the driver's face appearance and head/eyes activity. Different kinds of cameras and analysis algorithms have been employed in this approach: methods based on visible spectrum camera [13], [14]; methods based on IR camera [15] [16] [17] [18] and methods based on stereo cameras [19], [20]. Some of them are commercial products as: Smart Eye [17], Seeing Machines DSS [18], Smart Eye Pro [19] and Seeing Machines Face API [20]. However, these commercial products are still limited to some well controlled environments, so there is still a long way to go in order to estimate driver's distraction state. On the other hand, most works existing in the literature was carried out on visual distraction detection, less on cognitive and none on auditory and biomechanical distraction detection.

From this review of the state of the art we can conclude that in-vehicle and portable information and entertainment technologies are emerging rapidly, making it increasingly difficult to determine the scope of the potential distraction problem. 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 problem using realistic situations. There are different proposals to detect distractions but, to date, they are focusing in only some kind of distractions and they do not solve the problem in a general way.

Simulation is an optimal method of experimentation to acquire knowledge of driver's behaviour. The simulation methodologies applied in Europe to the road transport sector research are demonstrating their profitability and efficiency [21]. The main objective through the simulation, is to immerse the driver in his normal work environment. In order to do this, a cockpit fully equipped is required to perform the driving task. Moreover, it must be valid for the driver to react naturally to the task entrusted by the trainer. The driving simulator used for the development of this work, known as TUTOR [22], is a very high scope simulator for truck and bus, meaning that achieves a maximum level of immersion.

This paper presents a non intrusive approach to obtain driver's face pose estimation based on stereo image processing techniques. The paper is organized as follows: section II shows the experimental environment where the tests have been done. The algorithm for visual-based distraction detection is described in section III. Finally, the experimental results as well as the conclusions are detailed in sections IV and V respectively.

II. EXPERIMENTAL ENVIRONMENT

This section describes the environment where the tests have been carried out. It is divided into five main sections. First of them describes the physical simulator. Then, the camera vision system, experimental protocol and subjects selection are explained. Finally, the steps for test validation are exposed.

A. Truck Driving Simulator

The experiments were done in the facilities of CEIT [23] in San Sebastián (Spain), in a room with controlled light and sound environment.

The simulator [22] (Fig. 1) consists of a real truck cabin equipped with the common on-board devices: GPS, Tachograph, Hands-free and On-board Computer. These devices send information to a central host for a posterior analysis. Fig. 2 shows a representative diagram of the devices setup.

The visualization system is made up of three wall-backprojection system with a total surface of $22m^2$. The fact that the screens have no marked separation plus the geometry of the image system makes for a flawless overall impression. Moreover, two monitor screens are used as mirrors.

The cabin is assembled on a movement platform with 6 degrees of freedom on which drivers can feel the vehicle accelerating, braking, its centrifugal force, etc.

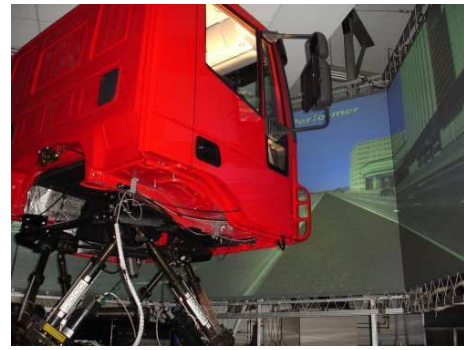


Fig. 1. Real truck cabin simulator

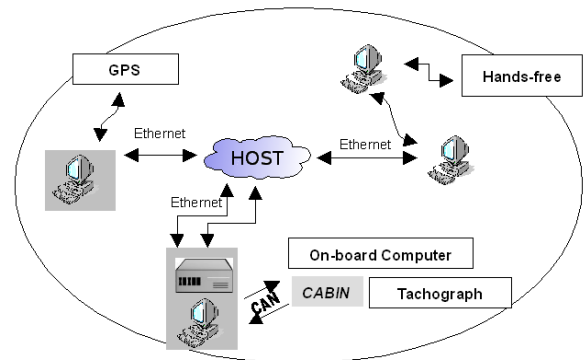


Fig. 2. Host and Devices Setup

Main computers are placed in the simulator Instructor Position (IP) located behind the cabin, which is surrounded by panels to create a closed environment (Fig. 3). The configuration of the simulator and the scenario can be changed from the IP. The information of previous tests can be consulted as well.



Fig. 3. Instructor Position

B. Camera Vision System

The designed hardware for image capture is divided into three parts: the stereo capture system, the illumination controller and the infrared illumination system. This system is located on the truck dashboard, in front of the driver pointing to his face. The cameras are separated a distance of 20cm and the driver is placed at 60 - 100 cm from them (Fig. 4).

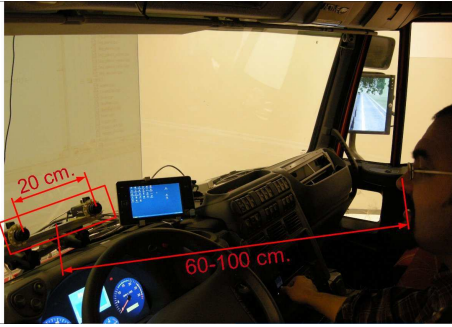


Fig. 4. Camera Vision System

C. Experimental Protocol

To design the experimental protocol we have based on the following initial hypothesis: "The potential driver distraction due to on-board devices is determined by the level of attentional demand required by them while driving, decreasing the effectiveness of the primary task: driving."

By analysing the professional drivers behaviour, the basic and most representative features in the context of this activity are identified [24]. Some scenarios, types of vehicles, incidents, on-board systems and critical situations are selected. Thus, the professional drivers behaviour should be generically represented.

Taking into consideration this basis, that involves observation and information recording during the task of driving, the next step is to define the basic simulation exercises.

These exercises are defined through the editing module from the simulator PI, we can set up scenarios, incidences, vehicle characteristics, load distribution, road conditions, weather, etc.

Tests have been designed with the goal of refuting the initial hypothesis of the research regarding the potential distraction of four different on-board systems which are commonly used in professional driving task. These devices are digital tachograph, GPS, hands-free and on-board computer.

Thus, these devices are on board in all the exercises but they will not always have a fundamental interaction to achieve the objective of the test.

Under these conditions, four exercises have been created (inter-city, mountain, urban and long-distance) based on three different scenarios (inter-city, mountain and urban). Each one includes different incidences: motor, tires or ABS breakdown and different vehicle incidences such as sudden brake of the precedent vehicle, broken down vehicles on the road, vehicles running a red light, etc. (see Table I) .

These exercises were implemented in 16 tests: five of them were based on the inter-city exercise, four on the mountain, three on the urban and the last four on the long-distance one.

The defined procedure to evaluate these tests consists on different drivers driving through different scenarios.

The first test of each exercise has been called "Control Test" which corresponds to the exercise undertaken by each driver in the different scenarios aseptically.

TABLE I
EXERCISES CONFIGURATION

Exercise	Incidences	On-board Devices
Inter-city	Vehicle incidences	GPS
	Motor breakdown	Hands-free
Mountain	Vehicle incidences	GPS
	Tire breakdown	Hands-free
		Tachograph
Urban	Vehicle incidences	GPS
	Motor breakdown	Hands-free
	ABS breakdown	Tachograph
Long-distance		On-board Computer
		Hands-free
	Vehicle incidences	Tachograph

It is important to have these control tests, because using them, the behaviour at different points of the scenarios in the subsequent tests of each driver can be compared and defined as distraction sources. Thus, a differential analysis of the produced incidents could be done.

Once the chain of exercises is finished, we have enough information about the drivers behaviour while driving in order to generate a distraction pattern for each one.

To analyse drivers behaviour a visual record of the driver is done.

D. Subjects

According to the previous considerations, the number of tests and their configuration, we have defined a minimum number of participants of 12 in order to have one participant for each test configuration to detect the dependent behaviour variables.

It is important to highlight that every participant needs to pass a test to exclude people with propensity to suffer simulator-sickness. Previous studies with similar conditions used groups from 7 to 30 participants [25] [26].

All subjects were informed of the purpose of the experiment and the security procedures in the simulator facilities.

E. Test Validation

The test model is validated following a three stages strategy:

- STAGE 1: A group of drivers with information about the technological tools and distraction sources while driving. In this stage, the exercises and tests have been designed basing on the analysis of task and taking into account the objectives and the underlying assumptions of the investigation. The main objective of this stage is the validation of the designed tests.
- STAGE 2: Professional drivers group. At this stage a group of 5 professional drivers who know the objectives of the research and the simulation environment advise the researchers to improve the exercises and tests to set up a simulation environment more realistic for them.

- **STAGE 3: Final drivers group.** A representative sample of drivers are selected for this group. It is composed of at least 12 drivers from different gender, age, experience, etc to obtain conclusions about distraction and driver's behaviour.

III. ALGORITHM FOR VISUAL-BASED DISTRACTION DETECTION

This section presents in detail the pose estimation algorithm using a fully automatic and incremental 3D face model creation. A block diagram of the algorithm is shown in Fig. 5. The method consists of the following steps:

- 0) **Calibration** of the stereo-rig. After this off-line process, all the following operations are fully automatic.
- 1) **Model creation.** Face is searched on the images using the Viola & Jones (V&J) algorithm. We use the Harris corner detector [27] to find interesting features on the driver's face. A 3D model of the face is created with the 3D coordinates of the points.
- 2) **Feature tracking.** 3D points are tracked over images. We use a single historic of 2D image patch projections from both right and left camera images in conjunction, along with its associated yaw rotation angle.
- 3) **Pose estimation.** We estimate the pose of the model by means of the POSIT algorithm within a RANSAC [28] process to reject tracking errors and outliers.
- 4) **Model correction.** After pose estimation, 3D model may be increased as required to previously occluded parts of the face. A background bundle adjustment [29] optimization refines the model as long as we add new 3D points to the model.

A. Face Model

The first step to accomplish the pose estimation is to obtain a 3D model of the face, made up of the 3D coordinates of the features which will be used for tracking and later model reconstruction. First, the face is detected using the Viola and Jones algorithm [30]. The first frame in which a frontal face is detected in both of the cameras is chosen for model creation.

Features are extracted using a Harris corner detector. If an initial Harris detector were applied over the whole detected face, resulting points would tend to group within the more contrasted areas. To have a spread distribution of initial features all over the face, the face is divided into 5x5 cells, and Harris points are taken from each cell (Fig. 6).

The next step is to accomplish the stereo matching of the features. This is explained in detail in sec III-C.

Having the projection of the point in both images, now we obtain the 3D coordinate of each model point out of the feature 2D projections. Fig. 7 shows the created 3D model. An extra filtering process is used to ensure that points for which the 3D coordinate are incorrectly computed are rejected from model. This filtering include aspects such a standard face size not bigger than 15x15x18 cm, located at around 60 to 90 cm far from the cameras and cylindrical shaped face constraints.

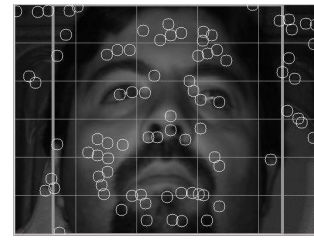


Fig. 6. Viola & Jones detected face, and initial distribution of feature candidates to extract the 3D model

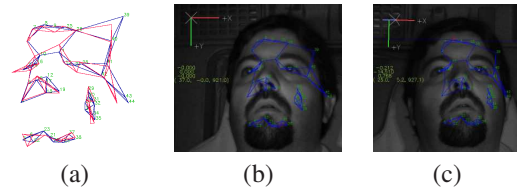


Fig. 7. (a) 3D view of the model. In red, correction done by Bundle Adjustment. (b) Projections over the left image and (c) right image.

B. Feature matching

Conditions for feature matching are very complicated in this kind of scenarios. A face does not typically present high contrasted features or corners. Moreover, the limited illumination leads to dark images, what makes features even less contrasted. Another important issue derived from low illumination is poorly focused images. The iris of the camera must be wide open to allow the maximum light into the sensor, what drastically reduces depth of field. This produces that some parts of the face, specially if the driver moves forward or backward, might not be well focused.

In such case in which the feature changes its appearance because of the angle of projection, it usually works better matching small patches of the image. On the other hand, if the image is not well focused, and to avoid incorrect tracking because of repetitive appearance in the face, it works better to use big patches.

To overcome this problem, the chosen matching method is based on the addition of the matching result for three different size patches around the feature. Thus, to obtain the correspondence of a 3D point from one image into another, three patches of different sizes are extracted around the feature point, and matched over the search area of the other image. We compute the most likely result as the maximum of the sum of the three matching results. Figure 8 illustrates the matching of a feature at model creation.

C. Feature tracking

The main problem to solve when a face is tracked is the changing appearance of the feature points as the face rotation angle changes. As features are not planar in shape, in general it is not a good solution to try a patch homography to correct appearance changes due to rotations. Moreover, this process is costly, and often need some a-priori information about the orientation of the patch in the 3D space. In our approach, we use the different view angles of the face from the two

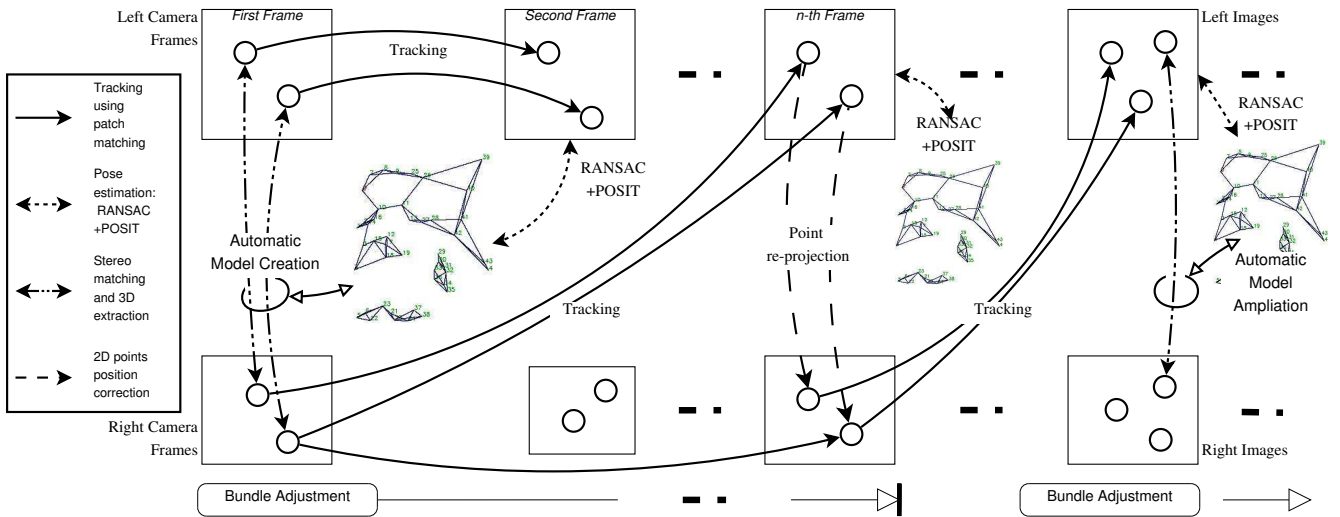


Fig. 5. Block diagram of the algorithm

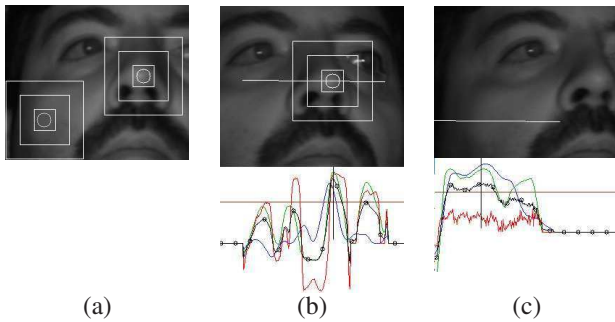


Fig. 8. Feature Points on first image (a), and its corresponding epipolar line on second image (b,c). The graph in (b) and (c) shows the result matching of patch size 10x10 (red), 25x25 (green), 50x50 (blue), and sum result (black), all restricted to the epipolar line. In (b) the matching is not correct, and point is discarded

cameras to try to advance the appearance that a feature will have after some rotation.

At initialization, a patch for each model point is stored with its associated projection angle for each image plane. The tracking is initially done with original image patches. But as the driver moves the appearance of the feature points changes. Let say the current rotation angle with respect to a camera is equal to a stored one. In that case the appearance of the 2D projection over the camera frame of the feature is almost the same to that stored patch associated to the angle. When the angle for any of the cameras approaches to any of the stored ones, the tracking switches its associated patch, previously stored, and not necessarily from the same camera. Thus, the tracking result is very accurate (see Fig. 9).

D. Pose Estimation

After the position of the tracking points has been updated for left and right frames, the 3D face pose is estimated from the 2D projection of each point. However, the matching process may not succeed for all points, and can result in

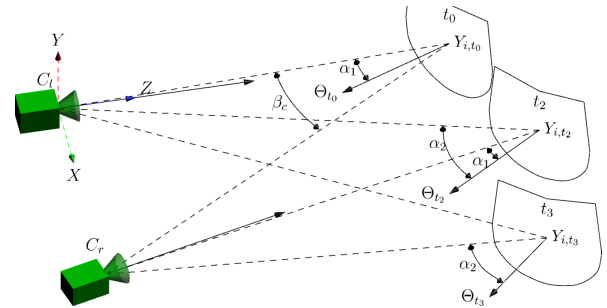


Fig. 9. Model is initially created with pose Θ_0 . C_l sees Y_i with an angle α_1 . After some rotation, at t_2 , now C_r sees Y_i with the same angle α_1 , and consequently with a very similar appearance to that saw by C_l at t_0 , which was stored. At the same time, C_r stores the new appearance. At t_3 , the camera C_r sees Y_i such as C_l did at t_2 , with and angle α_2

errors. Thus, a robust optimization method is required to estimate the best 3D face pose matching, that would detect as outliers the points that have been incorrectly tracked, so they can be safely discarded. The RANSAC algorithm is used to eliminate the outliers. 3D pose is obtained using DeMenthon's four point iterative pose estimation algorithm (POSIT) [31]. The POSIT algorithm calculates the pose of a 3D rigid object from its projection on a single image.

The pose is given as a translation T and a rotation R matrices, which indicate the position of the central point of the model with respect to the camera coordinate system, and its rotation from the initial given model.

On each RANSAC iteration, seven points are randomly selected from the model and used to calculate the pose using the POSIT algorithm. To detect the outliers, error is calculated as the euclidean distance between the tracking point and the reprojected using the estimated pose.

E. Model Corrections

The 3D model is initially created using a frontal view of the face. Consequently, for yaw rotations wider than $\pm 40^\circ$

approx, most of the points of the model are occluded. This makes necessary to augment the model with new points from parts of the face which were previously occluded. When the number of visible points from a camera falls below a threshold, usually 10, new points within the new areas of the face are searched and added to the model.

As many point candidates for addition may lay out outside the face itself. Those points are filtered attending to face shape and size constraints mentioned above. However, new 3D point coordinates will have an increased error derived from pose estimation at this moment. The 3D points taken during model creation are also subject to error derived from stereo correspondences. In order to get a better fitting of the model to the face and correct those errors, a bundle adjustment (BA) optimization is used to refine the 3D model. This corrects the 3D point coordinates and the poses in which any point has been added. To save on computational load, this stage is only applied at certain *keyframes*, t_k when a minimum movement has been detected in the pose, only during certain time after model creation or when a point has been added, until error falls below a threshold. Figure 7(a) shows the corrected model.

F. Gaze Focalization Interface

To study the influence of the different sources of distraction, a visual interface for gaze focalization representation has been designed. A projection of the obtained angles on the image plane is done in order to obtain the gaze focalization coordinates. This interface consists of two different parts. First of them is a visual interface which shows the gaze focalization over the real visual information at every frame. This way, it is possible to know where the driver was looking while an incidence was happening. This interface also classifies the actual direction into the different interest points as shown in Fig. 10.



Fig. 10. Gaze focalization visual interface

The second one is a graph which shows the coordinates where the driver was looking along the time. Using this interface is possible to study how long the driver was looking

to the different interest points in the cabin or how long the driver remains with his gaze fixed on a point.

IV. EXPERIMENTAL RESULTS

In this section the obtained experimental results are shown.

The algorithm has been tested over long sequences, more than ten minutes each one, of 12 different users. Two of them have been compared to its corresponding ground-truth. The sequences were recorded with professional drivers for all the exercises simulating very high immersion conditions and common driving disturbances, such as phone calls, handling the GPS, traffic conditions and takeovers.

The ground truth was obtained using a calibration pattern as a hat on the user's head, as shown in Fig. 11. This new form of evaluation was used because the traditional three spot lights on the driver's head projected corresponding flares on the windscreen, due to low illumination conditions inside the cabin, disturbing them.

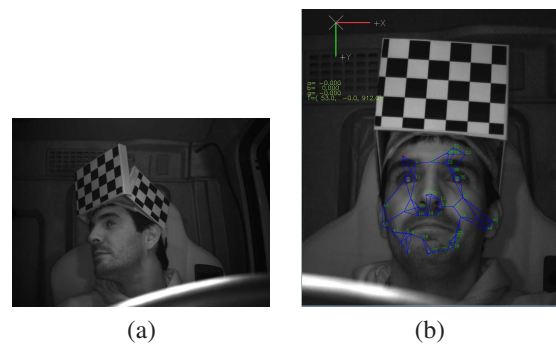


Fig. 11. Calibration pattern on the driver's head used as ground-truth. (a) 1392x1040 image, from left hand camera. (b) Cropped image from right hand camera, showing the 3D model. A condition is imposed at model creation to ensure that the calibration pattern does not make part in the 3D model.

Restrictions have been added to the model creation process to ensure that the hat does not make part of the model, so the algorithm evaluation is not influenced by the ground-truth process.

After the ground-truth has been evaluated, it is compared with the estimated pose from the algorithm. It can be observed in Fig. 12 how the accuracy increases after around 200 seconds of execution, moment at which the resulting residual error obtained by Bundle Adjustment is very small, and new executions of bundle adjustment do not increase the accuracy of the model. Under these conditions we stop the Bundle Adjustment optimization until new points are added to the model. The results in Table II compare the accuracy obtained with and without using bundle adjustment. The error results with bundle adjustment have been computed starting from second 200 of execution, after what no more points were added to the model.

The algorithm has been coded in c/c++, and tested in a 2.4GHz Core2 Duo processor. The most time consuming task, the patch matching, has been coded using the OpenCV library, which implements the patch matching using fast Fourier transforms what drastically reduces execution times.

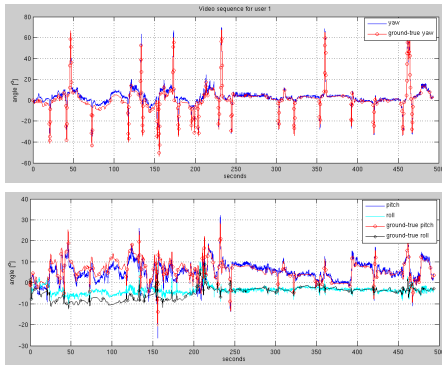


Fig. 12. Graphics with the yaw (up), pitch and roll (down) and ground truth of a section of a video sequence from user 1. Positive and negative peaks on yaw indicates when the driver is looking at the right or left mirror respectively. Peak on the pitch indicates when driver is looking to the on-board computer, the hands-free or the tachograph.

TABLE II

MEAN SQUARE ERROR ON POSE ESTIMATION ANGLES

Rotation	$\alpha < 15^\circ$	$\alpha < 30^\circ$	$\alpha < 45^\circ$	$\alpha \geq 60^\circ$
<i>yaw</i>	6.52	12.09	16.58	18.34
<i>yaw</i> with BA	1.64	4.47	6.01	8.30
<i>pitch</i>	3.82	7.8864	8.5947	-
<i>pitch</i> with BA	1.61	1.80	6.74	-
<i>roll</i>	5.27	16.62	-	-
<i>roll</i> with BA	1.16	4.06	-	-

The Bundle Adjustment is usually a time consuming task, so it is only executed at an interval of 20 frames, and in most of the cases does not extend more than 100 seconds after last points addition. We use the free software library SBA for this purpose [32]. Table III shows execution times for the algorithm. Mean times have been computed over 1 second of execution. An average frame rate of 26 fps has been achieved while the BA is working, and 31 fps when it is stopped. Although the processing time can peak up to 57 ms for some key frames, this is not an issue for working in real time, since the camera buffers filter these peaks.

TABLE III
PROCESSING TIMES

Task	Mean time	Max time
Model creation	453 ms	943 ms
Patch matching	16 ms	23 ms
POSIT + RANSAC	15 ms	18 ms
Points addition and BA	7 ms	16 ms
Whole process	38 ms	57 ms

To study the influence of the different sources of distraction the visual interface and gaze focalization graph (Fig. 13) is used. Analysing this information we can conclude that most of the time driver remains looking at the front, usually the horizon or the precedent vehicle. However, it is important to distinguish between these long periods of time (when the coordinates where the driver was looking change slightly) and periods when the driver has fixed gaze

(when the driver's attention level is low). Moreover, other elements such as mirrors or the on-board computer are the most common points where drivers look while driving, a decrease in the number of looks to them can be consider as a sign of attention level decrease.

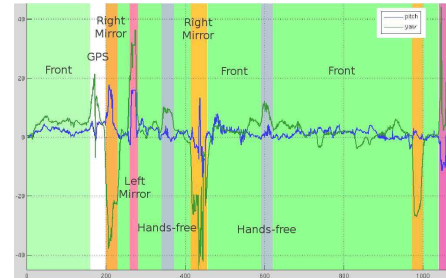


Fig. 13. Gaze focalization graph

Look to other interest points such as GPS or hands-free can produce distraction, thus it is important to study the reaction times after different incidences (sudden brake of the precedent vehicle, vehicles running a red light, etc), the driver behaviour during a call or while following the instructions from a GPS, etc.

Fig. 14 shows, as a example of driver's behaviour while driving, the reaction time from the moment an incoming phone call is produced until the driver picks up the phone and the coordinates where the driver is looking during it.



Fig. 14. Driver's head

The signal corresponding to the driver's gaze direction during the phone conversation is more chaotic than after the conversation. It is important to highlight that this particular call has been done during a high cognitive activity period as the looks to the mirrors and the high variance of the gaze direction shows. The high reaction time and the high variability in the gaze direction are due to distraction. In the final version more results about distraction will be included.

V. CONCLUSIONS AND FUTURE WORKS

A. Conclusions

This paper presents a face tracking and pose estimation algorithm to detect visual distractions, using a stereo vision

system. The algorithm is able to automatically construct a 3D model of the face, just requiring for the driver to look straight ahead for a few seconds. Tracking of feature points is carried out using mixed views from both cameras. Incorrectly tracked points are rejected using RANSAC, and 3D pose is recovered from the set of points using a fast algorithm, such as POSIT.

Thanks to model extension and Bundle adjustment the algorithm works reliably for the whole yaw rotation face range, $\pm 90^\circ$ degrees.

The Bundle adjustment method allows safely point addition to the model even assuming that there is an error on the pose estimation of the model to which the point is being added.

It has been demonstrate that the system it is valid to evaluate the effects of the different sources of distraction allowing to evaluate reaction times and driver behaviour while driving with different kinds of incidences.

B. Future Works

To improve the estimated driver's gaze direction a method to calculate eyes direction is being designed. To achieve this purpose the described matching method is used to locate the center of the pupils and their height and width. The fusion of both face and eyes direction will improve the accuracy at little changes on the gaze direction when the driver is moving only his eyes.

Another variables such as truck position in the lane are being studied to help to find the behaviour pattern during distraction.

Finally, with the designed interface the final conclusions about distraction and driver's behaviour while driving can be obtained. This way, a method to analyse distraction could be designed.

VI. ACKNOWLEDGEMENTS

This work has been financed with funds from the Ministerio de Ciencia e Innovación through the project DRIVER-ALERT (TRA2008 - 03600), as well as from the project CABINTEC (PSE-370000-2009-12).

REFERENCES

- [1] T. A. Ranney, E. Mazzae, R. Garrott, and M. J. Goodman, "Nhtsa driver distraction research: Past, present, and future," in *Transportation Research Center Inc., Tech. Rep.*, July 2000.
- [2] J. C. Stutts, D. W. Reinfurt, L. Staplin, , and E. A. Rodgman, "The role of driver distraction in traffic crashes," in *American Automobile Association AAA Foundation for Traffic Safety, Tech. Rep.*, 2001.
- [3] K. Young, M. Regan, , and M. Hammer, "Driver distraction: a review of the literature," in *Monash University Accident Research Centre, Tech. Rep.*, 2003.
- [4] T. A. Ranney, "Driver distraction: a review of the current state-of-knowledge," in *NHTSA. DOT HS 810 704.*, April 2008.
- [5] C. Berka, D. J. Levensowski, M. N. Lumicao, A. Yau, G. Davis, V. T. Zivkovic, R. E. Olmstead, P. D. Tremoulet, , and P. L. Craven, "Eeg correlates of task engagement and mental workload in vigilance, learning, and memory tasks," in *Aviation, Space, and Environmental Medicine*, vol. 78(5), May 2007, pp. B231–B244.
- [6] B. T. Skinner, H. T. Nguyen, , and D. K. Liu, "Classification of eeg signals using a genetic-based machine learning classifier," in *Engineering in Medicine and Biology Society: 29th Annual International Conference of the IEEE*, Aug 2007, pp. 3120–3123.
- [7] T. Wakita, K. Ozawa, C. Miyajima, K. Igarashi, K. Itou, K. Takeda, , and F. Itakura, "Driver identification using driving behavior signals," in *IEICE Trans Inf Syst*, vol. E89-D(3), 2006, pp. 1188–1194.
- [8] Y. Takei and Y. Furukawa, "Estimate of driver's fatigue through steering motion," in *IEEE International Conference on Systems, Man and Cybernetics*, vol. 2, 2005, pp. 1765–1770.
- [9] McCall, J.C., Wipf, D.P., Trivedi, M.M., Rao, and B. D., "Lane change intent analysis using robust operators and sparse bayesian learning," in *IEEE Transactions on ITS(8)*, vol. 3, Sep 2007, pp. 431–440.
- [10] "Volvo car corporation [online]," <http://www.media.volvocars.com>.
- [11] "Mercedes-benz [online]," http://www.emercedesbenz.com/Aug08/11_001331_Mercedes-Benz_To_Introduce_Attention_Assist_Into_Series_Production_In_Spring_2009.html.
- [12] "Lexus [online]," <http://www.testdriven.co.uk/lexus-ls-600h/>.
- [13] E. Vural, M. Cetin, A. Ercil, G. Littlewort, M. Bartlett, , and J. Movellan, "Drowsy driver detection through facial movement analysis," in *Springer Berlin / Heidelberg*, 2007.
- [14] M. C. Su, C. Y. Hsiung, , and D. Y. Huang, "A simple approach to implementing a system for monitoring driver inattention," in *Systems, Man and Cybernetics*, vol. 1, 2006, pp. 429–433.
- [15] Q. Ji and X. J. Yang, "Real-time eye, gaze, and face pose tracking for monitoring driver vigilance," in *Real-Time Imaging*, vol. 8, 2002, pp. 357–377.
- [16] L. Bergasa, J. Nuevo, M. Sotelo, R. Barea, and E. López, "Real-time system for monitoring driver vigilance," in *Intelligent Transportation Systems, IEEE Transactions*, vol. 7(1), 2006, pp. 63–77.
- [17] L. Bretzner and M. Krantz, "Towards low-cost systems for measuring visual cues of driver fatigue and inattention in automotive applications," in *Vehicular Electronics and Safety, IEEE International Conference*, Oct 2005, pp. 161–164.
- [18] J. Heinzmann, D. Tate, , and R. Scott, "Using technology to eliminate drowsy driving," in *SPE International Conference on Health, Safety, and Environment in Oil and Gas Exploration and Production*, Apr 2008, pp. 15–17.
- [19] "Smart eye [online]," <http://www.smarteye.se>.
- [20] "Seeing machines [online]," <http://www.seeingmachines.com>.
- [21] *TRAIN-ALL Integrated System for driver Training and Assessment using Interactive education tools and New training curricula for ALL modes of road transport*, Oct 2007.
- [22] "Tutor [online]," <http://www.landertsimulation.com/index.php?id=27&L=1>.
- [23] "Ceit [online]," <http://www.ceit.es/>.
- [24] U. of Plymouth, Ed., *How can the Instructor Improve the Long-term Education Process on the Simulator*. Institute of Marine Studies, 1998.
- [25] P.-H. Ting, J.-R. Hwang, J.-L. Doong, and M.-C. Jeng, "Driver fatigue and highway driving: A simulator study," *Physiology & Behavior*, vol. 94, no. 3, pp. 448 – 453, 2008. [Online]. Available: <http://www.sciencedirect.com/science/article/B6T0P-4RYNMH5-1/2/e19d7d5f224c51c59397cb8de12be573>
- [26] 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," *Conf Proc IEEE Eng Med Biol Soc*, vol. 2007, pp. 2303–6, 2007. [Online]. Available: <http://www.biomedsearch.com/nih/Nonintrusive-biosignal-measurement-system-in/18002452.html>
- [27] C. Harris and M. Stephens, "A combined corner and edge detector," in *Proc. Fourth Alvey Vision Conference*, 1988, pp. 147–151.
- [28] M. A. Fischler and R. C. Bolles, "Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography," *commun. ACM*, vol. 24, no. 6, pp. 381-395, 1981.
- [29] B. Triggs, P. McLauchlan, R. Hartley, and A. Fitzgibbon, "Bundle adjustment – a modern synthesis," in *Vision Algorithms: Theory and Practice*, ser. LNCS, W. Triggs, A. Zisserman, and R. Szeliski, Eds. Springer Verlag, Sep 1999, pp. 298–375.
- [30] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," in *IEEE Conference on Computer Vision and Pattern Recognition 2001*, 2001, pp. 511-519.
- [31] D. F. Dementhon and L. S. Davis, "Model based object pose in 25 lines of code," *int. J. Comput. Vision*, vol. 15, no. 1-2, pp. 123-141, 1995.
- [32] M. I. A. Lourakis and A. A. Argyros, "Sba: A software package for generic sparse bundle adjustment," *ACM Transactions on Mathematical Software*, vol. 1, no. 36, pp. 1–30, 2009.