Cost-Efficient Driver State and Road Conditions Monitoring System for Conditional Automation

Arman Allamehzadeh[1], Jesus Urdiales de la Parra[2], Fernando Garcia[2] *Member, IEEE* and Cristina Olaverri-Monreal[1*] *Member, IEEE*

***Abstract*— Driver State Monitoring Systems (DSMS) play a crucial role to determine whether the driver is prepared to take control of the vehicle. In this paper we present a cost-efficient, context-aware system based on smartphones that supports the driver in the driving task by monitoring their attention and detecting pedestrians, vehicles and drowsiness. The detection rates obtained, with 85% in the case of pedestrians and over 90% in the case of vehicles, along with the results of the driver state measured by eye movements proved the viability of the proposed approach to provide situational awareness in the novel paradigm of autonomous driving.**

I. INTRODUCTION

In conditional automation a driver's response to a Take Over Request (TOR) is expected in case of a sudden event [1]. Driver State Monitoring Systems (DSMS) therefore play a crucial role to determine whether the driver is prepared to take control of the vehicle.

According to a report from Pew Research [2], in addition to using mobile phones to make calls, over half of mobile phone users use their phone to retrieve information and over 40 percent for entertainment. For this reason, mobile phones are most of the time connected to the nearest mobile phone network tower currently providing coverage with uninterrupted registering of user location. In the context of digital technologies that result from pervasive computing, today's smart devices already integrate a variety of cost-efficient embedded sensors (i.e. accelerometer, digital compass, gyroscope, GPS, microphone, camera) that facilitate the acquisition of data which can be used to study driver behavior for road safety purposes [3]. According to the data from International Data Corporation [4], the smartphone market has been dominated by Android for the last four years. Relying on the sensor technology available in Android devices and to promote collision avoidance and road safety, this work implements a DSMS that also considers road conditions for later dissemination to the driver through appropriate warnings. We intend to support the driver in the driving tasks: control of speed, distance to leading car, traffic observation and other road users action prediction. Drowsiness and hypovigilance might lead to reduced situational awareness while manually operating a vehicle or being involved in other tasks in conditional automation. The system detects driver state by measuring eye movements, which can tell us how inattentive or drowsy the driver is.

Relying on the application in [5], an application to monitor driver attention through a context-aware system relying on the sensor technology that modern smart devices provide, we present in this paper a system with additional functionality and improved algorithms for the detection of pedestrians, vehicles and drowsiness.

II. RELATED WORK

Several companies have commercialized systems to assist the driver in the driving task. In the area of visual information processing, the company Mobileye offers detection of road users such as pedestrians or animals and also other objects such as street signs and traffic lights [6].

More cost effective mobile applications focus on face and eye detection. Most of them are intended for entertainment purposes so that they can be shared on social media applications [7] but others are a component of projects that aim to develop a robust eye localization platform based on low cost hardware [8]. Making use of the front camera of a mobile device, further applications scan the users eye retina patterns for identification purposes or detect user's eye movements to use as input modality to interact with a system [9].

In the field of Driving Assistance Systems the use of mobile phone sensors to acquire Floating Car Data (FCD) for the evaluation of traffic conditions and driving performance [3] has also been a topic of research. Some additionally use steering wheel angle, like the attention assist system by Mercedes-Benz based on visual driving data and acoustic feedback [10]; the driver alert control by Volvo [11], based on road features tracking and driving data or the Lexus driver monitoring system [12] and BMW driving assistant [13], that incorporate systems for eye tracking and automatic braking in case of an on road obstacle detection.

In line with this, cost effective mobile applications include the capability of detecting road traffic signs and monitoring the road alerting the driver about potentially unsafe situations, such as imminent vehicle collision, lane departure and speed limits [14], [15], [16].

One of the factors for driver's inattentiveness can be fatigue. The authors in [17] described several prevalent indications of fatigue in drivers such as yawning or slower reaction and responses. They also stated that different individuals showed different symptoms to varying degrees and therefore, there is no concrete method of measuring the level

[1] University of Applied Sciences Technikum Wien, Department of Information Engineering and Security, Vienna, Austria.
arman4u@gmail.com, olaverri@technikum-wien.at
[*] Corresponding author
[2] Intelligent Systems Lab (LSI) Research Group, Systems and Automation Department, Universidad Carlos III de Madrid, Spain.
jurdiale@pa.uc3m.es, fegarcia@ing.uc3m.es

of fatigue. It can therefore be understood that combined techniques of extracting fatigue related patterns based upon computer vision, driving performance and driver's physiological characteristics provide the most accurate results.

III. General Description

The system presented continues and extends the work presented in [5]. It put forth a novel architecture developed for mobile phones which is able to analyze road conditions and take action, requesting the driver to take control of the vehicle after gaze behavior analysis. The work presented in this paper goes beyond this paradigm, providing drowsiness detection and implementing advanced road user detection i.e. vehicles and pedestrians. Thus the application is able to analyze the road state, providing computer vision-based vehicle and pedestrian detection with accurate monocular-based localization. Furthermore, the application provides state-of-the-art drowsiness detection, improving driver behavior analysis in comparison with the previous version.

Pedestrian detection is based on the approach presented in [18] and adapted for current development. In addition, a novel vehicle detection based on a monocular camera and the internal sensor was developed. The distance estimation algorithm allowed the distance to the preceding vehicles to be provided. This distance can be used for further identification of risks, thus enabling the system to alert the driver of an emergency.

The device is configured to retrieve information from cameras, rear and front. The rear camera is facing the exterior of the vehicle and is used for road monitoring tasks, and the front camera is facing the interior and is used for driver monitoring.

IV. Road Monitoring Process: Pedestrians and Vehicles Detection

In this section, we present a novel vehicle detection system based on a low-cost mobile application, based on three different stages: vision-based detection, distance estimation and vehicle tracking. First, the vehicles are identified based on a robust vision-based algorithm. Second, the distance estimation is performed based on the well-known pinhole model and utilizes the internal sensors of the system. Finally, the tracking algorithm provides time consistency.

A. Vision-based Vehicle Detection

The vision-based vehicle algorithm is based on the use of Haar-Like features, as presented by Viola and Jones in [19] and performs calculations by means of fast operations of additions and subtractions of simple image features and a sequential cascade of classifiers for fast and reliable object detection. This system provides very good results with the detection of objects with stable forms by combining layers of simple classifiers.

The weak classifier that they proposed for a given feature follows the form:

$$h_j(x) = \begin{cases} 1 \ , \ if \ p_i f_i(x) < p_i \theta_i \\ 0 \ , \ otherwise \end{cases} \tag{1}$$

where $h_j(x)$ is the classifier, $f_i(x)$ is a given feature, θ_i is the threshold and p_i indicates the inequity direction.

B. Distance Estimation

Location estimation takes into account three paradigms. First, a pinhole-based estimation provides a basic distance estimation, assuming flat world. Later, the internal sensors of the system correct the distance estimation according to the position of the device. Finally, a tracking stage provides more accurate localization and time consistency, this tracking stage is composed by a tracking algorithm and data association.

1) Pinhole Model: Pinhole is based on simplification of the sensor of the camera as a single point in the space and the flat ground assumption. By means of these simplifications, it is possible to reach the model described in (2), which describes the relation between the world points in Cartesian coordinates (x,y,z) and the image points (u,v),

$$\lambda \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} f_u & 0 & u_0 \\ 0 & f_v & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} \tag{2}$$

where the camera image coordinates are represented by u and v, in pixels. (u_0,v_0) is the center of the image, f_u and f_v are the focal length for every coordinate in the image. x, y, and z are the Cartesian coordinates with the sensor of the camera as origin, and λ is a scaling factor.

The main limitation of the Pinhole model are the three solutions (x, y and z) which cannot be solved with the available information (u,v). This requires the fixation of one coordinate in order to provide a solution. In this case, it is assumed that the vehicles are located within the ground plane, and the mobile phone has a fixed location (at the windshield, assumed to be 1.4 meters, however it can be configured in the app). All this allows to assume all the detection to be located at z=-1.4 meters.

2) Distance Estimation Correction: Internal sensors included in smartphones allow the retrieval of the device's position, including rotation angles. Figure 1 describes how changes in these angles affect the estimation of distance to other vehicles. Equations (3), (4), (5) and (6) describe it mathematically.

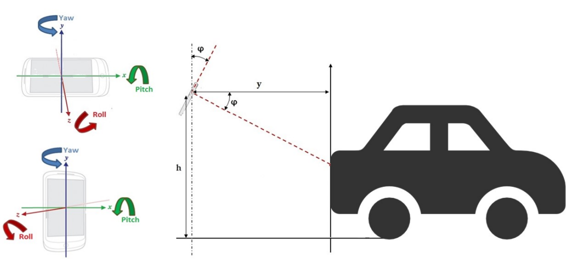

Fig. 1. Rotation angles according to the position of the device (left) and their effect in the vehicle detection (right).

$$\begin{bmatrix} x_f \\ y_f \\ z_f \end{bmatrix} = \Delta\Phi\Theta \begin{bmatrix} x_0 \\ y_0 \\ z_0 \end{bmatrix} \quad (3)$$

$$\Delta = \begin{bmatrix} cos(\triangle\delta) & 0 & sin(\triangle\delta) \\ 0 & 1 & 0 \\ -sin(\triangle\delta) & 0 & cos(\triangle\delta) \end{bmatrix} \quad (4)$$

$$\Phi = \begin{bmatrix} 1 & 0 & 1 \\ 0 & cos(\triangle\varphi) & -sin(\triangle\varphi) \\ 0 & sin(\triangle\varphi) & cos(\triangle\varphi) \end{bmatrix} \quad (5)$$

$$\Theta = \begin{bmatrix} cos(\triangle\theta) & -sin(\triangle\theta) & 0 \\ sin(\triangle\theta) & cos(\triangle\theta) & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (6)$$

where, Δ Φ Θ are the rotation matrices according to the camera axis, with $\triangle\delta$, $\triangle\varphi$ and $\triangle\theta$ corresponding to the increment of the Euler angles roll, pitch and yaw, respectively.

3) Tracking Algorithm: The tracking procedure is based on the use of a Kalman Filter (KF) and constant velocity model. This model has the limitation of the lineal movement definition, however the high frequency of the detection and the definition of the inaccuracies of the lineal model in the system error allows this limitation to be overcome. Furthermore, previous works [20] proved the viability of this simplification, which is relevant for use in a smartphone.

Equations (7) present the system error, which represents the linearization error by modeling the changes in speed (acceleration) as system noise, (8) is the measurement error.

$$Q = \begin{bmatrix} \frac{a^2dt^4}{4} & \frac{a^2dt^3}{2} & 0 & 0 \\ \frac{a^2dt^3}{2} & a^2dt^2 & 0 & 0 \\ 0 & 0 & \frac{a^2dt^4}{4} & \frac{a^2dt^3}{2} \\ 0 & 0 & \frac{a^2dt^3}{2} & a^2dt^2 \end{bmatrix} \quad (7)$$

where a is the maximum acceleration and dt is the time elapsed among image captures, calculated for each frame.

$$R = \begin{bmatrix} (\frac{x}{8})^2 & 0 \\ 0 & (\frac{y}{8})^2 \end{bmatrix} \quad (8)$$

Finally, the constant velocity model is represented in the following equations:

$$\hat{X} = \begin{bmatrix} x \\ v_x \\ y \\ v_y \end{bmatrix} \quad (9)$$

$$Y = \begin{bmatrix} x \\ y \end{bmatrix} \quad (10)$$

$$H = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \quad (11)$$

$$A = \begin{bmatrix} 1 & dt & 0 & 0 \\ 0 & 1 & 0 & dt \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (12)$$

where $\hat{X}$ is the state vector, with x,y the location in meters, and v_x and v_y the velocity in meters per second. Y is the measurement vector which retrieves only the location of the detected vehicle. H is the observation model, and A is the state transition matrix of the model.

4) Time Consistency: Time consistency is obtained thanks to an association technique, is used to associate new detection with previous detections. This phase was designed based on Nearest Neighbor (NN), as depicted in [20]. The NN algorithm was based on the use of the distance estimation obtained by the KF. The location of the new detection is obtained by the distance estimation algorithm. An assignment matrix is designed in each process and based in the minimization of the Euclidean distance.

V. Drowsiness Detection

Blink duration and interval, delay of lid reopening, and lid closure speed are indicators of sleepiness [21]. A psychophysiological measurement associated with the assessment of fatigue [22] is the blink rate per minute. According to [23], the normal average would be about 15 to 20 blinks per minute. In this work we rely on eye closed phase duration along with the the haar-cascade classifiers to estimate and detect drowsiness, as presented in subsection IV-A adapted for face, right eye and left eye detection.

To determine the driver status regarding fatigue and drowsiness, the application used the front camera of the mobile device to locate the driver's face position and evaluate the duration of the closed eye phase, according to the PERCLOS method as described in [24] and used in [25]. PERCLOS calculates the percentage of eye closure time. If the value exceeded the threshold, the system triggered an alarm to make the driver aware of the situation. The alarm did not stop until the calculated value drops under the threshold, i.e. the driver no longer showed signs of drowsiness. Threshold was set, as defined in [24] at 21%. Figure 2 shows the algorithm for drowsiness detection.

VI. System Evaluation Results

To evaluate the performance of our implemented application regarding object detection rate, accuracy and time, we performed several tests with a vehicle in which a mobile device was installed in the dashboard, at a distance lower than 55 cm. The tests were performed with two different Android smartphones:

A) Samsung Galaxy Fame (480p@25fps Primary=5MP, Secondary=VGA, RAM=512 MB)

B) Samsung Galaxy A5 with higher image processing performance and better camera resolution (1080p@30fps Primary=13 MP, Secondary=5MP, RAM=2 GB).

A. Pedestrian Detection Rate

For the evaluation of the performance of the pedestrian detection algorithm, 65 different scenarios were tested with a maximum number of 4 pedestrians per scenario in a detection range from 5 to 30 meters. Here the tests focused on the detection rate, while tests performed in [18] focused on the

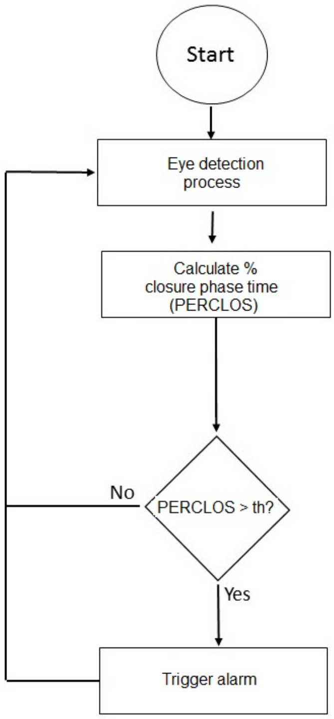

Fig. 2. Implemented Algorithm for Driver Drowsiness Detection

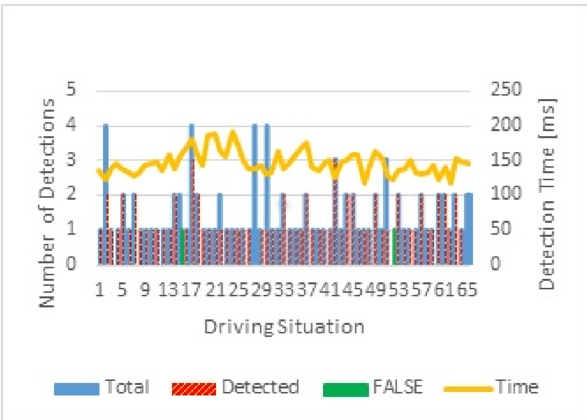

Fig. 3. Pedestrians detection rate and time for device B

distance measurement. According to the device performance, the detection rate with device B was higher. Figure 3 shows that using this device 85.5% of pedestrians were detected, and the rate of false detection constituted 2.4%. A reduction in the detection time was observed depending on the RAM and CPU power. The detection time ranged from 118 to 192 milliseconds, the average detection time being 146.8 milliseconds. Figure 4 shows some visual results of the tests.

In addition, we compared the results of device A using the single detection algorithm with those of the algorithm with the tracking stage. As illustrated by the following graphs, the detection process time benefits from the additional processing with the tracking stage in terms of prediction, correction, and data update. The mean time for the total process increased from 159.4 ms to 192.8 ms (Figure 5). However,

Fig. 4. Pedestrian detection with device B

the detection rate also increased from 45.7% to 78.2% and the number of false detected objects decreased from 17.3 to 2.5 percent (Figure 6). Thus the tracking stage increased the performance of the overall system. Furthermore, it was proved that the lower frame rate obtained with the A device limited the detection rate.

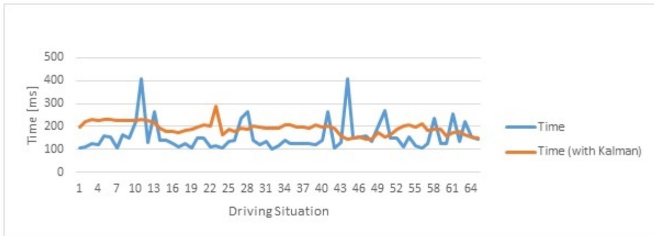

Fig. 5. Processing time comparison with and without applying tracking stage

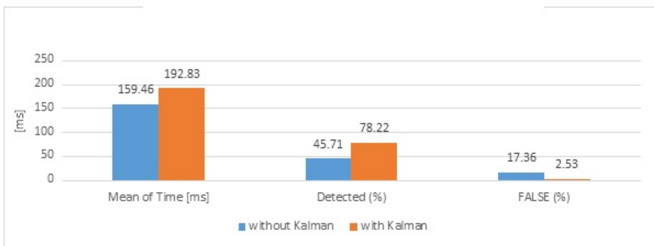

Fig. 6. Time performance of device A with and without applying tracking stage

B. Vehicle Detection

For the evaluation of the vehicle detection algorithm, we tested 55 different scenarios with a maximum number of 2 vehicles per scenario in a detection range from 5 to 30 meters. According to the device performance, the detection rate with device B was higher. For device B, it can be seen that 90.4% of vehicles are detected and the rate of false detection constituted 5.2% (Figures 7 and 8). A reduction in the detection time could be observed depending on the RAM and CPU power. The detection time ranged from 41 to 62 milliseconds, the average detection time being 49.7 milliseconds.

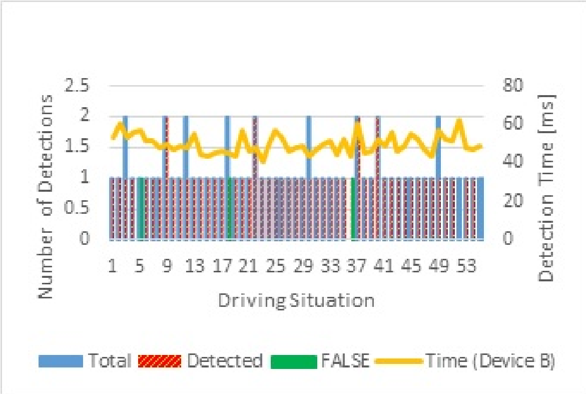

Fig. 7. Vehicles detection rate with device B

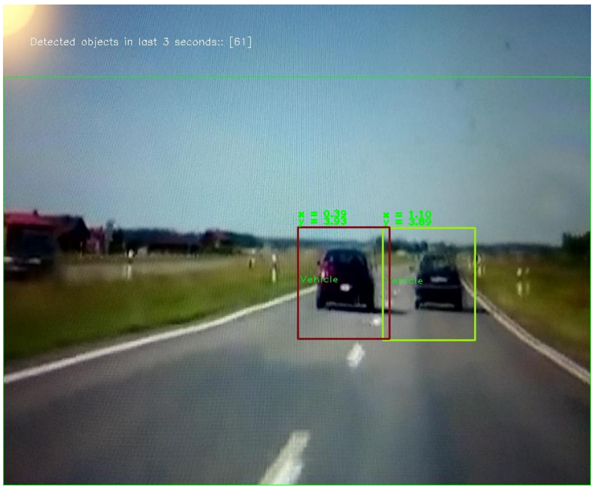

Fig. 8. Vehicles detection example with device B

C. Drowsiness Detection

To evaluate the performance of our implemented application regarding drowsiness detection, a test was performed on 5 drivers using a Samsung Galaxy A5 (Android version 6.0.1) with high image processing performance, and high camera resolution (1080p@30fps Primary=13 MP, Secondary=5MP, RAM=2 GB).

We evaluated the application in a 1-minute scenario for each driver, first in the morning (around 10 A.M) and then after a long period of working or in the evening (around 10 P.M). According to the device performance, the detection quality has a direct relationship with the camera quality. In the morning test the value indicating closed eye duration did not exceed 0,100 ms/min. The average calculated value for the 5 drivers was 0,084, the minimum being 0,067 ms/min (driver 4) and the maximum being 0,098 ms/min (driver 1).

As an example we show the data for driver 2 (Figure 10). In this case the closed eye duration phase value did not exceed 0,100 ms/min in the morning.However, in the evening it increased to almost 0,300 ms/min. This correlates with previous findings which showed that a lower blink frequency (as a consequence of a higher closed eye duration phase) is an indicator of drowsiness [24], proving the usability of the drowsiness detection using the smartphone.

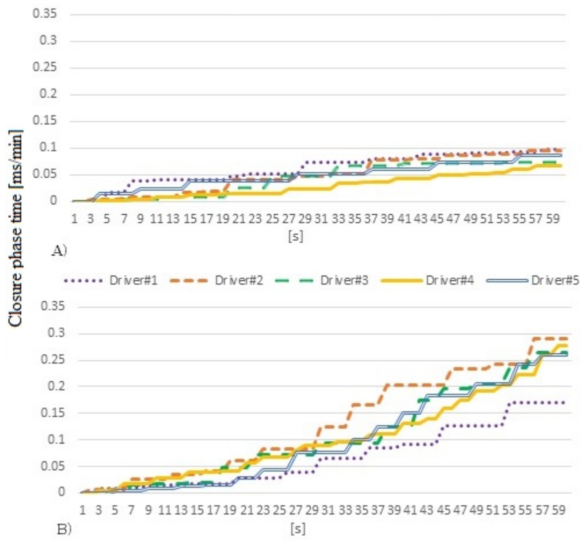

Fig. 9. a) closure phase time in morning conditions and b) closure phase time in evening conditions

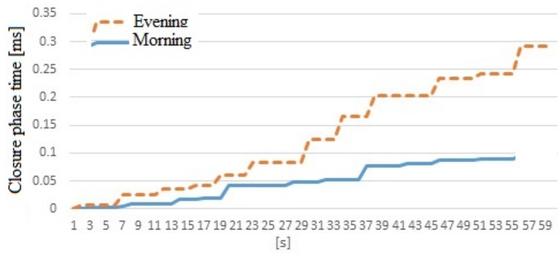

Fig. 10. Closure phase time in last 60 seconds for driver 2 in the morning and in the evening.

VII. Conclusion

The paper presents a novel approach for vision-based road user detection and distance estimation. Furthermore, it also provides vision-based drowsiness detection. The algorithms have been implemented for smartphone devices, adapting to its work limitations and taking advantages of the sensors included. The algorithms were included in a previously presented application, designed to monitor drivers during the autonomous driving experience and inform them to take control of the vehicle if the situation demands and they are in a condition to do so.

The detection rates obtained, 85% in the case of pedestrians and over 90% in the case of vehicles with a low false positive rate. Furthermore, the results of the drowsiness detection proved the performance of the algorithm. All the tests performed proved the viability of the smartphone and computer vision-based algorithms in order to implement an advanced application able to provide situational awareness in the novel paradigm of autonomous driving by means of a smartphone.

In future works distance estimation will be used to identify and adapt the alarms to the movement of the vehicle in real time. Moreover, further tests will be performed to the drowsiness algorithm, with longer driving times, under a simulator, providing further information of the performance of the system in real drowsiness scenarios.

ACKNOWLEDGMENT

This work is supported by Spanish Government through CICYT project (TRA2013-48314-C3-1-R and TRA2015-63708-R), Comunidad de Madrid through SEGVAUTO-TRIES (S2013/MIT-2713) and the KiTSmart Project, City of Vienna Competence Team for Intelligent Technologies in Smart Cities, funded by national funds through the MA23, Urban Administration for Economy, Work and Statistics, Vienna, Austria.

REFERENCES

[1] "International levels of driving automation for on road vehicles," http://www.sae.org/misc/pdfs/automateddriving.pdf/.

[2] L. Rainie, "Internet, broadband, and cell phone statistics," *Pew Internet & American Life Project*, vol. 5, 2010.

[3] J. Goncalves, J. S. Goncalves, R. J. Rossetti, and C. Olaverri-Monreal, "Smartphone sensor platform to study traffic conditions and assess driving performance," in *Intelligent Transportation Systems (ITSC), 2014 IEEE 17th International Conference on*. IEEE, 2014, pp. 2596–2601.

[4] IDC. (2016) Smartphone os market share. IDC. [Online]. Available: http://www.idc.com/promo/smartphone-market-share/os

[5] A. Allamehzadeh and C. Olaverri-Monreal, "Automatic and manual driving paradigms: Cost-efficient mobile application for the assessment of driver inattentiveness and detection of road conditions," in *Intelligent Vehicles Symposium (IV), 2016 IEEE*. IEEE, 2016, pp. 26–31.

[6] "Mobileye - our vision. your safety." http://www.mobileye.com .

[7] "Face detection application," https://play.google.com/store/apps/details?id=org.opencv.samples.facedetect .

[8] "Eye localization application," https://play.google.com/store/apps/details?id=org.opencv.eyelocvasko .

[9] "Face tracking application," https://play.google.com/store/apps/details?id=ro.raduturcas.facetrackingexample&hl=en .

[10] "Mercedes-Benz Attention Assist," http://m.mercedes-benz.de/de_DE/attentionassist/detail.html/.

[11] "Volvo Driver Alert Control (DAC)," http://support.volvocars.com/de/cars/Pages/ownersmanual.aspx?mc=Y555&my=2015&sw=14w20mc=Y555&my=2015&sw=14w20&article=2e82f6fc0d1139c2c0a801e800329d4e/.

[12] "Lexus driver monitoring system," http://www.lexus.com/models/RX/safe ty/.

[13] "JEDE FAHRT GENIESSEN," http://www.bmw.de/de/topics/faszination-bmw/connecteddrive/fahrer-assistenz/intelligentes-fahren.html.

[14] "Driver assistant application," https://play.google.com/store/apps/ details?id=com.driveassist.experimental .

[15] "Co driver application," https://play.google.com/store/apps/details?id=com.evotegra.aCoDriver .

[16] "On road application," https://play.google.com/store/apps/details?id=com.picitup.iOnRoad .

[17] Y. Dong, Z. Hu, K. Uchimura, and N. Murayama, "Driver inattention monitoring system for intelligent vehicles: A review," *Intelligent Transportation Systems, IEEE Transactions on*, vol. 12, no. 2, pp. 596–614, 2011.

[18] F. Garcia, J. Urdiales, J. Carmona, D. Martin, and J. M. Armingol, "Mobile based pedestrian detection with accurate tracking," in *2016 IEEE Intelligent Vehicles Symposium (IV)*, June 2016, pp. 44–48.

[19] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," in *Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on*, vol. 1. IEEE, 2001, pp. I–511.

[20] F. Garcia, D. Martin, A. de la Escalera, and J. M. Armingol, "Sensor fusion methodology for vehicle detection," *IEEE Intelligent Transportation Systems Magazine*, vol. 9, no. 1, pp. 123–133, Spring 2017.

[21] R. Schleicher, N. Galley, S. Briest, and L. Galley, "Blinks and saccades as indicators of fatigue in sleepiness warnings: looking tired?" *Ergonomics*, vol. 51, no. 7, pp. 982–1010, 2008.

[22] J. A. Stern, D. Boyer, and D. Schroeder, "Blink rate: a possible measure of fatigue," *Human Factors: The Journal of the Human Factors and Ergonomics Society*, vol. 36, no. 2, pp. 285–297, 1994.

[23] M. Cronin. (2012) Why do we blink so much? blinking provides mental rest, new research suggests. The Huffington Post. [Online]. Available: http://www.huffingtonpost.com/2012/12/30/whydoweblink-somuchmentalrest_n_2377720.html

[24] D. F. Dinges, M. M. Mallis, G. Maislin, and I. Powell, "Evaluation of techniques for ocular measurement as an index of fatigue and the basis for alertness management," No. HS-808 762, Tech. Rep., 1998.

[25] M. Flores, J. Armingol, and A. de la Escalera, "Driver drowsiness warning system using visual information for both diurnal and nocturnal illumination conditions," *EURASIP Journal on Advances in Signal Processing*, vol. 2010, no. 1, p. 438205, 2010. [Online]. Available: http://dx.doi.org/10.1155/2010/438205