# Aspects Concerning the Calibration Procedure for a Dual Camera Smartphone Based ADAS

Mihai Duguleana$^{(\boxtimes)}$, Florin Girbacia, Cristian Postelnicu,
Andreea Beraru, and Gheoghe Mogan

Department of Automotive and Transport Engineering, University Transilvania
of Brasov, Brașov, Romania
{mihai.duguleana, garbacia, cristian-cezar.postelnicu,
aberaru, mogan}@unitbv.ro

**Abstract.** We present the architecture of an Advanced Driver Assistance System (ADAS) based on common dual camera smartphones, emphasizing on the calibration procedure which is active during the initialization phase (prior to the actual driving). NAVIEYES project attempts to make use of the video information received from both the front and the rear cameras of the phone in order to infer and alert drivers upon potential dangerous situations. This study focuses on the information received from the front camera. 10 different mobile devices were tested, in order to choose the most powerful and ergonomic platform. A calibration experiment is carried by 22 subjects, and the first version of the application is tested using a driving simulator. The system is deployed on a real car, and several warning paradigms such as audio, video and mixed alerts are also analyzed. We present the HCI questionnaire, analyze the data and propose further developments.

**Keywords:** Personal navigation assistant · Smartphone · Dual camera · Driver assistant · ADAS

## 1 Introduction

Traffic accidents have always been a weak point of road transportation. A 2010 report from the World Health Organization (WHO) shows that more than 1.2 million traffic accidents occur each year [1]. In hope of diminishing the effects of this phenomenon, the United Nations General Assembly adopted resolution which proclaims the period 2011–2020 as the Decade of Action for Road Safety. Today we find ourselves at the middle of this period, and although several safety systems have been deployed on the latest car models, things are not improving radically. As an active component of the initiative, WHO has released in 2013 the Global status report on road safety [2]. Making the vehicles safer is an important objective resulted from this report, not only from a social and psychological point of view, but also from economical point of view, as the cost of traffic incidents (whether they are fatal or non-fatal) has been estimated to average 1.5 % of the gross national product in EU partner countries, totaling a global cost of approximately \$500 B per year [3]. WHO identified distracted driving and fatigue among the most common causes of accidents.

The project NAVIEYES [4] started in 2014 and aims to research new means of preventing at least a small part of potential dangerous driving situations. NAVIEYES proposes to improve classic Advanced Driver Assistance Systems (ADASs) by using computer vision capabilities to track driver's eyes and head, along with the traffic environment. The ADAS developed within NAVIEYES can be deployed on smartphones with dual cameras. The application will alert the driver for front obstacles (static or dynamic), change of direction (lane change), specific traffic signs or whether his drowsiness level is high. The core of the system is based on 3 concurrent modules: an estimator of eye gaze, head orientation and drowsiness level, a detector of road obstacles, lanes and traffic signs and an API necessary to integrate the first 2 modules. The outcome solution will work with any car kit that allows visibility for both smartphone cameras, as seen in Fig. 1.
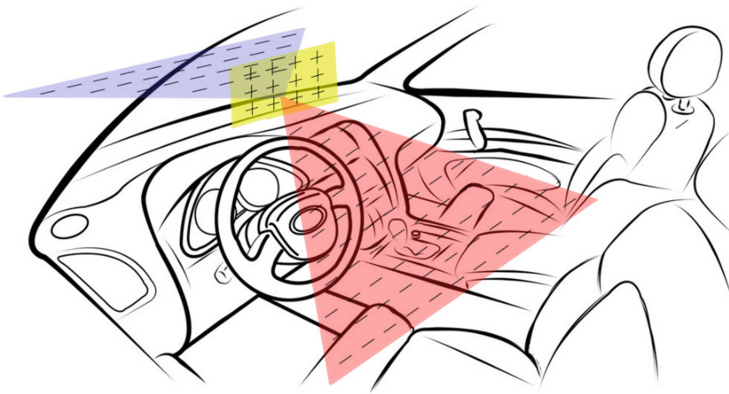


**Fig. 1.** Positioning the smart phone (yellow area), with the rear camera facing outside the car (blue area) and the front camera facing the car interior (red area) (Colour figure online).

## 2 Related Work

Concerning the information received from the front camera (the car interior and the driver), computer vision is at the moment used within actual ADASs for 2 not necessarily disjunctive purposes: evaluating the drowsiness level (the fatigue level) and analyzing the behavior of the driver. While the first plane of development focuses on studying eye movements, specifically eye blinks, the later includes the analysis of body postures, possibly face detection and recognition, head orientation, prediction of future actions and the assessment of driver's cognitive load among several others.

### 2.1 Drowsiness

Drowsiness appears due to several reasons, including long periods of driving, stress, diseases (such as sleep disorders, medication side effects), or even radio music. As some of the driver's actions are a direct result of the high levels of fatigue, and

additionally, as several activities are hard to categorize and include in structural patterns, most scientists which research driver information focus on calculating just the most common parameter from the first area: PERCLOS (PERCentage of eye CLOSure) [5]. Calculating PERCLOS based on video processing algorithms usually takes the following route:

- First, face detection will be computed. Ideally, face detection will make use of a fusion of methods such as Haar classifiers, filters, Viola-Jones, landmark model matching and others.
- After retrieving the image segment which contains driver's face, the eyes are usually classified and followed through consecutive frames, using other techniques such as neural networks (NNs), template matching, Haar classifiers, Dynamic Time Window, or particle swarm optimization. The biggest problem of this approach is the high sensitivity to even the slightest light variations.
- The number of blink occurrences is accounted over a period of time and compared with specific threshold.

Besides PERCLOS, another visual clue analyzed by researchers is yawning. Several studies inferred that yawning is triggered by low vigilance levels, during progressive drowsiness. However, the main issue when analyzing yawning is the weak quantification possibility. There is still a public debate on what exactly triggers yawns [6], and some studies also argue the precision of correlation between facial muscles and drowsiness (60 %) with respect to the correlation between blinking and drowsiness (>80 %) [7]. It has been proved that people may have high yawning levels without even experiencing drowsiness, and the other way around – sleep can occur without yawning. Another drowsiness cue is head movement. It has been shown that head movement distance and velocity has a strong correlation to somnolence [8]. Taken together, eye blinking, yawning and head movements provide critical data necessary to determine the drowsiness level of the subject.

## 2.2  Driver Behavior

The behavior of the driver can prove to be useful to ADASs. The focus of most studies is on inferring the eye gaze (if possible, the gaze intersection point) and the head orientation, or better said, driver's region of interest (ROI) and within this region, driver's point of interest (POI). Another important region is the focus of expansion (FOE). Since almost 2 decades ago, it has been concluded that at their simplest, drivers' fixation patterns on straight roads can be described as concentrating on a point near to the FOE (i.e. the horizontal line, where objects appear stationary). Researchers assumed that the reliance on the FOE is trivial, as it provides precise directional information to the driver and is the location near to which future traffic hazards are most likely to be first visible. Other researchers refer to the fixations on the tangent point (TP) - the inner lane marking (the boundary between the asphalted road and the adjacent green) bearing the highest curvature in the 2D retinal image, or in other terms, the innermost point of this boundary (see Fig. 2a). Other fixation strategy slightly different from TP relies on the retinal flow theory. As the driver moves through the environment, the objects

projected on his retina change with the movement, creating a "retinal flow" in a manner fairly similar with the optical flow computer vision algorithm. This flow depends on several parameters like the heading direction, car speed, the depth structure of the environment and whether objects themselves are static or dynamic. Basically, drivers fixate a spot on their future path (i.e. the middle of the lane) and track it for some time as they approach it. When the point comes too near to the car front, drivers will look for a new point to track. Depending on the curvature of the flow lines, drivers make steering decisions, as straight retinal flow lines emerge if the driver steers correctly (see Fig. 2b).
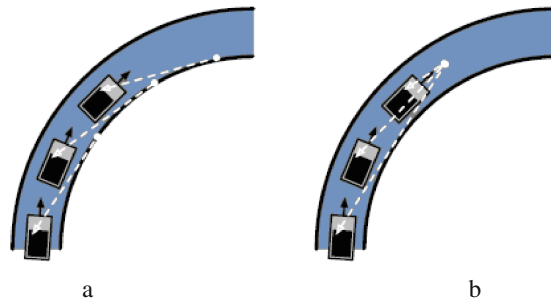


a                                        b

**Fig. 2.** (a) TP theory; (b) retinal flow theory

Processing a video stream from a single camera provides approximate results. Luckily, most assistance systems do not require detailed eye gaze direction, but only the coarse gaze direction, in order to reduce false stressful alarms. Coarse gaze direction can be computed based on the head orientation vector. Humans have a limited FOV, thus in order to gaze on some scene element, they will move their head to a comfortable position before orienting their eyes. If the coarse gaze direction extends to an angle higher than ±35 degrees, the driver will most likely rotate his head and from this point on, coarse gaze direction may be calculated by tracking the orientation of the head [9]. Several methods are used for head tracking. Most of the make use of information such as eye position and facial features. The active appearance model (AAM) was used to find facial features in various situations. Other methods are based on shape features, such as the cylindrical face model technique. The methods based on texture features work by finding driver's face (partial face) in the video stream and analyzing the intensity pattern of the facial image in order to estimate the head orientation. Among texture-based methods one can count the Principal Component Analysis (PCA), Kernel Principal Component Analysis (KPCA) and the Linear Discriminant Analysis (LDA). Furthermore, some scientists used Local Gradient Orientation (LGO) and Support Vector Regression (SVR) to estimate the driver's continuous yaw and pitch, while others analyzed the asymmetry of the facial image by using a Fourier transform to estimate the driver's continuous yaw.

## 2.3    Integrated Research

Both research areas (drowsiness and driver behavior) are integrated in just a few studies. Furthermore, only a few analyze the possibility of using a dual camera mobile phones as a sensorial system. Most of the work relies on using special laboratory equipment (head mounted eye trackers, infrared light cameras and so on). In [10], authors propose a similar system, with several drawbacks: the system is designed specifically for a test car and doesn't take into consideration the general case, in which a calibration procedure is required. This subject is precisely treated by this paper. In [11], researchers use only on the front camera to compute where the driver is looking, and produce a finite set: the road, top, left and right mirrors, street signs, car board or phone.

## 3    NAVIEYES Application Architecture

The core of the NAVIEYES system is based on 3 concurrent modules: an estimator of the eye gaze, head orientation and drowsiness level, a detector of road obstacles, lanes and traffic signs and an API necessary to integrate the first 2 modules. The outcome solution will work with any car kit that allows visibility for both smartphone cameras.

The first module handles the recording and analysis of the head and eye movement and provides important clues vital for understanding driver's intentions, for taking appropriate accident countermeasures. The output of this module is the gaze vector, the head orientation vector and PERCLOS. The first module is based on the video stream received from the front camera and requires a calibration procedure before actually retrieving data (see Fig. 3).
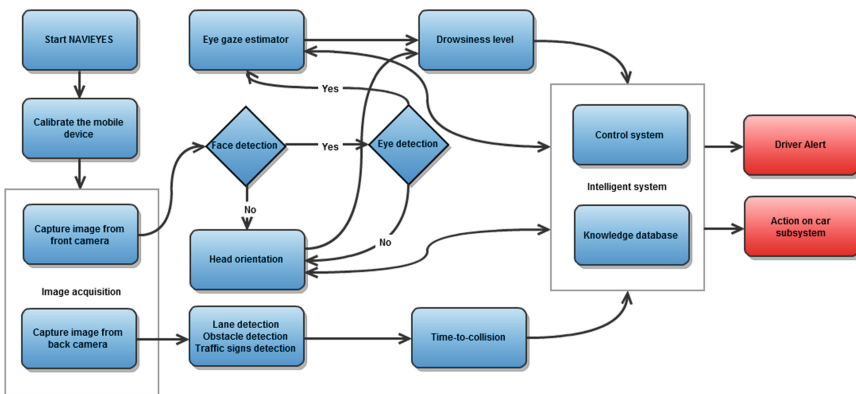


**Fig. 3.**  NAVIEYES application architecture

The second module analyses the exterior traffic environment, and produces 3 outputs: front collision warning (FCW), traffic sign detection (TSD) and lane departure warning (LDW). The second module receives the video stream from the rear camera. This assumes that the rear camera is correctly facing the car exterior environment.

The third module integrates the first 2 modules. The output from the first module is used to determine if the driver is actually paying attention to potential dangerous situations (front collision or lane switching). Although traffic signs are not considered in this category, based on the computed vectors (eye gaze and head orientation), the system can also infer if the driver actually sees the traffic sign.

## 4   Calibration Procedure

The calibration system starts as soon as the user enters a vehicle and opens the application. The calibration phase ensures that both the front and the rear cameras are positioned correctly (they are both able to receive the desired video streams). This assumes setting a correct position for the car kit and a video calibration of the device, by following these steps (see Fig. 4):
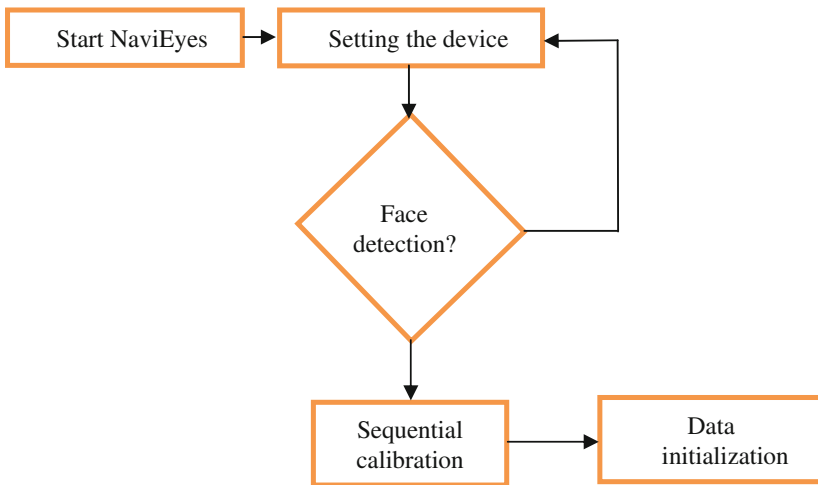


**Fig. 4.**  Calibration procedure

- The device is mounted into the car kit.
- The app tries to detect the face of the user.
- If the face is detected, the user is guided into a calibration sequence which quantifies what a look to left or right really means. The user is asked to maintain its head fixed, and to look at the maximum left and right. If not, the app asks the user to reposition the device, based also on the gyroscope sensor.
- If the calibration is a success (the eye gaze is established), previous session data is initialized (PERCLOS average, GPS position, total driving distance and others).

Before running the HCI experiment trials, 10 mobile phones were taken into account, based on their hardware specifications, their switching delay (the simultaneous use of both cameras is impossible by design of CameraService for Android, and a similar

problem exists on IOS) and on their face detection rate (using the basic OpenCv haar classifier). The following data resulted from our primary tests (Table 1):

**Table 1.** Smartphones tested for Navieyes deployment

| Model | Processor | Memory | Switching delay | Face detection overhead |
|---|---|---|---|---|
| Samsung Galaxy S3 | 1400 MHz | 1 Gb | 1.2 s | 0.3 s |
| Samsung Galaxy S4 mini | 1700 MHz | 1.5 Gb | 1 s | 0.2 s |
| Samsung Galaxy S4 | 1900 MHz | 2 Gb | 0.9 s | 0.2 s |
| Samsung Galaxy S5 | 2500 MHz | 2 Gb | 0.9 s | 0.2 s |
| HTC One Mini 2 | 1200 MHz | 1 Gb | 1.3 s | 0.2 s |
| HTC One M8 | 2300 MHz | 2 Gb | 1.1 s | 0.2 s |
| LG G3 | 2500 MHz | 2 Gb | 1.1 s | 0.2 s |
| iPhone 4S | 1000 MHz | 512 Mb | 1 s | 0.1 s |
| iPhone 5S | 1300 MHz | 1 Gb | 1 s | 0.1 s |
| iPhone 6 | 1400 MHz | 1 Gb | 0.8 s | 0.1 s |

We've concluded that all Android platforms are mostly taking around 1 s to switch between cameras, and their face detection overhead is 0.2–0.3 s, while iPhones switch cameras and detect faces faster. Thus, we've settled for using the Android based Samsung S3 within our experiments.

The calibration procedure was tested in a real vehicle, on 22 subjects (18 students, 3 Ph.D. students and 1 person from the administrative staff of our university). The car kit was already mounted, the subjects were just asked to insert the smartphone into it and to run the NaviEyes application (see Fig. 5).
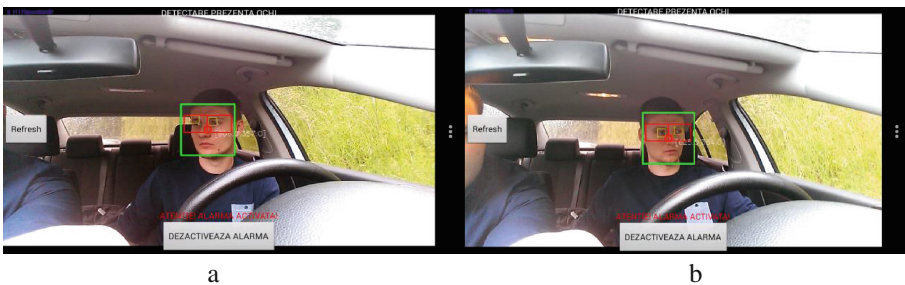


a                                        b

**Fig. 5.** Calibration procedure. Driver looks right (a) and left (b)

## 5   Virtual Experiment Setup

The same subjects were asked to use the ECA Faros Simulator and drive within an urban scenario with medium traffic for 3 min, on a designated route, while using the first version of the application. Each participant was involved in 3 trials, each using a different alert: sound alert coming from the phone, light alert coming from the phone and sound alert coming from the stereo speakers of the VR simulator. Subjects were asked to perform various activities within each trial, in no particular order: send a text message, look behind, and close their eyes for 3 s. The application was set to trigger the alarm after receiving 0.6 s of continuous invalid readings (no face detection, Fig. 6).



**Fig. 6.** Experiment setup: VR simulator and Samsung Galaxy S3 running Navieyes app

## 6   HCI Questionnaire

The HCI questionnaire proposed in this study was developed to acquire data on subjects' interaction with the application. Most questions from the quiz could be answered on a scale of 1 to 10, and were divided into 3 sections:

- The biographic section was designed to gather data related to the technological background of the participant. We tried to cluster subjects by their age and sex, their driving and computer skills and their experience with smartphones by using questions such as "How often do you drive?", "How often do you use smartphones?", "How familiarized are you with mobile apps?".

- The <u>calibration section</u> treats only aspects concerning the calibration procedure, with questions such as "How easy going was to set up the system?" or "How good do you find the detection rate?". We also addressed an open question: "How would you improve the calibration procedure?".
- The <u>usability section</u> treats the issues participants encounter during the VR experiment. In order to measure the usability of the application while driving on the car simulator, this section includes questions such as "How stressful is the alarm system no. 1 (phone sound alarm)?" or "How good is the 0.6 s timeframe?". The overall usability of the system is assessed with the question "How useful do you find the Navieyes application?" (Table 2).

**Table 2.** HCI questionnaire

| Section | Question | Answer |
|---|---|---|
| *Biographic section* | Age? | 20 years – 11 |
| | | 21 years – 7 |
| | | 25 years – 2 |
| | | 26 years – 1 |
| | | 37 years – 1 |
| | Sex? | M – 17 F – 5 |
| | How often do you drive? | 6.95 (0 – never, 10 – all the time) |
| | How often do you use smartphones? | 9.04 (0 – never, 10 – every day) |
| | How familiarized are you with mobile apps? | 8.22 (0 – never heard, 10 – always using) |
| *Calibration section* | How easy going was to set up the system? | 7.90 (0 – easy, 10 – hard) |
| | How good do you find the detection rate? | 7.81 (0 – bad, 10 – good) |
| | How would you improve the calibration procedure? | – |
| *Usability section* | How stressful is the alarm system no. 1 (phone sound alarm)? | 6.59 (0 – stressful, 10 – not stressful) |
| | How stressful is the alarm system no. 2 (phone light alarm)? | 7.40 (0 – stressful, 10 – not stressful) |
| | How stressful is the alarm system no. 3 (environment sound alarm)? | 7.5 (0 – stressful, 10 – not stressful) |
| | How do you find the 0.6 s timeframe? | 6.18 (0 – too long/short, 10 – perfect) |
| | How useful do you find the Navieyes application? | 8.72 (0 – useless, 10 – useful) |

The open question received advices on improving the detection rate, the sound alarm jingle or the application user interface.

## 7   Conclusions and Further Development

The centralized results from the HCI questionnaire conclude that Navieyes app will improve driver's road safety. According to question 13, the usability of the initial version of the application is rated at 8.72 on a 1-to-10 scale. Most study participants are rather young and have extensive experience with smartphones and mobile apps, according to the first section of the questionnaire. Due to the fact that only about half of them own a personal car, they however don't have too much driving experience according to question 3 (6.95). The calibration procedure was rated with 7.9 out of 10, and the most accepted form of alarm was the environment sound alarm.

As future development, we plan to further expand the application by integrating the data received from the rear camera, which will in term lower the stress emitted by the application alarm (as it will produce far less notifications).

## References

1. World Health Organization. Data and statistics of the World Health Organization. http://www.who.int/research/en/. Accessed 20 Feb 2015
2. World Health Organization. Global status report on road safety 2013. http://www.who.int/research/en/. Accessed 20 Feb 2015
3. Pedan, M., et al.: World Report on Road Traffic Injury Prevention. WHO Press, Geneva (2004)
4. NaviEyes site. http://navieyes.unitbv.ro. Accessed 20 Feb 2015
5. Dinges, D., Grace, R.: Perclos: a valid psychophysiological measure of alertness as assessed by psychomotor vigilance. Technical report, FHWA-MCRT-98-006, Federal Highway Administration, US Department of Transportation (1998)
6. Guggisberg, A.G., Mathis, J., Schnider, A., Hess, C.W.: Why do we yawn? The importance of evidence for specific yawn-induced effects. Neurosci. Bio-behav. Rev. **35**(5), 1302–1304 (2011)
7. Vural, E., Cetin, M., Ercil, A., Littlewort, G., Bartlett, M., Movellan, J.R.: Drowsy driver detection through facial movement analysis. In: Lew, M., Sebe, N., Huang, T.S., Bakker, E.M. (eds.) HCI 2007. LNCS, vol. 4796, pp. 6–18. Springer, Heidelberg (2007)
8. Van Den Berg, J.: Sleepiness and head movements. Ind. Health **44**(4), 564–576 (2006)
9. Cai, H., Lin, Y.: An integrated head pose and eye gaze tracking approach to non-intrusive visual attention measurement for wide FOV simulators. Virtual Real. **16**(1), 25–32 (2012)
10. You, C.W., et al.: Carsafe app: alerting drowsy and distracted drivers using dual cameras on smartphones. In: Proceeding of the 11th Annual International Conference on Mobile Systems, Applications, and Services, pp. 13–26, ACM, June 2013
11. Chuang, M.C., Bala, R., Bernal, E.A., Paul, P., Burry, A.: Estimating gaze direction of vehicle drivers using a smartphone camera. In: 2014 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), pp. 165–170, IEEE, June 2014