

Article

Optical Flow and Driver's Kinematics Analysis for State of Alert Sensing

Javier Jiménez-Pinto and Miguel Torres-Torriti *

Department of Electrical Engineering, Pontificia Universidad Católica de Chile, Vicuña Mackenna 4860, Casilla 306-22, Santiago, Chile; E-Mail: jejimenp@puc.cl

* Author to whom correspondence should be addressed; E-Mail: mtorrest@ing.puc.cl; Tel.: +56-2-354-4408; Fax: +56-2-552-2563.

Received: 30 January 2013; in revised form: 16 February 2013 / Accepted: 19 February 2013 / Published: 28 March 2013

Abstract: Road accident statistics from different countries show that a significant number of accidents occur due to driver's fatigue and lack of awareness to traffic conditions. In particular, about 60% of the accidents in which long haul truck and bus drivers are involved are attributed to drowsiness and fatigue. It is thus fundamental to improve non-invasive systems for sensing a driver's state of alert. One of the main challenges to correctly resolve the state of alert is measuring the percentage of eyelid closure over time (PERCLOS), despite the driver's head and body movements. In this paper, we propose a technique that involves optical flow and driver's kinematics analysis to improve the robustness of the driver's alert state measurement under pose changes using a single camera with near-infrared illumination. The proposed approach infers and keeps track of the driver's pose in 3D space in order to ensure that eyes can be located correctly, even after periods of partial occlusion, for example, when the driver stares away from the camera. Our experiments show the effectiveness of the approach with a correct eyes detection rate of 99.41%, on average. The results obtained with the proposed approach in an experiment involving fifteen persons under different levels of sleep deprivation also confirm the discriminability of the fatigue levels. In addition to the measurement of fatigue and drowsiness, the pose tracking capability of the proposed approach has potential applications in distraction assessment and alerting of machine operators.

Keywords: fatigue detection; alert state sensing; driver assistance; head tracking; PERCLOS; driver's kinematics

1. Introduction

Road traffic accidents are one of the main non-health related causes of death. The data and statistics of the World Health Organization [1] show that about 2.8% of non-health related deaths are due to suicide, violence and wars, while 2.1% are attributed to traffic accidents, even surpassing nutritional deficiencies, which account for about 0.9% of world deaths [2]. On the other hand, the social and economic cost of traffic incidents has been estimated to be 1% of the gross national product in low-income countries, 1.5% in middle-income countries and 2% in high-income countries, totaling a global cost of US\$518 billion per year [3]. Unlike many diseases and health problems for which there is no cure, traffic accidents can be reduced if proper education, law enforcement and engineering practices are implemented [4,5].

Several studies exist that analyze physiological cues associated with a driver's awareness and state of alert [6–8]. Measuring some of the cues, especially physiological ones, such as EEG, ECG, EOG, blood pressure and body temperature [9,10], may require invasive techniques, and despite some recent improvements in the development of highly sensitive and less intrusive electrodes for ECG monitoring [11], their use as a reliable metric is difficult, because signals like ECG often exhibit significant inter-individual variabilities that depend on factors, such as age, gender, spatial ability and intro-extroversion [8]. Other methods monitor the driver's steering performance (reaction rates and unexpected lane departures) to warn the driver. However, despite claims that these approaches have low false alarm rates, it is also known that these methods fail to predict micro-sleeps, and there is not enough evidence to support these methods as a reliable way of measuring the driver's state of alert [6,12,13]. Fortunately, there are many behavioral changes that provide reliable visual cues of the driver's state of awareness that can be measured in a non-invasive manner with image processing techniques, namely, eye-blinking frequency and percentage of eyelid closure over time (PERCLOS, [14,15]), yawn frequency, head movement and eye-gaze, among other facial expressions. The vision-based approaches must rely on specific techniques to detect the driver's head, face and eyes. Some methods employ intensity and color analysis techniques to segment the parts of the head from the image [13,16–21], while several other approaches rely on the Viola-Jones detector [22–28]. Some approaches only track the eyes, while others focus on particular facial cues, such as yawning [19,29]. A limitation of the approaches based on color analysis is their sensitivity to illumination conditions and the fact that they often cannot be applied at night [30,31]. This has motivated some researchers [30–34] to use near-infrared (IR) cameras, exploiting the retinas' high reflectivity to 850 nm wavelength illumination [35,36]. On the other hand, the performance of the approach is also determined by the type of classifier used to process the features extracted from the image. For example, some approaches employ neural-networks to classify segmented regions as the head and its parts [37,38], while others rely on a variety of template matching schemes [29,39–42]. For a recent survey on drowsiness detection systems, the reader is referred to [43].

This work presents a non-invasive sensing approach for driver fatigue and attention measurement, which is based on a standard charge-coupled device (CCD) camera with an 850 nm near-infrared (IR) filter and a circular array of IR LEDs. The proposed approach draws on ideas by the authors presented in [44], which introduces the use of face salient points to track the driver's head, instead of attempting to directly find the eyes using object recognition methods or the analysis of image intensities around the eyes, as the majority of the existing approaches to fatigue assessment. An advantage of salient points

tracking, as proposed in [44], is that the approach is more robust to occlusions of the eyes whenever they occur, due to the driver's head or body motion. On the other hand, the grid of salient points can be tracked with a low processing cost using the Lukas-Kanade algorithm for sparse optical flow computation. The measurement of the salient points' optical flow provides valuable information for computing changes in the driver's gestures, e.g., eyebrow raisings and yawning. However, it is to be noted that prior results have shown that eyebrow raisings and yawning do not have a sufficiently good correlation with fatigue and thus cannot be used as the main predictor of fatigue. Different works have studied the connection between yawning and vigilance showing that: (i) yawning indeed occurs during progressive drowsiness, which is compatible with the notion that yawning is triggered by states of low vigilance; however, yawns were not triggered nor followed by a specific autonomic activation [45,46]; and (ii) the correlation between facial muscle activity and drowsiness is lower (60–80%) than that of blinking (>80%) [47]. Moreover, it is possible to exhibit high levels of yawning without necessarily being in a hypovigilance state [48]. Therefore, facial muscle activity (including yawning and eyebrow raisings) offers little predictive information pertaining to sleep onset [14]. In fact, sleep can occur without yawning or even before any significant change in muscle activity or tonicity [14]. It has been shown in [49] that also head movement distance and velocity have a stronger correlation (>80%) to sleepiness than the correlations in [47] for changes in facial expression (60–80%). Because of these reasons, and the fact that the percentage of time that the eyes are closed (the eyelids cover the pupils at least 80% or more) over a given period of time (PERCLOS [14]) has a significantly stronger correlation to fatigue [15], efforts should be placed on improving head and eye tracking methods. Furthermore, recent works [50,51] confirm that among the different ocular variables, PERCLOS is the most effective to prevent errors or accidents caused by low vigilance states, thus confirming the original observations and findings reported in [14,15]. In this context, the contributions and novelty of this paper can be summarized as follows. A kinematic model of the driver's motion is introduced to obtain the pose of the driver described by five degrees of freedom (lateral tilt, nod and yaw of the head about the neck and frontal and lateral tilt of the torso). The use of the driver's kinematic model allows one to reach an outstanding performance, with an almost 100% tracking rate of the eyes. A high tracking rate is key to the computation of the PERCLOS, since computing the PERCLOS requires the knowledge of where the eyes are and whether they are open or closed. Another contribution of this work is the use of the driver's observed interpupillary distance (IPD) to estimate the distance from the driver's head to the camera (up to a scale factor), thus the approach yields the driver's motion in 3D space. It is shown that tracking in 3D space the back-projected salient points (from 2D image space to 3D space) is equivalent to tracking points on the 2D image space when the knowledge of the distance between the driver and the camera is available. Therefore, an equivalent result to that of tracking the salient points in 3D space is possible by tracking points in 2D space together with the computed driver-camera distance when the salient points are assumed to be a set of coplanar points lying on the facial tangent plane. Although the salient points do not belong to a coplanar plane; their difference in driver-camera distances are negligible, and therefore, the incidence of this assumption in the performance of the system is negligible for practical purposes. Furthermore, the driver's kinematic motion model allows one to implement an extended Kalman filter that simplifies the tracking of the points in the image space (only the five pose angles need to be estimated with the filter, instead of applying a filter to each of the salient points in the

image). Finally, the experiments involving a group of 15 persons—five of which were deprived from sleep for more than 24 h before the driving tests in the simulator; another five were asked to sleep only four hours the night before the test, and the remaining five had a full eight hours rest—demonstrate that the PERCLOS measure is strongly related to fatigue and somnolence and, thus, can be reliably used to warn drowsy and distracted drivers about their impairment and diminished ability to drive correctly.

In contrast to [22,31,32], the proposed approach only employs the Viola-Jones detector to find the driver's face at the initialization stage and only every certain number of frames whenever some reset conditions are satisfied (see Section 2.2.4). Moreover, the proposed approach relies on the tracking of a set of salient points, the driver's kinematic model and the measurement of the IPD, which combined allow for higher eye tracking rates despite occlusions and fast changes in pose. The idea of simultaneous shape modeling and tracking has also been suggested in [52], which uses also an incrementally built texture model, so that the tracker may operate without active illumination. However, the latter also does not include the driver kinematic model, nor uses the Lukas-Kanade to build and track a grid model like the one employed here, but rather builds an appearance and shape model, which allows for small track losses. Finally, it is to be noted that with the proposed approach in this work, it is possible to recover the driver's pose and 3D information using a single camera and the measurement of the IPD. To the best of our knowledge, this is an important aspect not considered yet in the existing published research.

A detailed description of the proposed approach is presented in the next section, followed by an explanation of the experimental methodology in Section 3. The results presented in Section 4 show that our approach yields high eye-tracking rates and reliable confirmation of the driver's state of alert, as inferred from the measured PERCLOS. The main conclusions are presented in the last section.

2. State of Alert Sensing

The proposed system design for sensing a driver's state of alert can be divided into three stages: (i) image capture; (ii) driver detection and pose tracking; and (iii) driver vigilance measurement. The image acquisition stage employs a standard low-cost charged-coupled device (CCD) surveillance camera with a near-infrared 850 nm filter and a circular array of infrared light emitting diodes (LEDs). The driver detection and tracking module employs a regular computer to process the images and obtain the driver's motion and facial cues, such as blinking, yawning and head pose. Finally, the driver vigilance measurement stage is responsible for computing the driver's drowsiness and attention levels using information about the percentage of time over a given period that the eyelids cover 80% or more of the pupils (PERCLOS [14,15]) and the head pose obtained in the previous stage. Other cues, such as blinking frequency and yawning, can also be computed; however, they are not taken into account for triggering alarms in the current system, because they were found to have weak correlations with vigilance and drowsiness levels in previous work by the authors [44].

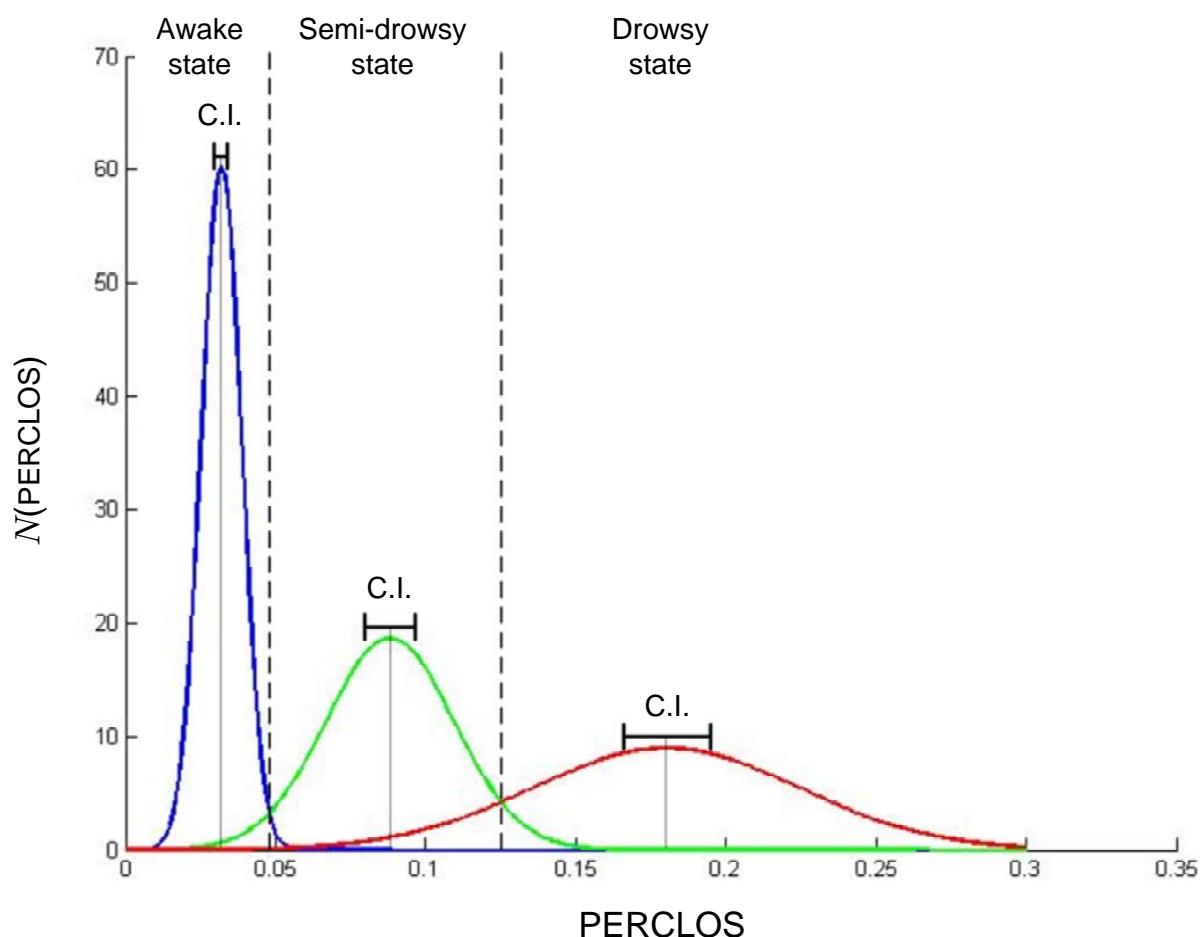
2.1. Image Capture and Preliminary Processing

The camera employed to acquire the driver's images has a 640×480 CCD array, a lens with focal length $f = 4$ mm and a circular arrangement of 26 infrared LEDs. The camera was modified to include an 850 nm band-pass filter lens that covers both the image sensor and the IR LEDs. The purpose of

the near-IR filter lens is two-fold. The filter lens serves to improve the rejection of external sources of infrared radiation and reduce changes in illumination produced by the sun as the car moves. On the other hand, using the filter facilitates the detection of the pupils, because the retina is highly reflective of the near-IR illumination of the LEDs, resulting in a “red-eye” effect similar to that of standard photography. A video of the images from the camera is generated using a composite-video to USB converter. The converter allows one to capture the frames directly on a PC at a rate of 30 fps.

The main tasks of the preliminary processing algorithms are the reduction of the image size to 384×288 and the execution of the Viola-Jones face detection algorithm when the system initializes to provide for a gross estimate of the head location. Reducing the images allows the system to achieve a frame rate of 16.5 fps using a 2.2 GHz CPU, which is a processing rate sufficient for on-line real-time operation of the system. It is to be noted that for fatigue detection purposes, any vision-based system should have a sampling and processing rate of at least 5–10-times per second, because: (i) eye blinks last between 0.1–0.4 seconds [53]; and (ii) microsleeps last from a fraction of a second up to 30 seconds, but more typically three to 14 seconds, according to the experimental studies in [54]. Furthermore, the resolution of the PERCLOS measurements computed at 16.5 fps is enough to obtain clearly distinguishable states of fatigue, as shown by our results (see, for example, Figure 1).

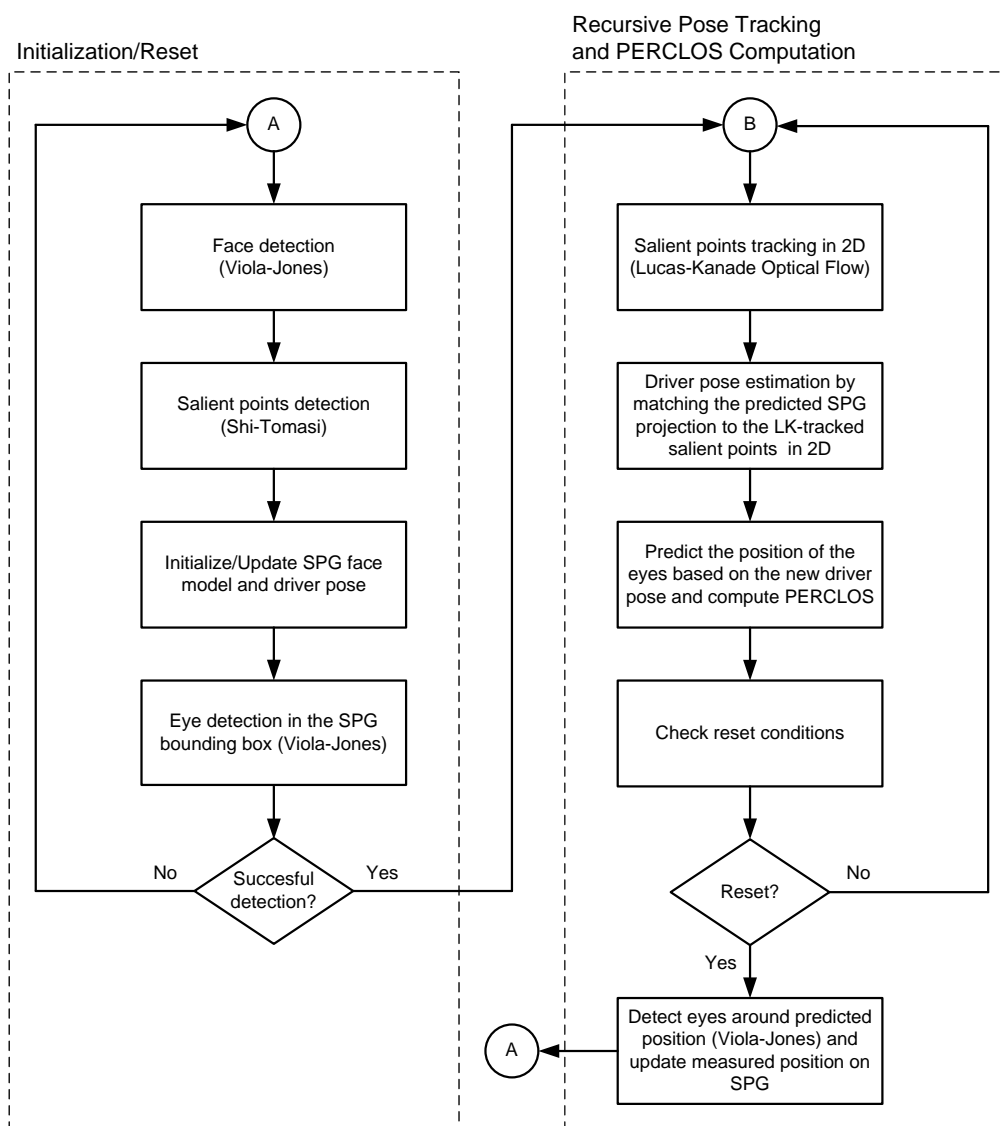
Figure 1. Normal distributions of the different states of alert calculated for the awake, semi-drowsy and drowsy subjects, showing the means, the confidence intervals (CI) and the identification threshold between one level of alertness and the other.



2.2. Driver Detection and Pose Tracking

Correct driver detection and tracking is key to the measurement of driver attention and fatigue. The driver detection and tracking process can be divided into an initialization phase, in which the driver's nominal pose is computed, and an iterative pose estimation and tracking phase, as shown in Figure 2.

Figure 2. Flow chart of the proposed method.

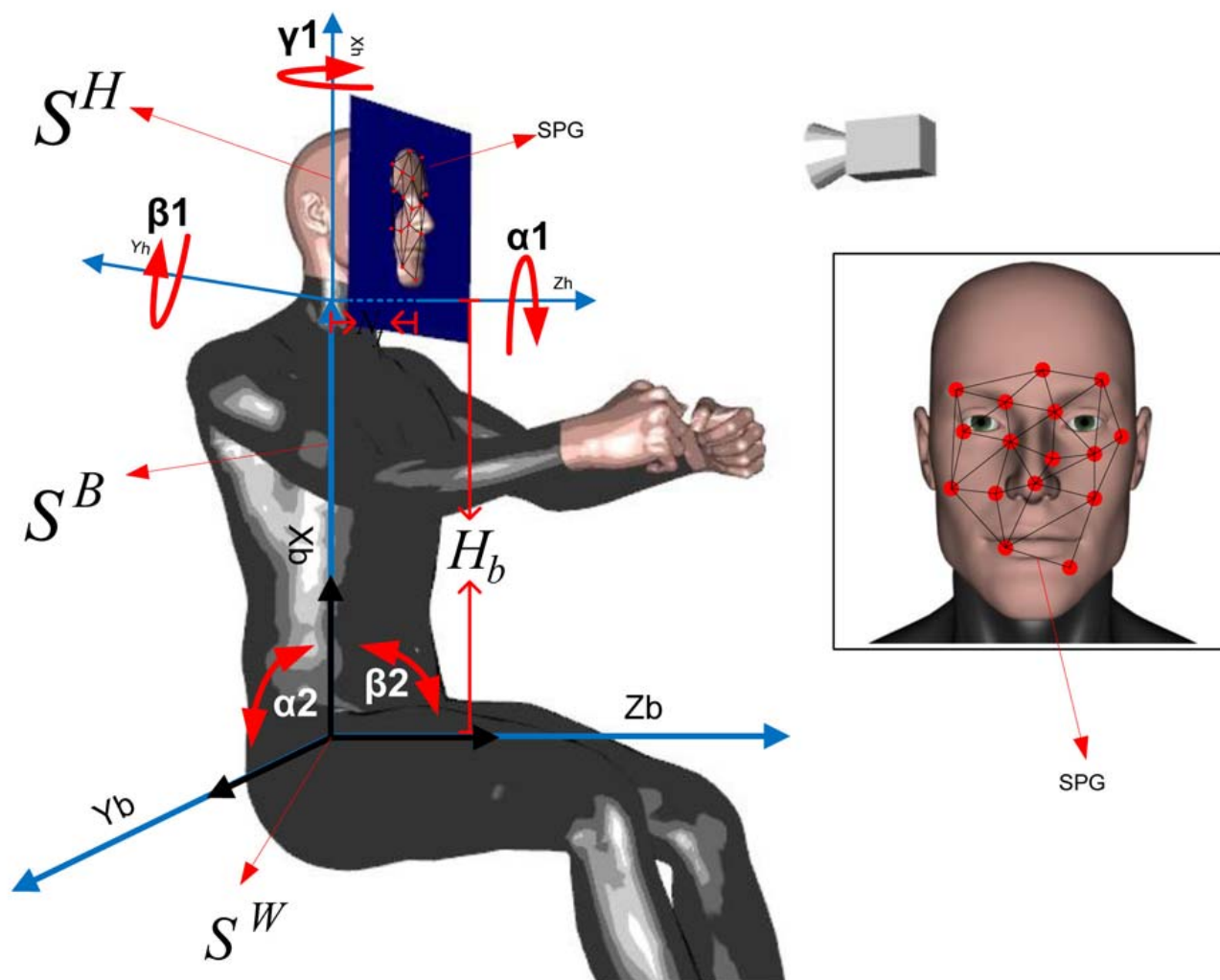


The initialization step starts with the detection of the driver's face using the well-known Viola-Jones object recognition method [55]. A bounding box for the head B^h is computed from the face position information. Distinctive corner points within the face that are good points for tracking regardless of their position, including points that change from face to face, such as eyebrow corners, freckles, moles, scars, beard, *etc.*, are found using the Shi-Tomasi detector [56]. A grid that will be referred to as the *salient points grid* (SPG) is next computed using the distinctive points. The SPG is modeled as a group of coplanar points in 3D space that make up a non-deformable mesh structure having five degrees of freedom (DOF), which replicate the driver's head-torso kinematics. The initialization process ends with the detection of the eyes using the Viola-Jones algorithm trained for such purpose. The search is carried

out in an area defined within B^h , where the eyes are expected to be found, on average, according to the head's anatomy.

Once the initial location of the SPG in 3D space is inferred using the driver's observed interpupillary distance (IPD), the driver's pose (position and orientation) is solved by computing an SPG motion, such that the projection of the SPG points from 3D space onto the 2D image plane coincides with the salient points tracked in the image using the Lukas-Kanade method [57]. The driver's kinematic model, the SPG and the driver's pose estimation method will be explained in greater detail in Sections 2.2.1–2.2.3, respectively. Finally, on each iteration, the driver's eyes are sought within expected regions, according to the newly determined head pose. Tracking the driver's pose significantly improves the ability to locate the driver's eyes, even if for several consecutive frames the eyes cannot be directly detected from the image analysis. Continuous tracking of the head, and, thus, the eyes, greatly simplifies locating the pupils with accuracy and determining whether they are visible or not due to blinking or head rotations, particularly sideways. The analysis of the PERCLOS [14,15] to provide an indication of the driver's state of vigilance will be explained in more detail in Section 2.4.

Figure 3. Driver's kinematics and the salient points grid (SPG) fixed to the driver's face.



2.2.1. Driver's Kinematic Model

The motion of the driver can be decomposed into head motion and torso motion, as depicted in Figure 3. The head motion is achieved by a complex system of muscles, tendons and ligaments in the neck that support and connect the cervical spine with the skull. The different forward/lateral extensor, flexor and rotator muscles allow the head to rotate sideways, bend forward, backward or lean sideways. These movements can be characterized by a 3-DOF spherical joint. On the other hand, the torso motion can be described by a two-DOF universal joint located at the waist that allows the driver to bend forward or backward (frontal tilt) or lean sideways (lateral tilt). For practical purposes, subtle pose changes that are possible with the complex musculoskeletal system of the back are assumed to have been taken into account as part of the waist frontal/lateral tilt or the neck motion.

The driver's upper body kinematics decomposed into torso and neck movements allows one to describe the motion of the head with reasonable fidelity for adequate tracking purposes. Thus the proposed motion model employs three coordinate frames: S^H , S^B , S^W for the head, body and world (fixed base) coordinates, respectively (see Figure 3). Obtaining the location of the head relative to the fixed world coordinate point is done by translation and rotation transformations between the coordinate frames.

For purposes of deriving the coordinate transformations between coordinate frames, first consider that S^H , S^B and S^W are spatially coincident, *i.e.*, are aligned and share the same origin point. The first step to express the position of points on the driver's face relative to S^W consists in computing a transformation to express face points with coordinates in the S^H as points relative to the body frame, S^B . This transformation takes into account the neck motion and yields the head pose relative to the torso by rotating S^H (and the points fixed to S^H) about the z -axis, y -axis and x -axis of S^B by angles α_1 , β_1 and γ_1 , respectively. The rotated S^H must be then translated by a distance H_b along the y -axis of S^B . These transformations yield the following homogeneous transformation matrix:

$$P_h = \begin{bmatrix} c(\alpha_1)c(\beta_1) & c(\alpha_1)s(\beta_1)s(\gamma_1) - s(\alpha_1)c(\gamma_1) & c(\alpha_1)s(\beta_1)c(\gamma_1) + s(\alpha_1)s(\gamma_1) & H_b \\ s(\alpha_1)c(\beta_1) & s(\alpha_1)s(\beta_1)s(\gamma_1) + c(\alpha_1)c(\gamma_1) & s(\alpha_1)s(\beta_1)c(\gamma_1) - c(\alpha_1)s(\gamma_1) & 0 \\ -s(\beta_1) & c(\beta_1)s(\gamma_1) & c(\beta_1)c(\gamma_1) & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (1)$$

where the notation $c(\theta) = \cos(\theta)$ and $s(\theta) = \sin(\theta)$ has been used for brevity.

The next step consists in computing the transformation matrix to express the torso motion relative to the base. To this end, S^B is rotated about the y -axis and z -axis of S^W by angles α_2 and β_2 , respectively, yielding the body homogeneous transformation matrix:

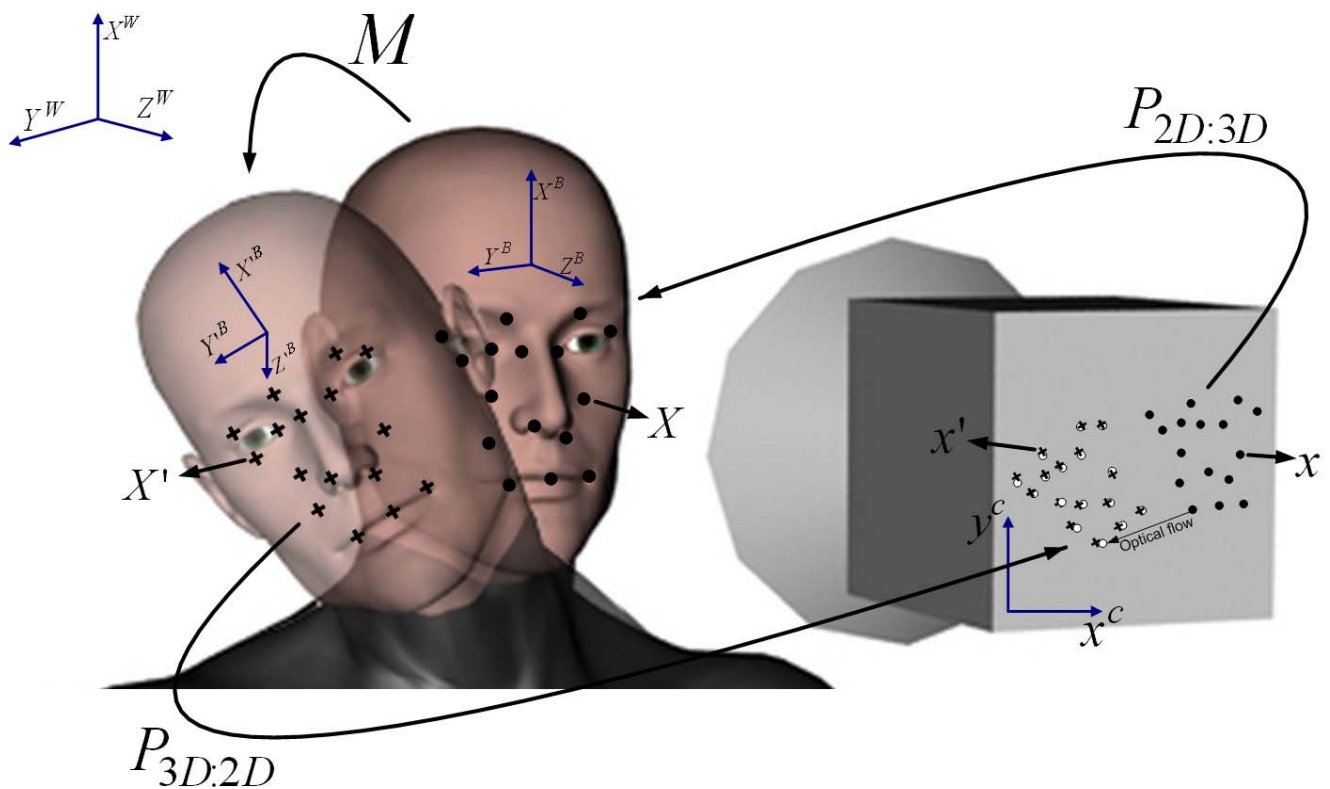
$$P_b = \begin{bmatrix} c(\alpha_2)c(\beta_2) & -s(\alpha_2) & c(\alpha_2)s(\beta_2) & 0 \\ s(\alpha_2)c(\beta_2) & c(\alpha_2) & s(\alpha_2)s(\beta_2) & 0 \\ -s(\beta_2) & 0 & c(\beta_2)c(\gamma_2) & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (2)$$

Finally, for a given driver's pose, described by angles $\alpha_1, \beta_1, \gamma_1, \alpha_2$ and β_2 , the location of any point of the SPG $X^H \in \mathbb{R}^4$ with the coordinates referred to S^H can be expressed as a point $X^W \in \mathbb{R}^4$ with the coordinates referred to, S^W using Equations (1) and (2) to calculate:

$$X^W = PX^H, \text{ with } P = P_b P_h \quad (3)$$

It is to be noted that the pose matrix P depends on the pose angles, $\alpha_1, \beta_1, \gamma_1, \alpha_2$ and β_2 . For two instants k and $k + 1$ corresponding to sampling times t and $t + T_s$ with sampling period $T_s > 0$, there will be two associated pose matrices P_k and P_{k+1} , and a transformation $M_{k+1} : P_k \rightarrow P_{k+1} \in \mathbb{R}^{4 \times 4}$ that maps points X_k^W of the SPG at time k onto points X_{k+1}^W at time $k + 1$. This transformation is illustrated in Figure 4 as mapping M .

Figure 4. Transformation of the salient points involved on the proposed approach.

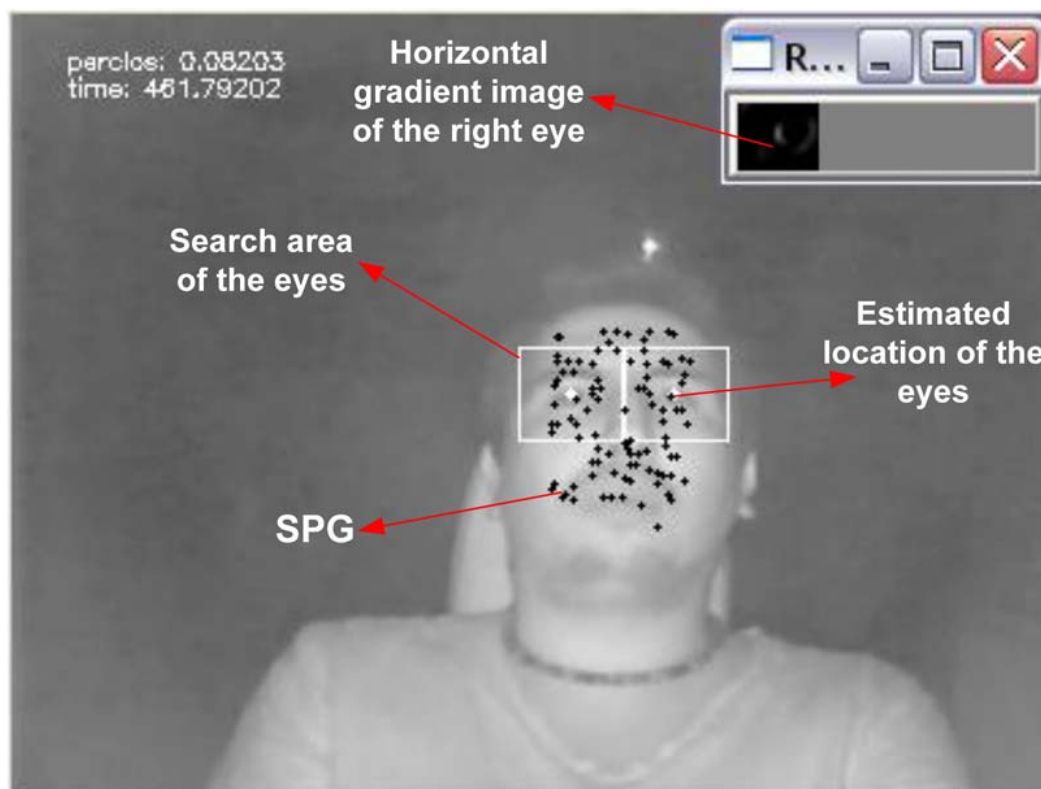


2.2.2. The SPG and the Perspective Projection Model

Visual tracking in 3D space of any object using a single-camera view is a challenging problem, because depth information is lost due to the perspective projection that maps scene points in 3D space onto image points in the 2D sensor plane of the camera. However, whenever some knowledge of the object's geometry and size is available *a priori*, it is often possible to recover 3D motion and pose information. The proposed driver tracking approach takes advantage of the fact that all salient points of the SPG belong to the driver's head, which for the purpose of the proposed approach, can be regarded as a *rigid object of standard size* located at a regular nominal distance from the camera. By *rigid object* it is meant that the skull as a structuring element is non-deformable. Hence, the eyes, the cheek bones and the nose preserve their relative distances with respect to each other. It is to be noted that locally

around the eyes and mouth, the face is a deformable (non-rigid) object that changes when the driver talks or makes gestures. However, unlike our prior work [44], here, we are not considering gestures, such as eyebrow raisings or yawning, as the small deviations of SPG points around the mouth can be handled correctly by the Lucas-Kanade tracker. Therefore, for the purpose of the proposed approach, changes in gesture can be neglected, and the SPG can be treated as a set of salient points that can be consistently tracked and that preserve their relative distance in 3D space, as illustrated in Figure 3 and shown for a real driver in Figure 5. By *rigid object* we do not mean the head is motionless or rigidly fixed. On the other hand, while there do exist correlations between the morphological characteristics of bones and the population that tend to follow geographic boundaries often coinciding with climatic zones, the size of the head changes little across different populations and phenotypes for people 18 years or older (see, for example, [58]). The average male head is around $20 \times 15 \pm 2.2 \times 0.7$ cm, while the average female head is $18 \times 13 \pm 1.2 \times 1.2$ cm. This ± 3 cm variation relative to the camera-head operating distance is negligible. Because of this reason, it is possible to consider the head as an object of *standard size*, that changes little from one driver to another, and due to its low variance, its size is very predictable. Furthermore, an important feature of the proposed approach is that it does not require the heads to be exactly equal, because the SPG is created on-line for each person.

Figure 5. Snapshot of the system running.



Among the salient points, the pupils are two points of special interest, because the IPD provides a reliable reference yardstick of *a priori* well-known and predictable size that is relatively invariant for adult people [59]. For females and males over 18 years of age, the IPD has respectively been estimated to be 62.3 ± 3.6 mm and 64.7 ± 3.7 mm; cf. [59]. If D_p is the driver's IPD and d_p is the length of the

projected IPD onto the image sensor, using the projective geometry equations, it is possible to find the distance, z_{hc} , between the driver's head and the camera as:

$$Z_{hc} = \frac{f}{d_p} D_p \quad (4)$$

where f is the focal length of the camera lens. The accurate knowledge of Z_{hc} in Equation (4) requires the exact knowledge of the D_p , f and the measurement of d_p . However, it is possible to show that even if an error in the knowledge of D_p introduces an error in the estimation of the absolute distance Z_{hc} between the driver and the camera, the relative displacement error between an initial and a final position is negligible. To this end, consider an erroneous measurement \tilde{Z}_{hc} , given by:

$$\tilde{Z}_{hc} = \frac{f}{d_p} \tilde{D}_p$$

due to an erroneous value \tilde{D}_p of the IPD. Then, the error $e_{Z_{hc}}$ between the true distance Z_{hc} and the erroneous measurement \tilde{Z}_{hc} is given by:

$$\begin{aligned} e_{Z_{hc}} &= \tilde{Z}_{hc} - Z_{hc} \\ &= \frac{f}{d_p} \tilde{D}_p - \frac{f}{d_p} D_p \\ &= \frac{f}{d_p} (\tilde{D}_p - D_p) \\ &= \frac{f}{d_p} e_{D_p} \end{aligned}$$

where $e_{D_p} = \tilde{D}_p - D_p$ is the error in the measurement of driver's IPD. The error $e_{Z_{hc}}$ represents the *absolute position error* in the measurement of the distance Z_{hc} . On the other hand, the *relative position error*, i.e., the error in measuring the displacement of the driver from an initial position Z_{hci} to a new position Z_{hcf} due to an error e_{D_p} in measuring the IPD is given by:

$$\begin{aligned} e_{\Delta z} &= \underbrace{\left(\frac{f}{d_{pf}} \tilde{D}_p - \frac{f}{d_{pi}} \tilde{D}_p \right)}_{\equiv \tilde{\Delta z}} - \underbrace{\left(\frac{f}{d_{pf}} D_p - \frac{f}{d_{pi}} D_p \right)}_{\equiv \Delta z} \\ &= f e_{D_p} \left(\frac{1}{d_{pf}} - \frac{1}{d_{pi}} \right) \\ &= f e_{D_p} \left(\frac{Z_{hcf}}{f D_p} - \frac{Z_{hci}}{f D_p} \right) \\ &= \frac{e_{D_p}}{D_p} \Delta z \end{aligned}$$

where $\Delta z = Z_{hcf} - Z_{hci}$ is the driver's displacement along the z -axis of the camera, $d_{pi} = \frac{f D_p}{Z_{hci}}$ and $d_{pf} = \frac{f D_p}{Z_{hcf}}$ are the projections of the interpupilar segment D_p when the driver is at an initial distance Z_{hci} and at a final Z_{hcf} of the camera, respectively.

From a practical standpoint, an error $e_{D_p} = \pm 3\sigma_{IPD} = \pm 11.1$ mm (three-times the standard deviation of the IPD, $\sigma_{IPD} \leq 3.7$, [59]), would yield an absolute position error $e_{Z_{hc}} = \pm 133.2$ mm when using a

lens of focal distance $f = 4$ mm and a typical camera with a pixel size of $\sim 5 \mu\text{m}$ ($d_p \sim 0.33$ mm). On the other hand, for the average driver with an IPD $D_p = 63.5$ mm, the scaling factor $\frac{e_{Dp}}{D_p}$ will be about ± 0.17 . Thus, considering that the driver typically sits about 50–80 cm from the camera, depending mainly on the arms' and legs' length, if the driver moves, for example, from an initial position $Z_{hcf} = 65$ cm to a position $Z_{hcf} = 55$ cm, then $\Delta z = -10$ cm and $e_{\Delta z} = \pm 0.17 \cdot 10 = \pm 1.7$ cm. Hence, the measured displacement will be in the interval $\tilde{\Delta z} \in (-11.7, -8.7)$ cm instead of -10 cm in the worst case. In practice, the video acquisition is done at a speed for which driver displacements Δz will be a few millimeters. Therefore, $e_{\Delta z}$ will be a few tenths of a millimeter between frames, which is negligible for practical purposes, even if in terms of absolute accuracy, and the measurement of the driver position with respect to his or her true position is offset by ± 13 cm in a worst case scenario. Moreover, since we are interested in measuring the relative displacement of the driver from the normal driving position, rather than calculating the exact distance between the driver and the camera with absolute accuracy, and considering that the displacement error is negligible, the method can be applied effectively for actual implementation. It is to be noted that if a developer wishes to have an accurate absolute position measurement, then either the IPD must be entered accurately as a parameter of the system or alternative position sensors, such as simple IR proximity sensors or highly accurate PSD-based sensors, should be included. Since the salient points belong to a 3D rigid object, their geometric constraints (e.g., relative distances among the points) are fully satisfied only in 3D space, but not preserved in the image plane, due to the perspective projection and the relative pose change between the driver and the camera. Hence, obtaining the motion of the head and change of driver's pose is possible by finding a transformation matrix $M : X \in \mathbb{R}^4 \rightarrow X' \in \mathbb{R}^4$ that corresponds to the motion of the SPG, such that the mapping of the initial's pose SPG onto the new pose SPG yields a new SPG, whose projection onto new salient points in the image matches the salient points already tracked in the 2D image plane. This idea is illustrated in Figure 4, which shows an initial head pose and a new head pose associated with a motion transformation M . The SPG for the initial pose is represented by dots, while the SPG for the new pose is represented by crosses in Figure 4. An inverse perspective matrix $P_{2D:3D} : x \in \mathbb{R}^3 \rightarrow X \in \mathbb{R}^4$ allows one to project the salient points in the 2D image back onto the initial head pose. The motion M can then be found as the one that produces a new SPG, whose standard perspective projection $P_{3D:2D} : X' \in \mathbb{R}^4 \rightarrow x' \in \mathbb{R}^3$ maps the new 3D SPG points correctly onto the new salient points in the 2D image found with the Lucas-Kanade optical flow computation method. The motion M is parameterized using the driver's kinematic model Equation (3), while the inverse projective transformation from camera to world coordinates ($P_{2D:3D}$) and *vice versa* ($P_{3D:2D}$) are found as explained next.

For clarity of exposition, it is convenient to introduce some notation first. Let X^W , X^H and X^C be any point in the set \mathcal{S} of salient points that conform the SPG, expressed as homogeneous coordinates in \mathbb{R}^4 referred to as S^W , S^H or S^C (the world, head or camera coordinate frame), respectively. Let $C \in \mathbb{R}^{4 \times 4}$ denote the homogeneous transformation matrix relating the camera position (translation) and orientation (rotation) with respect to S^W , *i.e.*, any point X^C with coordinates relative to the camera frame S^C can be expressed as a point X^W with coordinates relative to the world frame S^W according to:

$$X^W = CX^C \quad (5)$$

If the camera is located such that its coordinate frame S^C has a z -axis that: (i) points towards the driver; (ii) is parallel to the z -axis of S^W ; and (iii) is contained in the XZ plane of S^W , then:

$$C = \left[\begin{array}{ccc|c} R_{S_x^W, \pi} R_{S_z^W, \frac{\pi}{2}} & & & t_{3 \times 1} \\ \hline 0_{1 \times 3} & & & 1 \end{array} \right] = \left[\begin{array}{ccc|c} 1 & 0 & 0 & h_c \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & d_c \\ \hline 0 & 0 & 0 & 1 \end{array} \right] \quad (6)$$

where h_c is the height of the camera with respect to the seat (plane YZ of S^W in Figure 3), d_c is the distance from the camera's focal point to the z -axis of S^W and $R_{v, \theta}$ represents the rotation matrix about axis v by an angle θ , while $t_{3 \times 1}$ is the translation vector from the origin of S^W to the origin of S^C .

Points $X^C = [X_x^C, X_y^C, X_z^C, 1]^T$ in homogeneous coordinates of the camera frame S^C can be projected onto the camera's optical plane as points:

$$p = [p_x, p_y, 1]^T = \Pi X^C \quad (7)$$

using the standard pin hole camera model with a perspective projection matrix Π given by:

$$\Pi = \begin{bmatrix} \frac{f}{X_z^C} & 0 & 0 & 0 \\ 0 & \frac{f}{X_z^C} & 0 & 0 \\ 0 & 0 & \frac{1}{X_z^C} & 0 \end{bmatrix} \quad (8)$$

where f is the focal length of the camera and X_z^C is the distance between the point X^C and the camera's focal point measured along the z -axis of S^C . For simplicity of exposition, here Π is the perspective projection matrix for a coordinate frame S^C with origin located at the camera's focal point. If the origin of S^C is displaced from the focal point, then the last column of Equation (8) must also include the translation terms.

The distance Z_{hc} provides an initial value of X_z^C for all points in S . If the image coordinates (p_x, p_y) of point p , corresponding to point X^C and the distance X_z^C to point X^C are known, then it is possible to define an *inverse perspective mapping* $\Gamma : \mathbb{R}^3 \rightarrow \mathbb{R}^4$ that projects point p on the image plane back onto X^C as:

$$\Gamma = \begin{bmatrix} \frac{X_z^C}{f} & 0 & 0 \\ 0 & \frac{X_z^C}{f} & 0 \\ 0 & 0 & X_z^C \\ 0 & 0 & 1 \end{bmatrix} \quad (9)$$

such that:

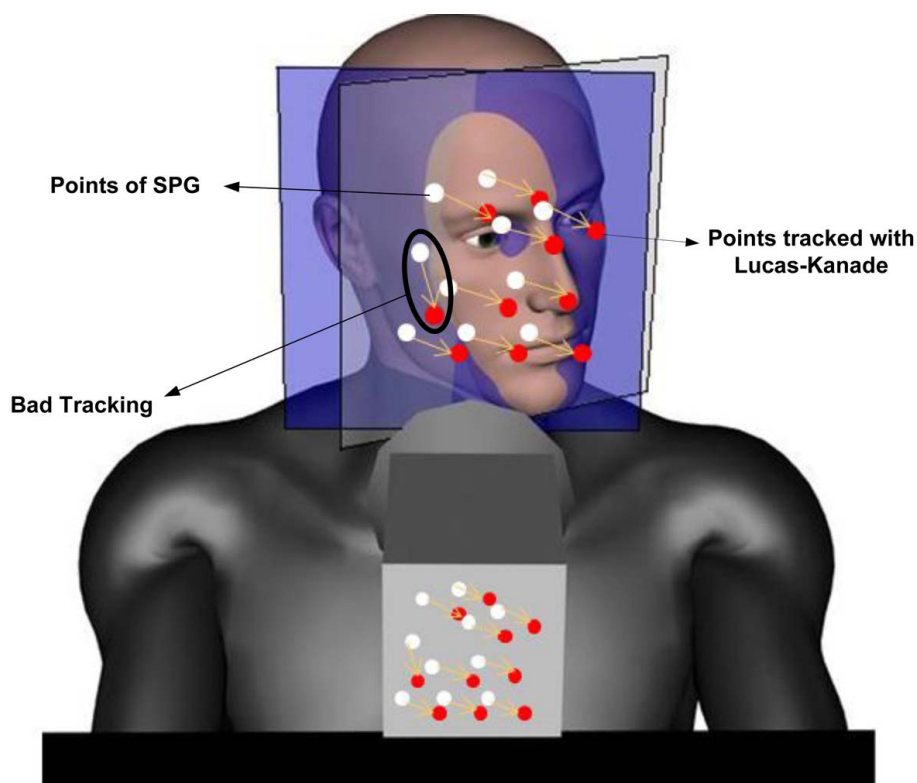
$$X^C = \Gamma p \quad (10)$$

The back-projection of points p onto points X^C of the SPG is illustrated in Figure 3 as a projection $P_{2D:3D}$ from 2D to 3D. On the other hand, the standard perspective mapping Π projecting 3D SPG points onto 2D points is represented in Figure 3 as the projection $P_{3D:2D}$.

2.2.3. Driver Pose Estimation

The driver's pose estimation problem consists in finding the pose angles α_1 , β_1 , γ_1 , α_2 and β_2 at time instant $k + 1$, given the knowledge of the pose at time k and the driver's motion M_{k+1} at instant $k + 1$ as measured from the image. The proposed approach to estimate and track the driver's pose angles employs the Lucas-Kanade's (LK) method to optical flow computation [57]. The LK method computes a set \mathcal{S}_k of salient points, $p_{j,k}$, $j = 1, 2, \dots, N$, in an image frame at instant k and tracks point-by-point yielding weights $w_{j,k+1}$ and a set \mathcal{S}_{k+1} of salient points $p_{j,k+1}$, $j = 1, 2, \dots, N$, in the image frame at instant $k + 1$ corresponding to the points $p_{j,k}$, $j = 1, 2, \dots, N$, in the previous frame, as illustrated in Figure 6. The velocity at which corresponding pixels move from $p_{j,k}$ to $p_{j,k+1}$, approximated by $\phi_{j,k+1} = (p_{j,k+1} - p_{j,k})/T_s$ for a sampling period T_s , is the so-called *optical flow* of the image's intensity at pixel $p_{j,k}$ at time instant $k + 1$. Each weight $w_{j,k+1}$, $j = 1, 2, \dots, N$, is a measure of the similarity between a pair of corresponding points $p_{j,k}$ and $p_{j,k+1}$ computed as the convolution of pixel neighborhoods surrounding $p_{j,k}$ and $p_{j,k+1}$. The weights $w_{j,k+1}$, $j = 1, 2, \dots, N$, provide a measure of the quality and reliability of the match and are particularly useful to discard points with lower weights, which are more likely to occur near the boundaries of the SPG when the head turns, as some points will become occluded. Bad tracking of some of the salient points, as depicted in Figure 6, may also occur when the points are occluded by an external object, like the driver's hand, or when weak saliency, due to low texture or contrast, makes correspondences ambiguous (non-unique).

Figure 6. Example of LK tracking between two frames, showing an incorrectly tracked point.



2.2.4. Tracking Reset Rules

Using the optical flow information $\phi_{j,k+1}$, $p_{j,k}$, $p_{j,k+1}$ and $w_{j,k}$, corresponding to the sets \mathcal{S}_k and \mathcal{S}_{k+1} of SPG points, the perspective and back-projection mappings Π and Γ , and the pose matrices P_k and P_{k+1} , it is possible to formulate the pose estimation problem considering that:

$$X_{j,k+1}^W = M_{k+1}X_{j,k}^W, \quad j = 1, 2, \dots, N \quad (11)$$

relates all SPG points $X_{j,k}^W \in \mathcal{S}_k$ to SPG points $X_{j,k+1}^W \in \mathcal{S}_{k+1}$ through the driver's motion M_{k+1} at instant $k + 1$. By Equation (3), the SPG points $j = 1, 2, \dots, N$ in world coordinates of frame S^W are related to the SPG points in the coordinates of the head frame S^H by:

$$\begin{aligned} X_{j,k}^W &= P_k X_j^H \\ X_{j,k+1}^W &= P_{k+1} X_j^H \end{aligned} \quad (12)$$

It is to be noted that the SPG points X_j^H have been written without a dependency on the time instant, because the SPG is assumed to be a rigid structure attached to the driver's head and, thus, move consistently with S^H . Although this assumption is violated for a limited number of points close to the mouth and eyebrows that move relative to S^H when the driver talks or makes gestures, the distance these points travel and their speed is negligible for practical purposes compared to that of S^H relative to S^W when the driver changes pose. Hence, replacing Equations (12) into (11) yields:

$$P_{k+1}X_j^H = M_{k+1}P_kX_j^H \Rightarrow P_{k+1} = M_{k+1}P_k \Rightarrow M_{k+1} = P_{k+1}P_k^{-1} \quad (13)$$

On the other hand, using Equations (5) and (10), the SPG points in camera coordinates S^C are related to the SPG points in world coordinates according to:

$$\begin{aligned} X_{j,k}^W &= CX_{j,k}^C = C\Gamma p_{j,k} \\ X_{j,k+1}^W &= CX_{j,k+1}^C = C\Gamma p_{j,k+1} \end{aligned} \quad (14)$$

Equations (13) and (14) allow one to rewrite Equation (11) in terms of the optical flow pair $p_{j,k}$, $p_{j,k+1}$, and the pose matrix P_{k+1} as:

$$C\Gamma p_{j,k+1} = P_{k+1}P_k^{-1}C\Gamma p_{j,k}, \quad j = 1, 2, \dots, N \quad (15)$$

and therefore, by Equations (10) and (7):

$$p_{j,k+1} = \Pi C^{-1}P_{k+1}P_k^{-1}C\Gamma p_{j,k}, \quad j = 1, 2, \dots, N \quad (16)$$

Defining the matrix:

$$T_{k+1}(\alpha_1, \beta_1, \gamma_1, \alpha_2, \beta_2) \stackrel{def}{=} \Pi C^{-1}P_{k+1}P_k^{-1}C\Gamma \quad (17)$$

as the transformation matrix that maps points $p_{j,k}$ onto $p_{j,k+1}$ due to a pose change from P_k to P_{k+1} dependent on the new pose angles α_1 , β_1 , γ_1 , α_2 and β_2 , at instant $k + 1$, the pose estimation problem can be formulated as the following optimization problem:

$$\Theta_{k+1}^* = \arg \min_{\alpha_1, \beta_1, \gamma_1, \alpha_2, \beta_2} \sum_i w_{j,k+1} \|T(\alpha_1, \beta_1, \gamma_1, \alpha_2, \beta_2)p_{j,k} - p_{j,k+1}\| \quad (18)$$

of finding the set of pose angles $\Theta_{k+1}^* = (\alpha_1^*, \beta_1^*, \gamma_1^*, \alpha_2^*, \beta_2^*)$ at instant $k + 1$ that minimizes the matching error between the pair of points $p_{j,k}, p_{j,k+1}$ delivered by the LK optical flow computation method considering the knowledge of the pose P_k obtained at time k in the previous iteration. It is to be noted that the transformation matrix $T_{k+1}(\alpha_1, \beta_1, \gamma_1, \alpha_2, \beta_2)$ in Equation (17) is a 3×3 identity matrix when the pose of the driver remains constant, *i.e.*, $P_k = P_{k+1}$, and the salient points in the image remain static, *i.e.*, $p_{j,k} = p_{j,k+1}, j = 1, 2, \dots, N$. This means that solving Equation (18) seeks to find the angles Θ_{k+1}^* for P_{k+1} that match precisely those of P_k , so that $P_{k+1}P_k^{-1} = I_{3 \times 3}$, ensuring that the value of the cost function is driven to zero.

The minimization problem Equation (18) is a nonlinear least squares problem, which can be solved by different gradient methods, Newton's method or direct search methods; see, for example, [60,61]. In our implementation, problem Equation (18) was solved using the Levenberg-Marquardt variant of the Gauss-Newton algorithm, as well as the direct search approach by Nelder-Mead. While the Nelder-Mead algorithm may converge to a non-stationary point, in practice, it converged faster than the Levenberg-Marquardt approach and was preferred for this reason. Despite that the theoretical convergence properties of the direct search approaches are often not satisfactory, algorithms, such as Nelder-Mead's, are known to work reasonably well for problems of relatively small dimension (up to 10) [60]. On the other hand, in the case of our problem, the angles are bounded and the initial driver position is known to be constrained to a specific range. This allows one to initialize the algorithm correctly without any danger of converging to a local minima. Moreover, tracking the angles with the extended Kalman filter (EKF) allows one to initialize the solution of Equation (18) at each iteration with the predicted values for angles. This ensures that the search for the solution starts at a close value with respect to the true pose that is being sought. It is also to be noted that the proposed approach considers reset conditions that allow one to restart the process of finding the driver's pose, so the actual risk of a permanently diverging solution is inexistent.

Considering that the pose estimation approach is incremental (*i.e.*, the proposed approach estimates a new pose starting from the pose estimated in the previous iteration, as explained in the preceding sections), small errors in the pose estimation occurring in some frames can accumulate over time. To prevent errors from accumulating, two reset conditions are implemented: (i) the system checks if the motion of the driver has not produced angles $\alpha_1, \beta_1, \gamma_1, \alpha_2$ or β_2 exceeding $\pm 20^\circ$ and (ii) the eyes are detected correctly using the Viola-Jones approach at least once every certain number of frames (our implementation checks for a correct eyes detection every 100 frames at 15 fps; the tracking of the eyes on the remaining frames relies on the salient points of the SPG, as explained in the next section). If any of the two conditions is not satisfied, then the tracking system and the pose vector is reset every n_s frames, until both conditions are satisfied again, in which case the tracking system and the pose vector are reset every n_l frames. In our implementation, n_s was set to 100 frames, while n_l was set to 5,000 frames, thus ensuring good tracking results by keeping the cumulative error to a minimum.

2.3. Eyes Location and Tracking

The eyes' location is initially obtained using the Viola-Jones approach in a sub-window within the SPG. The Viola-Jones detection approach is not used again during the normal operation of the system,

unless certain reset conditions occur (see Section 2.2.4). Once the initial location of the eyes is found, the Viola-Jones approach is not employed on each iteration for two reasons. First, the eye recognition becomes difficult or impossible under partial eye occlusions when the driver stares away from the camera, changes pose or temporarily moves his hands or an object in the line of sight of the camera. On the other hand, the Viola-Jones recognition approach is computationally more expensive than the tracking of the SPG points. Moreover, the SPG provides a set of reference points that allows one to locate the eyes relative to the SPG. Thus, if a few points of the SPG are lost due to occlusions or pose changes, the eyes can still be located relative to the remaining points in the SPG.

Consider the points $X_{j,k}^W, j = 1, 2, \dots, N$, in the SPG set \mathcal{S}_k at instant k expressed in the coordinates of S^W , and denote by $e_k^i \in \mathbb{R}^3, i = l, r$, the location of the left and right eye's pupil in homogeneous image coordinates. By Equation (10), the projection of $e_k^i \in \mathbb{R}^3$ onto the SPG, denoted by $E_k^{C,i} \in \mathbb{R}^4$, is given by $E_k^{C,i} = \Gamma e_k^i, i = l, r$. The location of the left and right eye relative to the SPG points can then be computed at every instant k as:

$$V_{j,k}^{C,i} = E_k^{C,i} - X_{j,k}^C, \quad i = l, r$$

Therefore, when the SPG points are obtained in iteration $k + 1$, the new location of the eyes can be estimated at $k + 1$ from the weighted average of relative displacement of the eyes relative to the new SPG points:

$$\begin{aligned} E_{k+1}^{C,i} &= \sum_j w_{j,k+1} \cdot (X_{j,k+1}^C + V_{j,k}^{C,i}) \\ &= \sum_j w_{j,k+1} \cdot (\Gamma p_{j,k+1} + V_{j,k}^{C,i}) \\ &= \sum_j w_{j,k+1} \cdot (\Gamma T_{k+1} p_{j,k} + V_{j,k}^{C,i}), \quad i = l, r \end{aligned}$$

where T_{k+1} is the transformation matrix defined in Equation (17) representing the motion of the driver at instant $k + 1$ according to the kinematic model Equation (3) and $w_{j,k+1}$ are the weights of the salient points computed by the LK method at time $k + 1$. Since $\Pi E_k^{C,i} = e_k^i, \forall k$, pre-multiplying by the perspective projection matrix Π , the last equation can be rewritten as:

$$\begin{aligned} e_{k+1}^i &= \sum_j w_{j,k+1} \cdot (\Pi \Gamma T_{k+1} p_{j,k} + \Pi V_{j,k}^{C,i}) \\ &= \sum_j w_{j,k+1} \cdot (\Pi \Gamma T_{k+1} p_{j,k} + e_k^i - p_{j,k}) \\ &= \sum_j w_{j,k+1} \cdot (e_k^i + (T_{k+1} p_{j,k} - p_{j,k})), \quad i = l, r \end{aligned} \quad (19)$$

since $\Pi \Gamma = I_{3 \times 3}$. It is to be noted that by Equations (16) and (17), $T_{k+1} p_{j,k}$ is the expected position of the j -th point of the SPG at time $k + 1$, i.e., $p_{j,k+1} = T_{k+1} p_{j,k}$, and therefore, $T_{k+1} p_{j,k} - p_{j,k} = p_{j,k+1} - p_{j,k} = \phi_{j,k+1}$ is the optical flow for the j -th SPG point. This result is important, because it implies that the eyes' location within the image can be updated and tracked using the weighted average of the optical flow $\phi_{j,k}$ of salient points instead of carrying out a more complex eyes recognition process.

This is also advantageous, because in addition to penalizing those points that have lower weights $w_{j,k+1}$, it means that instead of implementing a Kalman filter to track all the points in the SPG in 3D space, it should suffice to correctly estimate the pose values, $\alpha_1, \beta_1, \gamma_1, \alpha_2, \beta_2$. In our approach, an EKF is implemented using the kinematic model as a driving process to predict the pose angles together with the new measurements of the eyes' location given by Equation (19) to update the state estimate. It will be shown in the results section that by doing so, the eyes' can be tracked accurately with a high success rate, despite the driver's motion.

2.4. Driver Vigilance Measurement and Blink Detection

A driver's state of alert is a combination of factors that include fatigue, drowsiness and distraction from the driving task, while talking to other passengers or persons on a mobile phone. One of the indicators of distraction is the driver's pose, especially whenever the driver's head is not staring forward. While this information can be obtained with the proposed pose estimation and tracking approach, a more critical risk factor is fatigue and drowsiness, since it impairs the driver's attention and diminishes his or her ability to recover from wrong maneuvers until the necessary rest is taken. Thus far, the best measure of fatigue and drowsiness is the percentage of eye closure (PERCLOS) over some period of time [14,15,50,51]. More precisely, PERCLOS is calculated as the ratio between the amount of time the eyes are closed (pupils are 80% or more covered by the eyelids) with respect to the total time lapse:

$$\text{PERCLOS} = \frac{t_c}{t_c + t_o} \quad (20)$$

where t_c is the time the eyes are closed and t_o is the time the eyes are open. This measure is typically computed over a running window lasting one minute.

Several studies ([14,26,27,50,51,62]) have demonstrated that the PERCLOS measure has a high correlation with the level of drowsiness. One of the most important studies was carried out by the Federal Highway Administration of the United States [14] and showed that a person's PERCLOS increases directly with the level of fatigue. The test was made keeping ten subjects awake for 42 h and taking tests of PERCLOS and reaction time every two hours. The results show an average correlation between the reaction time and drowsiness of 0.878. Our results obtained from the tests carried out in a simulator are consistent with the previous studies about PERCLOS reported in the literature.

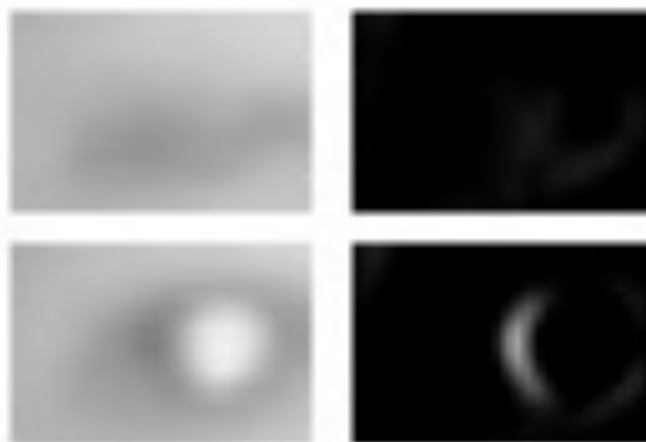
In order to detect blinks and determine whether the eyes are open or closed at every sampling instant k , a horizontal Laplacian filter is applied to a neighborhood $\mathcal{N}(e_k^i)$ of the image around the eyes central position e_k^i , $i = l, r$. The neighborhood $\mathcal{N}(e_k^i)$, $i = l, r$, has a width and height, respectively, equal to 18% and 33% of the bounding box for the driver's face SPG. The average $\bar{G}_{x,k}^i$ of the resulting horizontal gradient $G_{x,k}^i$ for the image subregion $I_{\mathcal{N}(e_k^i)} = \{I(p) | p \in \mathcal{N}(e_k^i)\}$, given by:

$$\bar{G}_{x,k}^i = \frac{1}{|\mathcal{B}_k^s|} \sum_{p \in \mathcal{N}(e_k^i)} G_{x,k}^i(p), \quad i = l, r \quad (21)$$

is calculated to determine the state of the eyes. When the eyes are open, the number of line segments in the vertical direction increases (pupils and corners of the eyes), and therefore, the horizontal gradient $G_{x,k}^i(p)$ contains more vertical edges. On the other hand, when the eyes are closed, only the horizontal

line of the eyelid is visible, and the response $G_{x,k}^i(p)$ to the horizontal gradient applied to $I_{\mathcal{N}(e_k^i)}$ is weaker, thus $\bar{G}_{x,k}^i$ decreases to a minimum when the eye closes. In order to establish whether a blinking has occurred, $\bar{G}_{x,k}^i$ is compared to a threshold η determined experimentally in such a way as to maximize the rate of detection while minimizing the rate of false alarms. Figure 7 shows a closed eye (upper-left), an open eye (lower-left), and the corresponding responses $G_{x,k}^i$ to the horizontal Laplacian filter. When the eyes are open, the lower right image clearly shows a more intense response on the vertical segments of the pupil, while the upper right image shows that the intensities in the response are weaker due to the lack of vertical edges.

Figure 7. Eye image and filter image for both open and closed eye.



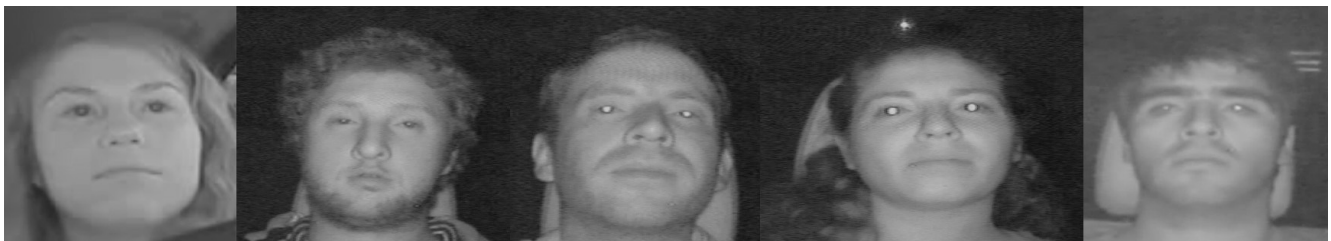
3. Testing Methodology

In order to validate the efficacy of the proposed approach, fifteen volunteers participated in the experiments carried out using a driving simulator. The subjects were divided into three groups for the purpose of comparing PERCLOS measures and reaction times at different levels of fatigue. Five individuals were fully rested (slept the regular 7–8 h), five individuals had minimal rest (slept at least 3.5 h, but not more than 4 h) and five individuals had no rest at all, *i.e.*, were asked not to sleep from one day to the next. The experiments were carried out on Saturdays between 9:00 and 10:00 AM for each participant with sleep deprivation, between 10:00 and 11:00 for participants with partial rest and between 11:00 and 12:00 for participants with full rest, thus requiring a month to collect the data from the fifteen subjects. All subjects were requested to have a regular 7–8 h sleep on the five days previous to the experiment and to record their sleep time from the time they went to bed until the time they woke up. The average sleep time for the group was 7.43 ± 0.61 h. The participants were asked to follow their regular work-day routines, including three meals. All participants declared to have no sleep disorders nor to be under any medication that could produce sleepiness. The fifteen volunteers were all first-time users of the simulator and had only five minutes to practice driving before the initial reaction-time measurement was carried out. The initial reaction time measurements lasted approximately another five minutes and were followed by the actual driving period of forty-five minutes, as explained below.

For the purpose of establishing the influence of drowsiness and fatigue on the tracking and PERCLOS measurements, while minimizing the influence of other factors, like age and phenotype, the participants

for this experiment were restricted to a group of similar characteristics consisting of drivers 24–26 years old (six females, nine males). The phenotypes of the group were similar, as shown in Figure 8. Skin colors were in the range from white to brown. Hair color or length was not an issue, since the approach employs a bounding box restricted to the face that encloses the eyes, eyebrows, nose and mouth with a small margin above the eyebrows and below the mouth. In each group, there were three males and two females. One of the three males in each group had a short beard. In each of the three groups, there was one driver that wore prescription polycarbonate glasses and one driver that wore prescription disposable contact lenses. The use of contact lenses had no visible effect on the reflection of the IR illumination. In fact, the transmittance spectrum for most disposable contact lenses is close to 90% for wavelengths above 400 nm, *i.e.*, they block UV light [63] and smaller wavelengths, but are almost transparent to light in the visible and IR spectrum. For uncoated polycarbonate and glass lenses, the transmittance spectrum is similar to that of disposable contact lenses (*cf.* vol. 1, ch. 51D, in [64], or [65]). Therefore, the minor reflections due to the anti-glare coatings were not an impediment to detect the eye blinks. Sunglasses and tinted lenses have lower transmittances for near-infrared, typically 40–60%, depending on the coating and tint. These type of tinted glasses were not considered for our experiments.

Figure 8. Five of the fifteen participants in the driving experiments under different levels of sleep deprivation.



The experiments started with a measurement of the participants' reaction time taken before driving. To this end, each subject was required to press a button as fast as possible whenever a green spot would turn red on the simulator's projection screen. This procedure was repeated fifty times using random amounts of time lasting from two to ten seconds between each reaction test.

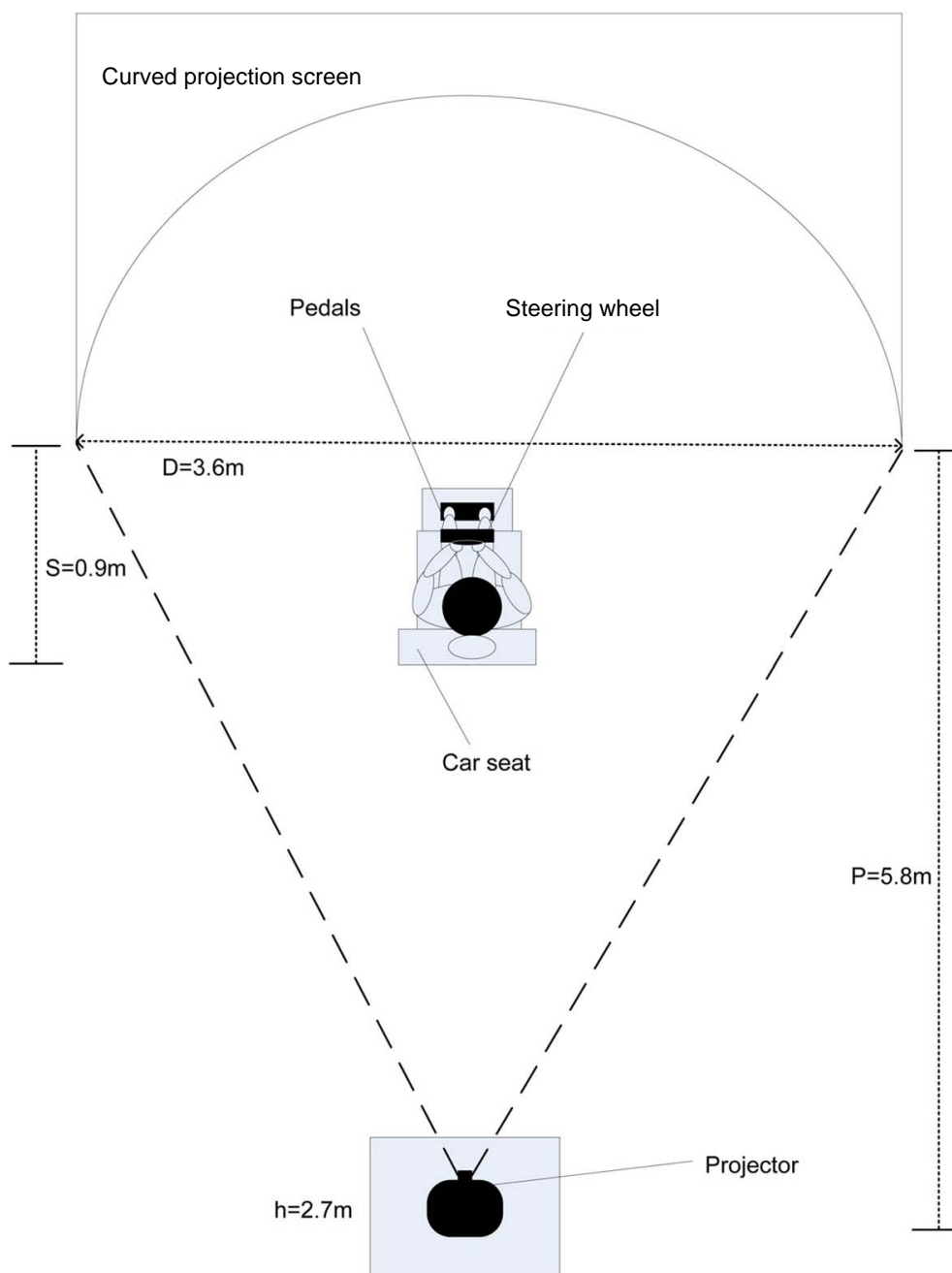
Once the tests to measure reaction time had been completed, each participant had to drive for forty-five minutes along a rather monotonous track scenario simulating a desert with hills and very few turns. The purpose of the chosen scenario was to induce drivers into falling asleep, while keeping visual distractions to a minimum to prevent arousing the driver's attention.

During the experiments, the driver's reaction time and driving behavior were simultaneously analyzed on-line and recorded with the data capture system implemented to that end. A snapshot of the software implemented to extract salient points and compute the PERCLOS measure is shown in Figure 5. Some of the participants in the driver's fatigue measurement experiment are shown in Figure 8.

The car simulator was built inside a closed lab with no external light sources using a Ford Escape 2009 seat and a Momo Racing Force Feedback Steering Wheel by Logitech, which included gas and brake pedals. A Viewsonic high resolution digital projector was used to project the scenes on a cylindrical projection plane, whose purpose was to immerse the driver into the virtual driving scenario and contribute to the realism perceived by the driver, due to the effects of video motion on the peripheral vision. In other

words, the curved backdrop surrounding the driver enhances the persons velocity sensation that would otherwise be very poor if a planar surface would have been used instead. The software employed to create the driving environment is the open source driving simulator Racer [66], which was configured to limit the driving speed to 100 km/h. A sound system was employed to generate the characteristic sound of a regular combustion motor vehicle.

Figure 9. Driving simulator layout.



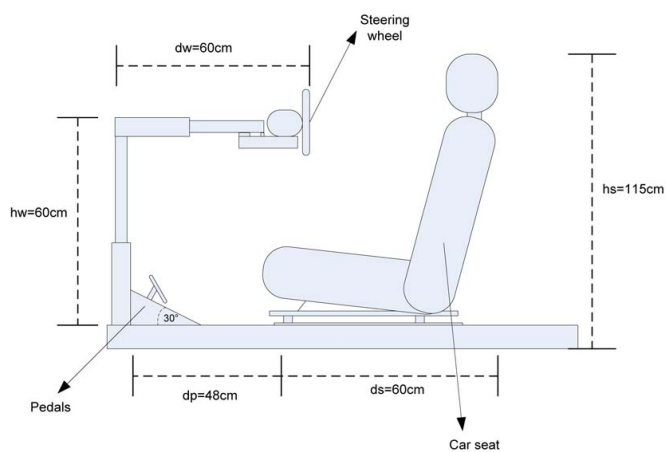
The layout of the simulator is illustrated in Figure 9, which shows the semicircular projection screen of 1.8 m radius, the projector located 5.8 m from the projection screen and 2.7 m above the ground, to avoid the car seat structure from casting shadows on the screen. The rear part of the seat structure is 0.9 m away from the center of the semicircular projection screen. This location ensures that the driver field of view subtends the whole projection screen and not just the central portion and also ensures that the

driver perceives the virtual world with a scale equivalent to that perceived from a real vehicle, as shown in Figure 10 for one of the driving experiments. From the seating position of the driver in Figure 10, the pavement below the seat and the shadow cast on the screen are not visible. This was possible locating the projector above the screen level and adjusting the keystone effect. The driving seat and its dimensions are shown in Figure 11.

Figure 10. Driving simulator during one of the experiments.



Figure 11. Dimensions and view of the driving seat and steering wheel structure.



4. Results

Eye tracking rates obtained with the proposed approach and two other comparison methods are summarized in Table 1. One of the comparison approaches presented in Table 1 is based on the direct detection of the eyes using the Viola-Jones recognition approach. The second comparison method is based on a salient points tracking approach, but without considering the driver kinematics nor its pose information. This approach had been proposed by the authors in [44] and significantly improved in this work by including the driver kinematics and the proposed scheme for tracking the SPG points, as shown by the results in the last column of Table 1. The results in Table 1 show that a relatively low ($38.03\% \pm 13.57\%$) tracking rate of the eyes is achieved by direct application of the Viola-Jones technique trained to detect eyes. Compared to the approach based on the direct eyes identification on every frame, the SPG tracking approach significantly improves the success of the eye tracking system with an average success rate of $97.10\% \pm 2.39$. Using the proposed method with the driver's kinematic model, an additional improvement in the tracking rate of 2.31% is possible, yielding on average a failure rate below 1% . The high tracking rate of the proposed approach ensures that the state of the alert measurement system would be able to compute the PERCLOS on practically every image frame, unlike the other approaches that are more sensitive to pose changes and rapid driver movements.

Table 1. Eye tracking results with the proposed approach.

Subject	Frames	Tracking Rate [%]		
		Viola-Jones Eyes Recognition	SPG Tracking without Driver Kinematics ([44])	SPG Tracking with Driver Kinematics
1	2768	22.24	93.39	100.00
2	6122	43.16	99.46	100.00
3	5219	21.18	98.58	100.00
4	3310	58.29	94.46	96.79
5	5253	44.33	98.50	99.48
Mean \pm 95% C.I.		38.03 ± 13.57	97.10 ± 2.39	99.41 ± 1.31

In addition to correctly tracking the driver's motion, an effective drowsiness warning system must be able to differentiate the driver's state of alert. To verify this requirement, the PERCLOS was computed for the different groups of awake, semi-awake and drowsy subjects using the proposed approach. The mean PERCLOS values computed for each group of drivers using windows of 60 seconds are summarized in Table 2 and clearly exhibit an increase for the groups with less hours of sleep. It is also possible to notice that the reaction time increases for the group of drowsy subjects. However, the average reaction time of the fully awake and the semi-drowsy subjects shows little change in contrast to the PERCLOS, which on average is more than doubled. On the other hand, it was observed that the reaction time does not directly correlate with the level of sleep, since some well-rested drivers had average reaction times larger than that of drivers in the drowsy group.

Table 2. PERCLOS average, standard deviation, 95% confidence interval for the different subjects (1–5 awake, 6–10 semi-drowsy and 11–15 drowsy) and reaction time, pose and pose rate-of-change for the ensemble of awake, semi-drowsy and drowsy subjects.

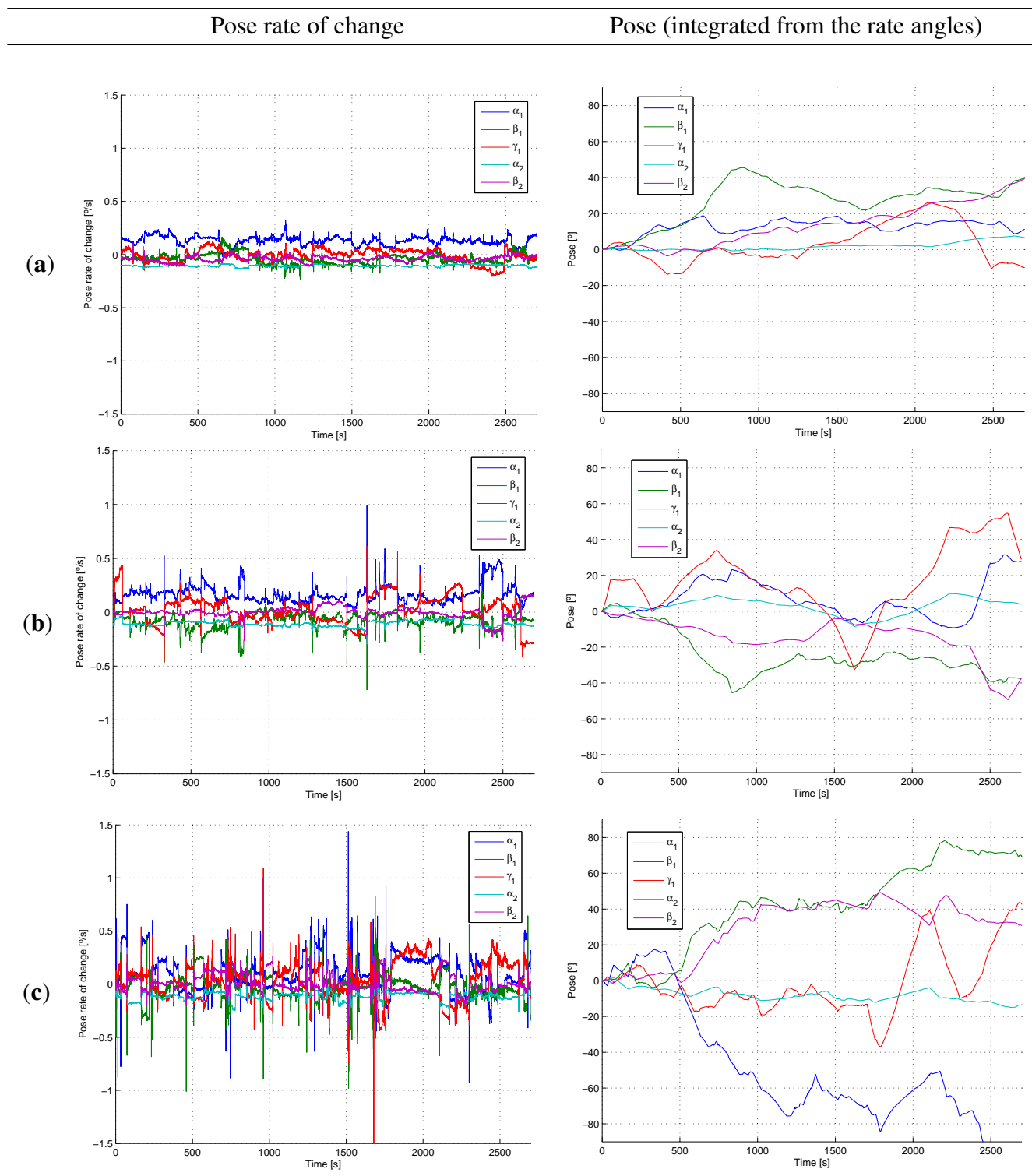
Subject A/S/D *	PERCLOS		
	Awake	Semi-Drowsy	Drowsy
1/6/11	0.0281	0.1089	0.1891
2/7/12	0.0329	0.0660	0.1768
3/8/13	0.0291	0.0905	0.2099
4/9/14	0.0444	0.1088	0.0975
5/10/15	0.0255	0.0663	0.2262
Mean	0.0320	0.0881	0.1799
Std. Dev.	0.0074	0.0214	0.0499
95% C.I.	±0.0021	±0.0084	±0.0146
Avg. Reaction Time [ms]	199.0	204.6	262.5
Avg. Pose Magnitude RMS [°]	61.7	51.1	67.7
Avg. Pose Rate RMS [°/s]	0.13	0.20	0.25

*Subject A = awake; S = semi-drowsy; D = drowsy.

Another indicator of drowsiness is the change in the driver's pose. Table 2 presents the average root mean square (RMS) value of the *pose magnitude* and the *pose rate magnitude* for the different group of drivers. The pose magnitude is computed as the Euclidean norm of the vector of pose angles $(\alpha_1, \beta_1, \gamma_1, \alpha_2, \beta_2)$. Similarly, the pose rate magnitude is calculated as the Euclidean norm of the vector containing the time derivatives of the pose angles. The pose magnitude RMS value for the awake and drowsy drivers was similar, and on average, larger than that of semi-drowsy drivers. However, there is a positive correlation between the mean PERCLOS measure and the RMS value of the rate of change in pose. This is consistent with the knowledge that a drowsy driver will attempt to make fast sudden corrections to deviations from the lane, and it is expectable that the driver will also try to regain the sitting pose quickly, while avoiding to fall asleep. It has been argued, see for example [6,12,13], that monitoring corrections in driving maneuvers and pose changes may not provide information sufficiently in advance to warn the driver. In fact, the evolution of the pose of the driver in time does not seem to provide an indication of fatigue as clearly as the rate of change of the pose, according to our results in Figure 12 for a selection of one awake, one semi-drowsy and one drowsy driver. Due to space limitations, it is not possible to include the plots for the fifteen subjects; however, the curves have similar evolutions for drivers within the same group (awake, semi-drowsy or drowsy). The first column of Figure 12 presents the rate of change of the pose angles, $\alpha_1, \beta_1, \gamma_1, \alpha_2$ and β_2 versus time for the awake subject 1 (first row), for the semi-drowsy driver 6 (second row) and the drowsy driver (last row). Clearly, the awake driver presents less sudden rapid motions than the semi-drowsy or the drowsy driver. Integrating the pose rate angles yields the curves in the right column of Figure 12. This second column corresponds to the pose without considering the reset conditions, and therefore, accumulates the measurement errors in the 45 minutes (2,700 seconds) of the experiment. The integrated pose curves for the drowsy driver deviate more from the starting pose than those of the semi-drowsy or fully awake drivers, due to the larger number of sudden pose corrections. We observed that awake drivers tended to seek a more comfortable

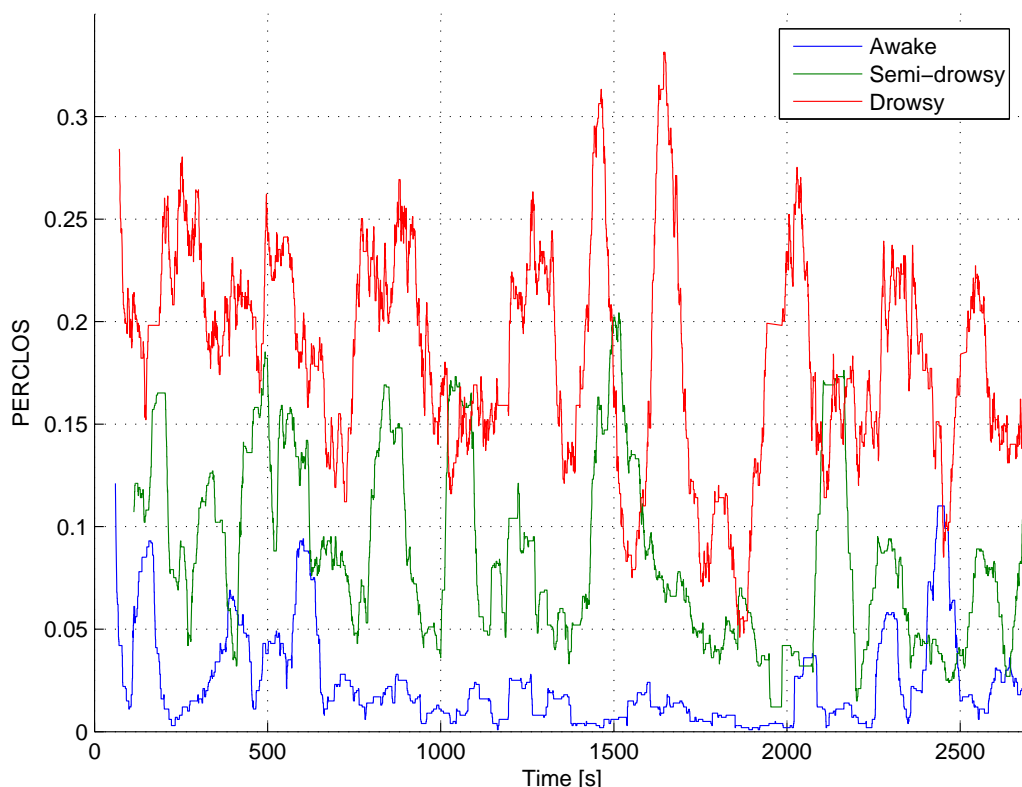
sitting position after a while of driving or simply changed position because of boredom. However, awake drivers kept their position for longer periods. On the other hand, drowsy driver's were struggling not to fall asleep, seemed also more concerned about not failing the test and, therefore, would move quickly to regain control of themselves.

Figure 12. Pose rate of change and pose computed from the integral of the pose rate without reset conditions for the awake (a); semi-drowsy (b) and drowsy (c) drivers, corresponding to test subjects 1, 6 and 11, respectively.



The PERCLOS curves for drivers 1, 6 and 11 are shown in Figure 13. Apart from the fact that the PERCLOS curve for the drowsy driver has an appreciably larger average value than that of the semi-drowsy and awake drivers, the PERCLOS of the drowsy driver increases precisely before instants in which the drowsy driver makes sudden motions (e.g., seconds 900, 1,500, 1,700, 2,000, 2,400). This can also be seen in Figure 14, which shows the evolution of the normalized magnitude (Euclidean norm) of the pose rate vector and the normalized PERCLOS measure for the drowsy driver 11. Both, the normalized magnitude of the pose rate vector and the normalized PERCLOS measure, have been smoothed using a moving average filter with a window spanning 80 seconds and normalized to values in the range $[0, 1]$ to facilitate the comparison. While there does not seem to be any straight forward connection between the amplitude of the peaks in the normalized PERCLOS and the amplitude of those in of the normalized pose rate magnitude, from Figure 14, it is possible to observe that the peaks in the normalized PERCLOS precede the majority of peaks of the normalized pose rate magnitude. This fact that was also observed for the other semi-drowsy and drowsy drivers strengthens the support for PERCLOS as a measure that has more predictive value than other physiological cues that can be measured in a non-invasive manner, such as the driver's pose variations or steering behavior.

Figure 13. PERCLOS for the awake (a); semi-drowsy (b) and drowsy (c) drivers, corresponding to test subjects 1, 6 and 11, respectively.



The previous results, together with the fact that the 95% confidence interval for the average PERCLOS value is very narrow (see Table 2), confirm that the PERCLOS measure is more reliable for correctly discriminating the different fatigue levels. This conclusion is also supported by the PERCLOS normal distribution curves plotted in Figure 1 for each group of drivers using the computed PERCLOS mean and standard deviation values. Figure 1 shows a clear difference in the mean PERCLOS for the different

levels of drowsiness with non-overlapping confidence intervals. The normal distribution curves for each group of drivers can be assumed to specify the probability distributions for each class and used to select the class for which the measurement has the highest probability of belonging to. The normal distribution curves were used to obtain the threshold values presented in Table 3, which are needed to classify the driver's level of alert.

Figure 14. Normalized pose rate magnitude and normalized PERCLOS measure for drowsy subject 11, both smoothed with a moving average filter using a window of 80 seconds.

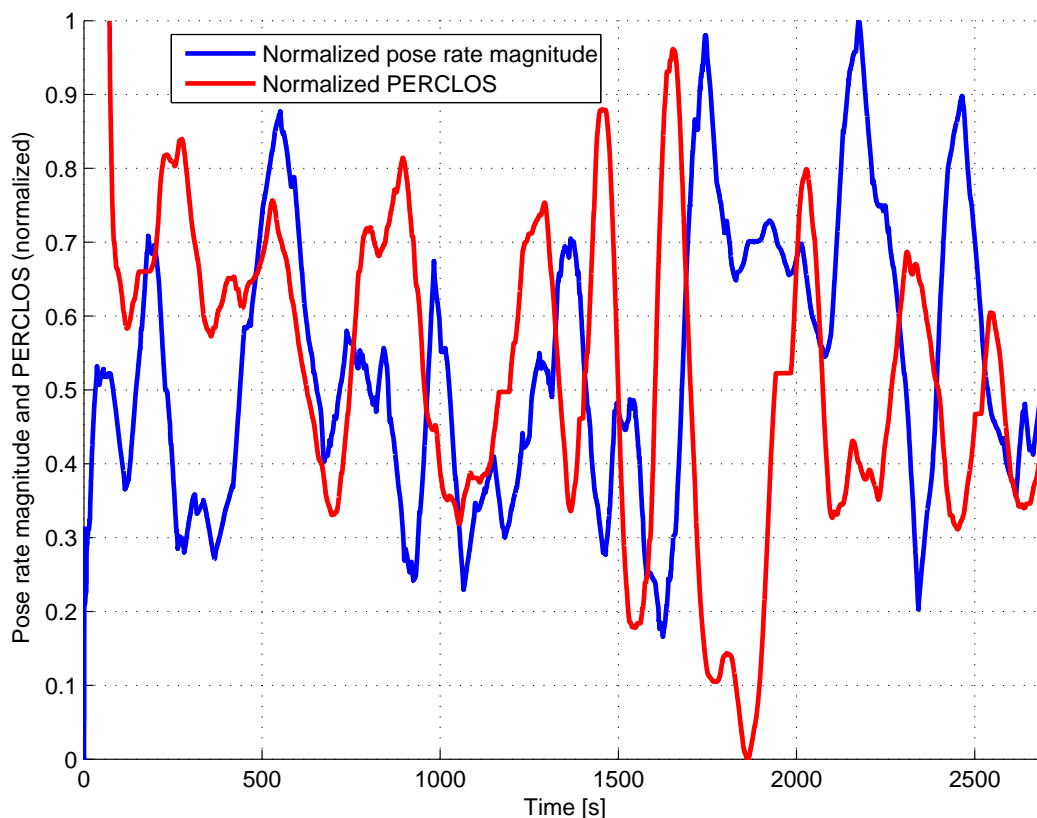


Table 3. PERCLOS threshold values for driver state classification using 60 seconds windows.

Driver state	Minimum PERCLOS	Maximum PERCLOS
Fully awake	0.000	0.048
Semi-drowsy	0.048	0.125
Drowsy-driver	0.125	1.000

5. Conclusions

A non-invasive sensing approach for driver fatigue and attention measurement was presented. The novelty of the approach is in the use of a kinematic model of the driver's motion and a grid of salient points tracked using the Lukas-Kanade optical flow method. The advantage of this approach is that it

does not require one to directly detect the eyes, and therefore, if the eyes are occluded or not visible from the camera when the head turns, the system does not lose the tracking of the eyes or the face, because it relies on the grid of salient points and the knowledge of the driver's motion model, which is useful for computing and predicting the pose of the driver. Another contribution of this approach is that it employs the observed interpupillary distance to estimate (up to a scale factor) how far the driver is from the camera. In other words, the approach does not require a stereoscopic system to resolve the relative motion of the driver. Moreover, the kinematic motion model for a driver with five degrees of freedom allows one to implement an extended Kalman filter that simplifies the tracking of the points in the image space. The results show that the tracking rate improves from $38.03 \pm 13.57\%$ to $97.10 \pm 2.39\%$, when the salient points are used instead of attempting to perform the eyes recognition using the Viola-Jones approach. An additional improvement from $97.10 \pm 2.39\%$ to $99.41 \pm 1.31\%$ is possible using the kinematic model with the extended Kalman filter.

The experiments performed involved a group of 15 subjects, five of which were asked to stay awake for more than 24 hours before the driving tests in the simulator, another five were asked to sleep only four hours the night before the test and five were asked to have a full eight hours rest. The computation of the percentage of time the eyes are closed covering at least 80% or more of the pupil (PERCLOS) for the different group of subjects delivers a measure that is consistent with the drivers' level of drowsiness. The results show that these three groups have a PERCLOS with a sufficiently small variance for classification purposes, *i.e.*, the PERCLOS measure can be used to effectively distinguish and detect the level of fatigue associated to the lack of rest. It was found that the group of subjects in the awake state presents a mean PERCLOS value of 0.0320 ± 0.0021 (C.I. 95%) with a standard deviation of 0.0074, while subjects in the drowsy state have a mean PERCLOS of 0.1799 ± 0.0146 (C.I. 95%) with a standard deviation of 0.0499, thus exhibiting a difference between the two states significant enough that can be used by the drowsiness and attention system to warn the driver about having reached dangerous fatigue levels, which could lead to an imminent accident, unless proper rest is taken.

In summary, the results demonstrate that the proposed system provides a solution for drowsiness and attention sensing that is reliable and more robust to occlusions or driver pose changes that often affect approaches based on the direct tracking of the eyes. In addition to the measurement of fatigue and drowsiness, the pose tracking capability of the proposed approach has potential applications in distraction assessment and alerting of machine operators, particularly of large construction and mining machinery, which is a subject of the authors' ongoing research. This study considered a group of similar participants in age and daily routine. The analysis of fatigue variation across gender, age or phenotype was not in the scope of the current work. These aspects, together with a detailed study of optimal IR illumination for people wearing tinted glasses, are also part of the authors' ongoing long-term research efforts.

Acknowledgments

This work was supported by Conycit of Chile under Fondecyt Grant 1106251.

References

1. World Health Organization. Data and statistics of the World Health Organization, 2012. Available online: <http://www.who.int/research/en/> (accessed on 11 March 2013).
2. World Health Organization. *The World Health Report 2004-Changing History. Annex Table 2: Deaths by Cause, Sex and Mortality Stratum in WHO Regions, Estimates for 2002*; WHO Press: Geneva, Switzerland, 2004.
3. Pedan, M., Scurfield, R., Sleet, D., Mohan, D., Hyder, A.A., Jarawan, E., Mathers, C., Eds.; *World Report on Road Traffic Injury Prevention*; WHO Press: Geneva, Switzerland, 2004.
4. World Health Organization. *World Report on Child Injury Prevention*; WHO Press: Geneva, Switzerland, 2008.
5. Powell, N.B.; Chau, J.K. Sleepy driving. *J. Med. Clin. North Am.* **2010**, *94*, 531–540.
6. May, J.F.; Baldwin, C.L. Driver fatigue: The importance of identifying causal factors of fatigue when considering detection and countermeasure technologies. *Transp. Res. Part F: Traffic Psychol. Behav.* **2009**, *12*, 218–224.
7. Yang, J.H.; Mao, Z.H.; Tijerina, L.; Pilutti, T.; Coughlin, J.; Feron, E. Detection of driver fatigue caused by sleep deprivation. *IEEE Trans. Syst. Man Cybern. Part A: Syst. Hum.* **2009**, *39*, 694–705.
8. Lal, S.K.L.; Craig, A. A critical review of the psychophysiology of driver fatigue. *Biol. Psychol.* **2001**, *55*, 173–194.
9. Zhao, C.; Zhao, M.; Liua, J.; Zheng, C. Electroencephalogram and electrocardiograph assessment of mental fatigue in a driving simulator. *Accid. Anal. Prev.* **2012**, *45*, 83–90.
10. Liang, W.C.; Yuan, J.; Sun, D.C.; Lin, M.H. Changes in physiological parameters induced by indoor simulated driving: Effect of lower body exercise at mid-term break. *Sensors* **2009**, *9*, 6913–6933.
11. Jung, S.J.; Shin, H.S.; Chung, W.Y. Highly sensitive driver health condition monitoring system using nonintrusive active electrodes. *Sens. Actuators B: Chem.* **2012**, *171-172*, 691–698.
12. Forsman, P.; Vila, B.; Short, R.; Mott, C.; van Dongen, H. Efficient driver drowsiness detection at moderate levels of drowsiness. *Accid. Anal. Prev.* **2013**, *50*, 341–350.
13. Dehnavi, M.; Eshghi, M. Design and implementation of a real time and train less eye state recognition system. *EURASIP J. Adv. Signal Process.* **2012**, *30*, 1–12.
14. Wierwille, W.; Ellsworth, L.; Wreggit, S.; Fairbanks, R.; Kirn, C. *Research on Vehicle Based Driver Status/Performance Monitoring: Development, validation, and Refinement of Algorithms for Detection of Driver Drowsiness*; Final Report DOT HS 808 247; Technical Report, National Highway Traffic Safety Administration, Washington, DC, USA, 1994.
15. Dinges, D.F.; Mallis, M.M.; Maislin, G.; Powell, J.W. *Evaluation of Techniques for Ocular Measurement as Index of Fatigue and the Basis for Alertness Management.*; Final Report DOT HS 808 762; Technical Report, National Highway Traffic Safety Administration: Washington, DC, USA, 1998.

16. Dong, W.; Wu, X. Driver Fatigue Detection Based on the Distance of Eyelid. In Proceedings of 2005 IEEE International Workshop on the VLSI Design and Video Technology, Suzhou, China, 28–30 May 2005; pp. 365–368.
17. Horng, W.B.; Chen, C.Y.; Chang, Y.; Fan, C.H. Driver Fatigue Detection Based on Eye Tracking and Dynamic Template Matching. In Proceedings of the IEEE International Conference Networking, Sensing and Control, Taipei, Taiwan, March 2004; Volume 1, pp. 7–12.
18. Qin, H.; Gao, Y.; Gan, H. Precise Eye Location in Driver Fatigue State Surveillance System. In Proceedings of IEEE International Conference on the Vehicular Electronics and Safety, ICVES, Beijing, China, December 2007; pp. 1–6.
19. Rongben, W.; Lie, G.; Bingliang, T.; Lisheng, J. Monitoring Mouth Movement for Driver Fatigue or Distraction with One Camera. In Proceedings of The 7th International IEEE Conference on Intelligent Transportation Systems, Washington DC, USA, October 2004; pp. 314–319.
20. Singh, S.; Papanikolopoulos, N. Monitoring Driver Fatigue Using Facial Analysis Techniques. In Proceedings of IEEE/IEEJ/JSAI International Conference on the Intelligent Transportation Systems, Tokyo, Japan, October 1999; pp. 314–318.
21. Tabrizi, P.R.; Zoroofi, R.A. Open/Closed Eye Analysis for Drowsiness Detection. In Proceedings of the First Workshops on Image Processing Theory, Tools and Applications, IPTA, Sousse, Tunisia, November 2008; pp. 1–7.
22. Flores, M.J.; Armingol, J.M.; de la Escalera, A. Driver drowsiness warning system using visual information for both diurnal and nocturnal illumination conditions. *EURASIP J. Adv. Signal Process.* **2010**, *3*, 1–19.
23. Flores, M.; Armingol, J.; Escalera, A. Real-Time Drowsiness Detection System for an Intelligent Vehicle. In Proceedings of the Intelligent Vehicles Symposium, Eindhoven, Netherlands, June 2008; pp. 637–642.
24. Hong, T.; Qin, H.; Sun, Q. An Improved Real Time Eye State Identification System in Driver Drowsiness Detection. In Proceedings of IEEE International Conference on the Control and Automation, ICCA 2007, Guangzhou, China, May 2007; pp. 1449–1453.
25. Lu, H.; Zhang, W.; Yang, D. Eye Detection Based on Rectangle Features and Pixel-Pattern-Based Texture Features. In Proceedings of International Symposium on the Intelligent Signal Processing and Communication Systems, ISPACS 2007, Xiamen, China, November 2007; pp. 746–749.
26. Sigari, M.H. Driver Hypo-Vigilance Detection Based on Eyelid Behavior. In Proceedings of Seventh International Conference on the Advances in Pattern Recognition, ICAPR '09, Kolkata, India, 4–6 February 2009; pp. 426–429.
27. Xu, C.; Zheng, Y.; Wang, Z. Efficient Eye States Detection in Real-Time for Drowsy Driving Monitoring System. In Proceedings of International Conference on the Information and Automation, ICIA 2008, Zhangjiajie, China, June 2008; pp. 170–174.
28. Zhang, Z.; Shu Zhang, J. Driver Fatigue Detection Based Intelligent Vehicle Control. In Proceedings of 18th International Conference on the Pattern Recognition, ICPR 2006, Hong Kong, China, August 2006; Volume 2, pp. 1262–1265.

29. Fan, X.; Yin, B.C.; Sun, Y.F. Yawning Detection for Monitoring Driver Fatigue. In Proceedings of the 2007 IEEE International Conference on the Machine Learning and Cybernetics, Hong Kong, China, August 2007; Volume 2, pp. 664–668.
30. Flores, M.; Armingol, J.M.; de la Escalera, A. Driver drowsiness detection system under infrared illumination for an intelligent vehicle. *IET Intell. Transp. Syst.* **2011**, *5*, 241–251.
31. Bergasa, L.; Nuevo, J.; Sotelo, M.; Barea, R.; López, M. Real-time system for monitoring driver vigilance. *Intell. Transp. Syst. IEEE Trans.* **2006**, *7*, 63–77.
32. Garcia, I.; Bronte, S.; Bergasa, L.M.; Almazán, J.; Yebes, J. Vision-based Drowsiness Detector for Real Driving Conditions. In Proceedings of the IEEE 2012 Intelligent Vehicles Symposium, Madrid, Spain, 3–7 June 2012.
33. Park, I.; Ahn, J.H.; Byun, H. Efficient Measurement of Eye Blinking under Various Illumination Conditions for Drowsiness Detection Systems. In Proceedings of the 18th International Conference on Pattern Recognition, ICPR 2006, Hong Kong, China, 20–24 August 2006; Volume 1, pp. 383–386.
34. Gu, H.; Ji, Q.; Zhu, Z. Active Facial Tracking for Fatigue Detection. In Proceedings of Sixth IEEE Workshop on the Applications of Computer Vision, (WACV 2002), Orlando, FL, USA, 3–4 December 2002; pp. 137–142.
35. Elsner, A.; Burns, S.; Weiter, J.; Delori, F. Infrared imaging of sub-retinal structures in the human ocular fundus. *Vision Res.* **1996**, *36*, 191–205.
36. Holz, F.G.; Spaide, R.F. *Medical Retina; Essentials in Ophthalmology*, Springer-Verlag: Berlin/Heidelberg, Germany, 2007.
37. D’Orazio, T.; Leo, M.; Spagnolo, P.; Guaragnella, C. A Neural System for Eye Detection in a Driver Vigilance Application. In Proceedings of the 7th International IEEE Conference on the Intelligent Transportation Systems, Washington, DC, USA, 3–6 October 2004; pp. 320–325.
38. Suzuki, M.; Yamamoto, N.; Yamamoto, O.; Nakano, T.; Yamamoto, S. Measurement of Driver’s Consciousness by Image Processing. A Method for Presuming Driver’s Drowsiness by Eye-Blinks coping with Individual Differences. In Proceedings of the 2006 IEEE International Conference on the Systems, Man and Cybernetics, Taipei, Taiwan, 8–11 October 2006; Volume 4, pp. 2891–2896.
39. Dong, W.; Qu, P.; Han, J. Driver Fatigue Detection Based on Fuzzy Fusion. In Proceedings of the Chinese Control and Decision Conference, CCDC 2008, Yantai, China, 2–4 July 2008; pp. 2640–2643.
40. Fan, X.; Yin, B.; Sun, Y. Nonintrusive Driver Fatigue Detection. In Proceedings of the IEEE International Conference on the Networking, Sensing and Control, ICNSC 2008, Sanya, China, 6–8 April 2008; pp. 905–910.
41. Ito, T.; Mita, S.; Kozuka, K.; Nakano, T.; Yamamoto, S. Driver Blink Measurement by the Motion Picture Processing and Its Application to Drowsiness Detection. In Proceedings of the 5th IEEE International Conference on the Intelligent Transportation Systems, Singapore, 3–6 September 2002; pp. 168–173.

42. Wang, Q.; Yang, W.; Wang, H.; Guo, Z.; Yang, J. Eye Location in Face Images for Driver Fatigue Monitoring. In Proceedings of the 6th International Conference on ITS Telecommunications, Chengdu, China, 21–23 June 2006; pp. 322–325.
43. Sahayadhas, A.; Sundaraj, K.; Murugappan, M. Detecting driver drowsiness based on sensors: A review. *Sensors* **2012**, *12*, 16937–16953.
44. Jimenez-Pinto, J.; Torres-Torriti, M. Face salient points and eyes tracking for robust drowsiness detection. *Robotica* **2012**, *30*, 731–741.
45. Guggisberg, A.G.; Mathis, J.; Herrmann, U.S.; Hess, C.W. The functional relationship between yawning and vigilance. *Behav. Brain Res.* **2007**, *179*, 159–166.
46. Guggisberg, A.G.; Mathis, J.; Schnider, A.; Hess, C.W. Why do we yawn? *Neurosci. Biobehav. Rev.* **2010**, *34*, 1267–1276.
47. Vural, E.; Cetin, M.; Ercil, A.; Littlewort, G.; Bartlett, M.; Movellan, J. Drowsy Driver Detection Through Facial Movement Analysis. In Proceedings of the 2007 IEEE International Workshop on Human-Computer Interaction, HCI'07, Rio de Janeiro, Brazil, October 2007; pp. 6–18.
48. Gutierrez-Alvarez, A.M. Do your patients suffer from excessive yawning? *Acta Psychiatr. Scand.* **2007**, *115*, 80–81.
49. van den Berg, J. Sleepiness and head movements. *Ind. Health* **2006**, *44*, 564–576.
50. Abe, T.; Nonomura, T.; Komada, Y.; Asaoka, S.; Sasai, T.; Ueno, A.; Inoue, Y. Detecting deteriorated vigilance using percentage of eyelid closure time during behavioral maintenance of wakefulness tests. *Int. J. Psychophysiol.* **2011**, *82*, 269–274.
51. Rosario, H.D.; Solaz, J.; Rodriguez, N.; Bergasa, L. Controlled inducement and measurement of drowsiness in a driving simulator. *IET Intell. Transp. Syst.* **2010**, *4*, 280–288.
52. Nuevo, J.; Bergasa, L.M.; Llorca, D.F.; Ocaña, M. Face tracking with automatic model construction. *Image Vision Comput.* **2011**, *29*, 209–218.
53. Schiffman, H.R. *Sensation and Perception. An Integrated Approach*, 5th ed.; John Wiley and Sons, Inc.: Hoboken, NJ, USA, 2001.
54. Paul, A.; Boyle, L.N.; Tippin, J.; Rizzo, M. Variability of Driving Performance During Microsleeps. In Proceedings of the Third International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design; Rockport, Maine, USA, 27–30 June 2005; pp. 433–440.
55. Viola, P.; Jones, M. Rapid Object Detection Using a Boosted Cascade of Simple Features. In Proceedings of the 16th IEEE Conference on Computer Vision and Pattern Recognition, Kauai, HI, USA, 8–14 December 2001; Volume I, pp. 511–518.
56. Shi, J.; Tomasi, C. Good Features to Track. In Proceedings of the 9th IEEE Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, June 1994; pp. 593–600.
57. Lucas, B.; Kanade, T. An Iterative Image Registration Technique with an Application to Stereo Vision. In Proceedings of the Imaging Understanding Workshop, Washington DC, USA, April 1981; pp. 121–130.
58. Ching, R.P. *Relationship Between Head Mass and Circumference in Human Adults*; Technical Report; University of Washington: Seattle, WA, USA, 2007.

59. Dodgson, N.A. Variation and Extrema of Human Interpupillary Distance. In Proceedings of SPIE: Stereoscopic Displays and Virtual Reality Systems XI, San Jose, CA, USA, January 2004; Volume 5291, pp. 36–46.
60. Bertsekas, D.P. *Nonlinear Programming*, 2nd ed.; Athena Scientific: Nashua, North Hampshire, USA, 1999; p. 780.
61. Press, W.H.; Teukolsky, S.A.; Vetterling, W.T.; Flannery, B.P. *Numerical Recipes 3rd Edition: The Art of Scientific Computing*, 3rd ed.; Cambridge University Press: New York, NY, USA, 2007.
62. Grace, R.; Byrne, V.; Bierman, D.; Legrand, J.M. and Gricourt, D.; Davis, B.; Staszewski, J.; Carnahan, B. A drowsy driver detection system for heavy vehicles. In Proceedings of the Digital Avionics Systems Conference, Bellevue, WA, USA, 31 October–7 November 1998; Volume 2, 1–8.
63. Harris, M.G.; Chin, R.S.; Lee, D.S.; Tarot, M.H.; Dobkins, C.E. Ultraviolet Transmittance of the Vistakon disposable contact lenses. *Cont. Lens. Anterior Eye*. **2000**, *23*, 10–15.
64. Stephens, G.L.; Davis, J.K. *Duane's Ophthalmology, Ch. 51D Ophthalmic Lens Tints and Coatings*, cd-rom 2006 ed.; Volume 1, Lippincott Williams & Wilkins: Philadelphia, Pennsylvania, USA, 2006. Available online: <http://www.oculist.net/downat0502/prof/ebook/duanes/pages/v1/v1c051d.html> (accessed on 11 March 2013).
65. Sperian Protection Americas Inc. *Uvex Lens Technology*; Sperian Protection Americas Inc., Smithfield, RI, USA, 2008; p. 23. Brochure.
66. Van Gaal, R.; Schuerkamp, U.; Pospisil, D.; Harrington, P. Racer Free Car Simulation, 2012. Available online: <http://www.racer.nl/> (accessed on 11 March 2013).

© 2013 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/3.0/>).