CrossMark

Eye blink detection for different driver states in conditionally automated driving and manual driving using EOG and a driver camera

Jürgen Schmidt[1,2] · Rihab Laarousi[1] · Wolfgang Stolzmann[1] · Katja Karrer-Gauß[2]

© Psychonomic Society, Inc. 2017

Abstract In this article, we examine the performance of different eye blink detection algorithms under various constraints. The goal of the present study was to evaluate the performance of an electrooculogram- and camera-based blink detection process in both manually and conditionally automated driving phases. A further comparison between alert and drowsy drivers was performed in order to evaluate the impact of drowsiness on the performance of blink detection algorithms in both driving modes. Data snippets from 14 monotonous manually driven sessions (mean 2 h 46 min) and 16 monotonous conditionally automated driven sessions (mean 2 h 45 min) were used. In addition to comparing two data-sampling frequencies for the electrooculogram measures (50 vs. 25 Hz) and four different signal-processing algorithms for the camera videos, we compared the blink detection performance of 24 reference groups. The analysis of the videos was based on very detailed definitions of eyelid closure events. The correct detection rates for the alert and manual driving phases (maximum 94%) decreased significantly in the drowsy (minus 2% or more) and conditionally automated (minus 9% or more) phases. Blinking behavior is therefore significantly impacted by drowsiness as well as by automated driving, resulting in less accurate blink detection.

Drowsiness affects humans' driving ability. This is why drowsy driving accounts for approximately 19%–24% of all fatal crashes in manual driving (Hell et al., 2012). To ameliorate the risk of drowsy driving, car manufacturers and scientists have been working on reliable systems to detect, warn of, and ultimately prevent this critical state before it begins. To determine the state of the driver online while driving, several physiological measurement techniques, such as electroencephalogram (EEG), electrocardiography (ECG), electrooculography (EOG), and performance-based measures (e.g., steering behavior), have been evaluated in the past (Dawson, Searle, & Paterson, 2014; Dong, Hu, Uchimura, & Murayama, 2011). The term *drowsiness* will be used in this work as it was by Sahayadhas, Sundaraj, and Murugappan (2012), as synonymous with *sleepiness*. An acceptable detection system for drowsiness has to fulfill several requirements that the aforementioned methods do not completely satisfy. Preferably, to be accepted by all drivers, the system should not be attached to the body. Current serial systems using either steering behavior or drivers' ability to stay in the lane to detect drowsy driver detection criteria meet this requirement (Daimler, 2008; Ford, 2010). However, with the development of automated driving functions, steering behavior is no longer a feasible detection method. Conditionally automated driving (CAD; based on the definition in SAE Standard J3016 [2014]), which is currently in development (Wei et al., 2013), gives drivers the freedom to let go of the steering wheel, relax their observations, and allow the algorithm to control the vehicle. Self-driving vehicles will become reality within the next 5 years (BMW Group, 2016). Among the

✉ Jürgen Schmidt
juergen.j.schmidt@daimler.com

1   Research & Development, Daimler AG, HPC X914,
    71059 Sindelfingen, Germany

2   Chair of Human–Machine Systems, Technische Universität Berlin,
    Berlin, Germany

technologies that will be needed to successfully deal with all constraints of future automated travel, the use of a camera to observe a driver's behavior is the most promising technology for detecting drowsiness (Jafar Ali, Sarkar, Kumar, & Cabibihan, 2012). Particularly, the analysis of eyelid movements and blinking as a drowsiness indicator has been the focus of many studies as reported in reviews by Stern, Boyer, and Schroeder (1994) and Dawson et al. (2014). Onset drowsiness influences the normal blinking behavior and can therefore be used to estimate the drowsiness state of the driver (Morris & Miller, 1996; Picot, Charbonnier, & Caplier, 2012). Unlike manual driving, CAD is an automated system offering the driver the possibility to change their gaze and eye closing behavior without negatively affecting driving safety. It is crucial to have an algorithm for detecting blinks or the eye aperture in both the manual and CAD phases to be able to distinguish alert and drowsy driver behavior with high accuracy, as drivers will be able to freely switch between the driving modes.

The goal of this study was to compare the performance of a blink detection algorithm in manual driving against its performance in conditionally automated driving, in phases during which the drivers were alert and phases during which they were drowsy. This should answer the open question concerning the degree of influence that CAD has on the detection rates of a blink detection algorithm and what influence a possible behavior change in blinking has on the blink detection in the future self-driving transportation. A second goal was to compare the performance of an EOG-based blink detection system, with the performance of a simultaneously recording camera-based detection system. By these means, it should be possible to determine which technology is best suited for assessing driver behavior in future studies with CAD.

Human blinking behavior

Blinks are usually associated with the urge to clean the eyes from particles and to spread tear film. Therefore, blinks are affected by humidity, temperature, chemical factors and air particles (Stern et al., 1994; Wolkoff, Nøjgaard, Troiano, & Piccoli, 2005). The time between blinks increases when an observer watches a visual display unit (Patel, Henderson, Bradley, Galloway, & Hunter, 1991) or is distracted by a demanding task (Wolkoff et al., 2005). On the other hand, time between blinks is reported to decrease with an auditory task during driving as compared to driving without an additional task (Tsai, Viirre, Strychacz, Chase, & Jung, 2007). This suggests a relationship between the type of the task and its effect on the driver's blinking frequency.

Studies that involve a long time-on-task, which evokes drowsiness, have reported changes in eyelid closure and gaze behavior. The palpebral aperture generally becomes smaller

"associated with a downward shift in gaze angle" (Lobb & Stern, 1986, p. 17). In addition, Stern et al. (1994) reported compelling evidence of an increasing blink rate with time-on-task in their literature review.

Blinks are categorized by their origin as voluntary, reflex, or spontaneous and are often accompanied by saccades and eye movements (Collewijn, van der Steen, & Steinman, 1985; Stern, Walrath, & Goldstein, 1984). Stern et al. (1984) further distinguished longer eye closures, such as "microsleeps," as a separate category representing nonblink closures. However, in this research, with the focus on driver drowsiness and the detection of blinks, microsleeps will be considered blinks.

Blink detection methods

EOG and video recordings are the main techniques used in driving/transportation studies to record the blinking behavior of a driver (Morris & Miller, 1996; Picot et al., 2012). Although EOG is considered to be the most reliable method due to its high frame rate, video-based assessment has gathered popularity for its practicability in the automotive industry due to its ability to measure contact free.[1]

To obtain a signal for the eye blink detection with EOG, several surface electrodes are positioned around the eyes. Since the cornea has a positive electric potential in reference to the fundus of the eye, a natural occurring eyelid movement during a blink affects the electric potential between the two electrodes positioned above and below the eye. A blink can thus be measured as change in the potential distribution of the eye (e.g., Jammes, Sharabty, & Esteve, 2008; Skotte, Nøjgaard, Jørgensen, Christensen, & Sjøgaard, 2007).

Using video recordings, eyelid movement is visible in the images and can be assessed using image processing methods. Different algorithms for that purpose are based on either the motion detection derived from differencing two consecutive images (e.g., Bhaskar, Keat, Ranganath, & Venkatesh, 2003; Chau & Betke, 2005; Fogelton & Benesova, 2016; Jiang, Tien, Huang, Zheng, & Atkins, 2013), a second-order derivative method of image differentiations (Gorodnichy, 2003), a state classification (e.g., Choi, Han, & Kim, 2011; Missimer & Betke, 2010; Pan, Sun, & Wu, 2008; Pan, Sun, Wu, & Lao, 2007), an evaluation of the color contrast or amount of visible color of specific eye regions (Cohn, Xiao, Moriyama, Ambadar, & Kanade, 2003; Danisman, Bilasco, Djeraba, & Ihaddadene, 2010; Lee, Lee, & Park, 2010), the distance between landmarks or arcs representing the upper and lower eyelid (Fuhl et al., 2016; Ito, Mita, Kozuka, Nakano, &

---

[1] Other traditional methods, such as electromyography or the use of magnetic coils or systems with mechanic connections to the lid, are rarely used in current driving studies, because they are not practicable. Therefore, they will not be considered in this study.

Yamamoto, 2002; Miyakawa, Takano, & Nakamura, 2004; Moriyama et al., 2002; Sukno, Pavani, Butakoff, & Frangi, 2009), the missing regions of the open eye like the iris or pupil due to their occlusion by the upper and lower eyelid (Hansen & Pece, 2005; Pedrotti, Lei, Dzaack, & Rötting, 2011), or a combination of the described methods (Sirohey, Rosenfeld, & Duric, 2002). Instead of measuring the real distance between the upper and lower eyelid, most of these algorithms use an indirect measure (motion detection, classification, color contrast, missing eye regions) to conclude whether the eye is closed. This is similar to the EOG technique, which also infers an eye closure indirectly from a potential difference between two electrodes.

The collected eye or eyelid movement signal of the EOG or the image processing is usually processed further in a second step. General approaches to detect blinks in the signal include the evaluation with thresholds (e.g., Divjak & Bischof, 2009; Grauman, Betke, Gips, & Bradski, 2001), filtering (e.g., Grauman et al., 2001; Jammes et al., 2008), derivation of the signals (e.g., Ebrahim, 2016; Torricelli, Goffredo, Conforto, & Schmid, 2009), transformation (e.g., Benoit & Caplier, 2010; Malik & Smolka, 2014), and valley/peak detection (e.g., Malik & Smolka, 2014; Radlak & Smolka, 2012). Depending on the signal quality and data processing method, more or less detailed information (start time, speed of the eye closure, or duration of the eye closure; see Picot et al., 2012) of the blinks can be parameterized.

Drowsiness detection and influence by the performance of the blink detection

By identifying changes in eye blink parameters, several studies have been successful using EOG or video recordings to estimate the drowsiness of a driver (EOG: Hu & Zheng, 2009; Picot et al., 2012; Video recordings: Bergasa, Nuevo, Sotelo, Barea, & Lopez, 2006; Friedrichs & Yang, 2010; Garcia, Bronte, Bergasa, Almazan, & Yebes, 2012). Unfortunately, they often fail to report the performance of the preprocessing algorithms used to detect blinks or the degree of eye closure, which implies a high reliability of the sensor system and error-free parameterization of the collected data. According to a study by Pedrotti et al. (2011), a commercially available eyetracker does not guarantee a correct blink detection rate of 100%. Consequently, only performance tests of the blink parameterization algorithms can reveal the influence of errors in the preprocessing of the detection of blinks. These evaluation tests for blink detection algorithms need to be carried out under the same constraints as those used in the application. Hence, the testing procedure of an eye blink algorithm for drowsy drivers needs to be evaluated in alert and drowsy driving phases. Because blinking as well as measuring systems are influenced in many ways (EOG: electromagnetic

compatibility, contact with the skin; Video recordings: position of the camera, frame rate of the video, glasses, eye physiognomy), the detection rate of blinks and the accurate measurement of the eyelid distance can differ completely in various experiments.

In a comparison of the detection rate of blinks between EOG and a remote eyetracker, Picot, Caplier, and Charbonnier (2009) used the blinks detected with a 250 Hz EOG system as a reference value for blinks detected with a remote eyetracker with different frequencies. In the study with data from 14 awake participants, they noticed a decreasing false detection rate with faster frame rates of the video (false detection rate with 30 Hz is greater than 100 Hz, >150 Hz, >200 Hz). Yet, the authors could not evaluate the difference between the various frame rates on the same video recordings. It is also unclear how many errors their EOG blink detection contained.

With respect to the detection of blinks in drowsy manual driving, several studies report a drop in the correct detection rate relative to the detection rate of the alert driving phases (Ebrahim, 2016; Jammes et al., 2008; Skotte et al., 2007). Following the higher rate of long eye closures observed for drowsy drivers during CAD by Schmidt, Braunagel, Stolzmann, and Karrer-Gauß (2016), a similar drop in the correct detection rate between the manual mode and the CAD mode can be expected and will be investigated in the following sections.

Method

Two experimental studies were conducted in a between-subjects design, using a manual and conditionally automated driving condition. None of the participants of the conditionally automated study were involved in the manual-driving study. In both studies, we intentionally used simulated evening conditions with a dark, heavily overcast sky and a monotonous roadside to induce drowsiness. The radio was switched off, and the use of any secondary devices was prohibited. The time for the whole process (introduction, pre-questionnaire, driving, and post-questionnaire) for each individual driver was limited in both studies to 4 h. The average driving time in the manual-driving study was 2 h 46 min for a distance of 335 km. In the CAD study, the participants drove on average for 2 h 45 min over 263 km. In both studies, half of the participants started the drive at 6:00 pm, and the rest at 10:00 pm. All of them had had a normal workday before the experiment. Participants had to rate their drowsiness level on the Karolinska Sleepiness Scale (KSS; Åkerstedt & Gillberg, 1990, Table 1) every 15 min (with a few exceptional extensions of up to 17 min in the CAD study). None of the participants in either study reported any known sleep disorder.

Table 1   Karolinska Sleepiness Scale (KSS)

KSS	Description
1	Extremely alert
2	Very alert
3	Alert
4	Rather alert
5	Neither alert nor sleepy
6	Some signs of sleepiness
7	Sleepy, but no effort to keep alert
8	Sleepy, some effort to keep alert
9	Very sleepy, great effort to keep alert

Manual-driving study

A total of 18 people participated in the manual-driving experiment (12 males and six females) and they ranged in age between 27 and 56 (mean [MN] = 34 years, standard deviation [SD] = 8 years). On average, the participants had possessed a driver's license for 16 years (max = 38 years, min = 8 years, SD = 9 years). The study took place in a Mercedes-Benz moving-base simulator with a dome on a hexapod platform and a 12-m axis for linear motion. A detailed Mercedes Benz S-class cabin was placed inside the dome and a 360° projection of the scenery was used throughout the drive to provide a realistic surrounding (Fig. 1). A supervisor was present in a separate observation room and had the ability to interact with the participant via intercom.

After a short pre-questionnaire and briefing, the participants started to drive on a circular two-lane highway 200 km long. Apart from two construction sites, located at Kilometer 62 and Kilometer 88, there were no particularly interesting landmarks on or alongside the road. Participants were instructed to drive at a speed of 130 km/h. On average, slower vehicles with a speed of 100 km/h had to be passed approximately every 2 min and the participants were passed by vehicles with a speed of 160 km/h approximately every 5 min. Furthermore, participants were requested by a recorded voice command to rate their subjective drowsiness level on the KSS.

The scale was attached to the inside of the car with several descriptions of the different drowsiness levels (Table 1). Previous evaluations based on recordings of manual driving conditions can be found in chapters 5–8 of the dissertation by Ebrahim (2016), the results of which are not part of this study.

CAD study

The test drives in the CAD experiment took place in a fixed-base driving simulator. All 46 participants completed the study (32 males and 14 females). Their ages ranged from 28 to 57 years (MN = 44 years, SD = 7.0 years). They had had a driver's license on average for 26 years (max = 42 years, min = 11 years, SD = 7.6 years). To simulate the driving environment, three flat screens (each 65 in. in diameter) were positioned 2.54 m in front the steering wheel of a car mockup, with the two outer screens tilted 144° relative to the one in the middle (Fig. 2).

An attendant supervisor was separated by a pin board from each participant and did not interact with them. Realistic sounds of headwind, tires, and engine were produced by speakers in response to the speed and acceleration of the simulated drive on a circular two lane highway with a length of 108 km. After arriving at the simulator site, the participants had to fill out a pre-questionnaire and were informed about the CAD function for the drive. The participants started the drive with a short manual introduction phase in which they got comfortable with the simulator and its dynamics by driving manually and performing several overtaking maneuvers. Since the implemented CAD system was new to all participants, a short practical introduction phase for the CAD function followed. During this time, the participants became familiar with all the prerequisites and limitations of the system and were guided by the examiner. The total duration of the manual and CAD introduction phase was 10 min.

To activate the CAD mode, the driver had to be driving in one of the two driving lanes and pull a lever on the side of the steering wheel. Additionally, this part of the route had to be "officially" approved for CAD (which was always the case,

Fig. 1   Mercedes-Benz moving-base simulator

Fig. 2 Mercedes-Benz fixed-base simulator

aside from small time periods explained later). Furthermore, the braking and acceleration pedal had to be released at the moment of CAD system activation and the driving speed had to be slower or equal to 110 km/h. To ensure the different modes were recognized correctly by the driver, different images in the middle of the speedometer display showed the current state of the CAD system (CAD not available, available, or active). Once switched on, the CAD function managed all lateral and longitudinal movements with a speed of 100 km/h. During an active CAD mode, the driver always had the possibility to take back control by pressing the brake or acceleration pedal, pushing the lever used for the activation of the system, or by changing lanes by steering. In each of these cases, the automated mode switched off immediately and the driver was back in charge of driving. Since CAD is limited to specific prerequisites that cannot always be guaranteed, it still relies on the driver reacting and taking back control if the system reaches these limits. For this reason, a basic reaction ability of the driver is always required. To ensure this, the driver had to confirm alertness requests within a given time of 5 s, similar to the alertness requests in trains, at intervals of either 30 s (23 participants) or 180 s (23 participants). If the driver failed to confirm the alertness request or the system reached its limit, a takeover request prompted the driver to take back the control of driving from the system. The CAD mode remained active during this time for up to 5 s (hand-over time) until the driver took over, and it switched off after 5 s if the driver did not take back control. During the drive, the participants faced several predefined situations, in which an ending of the availability of the CAD function on the highway was simulated. Directly after the maximum 5 s of hand-over time, different simulated road scenarios required adequate action by the driver. These situations, subsequently referred to as *takeover situations*, were relatively easy to handle and appeared after 30, 52, 70, and 92 min. An additional final situation was triggered dynamically. This situation was used to test the drivers' reaction ability by challenging them in their potentially severest state of drowsiness. To ensure that all participants entered the first four situations at the same time,

a leading vehicle drove in front of them, which the participants were asked not to overtake. Due to a blind spot on the side of the mockup that made it impossible to see overtaking vehicles, no faster cars overtook the driver during the experiment. Apart from the vehicles at the beginning, the leading vehicle, and the vehicles in the different situations, no vehicles were standing or driving in the same direction as the driver. The 15-min interval of the request to rate the KSS was occasionally prolonged by up to 2 min if it would have occurred during one of the predefined takeover situations. During the KSS query, the scale was presented to the drivers on the center screen and was switched off by the examiner immediately after the verbal KSS estimation of the driver. Unlike in the manual study, the scale only had annotations for the steps KSS = 1, 3, 5, 7, and 9, but not for the steps in-between, shown in Table 1. A detailed explanation of the situations and the reactions of the drivers can be found in a previously published article (Schmidt, Stolzmann, & Karrer-Gauß, 2016), as well as the evaluation of the process of drowsiness.

Signal measuring and use of existing signal processing methods

To record the driver behavior, the participants were equipped in both studies with an EOG measuring system called actiCAP (Brain Products GmbH, 2009) and the head-mounted eyetracker Dikablis (Ergoneers GmbH, 2016). Two electrodes were placed above and below the right eye to get the vertical EOG signal, two on the left and right side of the head for the horizontal EOG signal, and two more on the right and left mastoid bone to gather a reference signal and exclude noise. The camera of the head-mounted eyetracker was directed toward the left eye. Both eyes were used for the evaluation process (EOG on the right eye; head-mounted eyetracker on the left eye), under the assumption that both eyes would blink simultaneously (Collewijn et al., 1985). The two measuring devices had no visible effect on the driving, which was analyzed in Ebrahim (2016) and Schmidt, Stolzmann, and Karrer-Gauß (2016). The head-mounted eyetracker used infrared light to record the movements of one eye very precisely in the dark conditions of the experiments. Its resolution was 384×288 pixels. The width of the eye in the video image was approximately 75% and the height 50%. The camera was adjusted individually for each participant to a fixed position. This reduced the influence by head and body movements in the results of the blink detection. Furthermore, the use of a head-mounted eyetracker excluded effects by the necessary algorithmic detection of the eye region for remote eyetrackers. The EOG signals were recorded with 250 Hz and the eyetracker data with 25 Hz. For this study, the frequency of the EOG signal was sampled down to 50 Hz (f_1) and 25 Hz (f_2). Since the timestamp normally used for the synchronization was corrupted by lags caused by the signal

communication between the measuring devices, distinctive eye movements in the video were manually synchronized with peaks in the EOG signal together with the eyetracking data. The synchronized outcome was evaluated on several additional time frames from the entire drive that were distinct from the ones used for synchronization.

To detect eye blinks within the EOG signals, a signal processing procedure developed by Ebrahim (2016) was used. It was chosen due to its reported high recall and precision rate for blink events in the alert and drowsy driving phases. For a more detailed explanation, refer to pages 55 to 60 in the dissertation by Ebrahim (2016). Furthermore, the use of the algorithm for the EOG system and f_1 will be referred to as detection process a_1 and the use of the algorithm and f_2 as detection process a_2.

For the eye blink detection with the head-mounted eyetracker, an image processing algorithm developed by Fuhl et al. (2016) was used in combination with an algorithm for the pupil detection (Fuhl, Kübler, Sippel, Rosenstiel, & Kasneci, 2015). The algorithm of Fuhl et al. (2016) was chosen because the generated signal represents a direct measurement of the eyelid distance instead of an indirect measure based on motion detection, classification, color contrast or missing eye regions. Further, in contrast to the indirect measures by EOG, this algorithm can be later used in drowsiness detection algorithms to estimate the drowsiness level based on the direct distance of the eyelids. Since the algorithm developed by Fuhl et al. (2016) used data recorded with the same head-mounted eyetracker as in the presented manual and CAD studies, good signal estimations for the eyelid distances in the experiments were expected. Refer to the studies by Fuhl and colleagues (Fuhl et al., 2016; Fuhl et al., 2015) for more details of the image-processing algorithms. The output of the algorithm Fuhl et al. (2016) described the palpebral aperture in pixels according to the recorded video. An open eye was estimated with a value of approximately 144 pixels. The second algorithm by Fuhl et al. (2015) generated a binary signal for a detected or an undetected pupil.

Developed blink detection algorithm

Simple methods such as fixed thresholds for the classification of eye blinks without signal preprocessing did not detect blinks accurately for different participants in this study. The reasons for that are based on noise, misclassifications, and high inter- and intra-individual differences of the participants.

Therefore, a new blink detection algorithm was developed that included a preprocessing and several participant-dependent thresholds. The different steps of the process are outlined below. In the following, x_i describes the measured eyelid distance at time t_i in the unit pixel, MN the mean value, and SD the standard deviation.

1. Outlier removal A The raw signal of the eyelid distance included several extrema that were not plausible, possibly derived from noise or wrong interpretations of the image processing method. Outliers were individually defined for each participant as eyelid distance values outside the range between the 1st (eyelid distance: th_{outl}) and 99th (eyelid distance: th_{outh}) percentile of all eyelid distance values during the whole experiment.

All eyelid distance values $x_i \notin [th_{outl}\ th_{outh}]$ were replaced by $x_i = MN(x_j + x_k)$; $x_j, x_k \in [th_{outl}\ th_{outh}]$, and the associated times $t_j < t_i, t_k > t_i$ had minimal values of $t_i - t_j$ and $t_k - t_i$, respectively.

2. Outlier removal B Outliers with lower amplitudes were identified after the first step and replaced using neighboring data points similar to a moving average filter with a constraint. All eyelid distances x_l that were larger than $MN(x_{l-2} + x_{l-1} + x_{l+1} + x_{l+2}) + 3\ SD(x_{l-2}, x_{l-1}, x_{l+1}, x_{l+2})$ were replaced by $x_l = MN(x_{l-1} + x_{l+1})$. Similar to the first step, this replaced single values outside the 99th percentile of their four neighboring values, since an eyelid closure was longer than one sample.

3. Filtering (optional) The resulting signal was filtered using a third-order Savitzky–Golay filter and frame window size seven (Savitzky & Golay, 1964).

4. Identification of a participant-dependent eyelid distance Each participant had different values for the eyelid distances during driving. Therefore, a general eyelid distance for open eyes (th_o) was calculated for each driver using all eyelid distance values between Minutes 15 and 30 of driving. This excluded effects of the driver's customization to the environment and driving task. To account for the eyelid closure behavior in both manual and conditionally automated driving, the selection of th_o included consideration of the changing behavior of the drivers in CAD. Drivers might close their eyes for longer intervals or look downward. Thus, only the eyelid distance values above the average of all eyelid distance values were used to calculate th_o. Out of those values, th_o was defined as the mode to consider the most common eyelid distance position as the individual threshold for open eyes.

5. Eyelid movement minimum detection All eyelid closures contained an eyelid closing and opening phase. On this basis, all local minima in the signal lower or equal to th_o were identified during the entire drive and described by. $M = \{x_m \mid x_m$ is a local minimum $\wedge\ x_m \le th_o\}$; The corresponding times are summarized by $T = \{t_m \mid x_m \in M\}$. If a minimum contained more than one data point, the point with the lowest index was chosen.

6. Clustering of eyelid movements All detected eyelid closure values from the fifth step were clustered using a k-means clustering method with $k = 3$ (Arthur & Vassilvitskii, 2007). The two values separating the clusters are further described with th_u and th_d with $th_u > th_d$. A third value th_m was calculated with $th_m = MN(th_d + th_u)$. The smallest threshold th_d was interpreted as a distinctive threshold for a complete eye closure with the upper eyelid being very close to the lower eyelid: $Md = \{m_o \in M \mid m_o \leq th_d\}$. To distinguish blinks without a fully closed gap between the upper and lower eyelid, but with a small eyelid distance, Mm contained all local minima eyelid distances below or equal to th_m: $Mm = \{m_o \in M \mid th_d < m_o \leq th_m\}$. Minima larger than th_m, and lower than the upper threshold th_u had a larger gap between the eyelids $Mu = \{m_o \in M \mid th_m < m_o \leq th_u\}$. All values larger than th_u were very close to th_o and could be interpreted as smaller eyelid movements rather than blinks.

7. Determining the start and end values of the detected eyelid movements For all eyelid movements Md, Mm, and Mu, the eyelid distances at their start (Sd, Sm, and Su) and end (Ed, Em, and Eu) were determined. The start and end points are the local maxima of the eyelid distance signal before (Sd, Sm, and Su) and after (Ed, Em, and Eu) the eyelid distance signal intersected with the thresholds th_d, th_m, and th_u. If several minima of the same subgroup (e.g., md_1 and md_2 with $md_1, md_2 \in Md$) had the same eyelid distance value (here: $md_1 = md_2$) and start and end point (here: $sd_1 = sd_2 \wedge ed_1 = ed_2$ with $sd_1, sd_2 \in Sd$ and $ed_1, ed_2 \in Ed$), the ones with higher indices were discarded from the sets (here md_2, sd_2, and ed_2). If they had the same start and end point but different eyelid distance values (e.g., $md_1 < md_2$), the ones with the higher eyelid distances were discarded (here: md_2). Note that several eyelid movements of the set Md can lie in the boundaries of one single eyelid movement of the set Mm; similarly, several eyelid movements of the set Mm can lie in between one single eyelid movement of the set Mu.

8. Separate eyelid movements In case of overlapping sections between the ending of one eyelid movement and the start of the subsequent one caused by several equal values in the signal, the plateau was split into two equal parts with its center defining the end and the start point, respectively, of the two consecutive eyelid movements.

9. Approval of an eye closure The availability of the pupil signal from Fuhl et al. (2015) was used to check the visibility of the pupil in the samples in between the start and end point of the eyelid movements Md, Mm, and Mu, excluding the start and end samples. An eyelid movement was discarded if the pupil was visible in all investigated samples of the potential eye blink.

10. Amplitude clustering In this step the amplitude of the eyelid movements of the set Mu was calculated with the equation $Amp_q = \min(su_q - mu_q, eu_q - mu_q)$; $((su_q \in Su) \wedge (mu_q \in Mu) \wedge (eu_q \in Eu))$. The resulting values were sorted with a 3-means cluster algorithm (Arthur & Vassilvitskii, 2007). All eyelid movements in the group with the smallest amplitudes were excluded as potential blinks.

11. Data threshold If the eyelid distances in the minima (Md, Mm, and Mu) were greater than 100 pixels, they were excluded. 100 pixels corresponds approximately to an openness of the eye of 70%.

12. Merge eyelid movements A In the case that the number of smaller eyelid movement events between the start and end points of an event with a bigger eyelid distance was greater than or equal to one event, a check was performed to determine whether each eyelid movement event originated from noise. As the criterion, the biggest distance between the thresholds th_m or th_d and the eyelid distance in the minimum or minima of the set Mm or Md, respectively, was calculated. If the distance of one of the eyelid movements was smaller than or equal to $d = 7$ pixels (approximately a distance of 4% of eyelid closure in relation to the threshold), the associated minimum/minima of the set Mm, or, respectively Md was/were discarded and the start and end point of the eyelid event with the bigger eyelid distance (of the set Mu, respectively Mm) was retained. If the distance/s was/were larger than d, the associated eyelid distances in the sets Su, Eu and Mu, (respectively Sm, Em, and Mm) were eliminated and the event with smaller eyelid distances was retained.

13. Merge eyelid movements B All eyelid closure events in the three categories Md, Mm, and Mu were merged into one single set of blinks with the set of start point Sa, minima Ma, and end point Ea.

14. Combine eyelid movement events Due to occasional erroneous peaks in the eyelid distance signal, the described procedure could split one eyelid movement of one blink into several parts. To distinguish one single event from multiple ones, the pupil detection signal was checked during the time between the start and end samples extended on each side by one sample of two consecutive eyelid closure events. If the pupil was not detected between the two potential blinks, the two events were combined.

15. Adjust detected phases (optional) The signal contained noise that impaired the exact detection of the start and end points of a blink. To improve the detection, the raw signal was filtered using a Savitzky–Golay filter with polynomial order three and a frame window size seven (Savitzky & Golay, 1964). Similar to Step 7, a check was made afterward

to determe if each eyelid distance of the sets Sa and Ea resulting from Step 14 was still located on a local maximum of the filtered signal. If a start point was on a monotonically decreasing or an end point on a monotonically increasing slope, they were extended toward the start or end of the experiment until reaching a local maximum.

Due to the trade-off between good detection of blinks through several minima in the eyelid distance signal and the smoothing of the signal, the filtering of the raw signal after the removal of outliers in a late step (without using Step 3 – with Step 15; subsequently referred to as the *filtered method*; detection process: a_3) was additionally evaluated in a comparison with an algorithm in which we applied filtering of the data at an early step (with Step 3 – without using Step 15; subsequently referred to as the *raw method*; detection process: a_4). To evaluate the proposed detection processes a_3 and a_4 in combination with the chosen image-processing algorithm, the two proposed signal-processing algorithms by Bergasa et al. (2006) (detection process a_5) and Sukno et al. (2009) (detection process a_6) were implemented and tested on the same sequences as the other algorithms.

Definition of eyelid movements categorized as blinks

There have been different research approaches to measuring video-recorded eye blinks. Fogelton and Benesova (2016) compared different research groups, evaluating the different developed algorithms on the same data set recorded by Pan et al. (2007). The numbers of labeled ground truth blinks varied by nearly 7%. Differences in the lowest position of the upper eyelid during blinks were also identified by Jiang et al. (2013). To guarantee the precision and consistency of the labeling process and the characterization of each detected eyelid movement event of the algorithms implemented in this work, the lowest point of the upper eyelid was analyzed between the detected start and end samples of the detected eyelid movement events. To characterize the events, the middle and border of the pupil were used as the references for separating the eye into four regions where the border of the eyelid could reach its lowest point during an eyelid movement event (Fig. 3).

For this study, an eyelid movement categorized as a blink in the video data had to fulfill the following requirements:

1. The palpebral aperture had to gradually decrease at the beginning, reach its minimum, and then increase at the end of the eyelid movement, all between the determined start and end samples.
2. The eyelid had to cover enough of the pupil in the closed phase to block the vision temporarily—that is, the lowest point of the upper eyelid had to reach either region R_1 or R_2, defined in Fig. 3.
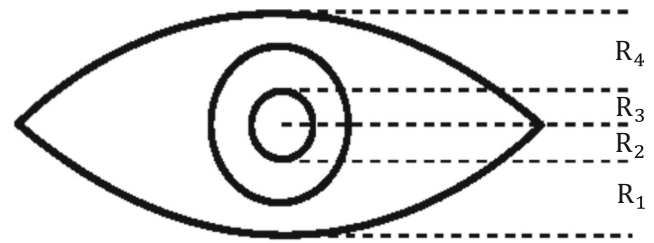

Fig. 3 Defined eye regions

3. The pupil had to be at least partially visible at both the start and end of the eyelid movement.

Depending on the maximum descent of the upper eyelid during the eyelid movement, the eyelid movements were further categorized as A_a (reached region R_1) and A_b (reached region R_2). Note that due to the complexity of the definition of the start and end of an eyelid movement, a detected eyelid movement was classified as a blink if Requirements 1–3 were fulfilled within the range of the detected start and end. To give a detailed overview of the specific types of other detections, they were defined in several subclasses:

A_c All eyelid closure events that fulfilled Requirements 1 and 3 but in which the lid only reached region R_3 during the detected boundaries were put in this class, counted as A_c events.

A_d Detected eye closures with the upper eyelid in region R_1 or R_2, but whose determined start and/or end sample failed one or both of Requirements 1 and 3, were put in this class. The eye closures had to fulfill the additional requirement: the real start and/or end fulfilling the requirements 1 and 3 had to be in a 200 millisecond time window towards the beginning and end of the detected start and end point.

A_e All detected blink-related events with the upper eyelid in regions R_1 or R_2 but that did not fulfill the additional requirement of eye closures for A_d were assigned to this class.

A_f All other events in which the eyelid only moved within or remained inside region R_4 (e.g., gaze movements) were put in this category, describing non-blink-related (A_f) events.

In case of multiple blinks inside one detected event, only one of the blinks (the one with the lowest eyelid distance) was assigned to one of the categories described above. All other blinks of a multiple-blink event were added to the undetected blinks and affected the detection rate.

Results

To obtain a detailed evaluation of the performance of algorithms based on various recording devices and signal-

processing approaches (EOG vs. head-mounted eyetracker unfiltered vs. head-mounted eyetracker filtered) in different driving modes (CAD vs. manual), signal frequencies (f_1 vs. f_2), and participant conditions (awake vs. drowsy), comparable driving sequences had to be analyzed for each case. In the experiments, an alert phase was defined as a section in which the participants reported a KSS ≤ 6, and a drowsy phase as one with KSS > 6. This was based on the instructions given the participants to rate themselves with a KSS of 8 or 9 if they were no longer able to drive. With respect to a future drowsiness detection algorithm based on blink detections, the driving phases with the state KSS = 7 were analyzed together with the phases in which the drivers had clearly rated themselves as no longer able to drive. In this way, the influence of rising drowsiness on blink detection in the transition phase from alert to drowsy was evaluated together with severe drowsiness in phases with a KSS of 8 or 9.

Due to missing data and errors in the data recordings, four participants in the manual-driving and five participants in the CAD study had to be excluded. One participant in the manual-driving experiment did not become drowsy; thus, 13 participants experienced the drowsy and 14 participants the alert phase in the manual-driving experiment. To examine a comparable number of participants in the CAD experiment, 16 participants were randomly chosen from the four groups (Group 1: Start at 6 p.m. + 30-s alertness requests; Group 2: Start at 6 p.m. + 180-s alertness requests; Group 3: Start at 10 p.m. + 30-s alertness requests; Group 4: Start at 10 p.m. + 180-s alertness requests). A further constraint on the selection of the participants of the CAD study was that they experienced an alert and a drowsy phase. For the analysis of all participants, 2 min from the alert and 2 min from the drowsy phase were chosen using a random algorithm (Wong & Easton, 1980). It was assured that the chosen parts from the manual-driving study did not occur during a KSS query or in a 2-min time window following the KSS task. For the CAD experiment, the 2-min sections in the alert and drowsy phase were chosen as continuous windows during a switch to CAD mode not occuring during a takeover situation, KSS query, or 2 min after them. Due to the short and frequently experienced alertness requests, no additional exclusion criterion was applied. This resulted in 118 min of data as input for the evaluation. Overall, 2,941 eyelid movements in categories A_a and A_b were assessed in the video of the head-mounted eyetracker. These are further described as *ground truth eye closures* in the two categories GT_a (upper eyelid reached region R_1) and GT_b (upper eyelid reached region R_2) (see Table 2).

Three detection rates were calculated to evaluate the algorithm: the true positive rate (TPR), the blink-related false

Table 2 Overview of the ground truth data

	Manual Driving		CAD	
	Awake	Drowsy	Awake	Drowsy
Number of 2-min sequences	14	13	16	16
Real eye closure events (GT_a and GT_b)	655	875	737	674

detection rate (FDR_{BR}), and the not-blink-related false detection rate (FDR_{NBR}), based on the classification provided above. These were defined as follows:

$$TPR = \frac{|A_a| + |A_b|}{|GT_a| + |GT_b|}$$

$$FDR_{BR} = \frac{|A_c| + |A_d| + |A_e|}{|A_a| + |A_b| + |A_c| + |A_d| + |A_e| + |A_f|}$$

$$FDR_{NBR} = \frac{|A_f|}{|A_a| + |A_b| + |A_c| + |A_d| + |A_e| + |A_f|}$$

$|A_z|$ and $|GT_z|$ define the numbers of all detected eyelid movements by the algorithms of the categories A_z and $GT_z (z = a, b, c, d, e, \text{or } f)$, which were labeled and categorized as stated above.

The distinction between the blink-related events and not-blink-related events was made because the blink-related events showed a strong similarity with blinks and contained a part of a true positive (part of a real blink). Furthermore, it could not be assured that blinks in the category A_c would be reassigned to the category A_b if a video recording with a higher sampling rate than the video recording of the head-mounted eyetracker recording with 25 Hz had been used for the labeling process.

The labeling process was started by labeling all detected eyelid movement events detected with EOG and sampling frequency f_1. To limit extensive labeling, the detected eyelid movements using the EOG algorithm with a lower frequency f_2 and the presented algorithm using the eyetracker data were checked for matching areas with a true positive from the start and end points generated by the first detection process a_1. If the complete intersection between the labelled eyelid movement (starting point s_l, end point e_l) with the corresponding boundaries from a_2(starting point s_{o2}, end point e_{o2}), a_3(starting point s_{o3}, end point e_{o3}), or a_4(starting point s_{o4}, end point e_{o4}) exceeded $P_1 = 80\%$ for eyelid movements of less than 500 ms, the event was categorized with the same label as the previously labeled event. If the time between s_l and e_l exceeded 500 ms, the time interval Δt of the intersection between the boundaries derived from a_1 and a_w (w = 2, 3, or 4) had to exceed

$P_2 = 100\% \left(1 - \frac{100 \text{ ms}}{\Delta t}\right)$ in order to adopt the same label. This was based on the natural time of an eye closure or opening, which Stern et al. (1984) described for longer eye closures. Eye movements with an intersection of less than P_1 or P_2 were labeled manually.

Since each of the research groups of the detection algorithms a_5 and a_6 defined blinks differently from the definition proposed in this article, each detected eyelid movement event of the algorithms a_5 and a_6 was labelled separately. The events were labelled as true positives dependent the individual definition of the authors. All other detected events that fulfilled the definition of a blink relateevent as described above were categorized as blink-related events if they differed from the individual definition of a blink.

As a summary for evaluation of the main research goals, the split as a function of measurement technique (EOG 25 Hz vs. eyetracker 25 Hz), driver state (awake vs. drowsy), level of automation (manual vs. CAD), and signal frequency (50-Hz EOG vs. 25-Hz EOG) are given in Table 3, showing $MN \pm SD$ of the detection rates of the eye closures. Since the detection rates with a_3 were best for the video recording, those results were selected for the subsequent evaluation. Two detailed tables showing the performance of all algorithms are included in the Appendix.

To evaluate the labeling process, all events of the basis detection with EOG and the sampling frequency f_1 were labeled three additional times by an independent labeler. The event-based matching was 98.5% for the manual drives and 95.6% for CAD.

Discussion

The results in Table 3 show that the detection rates TPR, FDR_{BR}, and FDR_{NBR} of eyelid movements vary between the different subgroups analyzed. A Friedman test showed significant differences between the TPR [$\chi^2(2) = 22.052$, $p < .001$] and FDR_{NBR} [$\chi^2(2) = 28.056$, $p < .001$]. In the direct comparison between the EOG system (f_2) with the best result for the head-mounted eyetracker operated at the same frequency (a_3), the rates are very similar with few exceptions. The results show a significant higher FDR_{NBR} for the video recordings ($Z = -3.2$, $p = .001$, $r = .42$). The Bonferroni correction was used for the signifince level of this and all subsequent post hoc tests. A larger difference between a_2 and a_3 is also visible in the detection rates between the awake CAD phases, favoring the video recordings. The difference was not significant anymore after correcting the significance level. This indicates that for a higher level of automation and for the purpose of simulator or real road driving studies, a camera might be better suited for blink detections than an EOG system under the same constraints. One reason could be the use of an indirect measuring technique rather than the two direct methods used with the video recordings (visibility of the pupil and eyelid distance). For manual driving, the EOG technique provides an advantage. During the labeling process, several problems related to each of the technologies were noted. EOG had problems with long eye closures, small amplitudes with vertical saccades resulting in signal courses similar to eyelid closures, and eye movements during eye closures. On the other hand, the image processing step of the eyetracker video for the eyelid detection had problems with disturbances by eyelashes, the blockage of the lower eyelid during the drop of the upper eyelid, and wrinkles in the skin or edges around the eye that were misinterpreted as eyelids. Shorter eyelid distances and longer eye closures are especially likely to occur during drowsy phases, contributing to the different problems of both measuring devices. This can be seen in the significant drop ($Z = -4.172$, $p < .001$, $r = .45$) of the correct detection rates from the alert to the drowsy phases (drop between 2% and 16%) independent of the measuring frequency (f_1 vs. f_2), measuring technology (EOG vs. video), and automation level (manual vs. CAD). Despite the different constraints in the experiments (speed, alertness, and takeover requests), the drivers were always aware of the relatively high travelling speed and their responsibilities in the CAD experiment. Particularly the short manual-driving portions during the takeover situations enhanced the preparation awareness for a takeover. The general and significant decrease in the correct detection rate from manual to automated driving (drop in the detection rate of the EOG and video based algorithm with 9 to 29%; $U = 1,832.500$, $p < .001$, $r = .49$) and significant increase in the FDR_{BR} ($U = 2,940.500$, $p = .004$, $r = .22$) cannot be explained by the experiments' constraints. Therefore, they imply a change in the behavior of the eyelid movements of the drivers, confirming the results of Schmidt, Braunagel, et al. (2016). This further shows that detection rates obtained during manual driving should not be applied universally to CAD mode. Comparing the two measurement

Table 3 Performance of the blink detection algorithms (in %) with their $MN \pm SD$

		a_1		a_2		a_3	
		Aw	Dr	Aw	Dr	Aw	Dr
Manual	TPR	94±5	92±13	91±7	89±13	91±6	85±20
	FDR_{BR}	6±11	2±2	5±8	2±2	5±8	5±7
	FDR_{NBR}	2±2	1±1	4±4	2±2	9±9	8±13
CAD	TPR	81±20	69±22	71±20	60±24	82±20	66±27
	FDR_{BR}	8±11	11±14	11±13	14±21	6±6	19±32
	FDR_{NBR}	4±6	4±6	5±8	2±4	6±6	4±6

Aw = Awake, Dr = Drowsy, TPR = True positive rate, FDR_{BR} = False detection rate blink-related, FDR_{NBR} = False detection rate not blink-related

frequencies (f_1) and (f_2), significantly better detections (rising from 3% and 10%; $Z = -4.736$, $p < .001$, $r = .62$) can be seen in the outcome of the blink detection with the higher frequency. This effect is greater for CAD and indicates a higher potential for detecting blinks using higher frequencies for the measuring system in CAD. On the basis of these results, along with the findings from Picot et al. (2009), this effect is also expected when using different measurement frequencies for video recordings. The four signal processing algorithms tested (a_3, a_4, a_5, and a_6) show significant differences in the accuracy of eye closure event detection (see the detailed results in the Appendix). A possible reason for that could lie in the adaption of the two re-implemented algorithms from Bergasa et al. (2006) and Sukno et al. (2009) to their own image-processing signals and study constraints, differences that do not apply to our present study.

The introduced subcategories of A_b and A_c give a valuable indication of the differences in the eyelid behavior during manual driving. Since the decrease in the rates of A_b and A_c from an awake to a drowsy driver is larger than the overall missed events and decrease in the correct detection rate, it can be concluded that alert drivers usually do not close their eyes during a blink as far as drowsy drivers. This might be due to the urge to look at the street and the attempt to reduce the vision lost during a blink to a minimum.

The other subcategories of the FDR_{BR} show the potential of improvement of the TPR in the blink detection process. The high rates for the CAD phases imply that the adapted blink behavior of the drivers during passive driving, with features such as longer shut times and eye closure and opening phases, causes more difficulties for the detection process.

Conclusion and future work

The goal of this article was to study different influences on the blink detection. The results indicate that the detection of blinks is primarily influenced by the level of automation, driver state, the measurement frequency, and the algorithms used. In this respect, the results quantify the influence on blink detection and show, how known research methods can be used to investigate the various influences. Furthermore, the introduced detection methods offer a new approach for detecting blinks and can be implemented in other studies using our detailed description of the signal processing steps. A detailed classification of eyelid movements reveals additional states in eyelid movement and behaviors of drivers in the tested circumstances. This could also be used as the common basis for comparing different approaches of blink detection in future studies. The video of a head-mounted eyetracker enabled a highly accurate labeling process, which is more difficult to obtain with a remote eyetracker. Using the head-mounted eyetracker, it was possible to show the results of the blink detection process without any influence of head movements

and to gain greater insight into the influences of the automation mode and the drowsiness state of the driver. These were the main reasons why a head-mounted eyetracker was selected as the exemplary video source for this study. The use of other signal-processing methods and their lower performance shows the dependency of single blink detection methods on specific image processing methods and the states of the drivers during the recordings of the studies. Therefore, adapting methods to different conditions and drivers is necessary for accurate blink detection. In contrast to the EOG method, the detection based on video images allows to extract more information from the video than just a single signal describing the eyelid distance. The variety of different video-based blink detection methods could be used to improve the correct detection rate by combining several of them in a single algorithm. A higher video-recording measuring rate could increase the accuracy of the detection as well. Aside from an improvement in the correct detection rate, future analyses should focus on the examination of driver behavior during CAD, especially during drowsy driving. To build on the detected blink sequences, further information about the driver behavior, such as the blink frequency, duration of the eyelid closure, amplitude, or velocity of the eyelid closing and opening, should be evaluated. In this way, drivers could be accurately classified as too drowsy to continue driving in the conditionally automated mode. Overall, we showed that video recordings can be used to detect driver behavior in the future modes of travelling as a replacement for the estimation by steering behavior. Further studies should examine the influence on the detection rates of a remote system.

Appendix

The performance of all algorithms is presented in Tables 4 and 5, together with the rates of each subgroup. Note that the subcategories of the TPR (TPR_{Aa} = part of the true positive rate with events labeled as A_a and TPR_{Ab} = part of the true positive rate with events labeled as A_b) and FDR_{BR} (FDR_{Ac} = part of the true positive rate with events labeled as A_c, FDR_{Ad} = part of the true positive rate with events labeled as A_d, and FDR_{Ae} = part of the true positive rate with events labeled as A_e) are rounded parts that add up to the TPR and FDR_{BR}, to show the specific shares of each defined classification of eyelid movement events. Note further that the labeling process for a_5 and a_6 did not distinguish whether the events belonged to any of the subgroups (e.g., A_a, A_c, A_e) since those are dependent on the definition proposed by the respective authors. All

other events remained in category A_f. A comparison of the four camera algorithms show significant differences for the TPR [$\chi^2(3) = 85.401$, $p < .001$], FDR_{BR} [$\chi^2(3) = 23.090$, $p < .001$], and FDR_{NBR} [$\chi^2(3) = 92.257$, $p < .001$]. Post-hoc tests revealed a significantly better detection rate for a_3 (a_3 vs. a_4: $Z = -3.344$, $p < .001$, $r = .44$; a_3 vs. a_5: $Z = -6.424$, $p < .001$, $r = .84$; a_3 vs. a_6: $Z = -5.596$, $p < .001$, $r = .73$), which was therefore selected for evaluation of the EOG algorithm.

Table 4 Detailed performance of blink detection algorithms a_1 to a_4 (in %), MN ± SD

		a_1		a_2		a_3		a_4	
		Aw	Dr	Aw	Dr	Aw	Dr	Aw	Dr
Manual	**TPR**	94±5	92±13	91±7	89±13	91±6	85±20	87±7	82±23
	TPR_{Aa}	69±21	84±16	68±21	82±15	65±22	76±19	63±23	73±21
	TPR_{Ab}	25±19	8±10	24±19	7±9	26±19	9±11	24±18	9±11
	FDR_{BR}	6±11	2±2	5±8	2±2	5±8	5±7	4±6	5±9
	FDR_{Ac}	6±10	0±1	5±8	0±1	5±7	1±2	3±6	0±1
	FDR_{Ad}	0±1	1±2	0±1	1±2	1±2	4±7	0±0	5±9
	FDR_{Ae}	0±0	0±1	0±0	0±1	0±0	0±0	0±0	0±0
	FDR_{NBR}	2±2	1±1	4±4	2±2	9±9	8±13	6±5	7±5
CAD	**TPR**	81±20	69±22	71±20	60±24	82±20	66±27	81±17	63±26
	TPR_{Aa}	71±18	64±24	62±17	57±24	71±17	58±25	71±15	55±24
	TPR_{Ab}	10±12	5±9	81±0	3±6	11±12	9±16	11±10	8±13
	FDR_{BR}	8±11	11±14	11±13	14±21	6±6	19±32	7±8	19±29
	FDR_{Ac}	1±2	0±1	1±2	0±1	2±3	1±1	2±3	1±2
	FDR_{Ad}	6±11	3±6	5±10	3±5	3±6	2±5	3±5	2±5
	FDR_{Ae}	1±2	7±13	4±10	11±18	1±3	16±31	2±5	16±29
	FDR_{NBR}	4±6	4±6	5±8	2±4	6±6	4±6	4±5	2±4

Aw = Awake, Dr = Drowsy, TPR = True positive rate, TPR_{Aa} = True positive rate in the eye region R_1, TPR_{Ab} = True positive rate in the eye region R_2, FDR_{BR} = False detection rate blink-related, FDR_{Ac} = False detection rate blink-related in the eye region R_3, FDR_{Ad} = False detection rate blink-related in the category A_d, FDR_{Ae} = False detection rate blink-related in the category A_e, FDR_{NBR} = False detection rate not blink-related

Table 5 Detailed performance of the blink detection algorithms a_5 and a_6 (in %), MN ± SD

		a_5		a_6	
		Aw	Dr	Aw	Dr
Manual	**TPR**	53±19	50±22	62±38	54±45
	FDR_{BR}	8±5	12±9	18±22	20±32
	FDR_{NBR}	51±28	37±25	19±27	13±20
CAD	**TPR**	52±26	33±24	47±46	22±38
	FDR_{BR}	18±20	28±27	18±31	26±38
	FDR_{NBR}	36±19	34±27	12±20	9±26

Aw = Awake, Dr = Drowsy, TPR = True positive rate, FDR_{BR} = False detection rate blink-related, FDR_{NBR} = False detection rate not blink-related

References

Åkerstedt, T., & Gillberg, M. (1990). Subjective and objective sleepiness in the active individual. *International Journal of Neuroscience, 52,* 29–37. doi:10.3109/00207459008994241

Arthur, D., & Vassilvitskii, S. (2007). *k*-means++: The advantages of careful seeding. In *Proceedings of the Eighteenth Annual ACM–SIAM Symposium on Discrete Algorithms* (pp. 1027–1035). Philadelphia: Society for Industrial and Applied Mathematics. doi:10.1145/1283383.1283494

Benoit, A., & Caplier, A. (2010). Fusing bio-inspired vision data for simplified high level scene interpretation: Application to face motion analysis. *Computer Vision and Image Understanding, 114,* 774–789. doi:10.1016/j.cviu.2010.01.010

Bergasa, L. M., Nuevo, J., Sotelo, M. A., Barea, R., & Lopez, M. E. (2006). Real-time system for monitoring driver vigilance. *IEEE Transactions on Intelligent Transportation Systems, 7,* 63–77. doi:10.1109/TITS.2006.869598

Bhaskar, T. N., Keat, F. T., Ranganath, S., & Venkatesh, Y. V. (2003). Blink detection and eye tracking for eye localization. In *IEEE TENCON 2003: Conference on Convergent Technologies for the Asia-Pacific Region* (pp. 821–824). Piscataway: IEEE Press. doi:10.1109/TENCON.2003.1273293

BMW Group. (2016). BMW Group, Intel and Mobileye team up to bring fully autonomous driving to streets by 2021 (Press release). Retrieved from https:/www.press.bmwgroup.com/global/article/detail/T0261586EN/bmw-group-intel-and-mobileye-team-up-to-bring-fully-autonomous-driving-to-streets-by-2021?language=en

Brain Products GmbH. (2009). Selecting a suitable EEG recording cap: Tutorial. Retrieved September 13, 2016, from www.brainproducts.com/downloads.php?kid=8

Chau, M., & Betke, M. (2005). *Real time eye tracking and blink detection with USB cameras.* Boston: Boston University, Computer Science Department.

Choi, I., Han, S., & Kim, D. (2011). Eye detection and eye blink detection using AdaBoost learning and grouping. In *2011 20th International Conference on Computer Communications and Networks—ICCCN 2011* (pp. 1–4). Piscataway: IEEE Press. doi:10.1109/ICCCN.2011.6005896

Cohn, J. F., Xiao, J., Moriyama, T., Ambadar, Z., & Kanade, T. (2003). Automatic recognition of eye blinking in spontaneously occurring behavior. *Behavior Research Methods, Instruments, & Computers, 35,* 420–428. doi:10.3758/BF03195519

Collewijn, H., van der Steen, J., & Steinman, R. M. (1985). Human eye movements associated with blinks and prolonged eyelid closure. *Journal of Neurophysiology, 54,* 11–27.

Daimler AG. (2008). *HighTech report 02* (Technical report). Retrieved from http://mediasite.daimler.com/Projects/c2c/channel/files/287920_Hightechreport_2_2008_en.pdf

Danisman, T., Bilasco, I. M., Djeraba, C., & Ihaddadene, N. (2010). Drowsy driver detection system using eye blink patterns. In *2010 International Conference on Machine and Web Intelligence (ICMWI)* (pp. 230–233). Piscataway: IEEE Press. doi:10.1109/ICMWI.2010.5648121

Dawson, D., Searle, A. K., & Paterson, J. L. (2014). Look before you (s)leep: Evaluating the use of fatigue detection technologies within a fatigue risk management system for the road transport industry. *Sleep Medicine Reviews, 18,* 141–152. doi:10.1016/j.smrv.2013.03.003

Divjak, M., & Bischof, H. (2009). Eye blink based fatigue detection for prevention of Computer Vision Syndrome. In *MVA: Proceedings of the 2009 IAPR Conference on Machine Vision Applications* (pp. 350–353). Tokyo: MVA.

Dong, Y., Hu, Z., Uchimura, K., & Murayama, N. (2011). Driver inattention monitoring system for intelligent vehicles: A review. *IEEE Transactions on Intelligent Transportation Systems, 12,* 596–614. doi:10.1109/TITS.2010.2092770

Ebrahim, P. (2016). Driver drowsiness monitoring using eye movement features derived from electrooculography (Doctoral dissertation). Stuttgart, Germany: Universität Stuttgart. doi:10.18419/opus-8820

Ergoneers GmbH. (2016). Dikablis (Product manual). Retrieved August 27, 2016, from www.irc-web.co.jp/vicon_web/news_bn/Dikabilis.pdf

Fogelton, A., & Benesova, W. (2016). Eye blink detection based on motion vectors analysis. *Computer Vision and Image Understanding, 148,* 23–33. doi:10.1016/j.cviu.2016.03.011

Ford. (2010). Driver alert (Press release). Retrieved September 12, 2016, from http://technology.fordmedia.eu/documents/newsletter/FordTechnologyNewsletter082010.pdf

Friedrichs, F., & Yang, B. (2010). Camera-based drowsiness reference for driver state classification under real driving conditions. In *2010 I.E. Intelligent Vehicles Symposium (IV)* (pp. 101–106). Piscataway: IEEE Press. doi:10.1109/IVS.2010.5548039

Fuhl, W., Kübler, T., Sippel, K., Rosenstiel, W., & Kasneci, E. (2015). ExCuSe: Robust pupil detection in real-world scenarios. In G. Azzopardi & N. Petkov (Eds.), *Computer analysis of images and patterns, 16th international conference, CAIP 2015, Valletta, Malta, September 2–4, 2015 (Lecture Notes in Computer Science 9256* (pp. 39–51). Heidelberg: Springer. doi:10.1007/978-3-319-23192-1_4

Fuhl, W., Santini, T., Geisler, D., Kübler, T., Rosenstiel, W., & Kasneci, E. (2016). Eyes wide open? Eyelid location and eye aperture estimation for pervasive eye tracking in real-world scenarios. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct publication* (pp. 1656–1665). New York: ACM.

Garcia, I., Bronte, S., Bergasa, L. M., Almazan, J., & Yebes, J. (2012). Vision-based drowsiness detector for real driving conditions. In *2012 I.E. Intelligent Vehicles Symposium (IV)* (pp. 618–623). Piscataway: IEEE Press. doi:10.1109/IVS.2012.6232222

Gorodnichy, D. (2003). Second order change detection, and its application to blink-controlled perceptual interfaces. In *2003 Proceedings of the International Association of Science and Technology for Development (IASTED) Conference on Visualization, Imaging and Image Processing (VIIP)* (pp. 140–145). Calgary: ACTA Press.

Grauman, K., Betke, M., Gips, J., & Bradski, G. R. (2001). Communication via eye blinks: Detection and duration analysis in real time. In *2001 I.E. Computer Society Conference on Computer Vision and Pattern Recognition: CVPR 2001* (pp. I-1010–I-1017). Piscataway: IEEE Press. doi:10.1109/CVPR.2001.990641

Hansen, D. W., & Pece, A. E. (2005). Eye tracking in the wild. *Computer Vision and Image Understanding, 98,* 155–181. doi:10.1016/j.cviu.2004.07.013

Hell, W., Schick, S., Gilg, T., Graw, M., Schönpflug, M., Adamec, J., … & Peldschus, S. (2012). Verkehrsunfallanalyse. In H.-T. Haffner, G. Skopp, & M. Graw (Eds.), *Begutachtung im Verkehrsrecht. Fahrtüchtigkeit: Fahreignung—Traumatomechanische Unfallrekonstruktion. Bildidentifikation* (pp. 219–325). Berlin, Germany: Springer. doi:10.1007/978-3-642-20225-4_3

Hu, S., & Zheng, G. (2009). Driver drowsiness detection with eyelid related parameters by support vector machine. *Expert Systems With Applications, 36,* 7651–7658. doi:10.1016/j.eswa.2008.09.030

Ito, T., Mita, S., Kozuka, K., Nakano, T., & Yamamoto, S. (2002). Driver blink measurement by the motion picture processing and its application to drowsiness detection. In *IEEE 5th International Conference on Intelligent Transportation Systems* (pp. 168–173). Piscataway: IEEE Press. doi:10.1109/ITSC.2002.1041208

Jafar Ali, M., Sarkar, S., Kumar, G. P., & Cabibihan, J.-J. (2012). A non intrusive human fatigue monitoring system. *International Journal of Future Computer and Communication, 6,* 284–288. doi:10.7763/IJFCC.2012.V1.76

Jammes, B., Sharabty, H., & Esteve, D. (2008). Automatic EOG analysis: A first step toward automatic drowsiness scoring during wake-sleep transitions. *Somnologie: Schlafforschung und Schlafmedizin, 12,* 227–232. doi:10.1007/s11818-008-0351-y

Jiang, X., Tien, G., Huang, D., Zheng, B., & Atkins, M. S. (2013). Capturing and evaluating blinks from video-based eyetrackers. *Behavior Research Methods, 45,* 656–663. doi:10.3758/s13428-012-0294-x

Lee, W. O., Lee, E. C., & Park, K. R. (2010). Blink detection robust to various facial poses. *Journal of Neuroscience Methods, 193,* 356–372. doi:10.1016/j.jneumeth.2010.08.034

Lobb, M. L., & Stern, J. A. (1986). Pattern of eyelid motion predictive of decision errors during drowsiness: Oculomotor indices of altered states. *International Journal of Neuroscience, 30,* 17–22.

Malik, K., & Smolka, B. (2014). Eye blink detection using Local Binary Patterns. In *2014 International Conference on Multimedia Computing and Systems (ICMCS)* (pp. 385–390). Piscataway: IEEE Press. doi:10.1109/ICMCS.2014.6911268

Missimer, E., & Betke, M. (2010). Blink and wink detection for mouse pointer control. In F. Makedon (Ed.), *PETRA '10: Proceedings of the 3rd International Conference on Pervasive Technologies Related to Assistive Environments* (p. 23). New York: ACM. doi:10.1145/1839294.1839322

Miyakawa, T., Takano, H., & Nakamura, K. (2004). Development of non-contact real-time blink detection system for doze alarm. In *SICE 2004 Annual Conference* (Vol. 2, pp. 1626–1631).

Moriyama, T., Kanade, T., Cohn, J. F., Xiao, J., Ambadar, Z., Gao, J., & Imamura, H. (2002). Automatic recognition of eye blinking in spontaneously occurring behavior. In *16th International Conference on Pattern Recognition* (pp. 78–81). Piscataway: IEEE Press. doi:10.1109/ICPR.2002.1047404

Morris, T. L., & Miller, J. C. (1996). Electrooculographic and performance indices of fatigue during simulated flight. *Biological Psychology, 42,* 343–360. doi:10.1016/0301-0511(95)05166-X

Pan, G., Sun, L., Wu, Z., & Lao, S. (2007). Eyeblink-based Anti-Spoofing in Face Recognition from a Generic Webcamera. In *2007 I.E. 11th International Conference on Computer Vision* (pp. 1–8). Piscataway: IEEE Press. doi:10.1109/ICCV.2007.4409068

Pan, G., Sun, L., & Wu, Z. (2008). *Liveness detection for face recognition*. Place: INTECH Open Access Publisher.

Patel, S., Henderson, R., Bradley, L., Galloway, B., & Hunter, L. (1991). Effect of visual display unit use on blink rate and tear stability. *Optometry and Vision Science, 68,* 888–892.

Pedrotti, M., Lei, S., Dzaack, J., & Rötting, M. (2011). A data-driven algorithm for offline pupil signal preprocessing and eyeblink detection in low-speed eye-tracking protocols. *Behavior Research Methods, 43,* 372–383. doi:10.3758/s13428-010-0055-7

Picot, A., Caplier, A., & Charbonnier, S. (2009). Comparison between EOG and high frame rate camera for drowsiness detection. In *2009 Workshop on Applications of Computer Vision (WACV)* (pp. 1–6). Piscataway: IEEE Press. doi:10.1109/WACV.2009.5403120

Picot, A., Charbonnier, S., & Caplier, A. (2012). On-Line Detection of Drowsiness Using Brain and Visual Information. *IEEE Transactions on Systems, Man, and Cybernetics—Part A: Systems and Humans, 42,* 764–775. doi:10.1109/TSMCA.2011.2164242

Radlak, K., & Smolka, B. (2012). A novel approach to the eye movement analysis using a high speed camera. In *2012 2nd International Conference on Advances in Computational Tools for Engineering Applications (ACTEA)* (pp. 145–150). Piscataway: IEEE Press. doi:10.1109/ICTEA.2012.6462854

SAE. (2014). Standard J3016_201401: Taxonomy and definitions for terms related to on-road motor vehicle automated driving systems (Issued January 16, 2014). Retrieved from standards.sae.org/j3016_201401/

Sahayadhas, A., Sundaraj, K., & Murugappan, M. (2012). Detecting driver drowsiness based on sensors: A review. *Sensors, 12,* 16937–16953. doi:10.3390/s121216937

Savitzky, A., & Golay, M. J. E. (1964). Smoothing and differentiation of data by simplified least squares procedures. *Analytical Chemistry, 36,* 1627–1639. doi:10.1021/ac60214a047

Schmidt, J., Braunagel, C., Stolzmann, W., & Karrer-Gauß, K. (2016). Driver drowsiness and behavior detection in prolonged conditionally automated drives. In *2016 I.E. Intelligent Vehicles Symposium (IV)* (pp. 400–405). Piscataway: IEEE Press. doi:10.1109/IVS.2016.7535417

Schmidt, J., Stolzmann, W., & Karrer-Gauß, K. (2016). Experimental evaluation of different request intervals for a driver alertness device for conditionally automated driving with induced drowsiness. In *Proceedings Automatisierungs—Assistenzsysteme und eingebettete Systeme für Transportmittel (AAET)*. Braunschweig, Germany: Gesamtzentrum für Verkehr Braunschweig.

Sirohey, S., Rosenfeld, A., & Duric, Z. (2002). A method of detecting and tracking irises and eyelids in video. *Pattern Recognition, 35,* 1389–1401. doi:10.1016/S0031-3203(01)00116-9

Skotte, J. H., Nøjgaard, J. K., Jørgensen, L. V., Christensen, K. B., & Sjøgaard, G. (2007). Eye blink frequency during different computer tasks quantified by electrooculography. *European Journal of Applied Physiology, 99,* 113–119. doi:10.1007/s00421-006-0322-6

Stern, J. A., Boyer, D., & Schroeder, D. (1994). Blink rate: A possible measure of fatigue. *Human Factors, 36,* 285–297. doi:10.1177/001872089403600209

Stern, J. A., Walrath, L. C., & Goldstein, R. (1984). The endogenous eyeblink. *Psychophysiology, 21,* 22–33. doi:10.1111/j.1469-8986.1984.tb02312.x

Sukno, F. M., Pavani, S.-K., Butakoff, C., & Frangi, A. F. (2009). Automatic assessment of eye blinking patterns through statistical shape models. In M. Fritz (Ed.), *Proceedings of Computer Vision Systems, 7th international conference, ICVS 2009, Liège, Belgium, October 13–15, 2009 (Lecture Notes in Computer Science 5815* (pp. 33–42). Berlin: Springer. doi:10.1007/978-3-642-04667-4_4

Torricelli, D., Goffredo, M., Conforto, S., & Schmid, M. (2009). An adaptive blink detector to initialize and update a view-based remote eye gaze tracking system in a natural scenario. *Pattern Recognition Letters, 30,* 1144–1150. doi:10.1016/j.patrec.2009.05.014

Tsai, Y.-F., Viirre, E., Strychacz, C., Chase, B., & Jung, T.-P. (2007). Task performance and eye activity: Predicting behavior relating to cognitive workload. *Aviation, Space, and Environmental Medicine, 78,* B176–B185.

Wei, J., Snider, J. M., Kim, J., Dolan, J. M., Rajkumar, R., & Litkouhi, B. (2013). Towards a viable autonomous driving research platform. In *2013 I.E. Intelligent Vehicles Symposium (IV)* (pp. 763–770). Piscataway: IEEE Press. doi:10.1109/IVS.2013.6629559

Wolkoff, P., Nøjgaard, J. K., Troiano, P., & Piccoli, B. (2005). Eye complaints in the office environment: Precorneal tear film integrity influenced by eye blinking efficiency. *Occupational and Environmental Medicine, 62,* 4–12. doi:10.1136/oem.2004.016030

Wong, C. K., & Easton, M. C. (1980). An efficient method for weighted sampling without replacement. *SIAM Journal on Computing, 9,* 111–113. doi:10.1137/0209009