

Drowsiness detection based on visual signs: blinking analysis based on high frame rate video

Antoine Picot, Sylvie Charbonnier and Alice Caplier

Gipsa Lab

Grenoble Université, Grenoble, France

name.surname@gipsa-lab.grenoble-inp.fr

Abstract—In this paper, an algorithm for drivers' drowsiness detection based on visual signs that can be extracted from the analysis of a high frame rate video is presented. A study of different visual features on a consistent database is proposed to evaluate their relevancy to detect drowsiness by data-mining. Then, an algorithm that merges the most relevant blink features (duration, percentage of eye closure, frequency of the blinks and amplitude-velocity ratio) using fuzzy logic is proposed. This algorithm has been tested on a huge dataset representing 60 hours of driving from 20 different drivers. The main advantage of this algorithm is that it is independent from the driver and it does not need to be tuned. Moreover, it provides good results with more than 80% of good detections of drowsy states.

I. INTRODUCTION

Drowsiness is the transition between awake state and sleep during which one's abilities to observe and analyse are strongly reduced. According to the National Highway Traffic Safety Administration (NHTSA), about 100,000 crashes are the direct result of driver drowsiness each year. This is the reason why more and more researches are made to build automatic detectors of this dangerous state.

According to Renner and Mehring [14], drowsiness can be detected both in brain activity which refers to the ability to process the information and in eyes activity which refers to the perception ability. This study focuses on the visual signs of drowsiness: blinks.

Electrooculogram (EOG), which is the measurement of the eye electrical muscles activity, has been widely used in the literature to estimate drowsiness ([3], [5]). EOG is the most reliable technique to detect and characterize blinks due to its high sample rate (from $250Hz$ to $500Hz$) and it is used as a reference by expert doctors to evaluate drowsiness ([6], [11]). Unfortunately, EOG requires at least three electrodes placed on the driver's skin, which is not pleasant. So, for obvious ergonomic reasons, research community has focused on the use of video to track the driver's eyes and face and thus detect whether he is drowsy or not.

Several features can be extracted from the video to estimate drowsiness such as blink duration, blink frequency, the PERcentage of eye CLOSure (PERCLOS) as well as driver's gaze or facial expressions ([10], [8], [2]). Most of these systems focus on too long blink durations to detect drowsiness ([7]). Johns [9] suggested to use the amplitude-velocity ratio as an universal drowsiness indicator. Most recently, Damousis et al. [4] proposed a fuzzy expert system merging blink duration,

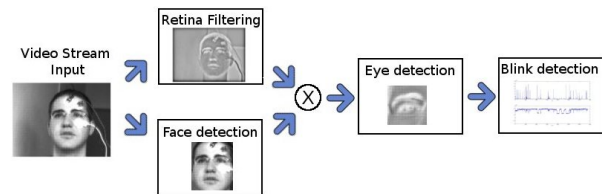


Fig. 1. Blink detection method

blink frequency and blink interval to estimate drowsiness. Nevertheless, blink video characterization is not very accurate due to the low frame rate of standard cameras ($30fps$) compared to EOG.

A recent study by Picot et al. [12] has shown that the use of a high frame rate video ($200fps$) can replace the EOG signal for the extraction of several blink features with the same accuracy. A drivers' drowsiness detection system based on these particular blinking features is proposed here. This system is presented on EOG because expertized high frame rate drowsiness video are still not available. Nevertheless, a huge dataset of 60 hours of driving EOG data is used here to develop this system. The database is first used to select the most relevant blinking features. A drowsiness detection method based on fuzzy logic is then proposed and validated on the database.

The outline of this paper is the following. In section II, a study of the different features for drowsiness estimation is processed by data-mining. This study leads to the proposition of a drowsiness detection system based on fuzzy logic. Results obtained on a database of 20 different drivers are presented and discussed in section III.

II. DROWSINESS DETECTION SYSTEM

The most relevant blinking features for drowsiness estimation are first selected from a large expertized dataset by data-mining. Fuzzy logic is then used to merge these different features and detect drowsiness.

A. Video based blink detection and characterization

The method to extract blink features from video has been presented in a recent study [12]. The general principle of this method is shown in Fig. 1.

Twelve features extracted with this method were compared to those extracted from the EOG. It has been shown that the following features: duration at 50% (D_{50}) which is the duration calculated from the half rise amplitude to the half fall amplitude, percentage of eye closure at 80% (P_{80}) which is the percentage where the eyes were closed at least 80%, the frequency (F) which is the number of blinks during a defined period, the amplitude-velocity ratio (A/PCV) which is the ratio between the amplitude of the blink and the maximum speed during the closing period and the energy of eye closure (E_C) which the maximum energy of moving edges during the closing period, can be estimated from a 200Hz video with the same accuracy as from EOG. In the following, a drowsiness detection using these features is presented. The analysis is done using features extracted from EOG but, considering the results obtained in [12], the detection could be implemented using features extracted from video and would reach the same performances.

The relevancy of these different features to estimate drowsiness is studied here by data-mining to find which features can be used to design the drowsiness detection system.

B. Material

The database used in this study is consistent. It includes 60 hours of driving data from 20 different drivers. Each subject has been recorded while driving on a simulator for 90 minutes, a first time perfectly rested and a second time suffering from sleep deprivation (the subject had slept for 4 hours only). Each recording includes four EEG channels and one EOG channel. Only the EOG channel is used in this paper. The EOG is recorded at 250Hz. Data acquisition was performed by the CEPA (*Centre d'Études de Physiologie Appliquée*), Strasbourg, FR.

Objective sleepiness was evaluated every 20s on each recording by an expert doctor using the scale described in table I [11]. As only drowsiness visual signs are studied here, the five expert decision levels are converted into a binary decision by considering as "drowsy" any decision superior or equal to 2. Otherwise, the driver is considered as "awake". This represents 6461 samples classified as "awake" and 1096 as "drowsy". Two thirds of this database is randomly selected. It is used to study the relevancy of the selected features. This represents 4384 "awake" samples and 728 "drowsy" samples which is large enough to extract information about the different features by data-mining. The last third (2077 "awake" samples and 368 "drowsy" samples) is used to validate the drowsiness detection system presented in II-D.

C. Feature selection

The purpose of this section is to study the relevancy of each feature mentioned in section II-A with respect to drowsiness. The mean value of each feature is computed every second on a sliding window. Three sliding windows' lengths (20s, 40s and 60s) have been studied. The drowsiness detection is processed on each feature by using a threshold. Two ways of choosing the threshold are discussed here. The first one is to use a

Objective sleepiness score	EEG activity in the band [4-12]Hz	Blinks and eye movements
0	negligible	Normal
1	less than 5s	Normal
2	less than 5s	Slow
3	less than 10s	Slow
4	more than 10s	Slow

TABLE I
OBJECTIVE SLEEPINESS CRITERIA

fixed common threshold s_F for all drivers. The second one is to compute the threshold according to equation (1) for each driver. In this equation, m is the mean value of the studied feature, σ its variance and k a multiplying constant. The values m and σ are learnt at the beginning of the recording when the driver is considered as awake. The threshold is noted s_{1min} when it is learnt on the first minute and s_{5min} when it is learnt on the first five minutes. This method has been inspired by the work of Svensson [15] on EOG.

$$s = m + k \cdot \sigma \quad (1)$$

This method provides a binary decision [awake; drowsy] every second while the database is expertized by time slices of 20s. To compare the results obtained with expert's decision, the following technique is used. Each 20s slice classified by the expert is directly compared to the system decision: if during the 20s interval, the system classifies at least 10s as "drowsy", then the decision for the slice is "drowsy". Else it is "awake".

The results are compared using a Receiver Operating Characteristic (ROC) curves, plotting true positive rate (TP_{rate}) in function of the false positive rate (FP_{rate}). The TP_{rate} is calculated as the ratio between the number of "drowsy" samples correctly classified by our system and the number of "drowsy" expertized samples. The FP_{rate} is computed as the ratio between the number of false "drowsy" detection on the number of "awake" expertized samples. The purpose is to have the highest TP_{rate} with the lowest FP_{rate} .

Fig. 2 shows the results for each feature and different thresholds. In this figure, the "square" markers correspond to the results obtained with a fixed threshold s_F , the "circle" markers to those obtained with a threshold s_{1min} learnt during the first minute and the "triangle" markers to those obtained with a threshold s_{5min} learnt during the first five minutes. These results are obtained by varying the different threshold values in a coherent range for each feature. The results obtained for D_{50} , P_{80} , F and A/PCV only are shown. Indeed, the E_C feature provides the worst results ($TP_{rate}=69,3\%$ and $FP_{rate}=32,3\%$ with an optimal threshold). Moreover, the results have shown that the best results are obtained with the 20s sliding window. So, only the results obtained with a sliding window of 20s are presented here.

It is obvious from Fig. 2 that results are better when a fixed threshold is used for all the features. This means that the first minutes of the recording provides a poor estimation

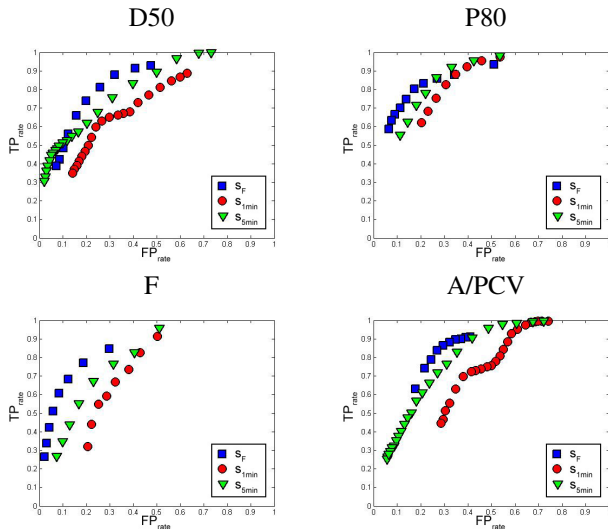


Fig. 2. Results obtained with the different features

Feature	Results	Threshold
D50	$TP_{rate} = 81.4\%$ $FP_{rate} = 24.9\%$	$s_F = 130ms$
P80	$TP_{rate} = 84,2\%$ $FP_{rate} = 22,3\%$	$s_F = 0.75\%$
F	$TP_{rate} = 78.3\%$ $FP_{rate} = 18,6\%$	$s_F = 0,1Hz$
A/PCV	$TP_{rate} = 78.5\%$ $FP_{rate} = 25.3\%$	$s_F = 17.25$

TABLE II

BEST RESULTS AND THRESHOLD OBTAINED FOR EACH FEATURE

of the “awake” state for the blinking features. Results obtained with s_{5min} are also much better for all the features than those obtained with s_{1min} which means that the first minute of video recording is not long enough of the driver ocular behaviour to be correctly assessed. Nevertheless, results obtained using s_{5min} are lower than those obtained using s_F . In the following sections, only fixed thresholds will be considered.

The best results as well as the best threshold s_F for each feature are presented in table II. This table shows that the best feature to detect drowsiness is P80 ($TP_{rate} = 83,2\%$ and $FP_{rate} = 21,3\%$). This is in concordance with the literature [10], [5]. Moreover, D50, P80, F and A/PCV only will be used to design the drowsiness detection system as they all provides good results.

D. Drowsiness detection using fuzzy logic

The method proposed to detect drowsiness use fuzzy logic to merge the four features considered. The first step of the method is to transform each feature into a fuzzy variable whose value is between 0 and 1. Each fuzzy variable is defined as the membership degree to the “drowsy” state: the closer to 1 the value of the variable, the higher the probability of being drowsy. The membership function is defined by equation (2). Coefficients a and b are respectively defined by equations (3) and (4) depending on the threshold s_F , selected in II-C for

each feature, in order to standardize the membership function. The values -0.25 and $+0.25$ have been chosen to centre the membership function on the considered threshold s_F .

$$f_{a,b}(x) = \begin{cases} 0 & \text{if } x \leq a \\ \frac{x-a}{b-a} & \text{if } a \leq x \leq b \\ 1 & \text{if } x \geq b \end{cases} \quad (2)$$

$$a = s_F - 0,25 \cdot s_F \quad (3)$$

$$b = s_F + 0,25 \cdot s_F \quad (4)$$

The membership degree of each feature D50, P80, F and A/PCV is respectively noted μ_{D50} , μ_{P80} , μ_F and $\mu_{A/PCV}$. Then, the fuzzy variables are merged at each time i using the equation (5). The fusion is computed at each instant as the average of the different membership degrees.

$$\mu_{fusion}(i) = \frac{1}{4} \sum_j \mu_j(i), \quad j \in \{D50, P80, F, A/PCV\} \quad (5)$$

The result is noted μ_{fusion} and its value is between 0 and 1. The relevancy of the detection depends on the value of μ_{fusion} : the closer to 1 μ_{fusion} , the more relevant the detection. The value 1 corresponds to the “drowsy” state and 0 to the “awake” state. To make the decision, the driver is considered as “drowsy” if μ_{fusion} is greater than 0,5. Otherwise, he is considered as “awake”. Actually, this function provides a compromise between the different features.

III. RESULTS AND DISCUSSION

The decisions of the drowsiness detection system are made every second. The system decision is directly compared to the expert’s one by 20s slices, in the same way as in the section II-C. The method has been tested on the third of the database which has not been used to choose the different thresholds. The different results are shown in Fig. 3. The “square” markers correspond to results obtained with the data fusion method, the “triangle” markers to those obtained with D50 only, the “star” markers to those obtained with P80 only, the “hexagram” markers to those obtained with F only and the “diamond” markers to those obtained with A/PCV only. Table III presents the TP_{rate} and the FP_{rate} related values.

First, results obtained using each feature alone proves that the thresholds used are robust. The results obtained on the last third of the database are indeed very similar to those found during the study of the different features on the two other thirds of the database. So, the method is independent of the driver since the same thresholds give good results on 20 different drivers. Moreover, results obtained with the fusion method are better than those obtained with only one feature. The fuzzy fusion gives $TP_{rate} = 81,7\%$ which is as good as the average TP_{rate} found with the features processed independently (around 80%). Indeed, good detections on the different features are corresponding which explains that the results of the fusion provides a TP_{rate} equivalent as those found for each feature independently. Furthermore, the $FP_{rate} = 13,4\%$

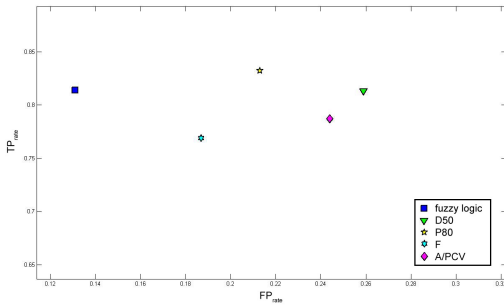


Fig. 3. Results of the drowsiness detection

	Fuzzy logic	D50	P80	F	A/PCV
TP_{rate}	81.7%	80.7%	82.8%	77.0%	78.6%
FP_{rate}	13.4%	25.1%	21.4%	19.3%	23.9%

TABLE III
RESULTS OF THE DROWSINESS DETECTION

obtained with the fusion is obviously better than those found for each feature. This means that the false alarms obtained with the different features are not occurring at the same time. Thus, when the different features are merged, the false alarms produced by one of the feature are eased by the others which explains the decrease of the FP_{rate} . So, the fusion of the different features make the detection more robust: it decreases the number of false alarms without changing the number of correct detections. These results are confirmed by a McNemar test comparing the performances obtained by fusion to those obtained with the features alone.

The results presented are also good compared to those found in the literature. Johns [9] obtained 75% of correct detections and 30% of false alarms while detecting drowsiness with the amplitude-velocity ratio only. This is coherent with the results found for the A/PCV feature ($TP_{rate} = 78.6\%$ and $FP_{rate} = 23.9\%$). A fuzzy expert system merging blink duration, blink frequency and blink interval to estimate drowsiness has been proposed by Damousis et al. [4]. They obtained 90% of correct detections with a fuzzy expert system which is slightly better than our results. Nevertheless they also had 30% of false alarms which is worse than the FP_{rate} obtained by the proposed method. This means that the features used in these paper for the drowsiness detection are more relevant. However, it is necessary to keep some reserves since these results have not been obtained on the same database. Nevertheless, the database used in this paper is large enough (60 hours of driving from 20 different drivers) to make a good assessment of the performances of our system.

IV. CONCLUSIONS AND FUTURE WORKS

An universal algorithm for the automatic detection of drivers drowsiness based on visual signs has been presented here. This algorithm is based on the fuzzy fusion of several blinking

features that can be extracted from a high frame rate video. The different features used in the system has been selected by data-mining on a consistent database. The algorithm has been tested on a large number of drivers and does not need to be tuned. Its validation on EOG signals gives good results with 81,7% of correct detections and only 13,4% of false alarms. These results should be equivalent when using a high frame rate camera since the features used here have been cautiously chosen and can be extracted from EOG and high frame rate video with the same accuracy.

The next step of this work is the achievement of an acquisition campaign of drowsy data with a high frame rate camera so as to validate the whole method proposed on video data. Once validated, this video based drowsiness detection system should be added to a EEG based drowsiness detection system ([13]) to merge the decisions to obtain a highly reliable automatic drowsiness detector.

ACKNOWLEDGEMENT

The authors are grateful to the Centre d'Études de Physiologie Appliquée (CEPA) in Strasbourg (FR) for providing the data and their help.

REFERENCES

- [1] A. Benoit and A. Caplier. Motion estimator inspired from biological model for head motion interpretation. In *WIAMIS*, 2005.
- [2] S. Boverie, A. Giralt, and J.-M. L. Quellec. Diagnostic fusion for in vehicle driver vigilance assessment. In *Proc. of the 17th IFAC World Congress*, Seoul, South Korea, 2008.
- [3] P. Caffier, U. Erdmann, and P. Ullsperger. Experimental evaluation of eyeblink parameters as a drowsiness measure. *European Journal of Applied Physiology*, 89:319–325, 2003.
- [4] I. Damousis, D. Tzovaras, and M. Strintzis. A fuzzy expert system for the early warning of accidents due to driver hypo-vigilance. *Personal and Ubiquitous Computing*, 13(1):43–49, 2009.
- [5] N. Galley, R. Schleicher, and L. Galley. Blink parameter as indicators of driver's sleepiness - possibilities and limitations. *Vision in Vehicles*, 10:189–196, 2004.
- [6] M. Gillberg, G. Kecklund, and T. Akerstedt. Sleepiness and performance of professional drivers in a truck simulator - comparisons between day and night driving. *Journal of Sleep Research*, 5:12–15, 1996.
- [7] W.-B. Horng, C.-Y. Chen, Y. Chang, and C.-H. Fan. Driver fatigue detection based on eye tracking and dynamk, template matching. In *Proc. of the IEEE Int. Conf. on Networking, Sensing and Control*, Taipei, Taiwan, 2004.
- [8] Q. Ji and X. Yang. Real-time eye, gaze, and face pose tracking for monitoring driver vigilance. *Real-Time Imaging*, 8:357–377, 2002.
- [9] M. Johns. The amplitude-velocity ratio of blinks: a new method for monitoring drowsiness. *Sleep*, 26:A51–A56, 2003.
- [10] R. Knipping. PERCLOS: a valid psychophysiological measure of alertness as assessed by psychomotor vigilance. Technical report, Federal Highway Administration (FHWA), 1998.
- [11] A. Muzet, T. Pbayle, J. Langrognet, and S. Otmani. AWAKE pilot study no.2: Testing steering grip sensor measures. Technical Report IST-2000-28062, CEPA, 2003.
- [12] A. Picot, A. Caplier, and S. Charbonnier. Comparison between EOG and high frame rate camera for drowsiness detection. In *Proc. of the IEEE Workshop on Applications of Computer Vision*, Snowbird, USA, 2009, to be published.
- [13] A. Picot, S. Charbonnier, and A. Caplier. *Recent Advances in Biomedical Engineering*, chapter Monitoring drowsiness on-line using a single encephalographic channel. IN-TECH, 2009, to be published.
- [14] G. Renner and S. Mehning. Lane departure and drowsiness - two major accident causes - one safety system. Technical report, Transport Research Laboratory, 1997.
- [15] U. Svensson. Blink behaviour based drowsiness detection method development and validation. Master's thesis, Linkping University, Dept. Biomedical Engineering, 2004.