FAST MICROSLEEP AND YAWNING DETECTIONS TO ASSESS DRIVER'S VIGILANCE LEVEL

Nawal Alioua[1], Aouatif Amine[1,2], Mohammed Rziza[1], & Driss Aboutajdine[1]
[1] LRIT, Unité Associée au CNRST, Mohammed V-Agdal University, Rabat, Morocco
[2] ENSA, Ibn Tofail University, Kenitra, Morocco
Email: nawal.alioua@yahoo.fr

Summary: Driver hypovigilance, often caused by fatigue and/or drowsiness, receives increasing attention in the last years; especially after it became evident that hypovigilance is a one of the major factor causing traffic accidents. Monitoring and detecting driver hypovigilance could contribute significantly to improve road traffic safety. This paper proposes fast methods to identify drowsiness and fatigue using respectively microsleep and yawning detections. In this study, the proposed scheme begins by a face detection using local Successive Mean Quantization Transform (SMQT) features and split up Sparse Network of Winnows (SNoW) classifier. After performing face detection, the novel approach for eye/mouth detection, based on Circular Hough Transform (CHT), is applied on eyes and mouth extracted regions. Our proposed methods works in real-time and yield a high detection rates whether for drowsiness or fatigue detections.

INTRODUCTION

The increasing number of traffic accidents due to a diminished driver's vigilance level resulting from sleep deprivation has become a serious problem for society. Statistics show that about 20% of all accidents are due to hypovigilante drivers (Bergasa et al., 2004). In fact, the hypovigilance reduces driver's capacity to react, judge and analyze information and it is often caused by fatigue and/or drowsiness. However, drowsiness and fatigue are different. Drowsiness refers to the inability to stay awake while fatigue refers to a cumulative process producing difficulty to pay attention. Therefore, it is very important to monitor driver's vigilance level and issue an alarm when he is not paying enough attention to the road. Several visual behaviors can be used to characterize hypovigilante driver. These visual cues include eye state (Papanikolopoulos and Eriksson, 2001), mouth motion (Wang and Shi, 2005), and head orientation (Smith et al., 2000). It is also possible to combine them in order to form a composite index that can robustly and consistently describe vigilance level. The eye state is often assumed to give a good indication of drowsiness level characterized by microsleep which is short period (2-6 s) during which driver rapidly closes the eyes and sleep. Driver's eye detection methods based on computer vision use camera to obtain facial information, extract the eye, and decide their openness degree. Many researchers use Percent of Eyelid Closure (PERCLOS) as a drowsiness indicator (Grace et al., 2001; Roman et al., 2001). Others researchers use the presence of iris to predict if the eye is open (D'Orazio et al., 2004). To extract information about driver fatigue level, some authors consider that a high yawning frequency is a strong fatigue indicator. Existing works use generally one criterion to determine driver's vigilance level. In this work, we propose a system that uses two criteria. The first criterion detects drowsiness by extracting microsleeps using the presence of the iris. The second one detects fatigue using yawning deduced from large mouth opening. Both methods are based on Circular Hough Transform (CHT) (Duda and Hart, 1972). The

organization of this paper is as follows. Section 2 explains the different steps of the proposed system. In Section 3, experimental results are exposed. Finally, discussion is presented.

METHOD

The aim of this work is to develop eye and mouth analysis algorithms for drowsiness and fatigue assessment respectively. The global system performs some steps before determining vigilance level. Firstly, the face is extracted from video frames. Then, localization of both eyes and mouth is performed. Finally, we apply the proposed methods based on CHT.

Face detection

The face is extracted from video frame to reduce search region and computational cost. We have used an existing method for face detection using the illumination insensitive features gained from the local Successive Mean Quantization Transform (SMQT) features and split up Sparse Network of Winnows (SNoW) classifier wich achive rapid detection. (Nilsson et al., 2007)

Eye and mouth localization

The current step is introduced for the same purpose as in the previous one. Also, it permits eliminating the possibility of confusing facial features. To isolate the eyes, we use gradient image to highlight the edge. After that, horizontal projection is computed on gradient image to detect upper and lower eye region boundaries. Finally, we apply vertical projection on resulting image to obtain right and left face limits and separate the eyes. For mouth localization, the right and left face limits are used to detremine right and left mouth boundaries. At last, horizontal projection on resulting gradient image is computed to find upper and lower mouth limits.

Driver's drowsiness analysis

This step is crucial since it detects microsleep in real time and issues immediately an alarm to avert the drowsy driver. To detect microsleep, we apply CHT on eye region images to identify iris. The eye is considered open if an iris is found. Since CHT extracts circles from edge images, the obtained results depend on applied edge detector. Some classic edge detectors such as Sobel, Prewitt, Roberts, Laplacian of Gaussian (LoG) and Canny was tested for extracting the iris edge on eye region images. Unfortunately, the obtained edges did not provide the desired form, i.e. a kind of circular form referring to the iris. In order to solve this problem, we propose a new iris edge detector more suitable to the eye's morphology.

Iris Edge detector. Our iris edge detector respects the eye's morphology. If we observe an open eye, we see three main components. The *pupil* which is the little black circle surrounded by the *iris*, the circle distinguished by the eye color and the white outer area which is the *sclera*. This structure enables extracting the iris edge from significant intensity variations between iris and sclera. Our iris edge detector considers only pixels x with grayscale intensity lower than an optimal threshold noted th_{edge}, which must be chosen to handle only with pixels appertaining to iris. For each x, a neighborhood containing n pixels at left and right of x is specified. The difference between x and its n right and left neighbors is then computed.
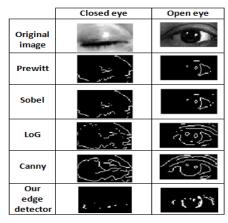
Figure 1. Iris edge detection

Left (*resp*. Right) edge: if n or $n-1$ left (*resp*. right) neighbors of x provide a difference with x higher than the high threshold noted th_{high} and also if n or $n-1$ right (*resp*. left) neighbors of x provide a difference with x lower than the low threshold th_{low}, we deduct that x is a left (*resp*. right) edge pixel of the iris and we put it at 1. In the case where the pixel appertains to the left edge, its left (*resp*. right) neighbors pixel's intensity is very higher (*resp*. similar). Inversely, when the pixel appertains to the right edge, the right (*resp*. left) neighbors pixel's intensity is very different (*resp*. similar). The th_{high} should distinguish difference between iris and sclera and th_{low} should respect similarity between iris pixels. Figure 1 shows examples of iris edge detection obtained by the proposed method compared to classic detectors. Note that these lasts cannot provide good iris edge since they detect components having circular form in closed eye while our detector identify such components only in iris edge.

Drowsiness detection based on eye state analysis using iris detection. Once the iris edge is found, we apply the CHT to obtain the iris radius. In the following, we present the CHT algorithm steps. Three edge pixels are randomly chosen at each iteration. If these pixels are not collinear and if the distance between each two pixels coordinates is higher than a fixed threshold th_{dist_px} , we compute the radius and center coordinates of the candidate circle defined by these three pixels. If the parameters of candidate circle are between two specific thresholds, they are assigned to an accumulator. After that, we compute the distance between the center and all edge pixels. If a distance is lower than a threshold th_{dist_ctr}, we increment the counter of the circle pixels. If this counter is higher than a threshold $th_{counter}$, we consider that the candidate circle can represents the iris and we keep the other pixels as a new edge. We repeat the previous steps until the edge contains few pixels or maximal iteration is reached. We conserve the circle having the highest radius after the end of the algorithm. Since driver's drowsiness is characterized by microsleep, we must find sleep intervals of at least 2 seconds. To reduce computational time, we beging by analyzing the left eye. If it is open, we proceed to fatigue detection. Otherwise, we analyze the right eye. If it is open, we switch to fatigue detection. Else, we increment the consecutive closed eye counter. We issue an alarm to avert the drowsy driver if the eyes remain closed for a certain period of time.

Driver's fatigue analysis

Driver's fatigue is characterized by a high yawning frequency. Since drowsiness often occurs after fatigue, we add this step to prevent the driver before microsleep. For this,we analyze the mouth state in order to find yawning that 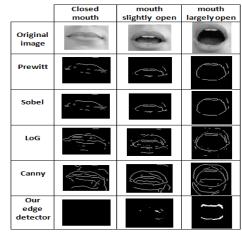is a large mouth opening which lasts from 2 to 10 seconds. To detect yawning, we apply CHT on mouth edge images. As we have already seen in (*C.1*), the well known edge detectors do not provide the desired results.

Figure 2. Large mouth opening edge detection

367

Then, we also propose an edge detector similar to the previous one that respects the large mouth opening structure.

Large mouth opening edge detector. The proposed edge detector is based on the mouth structure. If the mouth is widely open, a large dark area having a circular form appears. If we replace in (*C.1)* the term *"left"* by the term *"top"* and *"right"* by "*down*", we obtain the large mouth opening edge detector. In the other terms, the pixels appertaining to the down edge have very different down neighbours and similar top neighbours and vice versa for the top edge pixels. Figure 2 shows large mouth opening edge detection obtained by the proposed method compared to results of several classic edge detectors.

Fatigue detection based on mouth state analysis using yawning detection. The same algorithm as the one presented in (*C.2)* is used to detect large mouth opening based on CHT. The considered driver's fatigue indicator is yawning which is an interval of large mouth opening of at least 2 seconds. To reduce computational time, we analyze the mouth state if an eye is open. If the mouth is widely open, the counter of consecutive large mouth opening is incremented. When this counter exceeds the specified threshold, the counter of yawning is incremented. Once this second counter is higher than a fixed threshold, the system indicated that the person was suffering from fatigue, issuing a warning signal. We chose to analyze mouth state if an eye is open also because a driver that yawn having closed eyes does not pay attention and he is considered drowsy. Figure 3 illustrates our system for fatigue and drowsiness detection.
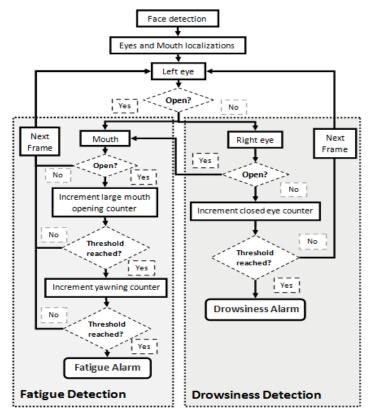

Figure 3. Driver's fatigue and drowsiness respectively detection system

RESULTS

We conduct several tests on real video sequences of different subjects and various lighting conditions to validate eye and mouth state analysis methods. All sequences are taken in a car using a low cost webcam at 30 frames per second (fps) providing images of resolution 640x480 and connected to a laptop by USB port, Figure 4 illustrates system equipments used for tests. Firstly, we evaluate eye state analysis method by analyzing all frames of the sequences to identify the presence of irises on cropped eye images. The second experiment is to evaluate mouth state analysis method in cropped mouth images. Subsequently, we introduced drowsiness

and fatigue analysis which led us to reduce the considered number of fps from 30 to 2 to meet the real-time constraints. In this experiment, automatic detection of face, eyes and mouth has been integrated but not evaluated in this work. The main purpose of the integration of face, eyes and mouth localizations is to take them into account in assessing the runtime system. The final experiment provides a comparison between our driver's drowsiness and fatigue detection system and other existing systems. All experiments are done on laptop having Intel Core 2 Duo Processor. Results are presented as statistical measures such *True Positive (TP)*, *False Positive (FP)*, *True Negative (TN), False Negative (FN)* and total number of samples *T*. We also use Correct Classification Rate and kappa Statistic (Fleissa et al., 1969). Note that Kappa Statistic noted κ is a measure of non-random agreement degree between observers or measurements of the same categorical variable. Table 1 is used to interpret kappa statistic according to its values.

Figure 1. Acquisition system

Table 1. Kappa statistic interpretation

Kappa Statistic	Interpretation
> 0.81	Almost perfect agreement
> 0.61 and < 0.8	Strong agreement
> 0.2 and <0.6	Moderate agreement
> 0.0 and < 0.2	Poor agreement
< 0	Disagreement

Evaluation of eye state analysis and mouth state analysis

The experiments concerning eye and mouth state analysis were made on real video sequences (V.) of different subjects in various lighting conditions. Table 2 shows statistical measures for each sequence of these experiments.

Table 2. Statistical measures of eye state analysis (left) and mouth state analysis (right)
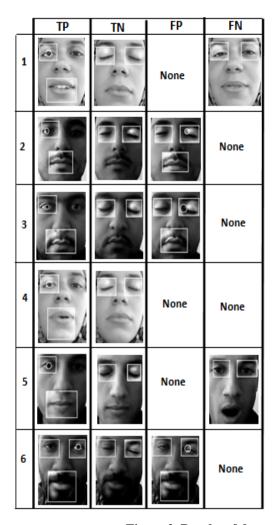
V.	TP	TN	FP	FN	T	CCR	κ
1	3336	233	17	14	3600	0.99	0.93
2	1760	36	4	0	1800	0.99	0.94
3	1772	18	2	8	1800	0.99	0.79
4	1482	273	9	38	1802	0.97	0.90
5	1762	24	2	14	1802	0.99	0.74
6	380	237	1	0	618	0.99	0.99
7	1636	135	14	17	1802	0.98	0.89
					Avr	0.99	0.88

V.	TP	TN	FP	FN	T	CCR	κ
1	154	321	7	0	482	0.98	0.98
2	213	649	3	0	865	0.99	0.99
3	137	687	15	0	839	0.98	0.94
					Avr	0.98	0.97

We observe, from Table 2, that the average (Avr) of CCR is 99% and the average of κ is 88% for eye state analysis. Also, the average of CCR is 98% and the average of κ is 97% for mouth state analysis. According to Table 1, both methods provide an almost perfect agreement between observers. In other words, the real class of the samples usually matches the estimated class proposed by the method. So, we deduct that our methods are very strong.

Evaluation of fatigue and drowsiness detections

We evaluate both eye and mouth state analysis on 6 video sequences captured in a car at 2 fps. Face, mouth and eyes are located automatically. Figure 5 shows respectively examples of iris and large mouth opening detections of *TP*, *FP*, *TN* and *FN* for the test sequences. Table 3 presents statistical measures. The first column of each statistic refers to the measure of iris detection and the second one refers to the measure of large mouth opening detection. Two columns are added representing video duration noted *vid.D* and execution time of the whole system noted *exec.T*. According to Table 3, the average of CCR is 97% and the average of kappa statistic is 94%. From Table 1, the combination of the two methods procures an almost perfect agreement between the observers. This means that both eye and mouth state analysis methods permit assignation of the right classes in the most cases. After comparing the two lasts columns, we deduce that the system respect the real time constraints because execution time and video duration are almost the same. Thus we deduct that the proposed system can be used to have an excellent and real-time estimation of driver's vigilance level.
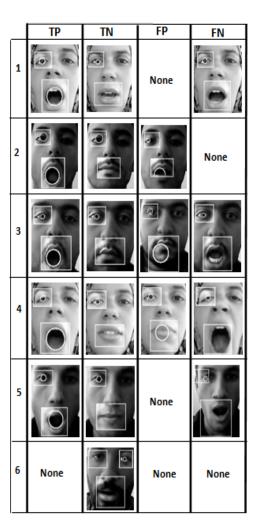

Figure 2. Results of drowsiness and fatigue detections

Table 3. Statistical measures of drowsiness and fatigue detections respectively

V.	TP		TN		FP		FN		T		CCR		κ		Vid.D	Exec.T
1	99	30	23	68	0	0	1	1	124	99	0.98	0.99	0.94	0.97	61	65
2	82	31	11	52	2	1	0	0	95	84	0.98	0.99	0.90	0.97	47	50
3	68	8	19	59	1	1	0	1	89	69	0.98	0.97	0.94	0.87	41	44
4	109	37	36	68	0	2	0	2	145	109	1.00	0.96	1.00	0.92	72	77
5	57	18	18	39	0	0	2	0	77	57	0.97	1.00	0.93	1.00	40	43
6	73	0	14	90	3	0	0	0	90	0	0.96	1.00	0.88	1.00	44	48
									Avr.		0.97		0.94			

The last experiment exposes a comparison between our system and other existing systems of driver's hypovigilance detection. The system depicted in (Tripathi and Rath, 2009) is also based on CHT and uses 173 images of ORL database for evaluation, this system provides success accuracy rate of 90.3%. The second system presented in (Hrishikesh et al., 2009) uses 70 images taken with an infra-red camera for tests and obtains a success rate of 90%. While the third system (Zhang et al., 2008), which is based on adaptive learning method to detect driver's eyes, uses 13000 real frames for tests and find an accuracy of about 95%. We deduct that the proposed system provides a very high success rate comparing to the mentioned existing systems.

DISCUSSION

This paper presents a new and fast approach based on CHT for microsleep and yawning detections to assess driver's vigilance level. The whole driver's drowsiness and fatigue detection system uses three steps: face detection using local SMQT features and split up SNoW classifier, eye/mouth localization applied on gradient image and eye/mouth state analysis to identify drowsy/tired driver. In the last step, we apply CHT on our edge detectors in order to find iris and large mouth opening. With 97% accuracy of CCR and rate of 94% of kappa statistic, it is obvious that our driver's hypovigilance detection system is robust compared to some existing systems. Note that, the eye (*resp.* mouth) state analysis method provides a detection rate of 99% (*resp.* 98%) and kappa statistic value attaining 88% (*resp.* 97%). As future works, we plan on one hand, to use multiple cameras in order to detect irises in various head orientations. On the other hand, we will introduce gaze tracking to detect driver's inattention, and to give us idea of where the driver is looking: at the dash display, at the roadside signs, or road ahead.

REFERENCES

Bergasa,L.,Nuevo,J.,Sotelo,M.,& Vazquez,M.(2004).Real-time system for monitoring driver vigilance. *IEEE Intelligent Vehicle Symposium*,78–83.

D'Orazio,T.,Leo,M.,Spagnolo,P.,& Guaragnella,C.(2004).A neural system for eye detection in a driver vigilance application.*The 7th International IEEE Conference on Intelligent Transportation Systems*, 320–325.

Duda,R. and Hart,P.(1972).Use of the hough transformation to detect lines and curves in picture. In *Commun.ACM*, 11–15.

Fleissa,J.,Cohen, J.,& Everitt,B.(1969).Large sample standard errors of kappa and weighted kappa.*Psychological Bulletin*, 323–327.

Grace,R.,Byrne,V.,Bierman,D.,Legrand,J.,Gricourt,D.,Davis, B.,Staszewski,J.,& Carnahan,B. (2001).A drowsy driver detection system for heavy vehicles. *Proceedings of the 17th Digital Avionics Systems Conference*,(2), I36/1– I36/8.

Hrishikesh,B.,Mahajan,S.,Bhagwat,A.,Badiger,T.,Bhutkar,D.,Dhabe,S., & Manikrao,L. (2009). Design of drodeasys (drowsy detection and alarming system).*Advances in computational algorithms and data analysis*, 75–79.

Papanikolopoulos,N. and Eriksson,M.(2001). Driver fatigue: a vision-based approach to automatic diagnosis.*Transportation Research Part C: Emerging Technologies*, 399–413.

Roman,B.,Pavel,S.,Miroslav,P.,Petr,V., & Lubomr,P.(2001). Fatigue indicators of drowsy drivers based on analysis of physiological signals.*Medical Data Analysis*, 62–68.

Smith,P.,Shah,M.,& da Vitoria Lobo,N.(2000).Monitoring head/eye motion for driver alertness with one camera. *ICPR00*, 636–642.

Nilsson,M.,Nordberg,J.,Claesson,I.(2007). Face Detection using Local SMQT Features and Split Up SNoW Classifier.*ICASSP 2007*.

Tripathi,D. and Rath,N.(2009). A novel approach to solve drowsy driver problem by using eye-localization technique using CHT. *International Journal of Recent Trends in Engineering*.

Wang,T. and Shi,P.(2005). Yawning detection for determining driver drowsiness.*IEEE Int. Workshop on VLSI Design and Video Tech.*, 373–376.

Zhang,G.,Cheng,B.,Feng,R.,& Zhang,X.(2008). A real- time adaptive learning method for driver eye detection.*Digital Image Computing: Techniques and Applications*, 300–304.