32nd Annual International Conference of the IEEE EMBS
Buenos Aires, Argentina, August 31 - September 4, 2010

Automated Video-based Measurement of Eye Closure for Detecting Behavioral Microsleep

Amol M. Malla, Paul R. Davidson, Philip J. Bones, *Senior Member, IEEE*, Richard Green and Richard D. Jones, *Senior Member, IEEE*

Abstract— **A device capable of continuously monitoring an individual's levels of alertness in real-time is highly desirable for preventing drowsiness and microsleep related accidents. This paper presents a development of non-intrusive and light-insensitive video-based system that uses computer-vision methods to measure facial metric for identifying visible facial signs of drowsiness and behavioral microsleep. The developed system uses a remotely placed camera with a near-infrared illumination to acquire the video. The computer-vision methods are then applied to sequentially localize face, eyes, and eyelids positions to measure ratio of eye closure. The system was evaluated in frontal images of nine subjects with varying facial structures and exhibiting several ratio of eye closure and eye gaze under fully dark and ambient lighting conditions. The preliminary results showed promising results with sufficient accuracy to distinguish between fully closed, half closed, and fully open eyes.**

I. Introduction

BEHAVIOURAL microsleeps (BMs) are brief episodes of behavioral signs of sleep in which an individual unintentionally stops responding to the task they are performing [1]. Lapse of responsiveness at the wrong moment in high risk jobs such as commercial truck and bus drivers, pilots, air-traffic controllers, and medical house-staffs can lead to disastrous consequences, including multiple fatalities [2][3]. A non-intrusive device capable of identifying signs of drowsiness and onset of BM will provide a valuable early warning system to prevent accidents.

Some of the visible facial signs of drowsiness and BM are prolonged eyelid closure [1], change in eye-blink parameters such as reduced spontaneous-blink rate, increased blink duration and amplitude, slow eye movements, reduced facial tone, and head nods [1][4][5][6][7]. One of the most reliable and valid visual measure of a driver's alertness considered by the American Federal Highway Administration is PERCLOS PERCLOS [7][8] which uses measure of percentage of eye closure over a 1 min window.

To monitor the facial signs of drowsiness and BMs, we devised a low-cost camera and near-infrared (NIR) illumination-based system to non-intrusively monitor the subject under any lighting condition. We then employed signal processing and computer vision techniques to detect the position of face, eyes, and eyelids within a frame of the video. The eyelids positions were then used to calculate the ratio of eyelid closure, which in turn can be used as a metric for drowsiness and BM detection.

II. Data collection and annotation

A. Data collection

Videos of nine subjects imitating signs of drowsiness and microsleep were recorded under dark and ambient recording conditions. There were 5 males and 4 females subjects, out of which 5 of them were Asian and 4 were Europeans. Also 4 of the subjects wore glasses.

Videos were collected using QuickCam 4000 webcam whose IR block filter was removed. The videos with 640 x 480 pixel image resolution were collected at 30 fps. A 7 cm diameter ring of 6 infra-red emitting diodes (IRED) were placed concentric to the camera to illuminate the scene. During the recording, the subjects were asked to perform various actions, including slow eyelids closure imitating drowsiness, gaze in different directions and varying degrees with and without head movement, and also droopy head-nods followed by a quick recovery of head posture to simulate microsleep.

B. Annotation

From the videos of each subject, set of 66 frontal face frames (33 frames for dark and ambient lighting conditions each) were selected for manual annotation. For each lighting conditions, 3 frames for each of 5 levels of eyelid closers and 6 different gaze directions were selected. Fig. 1 shows an example of manual annotation carried out in the 594 frames. Each frame were manually annotated with the coordinates of nose, eyebrows, center and radius of iris, area of visible sclera, apex positions of eyelids, and corner of the eyes to form the reference data.

Manuscript received 23 April 2010. This work was supported by Christchurch Neurotechnology Research Program, Christchurch, New Zealand.

All authors are members of the Van der Veer Institute for Parkinson's and Brain Research, Christchurch, New Zealand.

Amol M. Malla is with Electrical and Computer Engineering, University of Canterbury, Christchurch, New Zealand e-mail: amolmalla@hotmail.com.

R. D. Jones is with Medical Physics and Bioengineering, Christchurch Hospital, and Medicine, University of Otago, and Electrical and Computer Engineering, University of Canterbury, Christchurch, New Zealand (phone: 643-378-6077; fax: 643-378-6080; e-mail: richard.jones@vanderveer.org.nz)

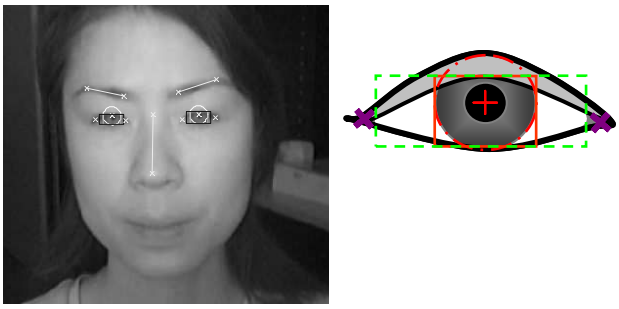

Fig. 1. Annotation of the eye and other facial features. In total 594 frames comprising of 66 frames for 9 subjects were annotated.

III. AUTOMATED EYE CLOSURE MEASUREMENT METHOD

A problem with a remote camera-based system compared to head mounted system is that the position of features of interest like the face, eyes, and eyelids must first be isolated from rest of the image background. We took an approach of narrowing down the search region of interest (ROI) within an image by sequentially detecting the face, eyes, and eyelids within an image and then using the eyelids position to measure the degree of eye closure. An advantage of this approach is that the ROI substantially decreases with each step, which speeds-up computational time. However, a major disadvantage is that each step relies on the accuracy of preceding facial feature detection algorithm.

A. *Face ROI detection and stabilization*

To localize a face ROI (fROI) within an image, we used the Haar-object detection algorithm [9] and a face classifier implemented by [10] in OpenCV libraries [11]. Fig. 2 shows an example of a fROI detected with the Haar-face detection algorithm.

Fig. 2. An example of a face region of interest (fROI) detection with Haar-face detection algorithm implemented in OpenCV.

Although the Haar-face detection algorithm correctly enclosed a face within the localized fROI in majority of frames, there were fluctuations in position and size of fROI between consecutive frames in a given short sequence of video even when the subjects were relatively stationary.

Kalman filter [12] was used to improve the stability of the detected fROI. Kalman filter acts as the low-pass filter to stabilize the fROI and also correctly tracks the fROI when the Haar-face detection algorithm fails. Kalman filter stabilizing effect on the size parameter of fROI and also prediction of missing fROI in frames between frame-number 1112 and 1123 is shown in Fig. 3.

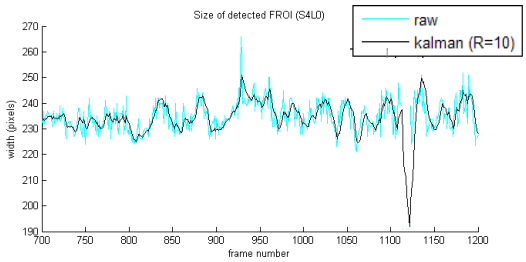

Fig. 3. Raw and Kalman filtered size parameter of fROI. Kalman filter stabilizes and also predicts missing fROI.

B. *Anthropomorphic estimation of eye region of interest*

The left and right eye regions of interest (eROI) were derived by scaling the fROI with anthropometric factors. The anthropometric factors define the position and size of the eROI as proportional constants relative to position and size of the fROI. The proportional constants were derived by enclosing the population distribution of manually annotated center of eyes (COE) in the reference database relative to corresponding fROI. Hence, an eROI represents a search ROI for COE in a given fROI. Fig. 4 shows examples of eROI estimation. The performance of eROI estimation was dependent on the precision of fROI detection. The detection of the fROI after the Kalman filter stabilization was reliable in majority of frontal face images. Therefore, the eROI was also localized to reliably include the COE in most frontal face images. In addition, the eROI was also correctly estimated in some images with slight head rotation as shown in Fig. 4.

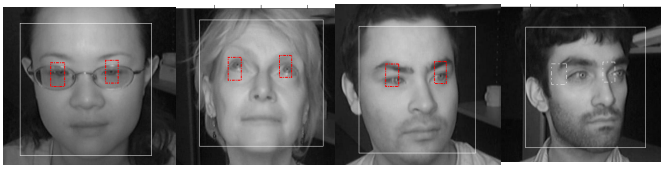

Fig. 4. Localization of eROI based on anthropometric scaling of fROI performed well in most frontal face images and also in images with small head rotation. Incorrect detection of fROI resulted in poor estimation of the eROI as shown in bottom right image.

Reliable estimation of eROIs in most frontal images is an acceptable result for application of microsleep detection as the subjects are most likely to be facing straight ahead during the event of drowsiness and microsleep. Also subjects with different head orientation are likely to be un-attentive on the task but awake.

C. *Centre of eye detection*

In addition to COE, an eROI can also include eyebrow. It was important to differentiate an eye from other facial features within an eROI to reliably detect the eyelid position. A template matching method was used to localize COE within an

within an eROI. An eye template image (T) with annotated COE at the center of image was formed by taking an average intensity of 378 sub-images of fully opened eyes in the reference database. The T image shown in Fig. 5 encodes the shape and contrast in intensity of eye and its surrounding.

The T was cross-correlated with the sub-images ($S_{X,\,Y}$) of same size derived by placing each pixel within the eROI at the center of sub-image. The correlation coefficient (r_{xy}) for each pixel in an eROI form a correlation matrix (CM) for the eROI. The r_{xy} provides a normalized comparison without bias due to changes in intensity levels of the two input images.

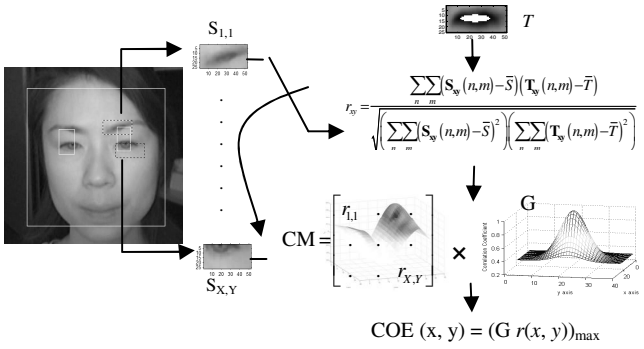

Fig. 5. Illustration of template matching method used for COE detection. The correlation matrix in this example is also shown as the 3D mesh plot.

Performance of COE detection algorithm was further improved by weighting the CM so that it emphasizes the most likely position of COE in an eROI. The population distribution of annotated COE with respect to the centre of the eROI in the reference database showed a Gaussian distribution. Hence, the CM was weighted with 2D circularly symmetric Gaussian function (G). The pixel with the maximum r_{xy} was selected as the COE.

Fig. 6 shows an example where the COE in right eROI is corrected by weighing the CM with G, where as previously correctly estimated COE in left eROI was not affected.

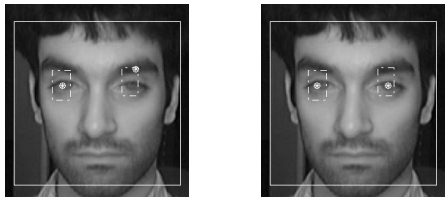

Fig. 6. Example of COE detection and its improvement by weighting the correlation matrix with 2D circularly symmetric Gaussian function.

The COE detection algorithm was applied to 924 eROIs of subjects in the reference database to evaluate its performance. The mean and standard deviation (SD) of median and 90th percentile magnitude error in the x and y coordinates of estimated COE in each subjects is shown in TABLE 1. The COE detection algorithm performed well in estimating the y-coordinate of COE, but the estimation of the x-coordinate was poor. This is likely due to better encoding of vertical change in grayscale intensity by the eye template.

TABLE 1
PERFORMANCE OF COE DETECTION ALGORITHM

	Mean ± SD magnitude error (pixels)	
COE	Median	90th percentile
x	3.7 ± 6.4	5.9 ± 1.7
y	1.7 ± 0.6	3.2 ± 0.9

Good estimation of y-coordinate of COE is critical for eye closure measurement application because it defines a vertical position to start the search of y-coordinates of upper and lower eyelids.

D. Eyelids detection

The vertical positions of apices of upper and lower eyelids (UEL_y and LEL_y) were detected by respectively searching for the change in average image intensity above and below the estimated COE. The change in image intensity from dark eye region to light skin region was captured by a simple vertical integral projection (VIP) method, where average image intensity of all the columns in each row in an eROI was calculated. As marked by solid lines in the VIP in Fig. 7, the steepest negative gradient above and the steepest positive gradient below the COE respectively aligns with the annotated UEL_y and LEL_y. Hence, the UEL_y and LEL_y in an eROI were detected by calculating the 1st derivative of VIP (VIP') and locating the local minima and local maxima on the respective side of the COE.

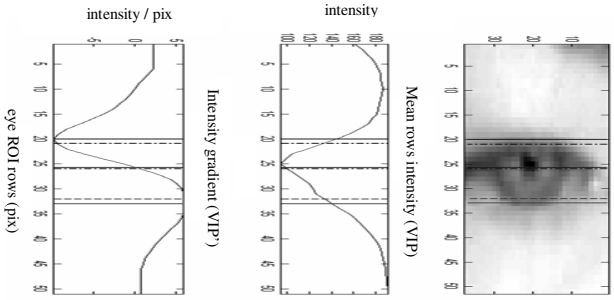

Fig. 7. Right to left: eROI, vertical integral projection (VIP), and 1st derivative of VIP (VIP'). The solid lines are the manually annotated positions and dashed lines are the estimated positions of the UEL_y, COE, and LEL_y. The respective UEL_y and LEL_y in an eROI were detected by locating the local minima above and local maxima below the estimated COE y-coordinate in VIP'.

As presented in TABLE 2, estimation of the UEL_y was relatively accurate but the accuracy of LEL_y estimation was poor. An average height between UEL_y and LEL_y of fully open eyes in the reference database was 14.2 pixels. The LEL_y detection was poor because fewer eyelashes at the lower eyelid produced poor contrast in image intensity. Also, during upward gaze, the contrast at the edge of iris and sclera forms the first local minima below COE in VIP', which is incorrectly detected as the LEL_y. Although the LEL_y detection algorithm needs further refinement, it performed well in general as suggested by low mean of median error magnitude. Particularly, LEL_y was accurately detected in majority of

majority of frames with fully open eyes.

TABLE 2
PERFORMANCE OF UEL_y AND LEL_y DETECTION METHOD

	Mean ± SD magnitude error (pixels)	
	Median	90^{th} percentile
UELy	1.4 ± 2.0	2.7 ± 4.2
LELy	2.1 ± 2.9	6.9 ± 2.5

E. Measurement of eye closure

The eye closure was measured by first determining the average height of an open eye ($\hat{H}$) for each individual subject and then taking the ratio of closed portion of the eye ($\hat{H}$-h) to the $\hat{H}$. As part of the system initialization process, $\hat{H}$ for each subject was determined by averaging the vertical distance between estimated UEL_y and LEL_y in known frames with fully open eyes. During initialization, the average vertical distance (D_{open}) between UEL_y and center of fROI ($fROI_c(y)$) in proportion to height of the fROI is also determined. As illustrated in Fig. 9, the D_{open} is used to calculate the position of upper eyelid when the eyes are open (UEL_{y_open}) in each image based on height of fROI. Then ratio of eye closure (EC) in a frame is calculated by taking the ratio of difference between UEL_{y_open} and estimated UEL_y to $\hat{H}$. Fig. 8 shows and example of eye closure measurement performance.

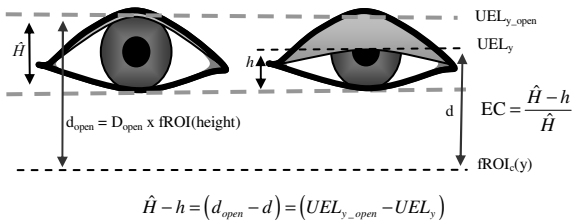

Fig. 9. Illustration of calculating the ratio of eye closure (EC) in an image.

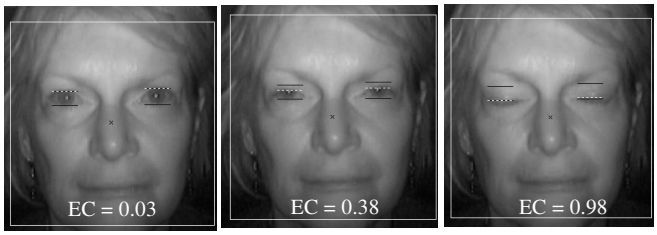

Fig. 10. Examples of eye closure measurement in fully open, partially open, and closed eyes respectively from left to right.. The number in each image represents the calculated ratio of eye closer (how much the eye is closed) for the image.

Currently, the EC measurement system has an average median error magnitude of 0.15, which is sufficient to differentiate the eyes into 3 categories of eye closures, i.e., fully open (0.0), half open (0.5), and fully closed (1.0). However, the average 90^{th} percentile error magnitude of EC is 0.42, which is unreliable for differentiating eyes into any particular category.

IV. DISCUSSION

The individual performance of the face, COE, and eyelid detection methods used in this project were encouraging. However, the compound errors of each facial detection method resulted in unsatisfactory performance of final eye closure metric for monitoring alertness. This is a disadvantage of the passive top-down system development approach adopted in this project.

Particularly, the poor detection of lower eyelid substantially reduced the reliability of eye closure measurement. Hence, further investigation to improve the accuracy and reliability of the lower eyelid detection method will be desirable for improving the eye closure measurement system so that it can be reliably used for monitoring behavioral microsleep.

REFERENCES

[1] M. T. R. Peiris, R. D. Jones, P. R. Davidson, G. J. Carroll, and P. J. Bones, "Frequent lapses of responsiveness during an extended visuomotor tracking task in non-sleep-deprived subjects," *Journal of Sleep Research, 15*, 2006, pp. 291-300.

[2] J. L. Kolstad, "Grounding of the US Tankship Exxon Valdez on Bligh Reef, Prince William Sound Near Valdez, AK, March 24, 1989. Washington, DC: National Transportation Safety Board", Washington, 1990.

[3] L. Trosvall, and T. Akerstedt, "Sleepiness on the job: continuously measured EEG changes in train drivers," *Electroencephologram Clinical Neurophysiology, 66*, 1987, pp. 502-511.

[4] Q. Ji, & X. Yang, "Real-time eye, gaze, and face pose tracking for monitoring driver vigilance," *Real-Time Imaging, 8*, 2002, pp. 357-377.

[5] S. K. L. Lal, and A. Craig, "A critical review of the psychophysiology of driver fatigue," *Biological Psychology, 55*(3), 2001, pp. 173-194.

[6] B. S. Oken, M. C. Salinsky, and S. M. Elsas, "Vigilance, alertness, or sustained attention: physiological basis and measurement," *Clinical Neurophysiology, 117*(9), 2006, pp. 1885-1901.

[7] W. W. Wierwille, and L. A. Ellsworth, "Evaluation of driver drowsiness by trained raters," *Accident Analysis & Prevention, 26*(5), 1994, pp. 571-581.

[8] M. M. Mallis, "Evaluation of Techniques for Drowsiness Detection: Experiment on Performance-Based Validation of Fatigue-tracking Technologies," 1999, *Drexel University.*

[9] P. Viola, and M. Jones, "Rapid object detection using a boosted cascade of simple features," *Paper presented at the Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition.*

[10] R. Lienhart, & J. Maydt, "An extended set of Haar-like features for rapid object detection," *Paper presented at the Proceedings of the International Conference on Image Processing. 2002.*

[11] OpenCV. Open Source Computer Vision Library Reference Manual, 2001.

[12] G. Welch, and B. Gray, *An Introduction to the Kalman Filter*. Chapel Hill: University of North Carolina, 2006.