A Real Time Eye-motion Monitoring System

K.K. Delibasis[1], P.Asvestas[1], G.K. Matsopoulos[1], T. Economopoulos[1], Nicholas Assimakis[2]

[1]Department of Electrical and Computer Engineering
National technical University of Athens, Greece
Athens, Greece
gmatso@esd.ece.ntua.gr

[2]Department of Electronics
Technological Educational Institute of Lamia
Lamia, Greece
assimakis@teilam.gr

Abstract—**A real time image based system is presented that is able to measure eye movement related quantities at high frame rate. The system's hardware component allows for high frame rate image acquisition under near infrared constant lighting conditions, whereas the software component is capable of performing real time image processing to determine the values of the required eye-related parameters. The measured quantities include blink parameters, eye-lid parameters and pupil diameter. The robustness of the proposed system is tested for different types of iris color, as well as under normal lighting conditions and near Infra-red light. The results are validated by off-line observer based inspection of the video sequences.**

Keywords-component; pupil extraction, image processing, blink measurements, lid measurements, real time estimations

I. INTRODUCTION

A number of systems have been proposed that assess mental and behavioural state of a subject, such as drowsiness, fatigue and specific psychiatric conditions, based on parameters extracted from EEG recordings [1], oculography [2] and video recordings [3]. Previous commercial [4] [5] and experimental systems [6] [7] exploited the use of standardized camera, acquiring images of the subjects' eye or whole face to detect drowsiness level. The main principles of video-based systems include the use of a camera and controlled illumination conditions, for image acquisition and software modules for the processing of the digital image sequences. The majority of the video based systems acquire images of the human eye and extract eye-motion parameters by involving image processing techniques, such as segmentation of objects like the pupil and iris. Biometric related applications also include normalization and generation of the human iris template to be matched with stored templates existing in a database [8]. A comprehensive review on the acquisition and performance of iris-based systems as reported in the literature can be found in [9].

The aforementioned systems are based mainly to iris recognition, their performance in identifying a subject's iris varies from 2 to 7 seconds, approximately, whereas substantial constrains on subjects position and motion during recognition process are imposed. Therefore, real time calculations at high frame rate are not possible.

The problem of image-based eye detection and tracking has been addressed by many researchers in the literature, employing colour and edge information [10], [11], a deformable eye manually initialized [12], two-dimensional feature correspondence to solve the two-view transfer problem [13], iris centre tracking based on the Kalman filter [14]. Most of the aforementioned methodologies are laboratory oriented and their application under field condition is limited. Most of the video-based systems mentioned above operate at low frame rate, their application under field condition is limited, whereas the required execution time for iris recognition systems is much prolonged, In addition substantial constrains on subjects position and motion during recognition process are imposed. Therefore, real time calculations at high frame rate are not reported in the literature.

In this paper, a prototype system hardware and software system for estimating eye parameters that can be used for the detection of the drowsiness is proposed. The hardware is based on the use of a CMOS array sensor placed in front of the subject, capable of acquiring eye region image data at frame rate as high as 100 frames per second (fps). The system operates in constant illumination conditions using a NIR light source in conjunction with NIR filter attached to the camera. The lens is connected to the camera through an optical bundle 'driving' the image data to the camera and providing flexibility of usage. The camera is connected to a processing unit where all developed image processing software modules for estimating specific eye related parameters are embedded. Section II provides a detailed description of the proposed system, including technical specifications of the hardware part as well as the software modules and algorithms developed for real time image processing of the acquired data and the calculation of the eye parameters. In Section III, results from the application of the system are presented in terms of the software performance and mainly on overall system performance under different conditions and subjects. Finally, in Section IV, issues related to the current system design and performance, as well as future work are discussed.

II. THE PROPOSED SYSTEM

A. *Hardware system component*

The hardware system that has been implemented performs the image acquisition at high frame rate, so that reliable eye-related measurements are possible. More specifically, the eye-

978-1-4244-4530-1/09/$25.00 ©2009 IEEE

based monitoring hardware system comprises from the following components:

1. The CMOS array sensor: A ½-inch CMOS chip is selected, operating at 100 frames per second (fps), in order to capture the fast eye-lid movements. Taking into consideration the physiology of eye movement and the experimental setup, a shutter speed of 4,000 - 8,000 sec^{-1} is required for the eye saccadic movement, whereas a shutter speed of 8,000 - 12,000 sec^{-1} is required for the eye lid movement. [15]. According to the above requirements, a ½-inch CMOS monochrome camera (MV-D640 CMOS Area Scan Camera by Photon-Focus) was finally selected with small dimensions and light weight, pixelation of 640×480 pixels, pixel size equal to 9.9 µm x 9.9 µm, dynamic range equal to 60dB, global shutter, minimum exposure time of 40µsec, maximum frame rate: up to 200fps at 640×480 pixels. A camera link cable and the video grabber to acquire the data were also included.
2. The lens: taking into consideration the camera – object distance and the eye dimensions, the required focal length is estimated equal to 17.1mm for the any ½ inch CMOS camera. Considering a narrow Depth of Field (DoF) that keeps all parts of the eye in focus (including eye lids as well as the far side of the cornea) and diameter of the Circle of Confusion – CoC equal to the sensor pixel size, an f/number equal to f/5 was determined [16], [17].
3. NIR components: A long pass optical filter in the Near Infra Red (NIR) area of the spectrum with cut-off frequency at 830nm was selected (LP830 IR filter). It was combined with an NIR light source, consisting of 18 infrared (IR) light emitting diodes (LEDs) (model VTE1163H from PerkinElmer Optoelectronics), placed at an approximate distance of 10 cm from the eye region with continuous power dissipation of 200 mW each and peak emission wavelength at 880nm. Both the NIR filter cut-of frequency and the light source peak emission wavelength were reasonably close to the 836 nm atmospheric water absorption line, which achieves an almost constant illumination, unaffected from external light conditions.
4. Optical connector: The 1m long optical bundle connector (IG-154 model, Schott North America Inc) was of 51 mm bend radius, with 4×4 mm^2 format size, providing resolution of 50 lines per mm. The optical bundle fibre is connected to the lens and the camera body with C-mount adapters.

The acquired image data are transferred to the computer platform for processing and storage via a video grabber (microEnable III) supporting up to 240 Mbytes continuous data transfer over 64bit PCI protocol for the acquisition and digitization of the image data

B. *Structure of the software system component*

The proposed system consists of a number of modules for image segmentation as well as the measurement of the required parameters. The task of segmentation consists of edge detection, extraction of pupil and iris boundary and extraction of upper eye-lid boundary. All modules are executed on a PC-based computer platform.

The software modules of the proposed system utilize the high frame rate to narrow the parameter range in subsequent frames, using the already obtained segmentation results.

1) *Pupil and Iris segmentation*

The detection of edges in each frame is a prerequisite for any further segmentation. The Canny edge detector was used for the detection of the edges of the image, using a standard deviation of the Gaussian equal to 2 and low and high thresholds for the non-maximum suppression equal to 0 and 0.05, respectively. The values of the parameters were determined after experimentation. This step provides the binary edge image, $E(x,y,t)$, whose pixels are set to 1 if they are determined as edge pixels, as well as the gradient, $\mathbf{G}(x,y,t)=(G_x(x,y,t),G_y(x,y,t))$ of the frame $I(x,y,t)$ at each pixel (x,y), where G_i is the image gradient along the x and y direction (i=x and y respectively), calculated as $G_i = g_i \otimes I$, with the $\otimes$ denoting the convolution operation and g_i is the partial derivative of a Gaussian function along the direction i.

It is assumed that the pupil and iris have circular shape, therefore the extraction of their boundary is based on a modified version of the Hough Transform (HT) for detecting circular objects, as described in [18]. The proposed method uses the gradient $G(x,y,t)$ and the edge image $E(x,y,t)$ obtained from the edge detection step, and operates using the following parametric definition of the required shape.

$$\begin{aligned} x(\theta) &= x_p + R_p \cos\theta \\ y(\theta) &= y_p + R_p \sin\theta \end{aligned}, \ \theta \in [0,2\pi) \tag{1}$$

where (x_p, y_p) and R_p are the center and the radius of the pupil, respectively. The algorithm operates as following: for each edge pixel (x,y), the value of the parameter θ is determined by requiring the gradient direction to coincide with the direction of the normal vector. The value of pixel coordinates (x_p,y_p), calculated as following, are used to define the pixel of the accumulator array that is increased by one:

$$\begin{aligned} x_p &= x(\theta) - R_p \cos\theta \\ y_p &= y(\theta) - R_p \sin\theta \end{aligned} \tag{2}$$

The algorithm is executed for different values of the required radius R_p and the maximum accumulator value is detected for each radius. The global accumulator maximum over all radii in the allowed radius range, determines the coordinates of the circle centre and its radius.

The minimum and maximum values of the pupil parameters that are used in the HT, are defined as following: for the first frame the x_p, y_p parameter range spans the whole image, and the R_p parameter varies between 7 and 15 pixels, whereas in the subsequent frames the allowed range for x_p, y_p and R_p spans 5 pixels, centered on the values determined in the previous frame.

The subsequent iris segmentation is performed by using the same algorithm (HT for detection of circular shapes). The allowed range of the iris center coordinates x_i, y_i and the iris radius R_i is defined for the first frame with respect to the already determined pupil parameters (± 3 pixels for the iris center coordinates with respect to the pupil center coordinates and between 2 and 7 pupil radii for the iris radius) and for the subsequent frames with respect to the iris parameters of the previous frame. In this way, execution time is kept minimum without compromising system robustness.

2) Upper eye-lid boundary extraction

Boundary extraction of the upper eye-lid is required for eye-lid speed measurements. In this work, the eye-lid is modeled by a parabola with the apex at (x_0, y_0), focus $(x_0, y_0 + p)$ and equation of directrix $y = y_0 - p$, explicitly represented as following:

$$(x(t), y(t)) = (x_0 + 2pt, y_0 + 2pt^2) \qquad (3)$$

The modified version of the Hough transform (HT) as described in [Nixon] is used to determine the parabola parameters x_0, y_0, p following the approach that was briefly described above for circle detection.

In order to accelerate the execution of the HT for parabola, the following modifications have been made: only edge pixels within a square enclosing the iris (as determined for the current frame), with vertical image gradient component greater than horizontal image gradient component, are allowed to contribute to the HT accumulator array. The allowed parabola parameter range in subsequent frames is defined as a function of the parameter values determined in the previous frame.

Fig. 1 shows typical results from the lid detection algorithm for three different lid positions. Fig. 2 shows the combined results of pupil, iris and upper eye-lid segmentation for two frames corresponding to fully open and partially closed eye.

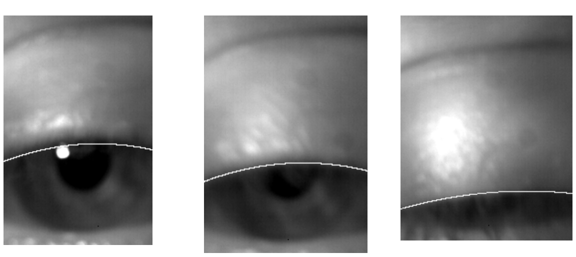

Figure 1. Typical result from the Hough Transform based lid detection algorithm for opened, semi-opened and closed eye.

Identify applicable sponsor/s here. *(sponsors)*

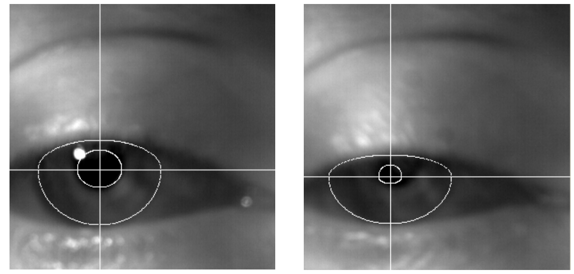

Figure 2. Combining pupil, iris and upper eye-lid segmentation for two frames corresponding to fully open and partially closed eye.

C. Eye-related measurements

Having segmented the objects of interest in each image frame, the calculation of the eye-related measurements proceeds as following:

1) Blink parameters

The blink parameters, which include the number of blinks, the interval between blinks and the blink duration, are calculated using the number of pixels of the segmented pupil. The number of pixels is measured according to the following algorithm:

1. The square region P is determined that is centered on the pupil center and contains the pupil.
2. Pixels that belong to region P are considered to belong to one of the three following classes: dark pupil, "background" iris and possible reflection inside the pupil. Therefore the histogram clustering approach described in [19] is applied and the pupil pixels are counted.
3. An eye blink is identified by the system if the number of segmented pixels in the current frame is less than 20% of the average number of segmented pixels in last k frames, k=30.

2) Eye-Lid measurements

Eye-lid measurements include the following: Lid closure/opening speed defined as speed of lid motion in every frame (pixel / frame), Lid closure/opening duration defined as number of frames during the lid closure / opening phase. All measurements are performed by using the coordinates of the apex of the parabola that models the eye-lid, which corresponds to the center of the eye-lid.

3) Pupil diameter measurements

Having calculated the blink parameters, pupil diameter can be trivially derived using the pupil range that was determined by the modified HT.

Fig. 3 shows the changes of the pupil area in number of pixels (thin curve) as well as the mean value of the pupil area over the previous 30 frames (thick curve). The corresponding image frames showing the eye condition during a blinking are also added.

In order to convert the eye-lid measurements from pixels/sec to mm/sec, the pixel size is required, which is obtained by using the results from the iris segmentation, since the iris is the only part of the imaged eye that does not change in size.

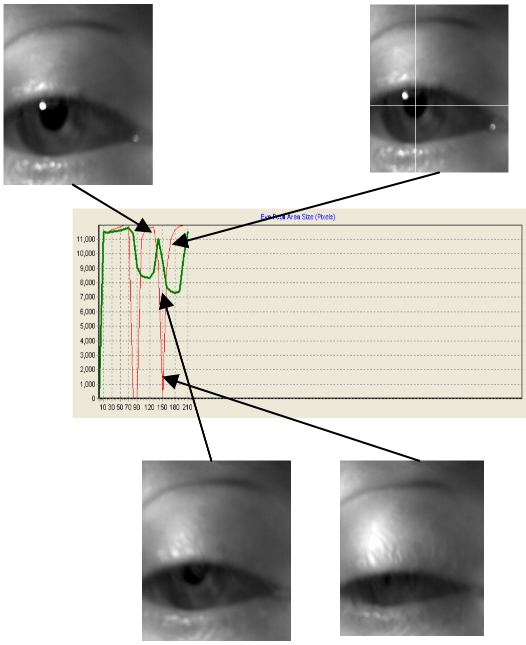

Figure 3. The changes of the pupil area in number of pixels are shown (thin curve) as well as the mean value of the pupil area over the previous 30 frames (thick curve). The corresponding image frames showing the eye condition during a blinking are also added.

III. RESULTS

Initial tests have been carried out by the implementation of the proposed system. The CMOS camera has been mounted on the head and image frames are then acquired in real time. The experiments for system verification included testing the performance of the image processing modules in estimating the eye related parameters within the laboratory environment as well as testing the system under different conditions and subjects.

The evaluation of the results was made off-line by observing the median frame number for each blink in the camera captured videos. The median frame number of fully closed eye was selected for the observer-based offline verification, since manual detection was more accurate than detecting the beginning and the end blink frame. No observer-based lid measurements were made due to impracticality issues. Results on the performance of the system for a typical brown eye subject without glasses, under NIR light, obtained indoors are presented in Table I. In order to assist comparison, the median blink frame was also derived by the parameters obtained by the proposed system.

The performance of the developed sensor was tested under two different conditions: normal lighting and NIR lighting, on a small-scale test involving 50 subjects (male 74% and female 26%), with an ID number assigned to each of them for further evaluation of the obtained results. The testing scenarios included: tests under different light conditions (with normal light and with NIR light source) and tests with different eye-colored subjects. These tests were conducted indoors (in the laboratory environment) involving all 50 subjects. In all tests, the CMOS camera was operated at 100fps. Each subject had its eye imaged for approximately 100 blinks. Blink detection was considered successful for a subject if at least 98 blinks were detected by the system. Table II summarizes the performance of the proposed system for all subjects under different testing conditions.

TABLE I. BLINK ESTIMATES USING PROPOSED SYSTEM, COMPARED TO THE OBSERVER-BASED OFF LINE PROCESSING, FOR INDOOR EXPERIMENT USING NIR LIGHT SOURCE, NIR FILTERS ATTACHED TO THE CAMERA AND FOR SUBJECTS WITHOUT GLASSES

Blink No	**Measurements by the proposed system**			**Observer-based, off line measurements**	
	Blink Interval	***Blink duration***	***Median frame of blink***	***Blink detected***	***Median frame of blink***
1	171	20	181	Y	184
2	36	30	242	Y	239
3	46	26	316	Y	314
4	78	16	415	Y	412
5	94	27	531	Y	529
6	69	24	625	Y	624
7	75	12	718	Y	716
8	228	30	967	Y	970
9	181	29	1.178	Y	1179
10	423	21	1.626	Y	1626
11	44	28	1.694	Y	1694
12	86	17	1.803	Y	1803
13	71	15	1.890	Y	1891
14	156	27	2.067	Y	2066
15	12	27	2.106	Y	2107
16	536	35	2.673	Y	2675
17	65	11	2.761	Y	2761
18	58	18	2.833	Y	2831
19	54	20	2.906	Y	2907
20	33	36	2.967	Y	2970
21	125	24	3.122	Y	3121
22	179	13	3.320	Y	3322
23	142	25	3.481	Y	3483
24	264	18	3.766	Y	3768

The system achieved 100% performance rate for the cases of indoor use, with the existence of NIR light source in conjunction with NIR filters attached to the camera, irrespectively of eye colour. The system had poor performance (40%) when normal light was used as a light source.

TABLE II. THE PERFORMANCE OF THE PROPOSED SYSTEM UNDER DIFFERENT LIGHTING CONDITIONS, FOR SUBKECTS GROUPED BY EYE COLOR. SUCCESSFUL SUBJECTS ARE GIVEN, OVER THE TOTAL NUMBER OF PARTICIPATING SUBJECTS. A SUBJECT IS CONSIDERED SUCCESSFUL IF ACCURACY OF BLINK DETECTION IS NOT LOWER THAN 98%.

Eye Type	**Test conditions**			
	Indoors, normal light		***Indoors, NIR light***	
	Subjects	***Accuracy %***	***Subjects***	***Accuracy %***
Dark Brown	7/15	47	15/15	100
Brown	12/26	50	26/26	100
Blue	1/6	17	6/6	100
Green	0/3	0	3/3	100
Total	20/50	40	50/50	100

The proposed system relies on the real time image processing modules to estimate the specific eye related parameters. All the software modules were developed for the specific application under the constrain for minimum execution time. The eye segmentation algorithm computes the blink parameters in 3.29 ms per frame (frame size 640×480 pixels) while the eye lid detection algorithm computes the lid parameters in 7 ms per frame in a high end Intel Core II PC. These execution times can be compared favorably to similar approaches, such as [20] where frame processing requires 4-7 sec, [9] that reported a rate of frame processing of 15fps, [10] which reported the use of an automatic method for detecting eye states based on color and edge information supporting a rate of 20 fps, [4] [5] [6] [7] that reported eye of face video analysis at 30 – 60 fps.

IV. DISCUSSION

In this paper, a prototype system was presented for the estimation of specific eye parameters that can be related with the detection of the drowsiness. Blink, lid and pupil parameters are estimated in real time using the developed image processing modules. Compared to other similar systems, the proposed prototype has several advantages including high frame acquisition rate, and low processing time for estimating the eye parameters (~ 10 ms per image frame).

We are currently working towards evaluation of the usefulness of the proposed system in a number of psychiatric conditions, as well as its comparison and/or integration with eye-parameter measuring systems, based on signal acquisition and processing, such as electro-occulography.

ACKNOWLEDGMENT

The proposed eye motion monitoring system was developed within the SENSATION Integrated Project (Sixth Framework Program Priority IST 2.3.1.2 - Micro and Nano sensors).

REFERENCES

[1] C.E. Vasios, G.K. Matsopoulos, E.M. Ventouras, C. Papageorgiou, V.P. Kontaxakis, K.S. Nikita, N. Uzunoglu, "Multivariate Autoregressive Modeling Combined with Simulated Annealing Optimization for Classifying Sources of Event Related Potentials", International Journal of Bioelectromagnetism, vol. 7, no. 2, pp. 78-81, 2005.

[2] Yamada F.: Frontal midline theta rhythm and eye blinking activity during a VDT task and a video game: useful tools for psychophysiology in ergonomics. Ergonomics, **41**:678-688 (1998).

[3] Forczmanski P., Kukharev G.: Comparative analysis of simple facial features extractors. J. Real-Time Image Proc. **1**:239-255 (2007).

[4] Siemens VDO: Electronic warning system to keep drivers alert. http://www.siemensvdo.com/press/releases/commercialvehicles/2006/sv-200609-005-en.htm.

[5] Smarteye Inc.: SmarteyePro. http://www.smarteye.se

[6] Wierwille W.: Historical perspective on slow eyelid closure: Whence PERCLOS. Ocular Measures of Driver Alertness: Technical Conference Proceedings, U.S. Department of Transportation, Herndon, Virginia, (1999).

[7] The John Hopkins University: Drowsy driver detection system. Technical Report, www.jhuapl.edu/ott/technologies/featuredtech/DDDS.

[8] Daugman J.: Biometric personal identification system based on iris analysis. U.S. Patent 5 291 560 (1994).

[9] Matey J.R., Naroditsky O., Hanna K., Kolczynski R., Lolacono D.J., Mangru S., Tinker M., Zappia T.M., Zhao W.Y.: Iris on the move: Acquisition of images for iris recognition in less constrained environment. Proc IEEE, **94**(11):1936-1947 (2006).

[10] Liu, H., Wu, Y., Zha, H.: Eye states detection from color facial image sequence. In: SPIE International Conference on Image and Graphics, **4875**:693–698 (2002).

[11] Sirohey, S., Rosenfeld, A., Duric, Z.: A method of detecting and tracking irises and eyelids in video. Pattern Recognit. **35**(6):1389–1401 (2002).

[12] Deng Y. and Lai F.: Region-based template deformation and masking for eye-feature extraction and description. Pattern Recognit. **30**(3):403-419 (1997).

[13] Yu L.H., Eizenman M.: A new methodology for determining point-of-gaze in head mounted eye tracking systems. IEEE Trans. Biomed. Engin. **10**(51):1765-1773 (2004).

[14] Xie X., Sudhakar R., Zhuang H.: Real-time eye feature tracking from video image sequence using Kalman filter. IEEE Trans. Syst. Man, Cybern. **25**:1568-1577, (1995).

[15] Barber P. J., Legge D., Perception and Information chap. 4, Information Acquisition, pp 51-66 Methuen, London.

[16] Ray, Sidney F. Applied Photographic Optics, 3rd ed. Oxford: Focal Press 2002, ISBN 0-240-51540-4.

[17] Lewis Larmore, Introduction to Photographic Principles. 2nd ed. New York: Dover Publications, Inc1965.

[18] M. Nixon, A. Aguado, Feature extraction and image processing, Newnes 2002, ISBN 0 7506 5078 8, pp.173-198.

[19] Du-Ming Tsai and Ying-Hsiung Chen, A fast histogram-clustering approach for multi-level thresholding, Pattern Recognition Letters 13 (1992) 245-252.

[20] Tian, Y., Kanade, T., Cohn, J.F.: Dual-state parametric eye tracking. In: International Conference on Automatic Face and Gesture Recognition (2000).