

Enhanced Eye-Tracking Data: a Dual Sensor System for Smart Glasses Applications

Paweł Krzyżanowski, Tomasz Kocejko, Jacek Ruminski, Adam Bujnowski

Gdansk University of Technology, Department of Biomedical Engineering,

Faculty of Electronics, Telecommunications and Informatics, Poland

Email: pawkrzyz@pg.gda.pl, tomkocej@pg.gda.pl, jacek.ruminski@pg.gda.pl, bujnows@biomed.eti.pg.gda.pl

□ *Abstract* – A technique for the acquisition of an increased number of pupil positions, using a combined sensor consisting of a low-rate camera and a high-rate optical sensor, is presented in this paper. The additional data are provided by the optical movement-detection sensor mounted in close proximity to the eyeball. This proposed solution enables a significant increase in the number of registered fixation points and saccades and can be used in wearable electronics applications where low consumption of resources is required. The results of the experiments conducted here show that the proposed sensor system gives comparable results to those acquired from a high-speed camera and can also be used in the reduction of artefacts in the output signal.

I. INTRODUCTION

THE rapid development of wearable electronics has meant that this technology is ever-present in everyday life; the number of such applications and solutions for human-machine interactions is constantly increasing. However, the miniaturization of these devices and the growing array of functionalities require the development of new interfaces. For many applications, for example septictype interactions [1], the eye-tracking interface is a very convenient option. With just a glance from the user, it is possible to execute commands. This technology has many possible applications [2, 3]; however, it is still under continuous study [4, 5]. In addition, research to enhance video-based gaze tracking is needed to increase both the quality and the amount of data acquired by the system [6] as well as to increase the speed of real-time processing [7].

Since modern wearable electronics like smart glasses are multitasking devices, the design of interfaces within the constraints of the processor unit's computational power is a challenge. In modern eye-tracking interfaces, the camera frame rate is a factor that can increase the potential number of applications of the interface. With the advanced image processing required for reliable estimation of eye and gaze

position, the eye-tracking interface is the device which consumes the majority of the power resources. In addition, a need for high computational power and the excessive power consumption are difficult requirements to fulfil.

In gaze-tracking interfaces, it is important to achieve a correct image acquisition, an accurate calculation of the fixation point and the appropriate processing of these data. In practical applications, the time required by the CPU/GPU to process the data is significant [8] and the reduction of this time results in lower consumption of power, which is also important when considering the design of a battery charging circuit. It should be emphasized that studies which have presented new integrated circuits (ICs) have achieved effective results with software and hardware data processing, but these ICs are not in widespread use [9].

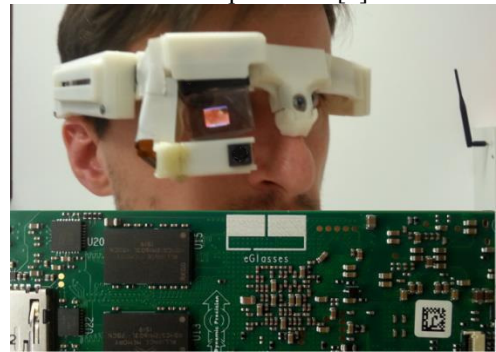


Fig. 1 Prototype of the eGlasses platform

The eGlasses (Fig. 1) electronic eyewear is a multisensory platform capable of running on both Linux and Android operating systems. The eye-tracking module is one of the eGlasses input interfaces. It allows operation by graphic user interface, controlled by gaze, and also provides data that can be used in interaction with everyday objects. The core elements of the eye-tracking module hardware are infrared LED-based eye surface lights, the camera that registers eye movements (eye camera) and the camera that registers the scene images (scene camera). In the current prototype, the eye-tracking module allows a 30 Hz sampling frequency at a resolution of 320x240 pixels. This sampling rate significantly decreases for higher resolutions (e.g., it is only 15 Hz for an image of 640x480 pixels resolution). The

□ This work has been partially supported by NCBiR, FWF, SNSF, ANR and FNR in the framework of the ERA-NET CHIST-ERA II, Project eGLASSES – The interactive eyeglasses for mobile, perceptual computing and by statutory funds from the Faculty of Electronics, Telecommunications and Informatics Faculty, Gdansk University of Technology.

use of an optical sensor with low power consumption allows a high frame rate to be maintained, while preserving a high resolution in the eye-tracking camera. This study presents a technique for increasing the eye-tracker sampling frequency by combining the low-rate camera with the additional optical sensor. So far such an idea has been used mostly in non-wearable electronic equipment to minimize errors caused by head rotation [10].

The rest of the paper is organized as follows: Section 2 describes the methods, experimental stand and algorithm. The experimental set-up and results are presented in Section 3. Section 4 contains a discussion and potential applications for this new concept. Section 5 presents the conclusion.

II. METHODS AND MATERIALS

The underlying concept of this research was to use an eye-tracking camera and an additional simple optical sensor to increase the number of registered pupil positions. The principle of operation of the additional sensor was based on measurements of eyeball displacements. To evaluate this idea, a custom-built eye-tracker allowing measurement of pupil movements with a 180 Hz sampling rate was utilized. The optical sensor was then added to complement the eye-tracker. An appropriate test stand was built to conduct the experiments in a controlled environment.

A. Experimental Stand and Data

The experimental hardware equipment consisted of the three main components: the eye-tracking camera, the optical motion sensor, and a model of an eyeball made of plastic. The model was of spherical shape and contained an artificial pupil (Fig. 2).

The stand was equipped with two bearing servomotors (HD3688MG) and the frame was constructed using 3D printer technology. The construction had two degrees of freedom: horizontal (x) and vertical (y). The servomotors were controlled by a PWM signal from the microprocessor. This approach allowed independent movement in both the x- and y-directions.

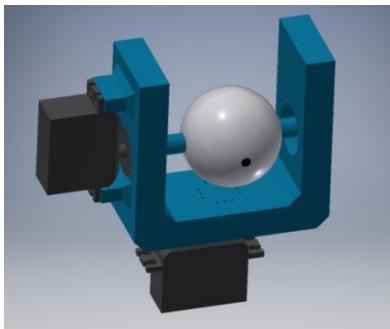


Fig. 2 The 3D frame (Fr) with the eyeball model controlled by servomotors (Sv)

The servomotors allowed the eyeball position to be changed up to 25 times per second. With this frequency it was possible to simulate saccades that lasted 40 ms. Healthy eye movements were therefore able to be simulated [11]. The camera captured the pupil position with a frequency of

180 Hz. The basic parameters of the eye-tracker are presented in Table 1.

TABLE I
BASIC PARAMETERS OF THE EYE TRACKER USED

Parameter	Value
Eye-tracking technique:	Dark Pupil, Pupil Centre Detection
Eye-tracking:	Monocular (left eye)
Data rate:	180 Hz
Accuracy:	0.7°
Precision:	0.5°
Light condition restrictions:	No
Eye-tracking camera resolution:	320x240 px

The eye camera was connected to a PC computer with the eye-tracking software installed and running on Linux OS. The main classes of the eye-tracking software covered eye and scene camera image acquisition, pupil-detection algorithms, video for Linux camera support, gstreamer and opencv camera support and fixation and gaze-tracking algorithms. The eye-tracking software offered three algorithms for pupil centre detection; this study was conducted using the most accurate one, based on the selection of pupil boundary candidate points and ellipse fitting.

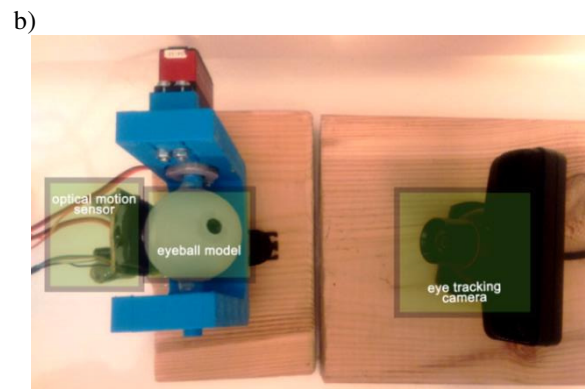
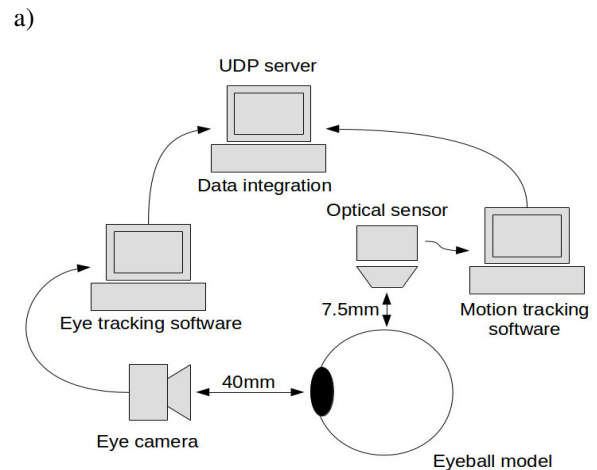


Fig. 3 Configurations of a) the experimental stand and b) its implementation

To track the eyeball motion, the PAN3101, a low cost CMOS optical sensor integrated with DSP, was used. The

DSP serves as an engine for the estimation of non-mechanical motion. The optical navigation technology used in this optical sensor enabled investigators to measure changes in position by optically acquiring sequential surface images (frames) and mathematically determining the direction and magnitude of movement. The current x and y movements can be read from the registers via a serial port. The sensor required additional optics. The one used in this research (HDNS-2100) allowed application of the sensor at a distance of 7.5 mm from the eyeball model. In fact, this is currently a primary impediment to *in vivo* use of this sensor. However, it appears to be possible to increase the distance between the optical sensor and the eyeball, although this would require some changes to the angle between the mirrors and the focal length of the lens. The IR LED diode power should also be adjusted in accordance with safety limits [12]. During experiments, the sampling rate of the sensor was set to 125 Hz. The optical sensor was connected to the additional PC unit. All data acquired by the sensors were transferred using the UDP protocol to the external server for further integration and processing.

There were certain discrepancies in the parameters measured by each sensor. The camera was used to acquire pupil positions frame by frame in an absolute coordinate system, while the optical sensor measured the total eyeball movement (total shift between the last two frames). The shift was the sum of the movement before and after camera sampling, so the value of this shift had to be split proportionally between the frames (Fig. 4).

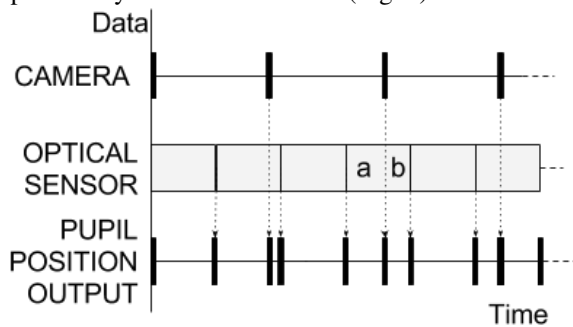


Fig. 4 Output pupil positions and incoming data from both sensors: at a point (the eye camera) and proportional to a shift in time (the optical sensor)

B. Data Acquisition and Fusion Algorithm

The concept underlying of this project was to insert a high frequency signal into a low frequency one (Fig. 5). The algorithm developed for this comprised two major parts. The first was a calibration procedure, while the second was devoted to calculating the additional pupil positions from data registered by the optical sensor, with reference to those captured by the camera. A description of the algorithm is as follows:

1. Start, set the first pupil position from the camera and reset the shift register in the optical sensor (timestamp = 0).

2. Get data from the optical sensor (timestamp, x and y movement).
3. Calculate a new pupil position by adding to the previous pupil position the scaled shift obtained from the optical sensor.
4. Repeat points 2-3 until new data from the eye-tracker is ready.
5. Get data from the eye camera (pupil position and timestamp t_1) and update to a new fixation point.
6. Get data from the optical sensor and split information proportionally to time before and after timestamp t_1 .
7. Calculate a new pupil position based on data after t_1 .
8. Go to point 2.

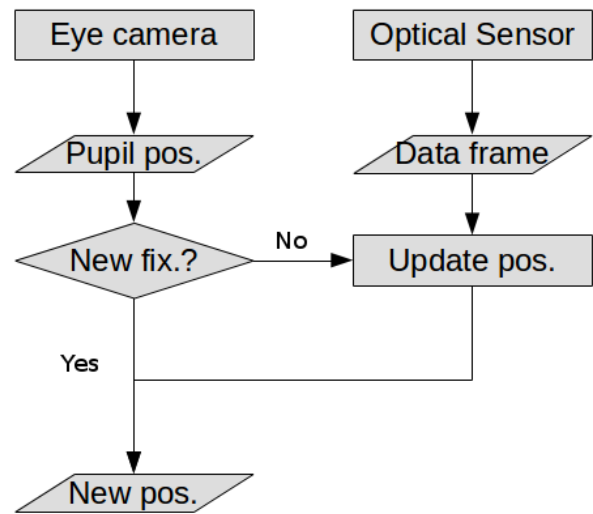


Fig. 5 Block diagram of the proposed algorithm

C. Experiments

To evaluate the concept of extending a set of pupil positions recorded by the eye tracking camera using data acquired by the optical sensor, several experiments were conducted. In the first experiment, a set of pupil positions were recorded over a certain period by both the reference eye-tracking camera (180 Hz sampling rate) and the optical sensor (125 Hz sampling rate). This was done in order to check the similarities of the paths recorded by the sensors. Pearson correlation coefficients were calculated separately for the x and y coordinates:

$$corr_x = \frac{\sum_i (px_i - p\bar{x})(mx_i - m\bar{x})}{\sqrt{\sum_i (px_i - p\bar{x})^2} \sqrt{\sum_i (mx_i - m\bar{x})^2}}, \quad (1)$$

$$corr_y = \frac{\sum_i (py_i - p\bar{y})(my_i - m\bar{y})}{\sqrt{\sum_i (py_i - p\bar{y})^2} \sqrt{\sum_i (my_i - m\bar{y})^2}}, \quad (2)$$

where px_i , py_i are the x and y coordinates of a pupil position, mx_i , my_i are the x and y coordinates of the sample registered by the optical motion sensor, $p\bar{x}$, $p\bar{y}$ are the average values of pupil position samples along the x - and y -axes, and $m\bar{x}$, $m\bar{y}$ are average values of eyeball positions along the x - and y -axes, registered by the optical motion sensor.

Following this, a set of experiments exploring different patterns of movement were performed. Square-shaped

movements (Fig. 7) were used to calculate the proper constant ratio between the measurements from both sensors. A set of pupil positions from the camera was recalculated by using perspective transformation to compensate for the viewing angle. After calibration, various sequences of the eyeball movement were tested (ellipse, triangle and random) with a fixation duration of between 140 and 530 ms. Each session lasted for ten repeated sequences, and the measurements were stored with a synchronized timestamp.

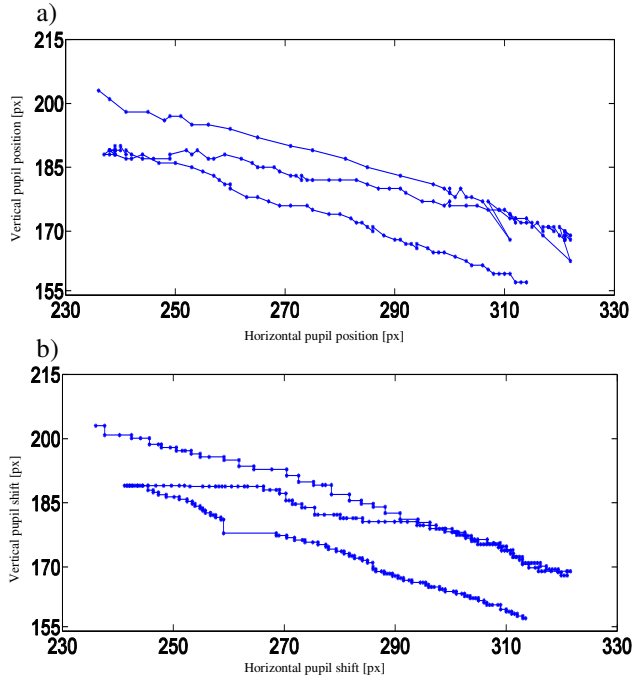


Fig. 6 Similarities between the eyeball movement recorded by the camera and optical sensor: a) The plot of consecutive pupil centre positions detected by the eye tracker b) The plot of the eyeball motion (x, y) coordinates detected by the optical sensor

The sampling rates remained the same; 180 Hz for the camera and 125 Hz for the optical sensor. Next, the original 180 Hz raw data were reduced to simulate recording with 5, 10, 30 and 60 Hz sampling rates. This allowed the creation of a set of data that could be extended by the optical sensor readings and which fully corresponded to the reference data. The results were compared to both: the high-speed 180Hz camera reference as well as the theoretical, programmed ideal eyeball movement. The signals from the optical sensor, the camera and the new combined set had different lengths, and all data were therefore resampled to 1ms using interpolation. The similarity between signals was checked by calculating the mean square error:

$$MSE_x = \frac{1}{n} \sum_{i=1}^n (\hat{X}_i - X_i)^2, \quad (3)$$

$$MSE_y = \frac{1}{n} \sum_{i=1}^n (\hat{Y}_i - Y_i)^2, \quad (4)$$

where n is number of samples, \hat{X}_i and \hat{Y}_i are signal pupil positions in the horizontal and vertical dimensions, and X_i and Y_i are pupil positions from the reference signal (ideal movement or high speed camera).

III. RESULTS

The similarities between the pupil positions recorded by the camera and the eyeball movement acquired by the optical sensor are presented in Fig. 6. The calculated correlation between waveforms recorded by the eye camera at 180 Hz and the optical sensor at 125 Hz were strongly positive:

$$corr_x = 0.91$$

$$corr_y = 0.89$$

In the next experiment, data square movement was implemented to calculate the ratio between sensors. The results are presented in Fig. 7 and numerical values are given in Table 2.

Following this, experiments were performed to check for dependencies between the information gain and the signal variability. The eyeball motion was programmed for fixation duration of 140 ms and data from 30, 10 and 5 Hz camera frequencies were enhanced by the optical sensor. The calculated pupil positions were compared to both reference signals: the high-speed camera and the theoretical route. For greater clarity, the results are presented in Figs. 8-10 and in Table 3. The axes of the coordinate system are pixels (except for the optical sensor figure) and the values plotted are the centred pupil positions.

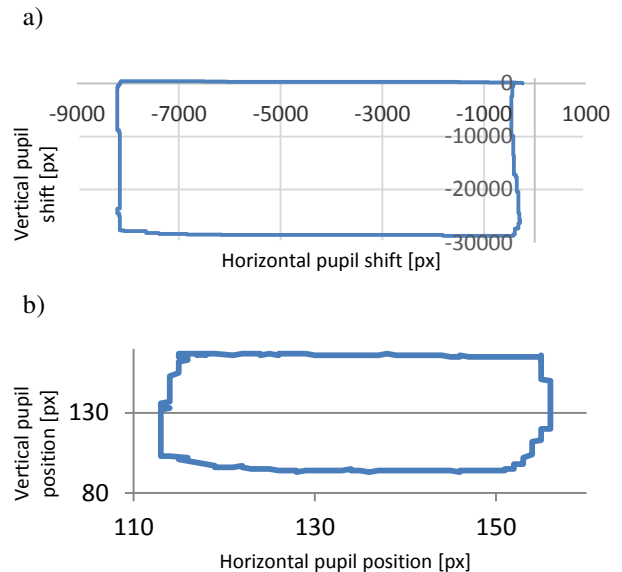


Fig. 7 Calibration of data from both sensors: a) movement recorded by the mouse sensor; b) pupil positions recorded by the high-speed camera after perspective transformation

TABLE II
CALCULATED SCALE CONSTANTS

Name	Value
X ratio:	$4.58 \cdot 10^{-3}$
X standard deviation:	$1.59 \cdot 10^{-4}$
Y ratio:	$2.49 \cdot 10^{-3}$
Y standard deviation:	$2.08 \cdot 10^{-5}$

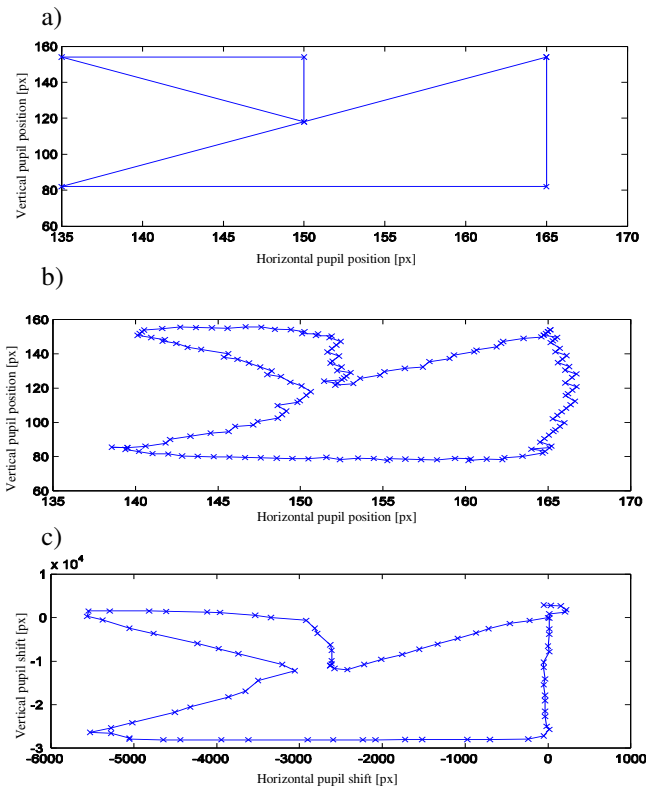


Fig. 8 Raw data before processing: a) ideal movement from the eyeball model; b) pupil positions recorded by the camera; c) eyeball shift measured by the optical sensor

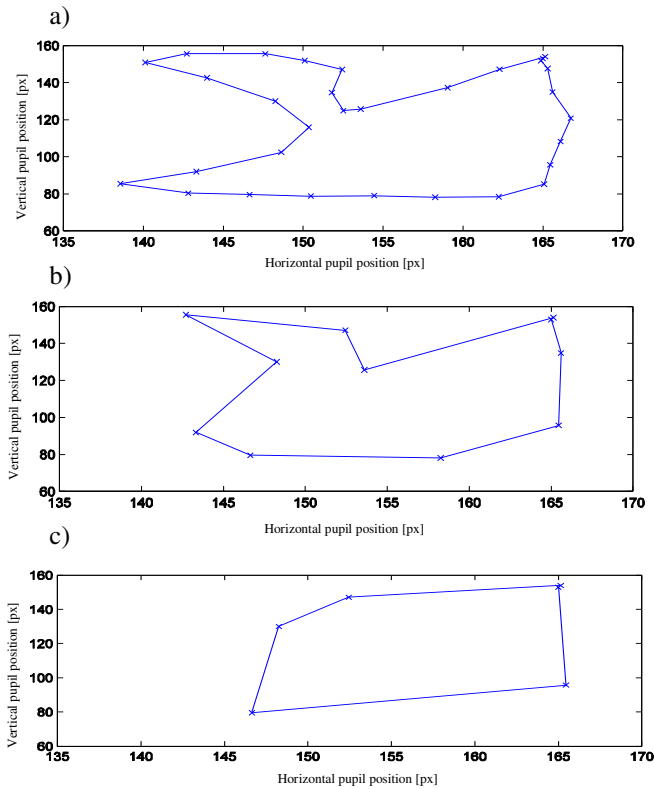


Fig. 9 The effect of decreasing information during reducing of the camera sampling rate: a) 30 Hz camera showing good reproduction of the signal; b) 10 Hz camera showing sufficient reproduction but with some information already lost; c) 5 Hz camera showing high distortion of signal

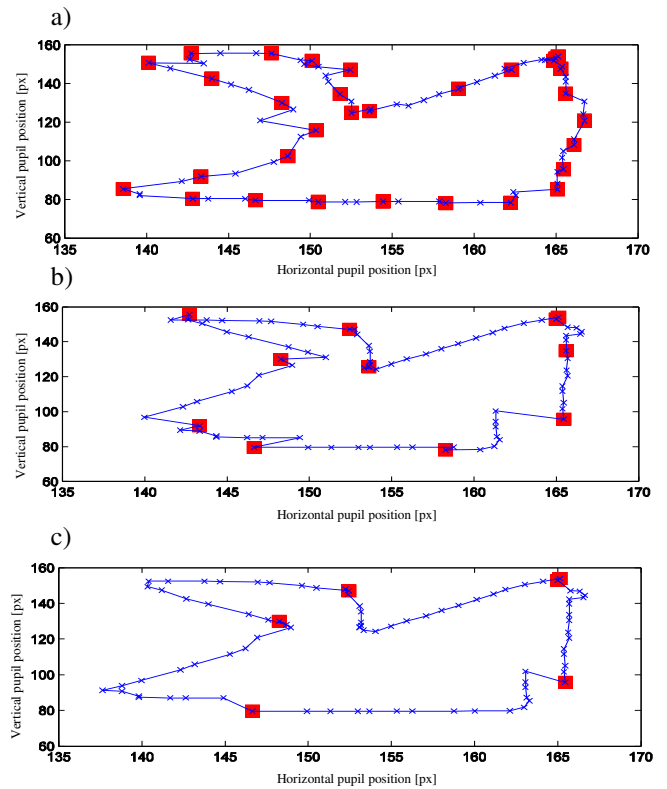


Fig. 10 Enhanced signal after combining data from both sensors. Red squares are pupil positions from camera, and blue crosses are fixations calculated by the system a) 30Hz camera showing no new data compared to the camera system only; b) 10Hz camera signal showing some curves between points; c) 5Hz camera showing the largest improvement compared to using only the camera system

In the last experiment (Fig. 11), a random motion was tested and the system also gave additional data where the sampling rate was insufficient.

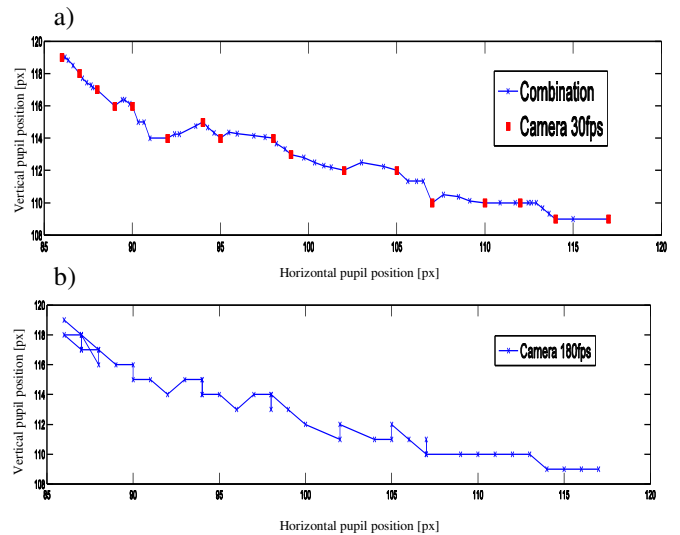


Fig. 11 Experiment combining data from both sensors: a) Low speed camera (red squares) and calculated pupil positions (blue 'x') from optical sensor; b) reference high-speed camera

Table III
EXPERIMENT WITH ENHANCED DATA USING THE DUAL SENSOR SYSTEM

Fixation period: 140 ms	MSE			
	Camera - Reference		Combined - Reference	
	X	Y	X	Y
Shape 60 Hz - 180 Hz cam	0.09	0.30	0.25	0.81
Shape 60 Hz - ideal move	7.91	26.00	8.32	24.25
Shape 30 Hz - 180 Hz cam	0.16	0.63	0.71	2.91
Shape 30 Hz - ideal move	8.14	25.73	8.65	19.76
Shape 10 Hz - 180 Hz cam	3.56	20.33	3.40	16.16
Shape 10 Hz - ideal move	8.83	24.42	9.28	15.89
Shape 5 Hz - 180 Hz cam	12.65	124.06	2.66	15.55
Shape 5 Hz - ideal move	30.71	113.25	9.00	20.57

IV. DISCUSSION

These experiments confirmed the assumption that data captured with the eye-tracking camera can be expanded using additional samples gathered by a faster optical sensor.

Some older studies state that the minimum sampling frequency to record saccades should be 300 Hz [13], but newer research shows that a 50-60 Hz sampling rate is sufficient [14]. The additional data acquired from the optical sensor (which can be set up to 3 kHz) should be sufficient for measurement of both fixations and saccades. The system developed here can be used in applications where a low computing time is required, e.g., in systems containing many peripherals that involve computational complexity [15].

Analytical study of the data obtained by means of the dual sensor system gives the expected results. The eyeball movement direction changed 9 times per second. When the camera sampling rate was three times larger than the signal variability, there was no new information gain after combining the data. When the signal fluctuates fast enough when compared to the camera sampling rate, the information gain increases. This can be used to find the optimal usage of computational resources with an acceptable quality of eye tracking in any mobile device.

The additional value of this system emerges from the first and last experiments conducted (Fig. 6 and Fig. 11). In these recordings, the camera eye-tracking algorithm shows small fluctuations or artefacts. In contrast, the optical sensor during the same timestamp shows no distortions. This feature can be used as redundancy and may result in increased reliability of the system.

V. CONCLUSIONS

The study performed here shows that it is possible to extend the number of pupil positions captured by the camera based on data provided by a much faster optical sensor. A high level of similarity between the expanded and reference datasets proves that the proposed algorithm is correct and reliable. The algorithm developed here may influence the implementation of the eye-tracking procedures utilised in wearable electronic devices such as smart glasses. The possibility of extending the number of pupil positions using the optical sensor allows for a significant reduction in the

eye-tracking camera sampling rate, and thus enables reduction of the required computational power and power consumption.

REFERENCES

- [1] K. Czuszynski, J. Ruminski, T. Kocejko and J. Wtorek, "Septic safe interactions with smart glasses in health care," in Engineering in Medicine and Biology Society, Milan, 2015, pp. 1604-1607, 10.1109/EMBC.2015.7318681
- [2] T. Kocejko, A. Bujnowski and J. Wtorek, "Eye-mouse for disabled," in Human-Computer Systems Interaction, Springer Berlin - Heidelberg, 2009, pp. 109-122, 10.1109/HSI.2008.4581433
- [3] T. Kocejko, A. Bujnowski and J. Wtorek, "Complex human computer interface for LAS patients," in Human System Interactions, IEEE Press, pp. 272-275, 10.1109/HSI.2009.5090991
- [4] Z. Yücel, A. Salah and C. Mericli, "Joint attention by gaze interpolation and saliency," in Cybernetics vol. 43, 2013, pp. 829-842, 10.1109/TSMCB.2012.2216979
- [5] B. R. Pires, M. Devyver and A. Tsukada, "Unwrapping the eye for visible-spectrum gaze tracking on wearable devices," in Applications of Computer Vision, 2013, pp. 369-376, 10.1109/WACV.2013.6475042
- [6] F. Li, S. Kolakowski and J. Pelz, "Using structured illumination to enhance video-based eye tracking," in Image Processing, ICIP, San Antonio, 2007, pp. 373-376, 10.1109/ICIP.2007.4378969
- [7] M. Mehrubeoglu, L. M. Pham and H. T. Le, "Real-time eye tracking using a smart camera," in Applied Imagery Pattern Recognition Workshop (AIPR), IEEE, Washington, 2011, pp. 1-7, 10.1109/AIPR.2011.6176373
- [8] A. Al-Rahayfeh and M. Faezipour, "Enhanced frame rate for real-time eye tracking using circular Hough transform," in Systems, Applications and Technology Conference, Farmingdale, 2013, pp. 1-6, 10.1109/LISAT.2013.6578214
- [9] D. Kim, J. Cho and S. Lim, "A 5000S/s Single-Chip Smart Eye-Tracking Sensor", in: Solid-State Circuits, ISSCC 2008, pp. 46-59, 10.1109/ISSCC.2008.4523049
- [10] D. Beymer and M. Flickner, "Eye gaze tracking using an active stereo head," in Computer Vision and Pattern Recognition vol.2, Madison, USA, 2003, pp. 451-458, 10.1109/CVPR.2003.1211502
- [11] W. M. King, S. G. Lisberger and A. F. Fuchs, "Oblique Saccadic Eye Movements of Primates," in Journal of Neurophysiology vol. 56, 1986, pp. 769-784,
- [12] ICNIRP guidelines on limits of exposure to laser radiation of wavelengths between 180 nm and 1,000 µm, in Health Physics 105(3), 2013, pp. 271-295.
- [13] M. Juhola, V. Jantti and I. Pykko, "Effect of sampling frequencies on computation of the maximum velocity of saccadic eye movements", in Biological Cybernetics, Springer-Verlag, 1985, pp. 67-72, 10.1007/BF00337023
- [14] R. Wierst, M. Janssen and H. Kingma, "Measuring Saccade Peak Velocity Using a Low-Frequency Sampling Rate of 50 Hz", in Biomedical Engineering, 2008, pp. 2840-2842, 10.1109/TBME.2008.925290
- [15] T. Kocejko, J. Ruminski, J. Wtorek and B. Martin, "Eye tracking within near-to-eye display," in Human System Interactions, Warsaw, 2015, pp. 166-172, 10.1109/HSI.2015.7170661