Proceedings of the 21st International Conference on Global Research and Education (Inter-Academia 2024)

A Human Eye Model for Investigating the Metrological Properties of Video Eye Trackers

A. Ostaszewska-Liżewska[a,*], A. Pakuła[b] and K. Szykiedans[b]

[a] *Faculty of Mechatronics, Institute of Metrology and Biomedical Engineering, Warsaw University of Technology, św. A. Boboli 8, 02-525 Warsaw, Poland*
[b] *Faculty of Mechatronics, Institute of Micromechanics and Photonics, Warsaw University of Technology, św. A. Boboli 8, 02-525 Warsaw, Poland*

Doi: 10.12693/APhysPolA.146.623 *e-mail: anna.ostaszewska@pw.edu.pl

This article presents the development and testing of a model of the human eye for use in video eye-tracking systems. The research focuses on creating a universal model to simulate eye features such as pupil diameter, eyelid shape, and iris colour and generate the necessary pupil and corneal reflections required for accurate eye-tracking measurements. The study involves modifying existing models, particularly OEMI-7, to enhance its compatibility with different eye-tracking methods. Experimental validation was conducted using a robotic manipulator to simulate eye movements, showing the significance of the model for testing and standardising eye-tracking systems across various applications.

topics: video eye tracking, artificial eye, accuracy and precision, OEMI-7

1. Introduction

1.1. Eye-tracking methods and application

An eye tracker (oculograph) is a device that enables the measurement and recording eye movements to compute pupil position, fixations (gaze point), saccades and velocity [1]. Some eye trackers enable the measurement and recording of pupil size or blink rate. Currently, the most popular are non-contact techniques, based solely on analysing images from one or more cameras, referred to as *video-oculography* (VOG). The eye position is calculated using feature-based, appearance-based, or event-based methods. Feature-based methods involve tracking selected attributes or markers from the image, such as the pupil's edges, the corners of the eyes or pupil and corneal reflections. In order to facilitate feature extraction, *infrared* (IR) diode illumination and infrared cameras are often used. Appearance-based methods focus on the direct analysis of the entire image of the eye without extracting individual features. They often use machine learning techniques, especially deep learning, to recognise patterns directly from the raw image data. Event-based methods are a cutting-edge area of research. These methods leverage the unique capabilities of event (neuromorphic) cameras to capture rapid changes in pixel intensity at high temporal resolution and with low latency [2]. Hybrid methods are also used to improve the accuracy and precision of eye tracking. They are a combination of selected methods from those described. Ongoing work is on new eye-tracking methods aimed at miniaturising the equipment, increasing temporal and spatial resolution, and improving accuracy and precision. Development in this field expands the areas of eye-tracking applications. Currently, eye trackers are used to facilitate interaction in gaze-controlled systems within human-machine interfaces [3], such as those dedicated to people with disabilities, or to enhance immersion in video games [4] or safety systems [5]. Eye trackers are also employed in research ranging from consumer and marketing studies to psychological [6] and psychomotor assessments [7]. As technology advances, eye trackers offer increasingly higher temporal and spatial resolution, enabling their application in medicine [8], especially in neurological diagnosis [9]. As application areas expand, the requirements for measurement quality also increase.

1.2. Eye tracker tests and validation

To select an eye tracker for a given application [10] or to validate it during development work [11], it is necessary to determine its metrological parameters on the way of conducted test series [1]. Studies involving observers are subject to errors caused by involuntary eye movements during fixation (tremor, microsaccades, drift), natural

individual differences in eye structure (either congenital or resulting from surgery), facial structure (drooping eyelids, dark eyelashes), iris colour, visual correction (glasses or contact lenses), and individual patterns of motor reflexes during image analysis [12].

1.3. Human factor elimination

To eliminate the influence of factors related to variance in individual characteristics and the inability to perform repeatable eye movements, a common practice is to conduct measurements using dedicated test setups that replace human participants. Such studies are primarily conducted by manufacturers during research and development of specific products [12] and by independent researchers who examine the characteristics of commercially available devices or those they are developing [13]. Manufacturers typically do not disclose the details of their research, and the setups described by researchers are usually dedicated to the specific method being investigated. There is a lack of a universal measurement station that allows testing devices that use different video eye-tracking methods.

To ensure the execution of a repeatable and reproducible measurement procedure, enabling its standardisation and inter- and intra-laboratory comparison, it is necessary to develop a universal automated setup that allows measurement using any of the methods employed. The key component of such a setup would be a human-eye model. This article presents a proposed model of the human eye that reproduces the eye characteristics and physical phenomena currently used in video eye tracking to calibrate, test, and validate eye-tracking systems without the need for human intervention.

2. Universal artificial eye

2.1. Assumptions

The construction and materials used to create the eye model should be compatible with the detection and tracking method used by the examined eye tracker. Appearance-based methods have the lowest requirements because the algorithm will recognise as an eye a sphere with a flat image of the iris and pupil on a white background, allowing the IR beam from the external surface to be reflected. This can be achieved using a wide range of market offers aimed at ophthalmology training such as Silicone Human Eye Replica Gulden Ophthalmics [14]. Providing eyelids with eyelashes will also give the possibility to use some feature-based methods, which utilise artefacts such as eye corners, dark pupil disc, points of intersection of the eyelid with the iris disc,

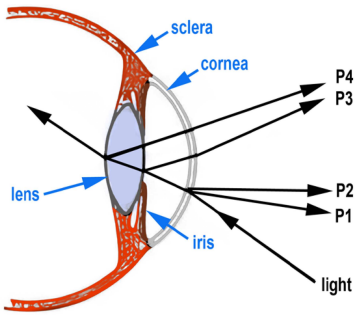

Fig. 1. Purkinje reflections: P1 (also known as glint, rightest, from the anterior corneal surface), P2 (from the posterior corneal surface), P3 (from the anterior lens surface), P4 (from the posterior lens surface).

etc. However, for most of the latter methods and event-based methods, pupil and corneal reflections will be necessary to get bright and black pupil-effect and Purkinje reflections.

The bright pupil effect occurs when the light source is aligned with the eye's optical axis. The light then reflects off the retina, resulting in an easy-to-track, bright, high-contrast image of the pupil, regardless of the iris colour. Otherwise, the dark pupil effect is visible with the light reflections from various eye tissue layers, known as the Purkinje reflections shown in Fig. 1.

Although reflections are also created in natural light, most eye trackers are based on eye features and are equipped with near-infrared illuminators and IR cameras (750 nm). This wavelength of light is invisible to humans, so it does not disturb their natural behaviour. At the same time, it provides sufficiently high contrasts and allows measurement regardless of the ambient lighting conditions. The eye position is calculated based on selected features, e.g. the pupil centre position and P1 (glint). The distance between these two points is mapped to the gaze direction and expressed in angular values or pixels representing the fixation point on the observed screen. Modern eye trackers can switch (often automatically) between dark and bright pupil detection modes, depending on environmental conditions. The simultaneous use of both the bright and dark pupil effects is used in event-based vision, where events in the pupil area are actively generated by alternately blinking two near-infrared illumination sources [2].

Therefore, the universal artificial eye for feature and event-based methods cannot be an opaque sphere of uniform material with just an image of the eye. Such a model must reproduce the spatial distribution of materials with different refractive indices so that pupil–corneal reflections are produced both in natural and IR light and change their mutual positions with eye-ball rotation, mimicking the human organ.

2.2. Consumer-ready artificial eyes

No comprehensive eye model is available for testing, developing, and validating eye-tracking techniques, irrespective of the specific eye detection and tracking algorithm used. Several eye-tracking equipment manufacturers offer their eye models, such as Tobii (Fig. 2a), SR Research EyeLink (Fig. 2b), and SMI SensoMotoric Instruments (Fig. 2c).

While these devices are readily adaptable, their design is relatively simple, as they consist of a reflective surface that generates a glint (P1) and a dark circle inside the artificial eye, miming the contrast observed between the iris and pupil when the dark pupil effect is employed. Nevertheless, it is impossible to generate the bright pupil effect. Therefore, these materials replicate only fundamental reflective characteristics, and their attributes are fine-tuned to the specific system provided by the manufacturer. Disassembling these structures to facilitate simple adaptations, such as the pupil size or iris colour changes, is not feasible. A more accurate depiction of the eye's internal structure is provided by phantoms specifically designed for ophthalmologists' training. However, they do not deliver a realistic external appearance, which is necessary for compatibility with appearance-based techniques. An exemplary model is the WHOLE Model Eye (Rowe Technical Design, Inc.) [15], which operates without the sclera surrounding the iris. Another example is the Retinoscopy Practice Eye [16], which does not include the iris or sclera.

In this context, the Ocular Instruments OEMI-7 artificial eye [17] is notable for including an anterior chamber, a crystalline lens, a reflecting retina, a sclera, and an iris. By its intricacy, it closely emulates the structure of the human eye, enabling the capture of all Purkinje reflections and the distinct bright and dark pupil effect (Fig. 3).

2.3. OEMI-7 disparities

Despite its advantages, the OEMI-7 model necessitates specific modifications due to its original purpose of ophthalmic imaging training. Moreover, significant differences from the human eye affect reliability in eye-tracking applications. Preliminary experiments performed with the Gazepoint GP3 HD eye tracker reveal that OEMI-7 lacks the expected attributes of a human eye. It produces a different picture than the human eye (Fig. 4), to the point where the eye tracker fails to identify it as a human eye.

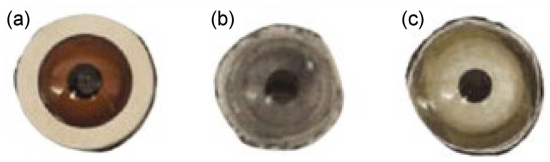

Fig. 2. Ready-to-use eye models: (a) Tobii, (b) SR Research (EyeLink) and (c) SMI (SensoMotoric Instruments.

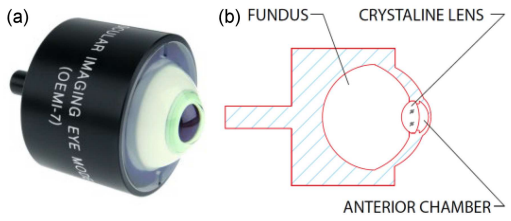

Fig. 3. OEMI-7 Ocular Instruments artificial eye: (a) a general view, (b) the cross-section [17].

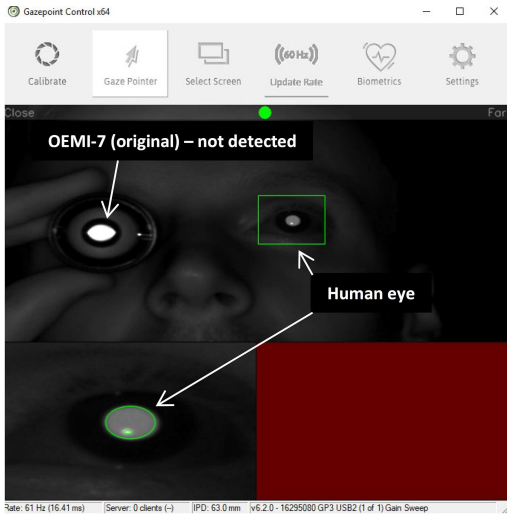

Fig. 4. OEMI-7 in comparison with the human eye, Gazepoint Control software interface.

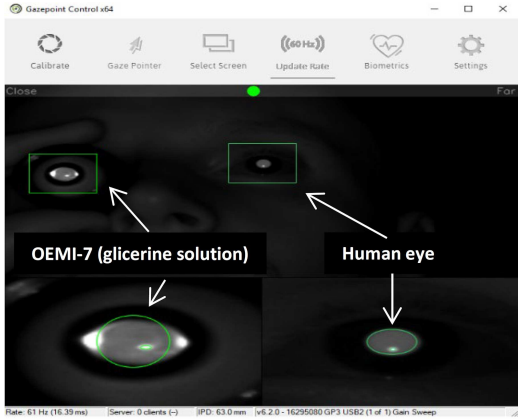

Fig. 5. OEMI-7 modified with a glycerine solution filling in comparison with the human eye, Gazepoint Control software interface.

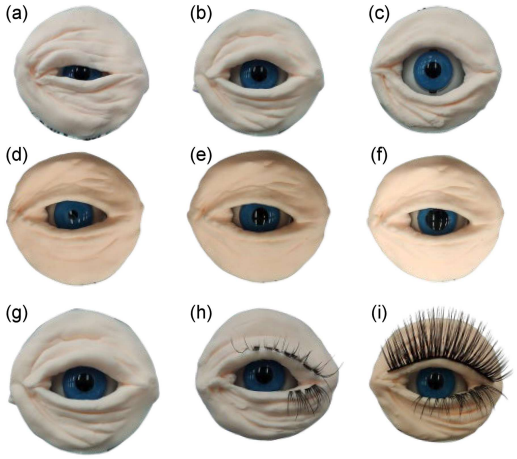

Fig. 6. Modified artificial eyes set: eyelid (a) drooping, (b) normal, (c) wide-opened; pupil diameter: (d) 3 mm, (e) 5 mm, (f) 7 mm; eyelashes (g) none, (h) normal, (i) long.

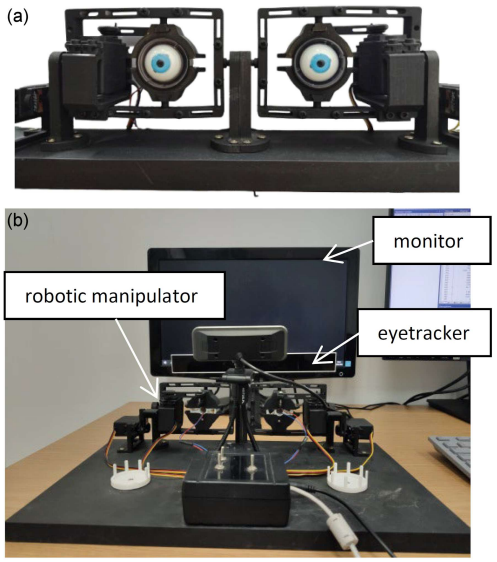

Fig. 7. The testing platform (a) robotic manipulator, (b) test setup.

Several factors, such as the refractive index, reflectance of internal structures, and pupil diameter, contribute to this condition. The initial material used to manufacture the lens of the OEMI-7 bears a refractive index of 1.485, surpassing the refractive index of the lens found in the human eye (about 1.406) [18]. The artificial lens exhibits a reflectance from its back surface that is 4.8 times greater than a genuine human lens [19]. Additionally, the pupil's diameter was set to 7 mm, exceeding the average size of a human pupil under standard room lighting conditions (about 5 mm). Other limitations not observed in GP3 trials that could restrict the model's functionality include the absence of eyelids providing a realistic appearance and the inability to choose the iris colour — the only available one is brown.

2.4. OEMI-7 modification

To adjust the refractive index of distilled water in the anterior and posterior chambers of OEMI-7, it was substituted with a 60% glycerin–water solution to achieve a refractive index that is more comparable to that of a human eye (1.413). The source of the excessively intense reflection from internal structures was the fluorescent layer covering the element, imitating the retina. In order to maintain the integrity of the original OEMI-7 file, a cornea model was generated using the Autodesk Fusion software and produced using PLA 3D printing technology on a PRUSA MK4 machine [20]. The described improvements enabled the video eye tracker to be calibrated and operated without failure (Fig. 5).

To facilitate a comprehensive comparative analysis of eye trackers, variations among individuals in the Caucasian population were identified that may impact the degree of accuracy and precision in the measurement. The selection criteria included iris colour (brown, blue, green), eyelid shape (normal, drooping, wide open), and eyelash type (none, normal, thick). In consideration of the inherent response of the pupil to lighting conditions during measurement, the diameter of the pupil was also considered, selecting values of 3 mm (eye in a bright room), 5 mm (optimal lighting), and 7 mm (in the absence of light). Irises of three different colours and three different pupil diameters were manufactured using the silicone mould casting technique. Eyelid variations were fabricated using polymer modelling clay, while eyelashes were created using preconfigured semi-finished products. Thus, each of the noted characteristics was manifested at three levels (Fig. 6).

2.5. Experiment plan

The experiment aimed to validate the proper functioning of the modified eye model by calibrating the eye tracker and collecting a series of measurements of fixation points and pupil diameter. The objective was to assess the percentage of data loss, point of gaze and pupil diameter measurement precision for the entire monitor plane. From the entire set of possible configurations, the following were selected for pilot testing: pupil diameter of 5 mm, blue iris and two types of eyelids without eyelashes, i.e., normal (Fig. 6e) and drooping (Fig. 6a), to check the effect of its shape on the results.

The testing platform was set up with a 1350 px × 780 px monitor. The GP3 HD eye tracker was positioned following the manufacturer's guidelines,

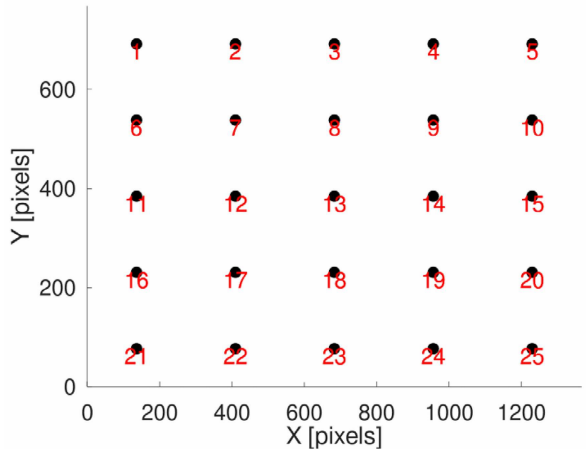

Fig. 8. The consecutive stimuli matrix.

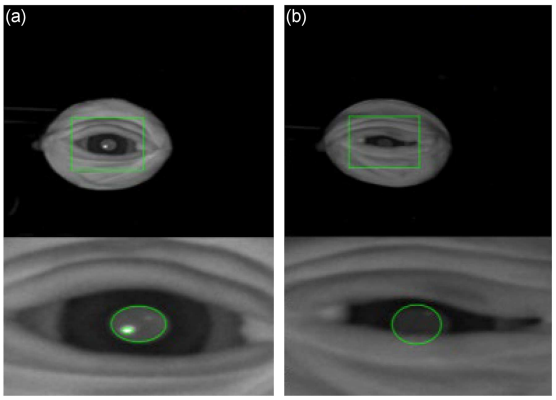

Fig. 9. OEMI-7 with modified retina element and glycerine solution filling.

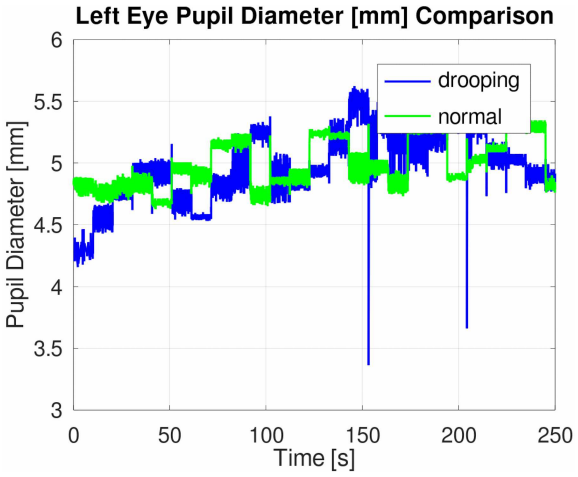

Fig. 10. The results of pupil diameter measurements during the study for normal and drooping eyelids.

directly below the edge of the monitor. The artificial eyes were rotated using a robotic manipulator [21], allowing for precise positioning of the eyes with a standard deviation of ±0.5 degrees (Fig. 7).

The eyes were positioned 0.8 m away from the plane of the monitor and eye tracker. The sampling rate of the eye tracker was configured at 60 Hz.

A set of 25 uniformly spaced points on the screen served as the stimulus. Successive data points were shown in the sequence indicated in Fig. 8, with each point lasting 10 s, resulting in a total test duration of 250 s. The ocular models were manipulated to mimic the visual attention of a human observer towards the consecutively shown stimuli. In accordance with the manufacturer's recommended procedure, the eye tracker was calibrated at 25 points with the same position as the stimuli prior to the measurements. Subsequently, the measurements were conducted using triplicate repetition.

3. Results and discussion

After the described modifications, the artificial eye looked like a human organ in the infrared camera and was detected without any problems immediately after activating the eye tracker (Fig. 9).

As the artificial eyelid did not reproduce blinks, the percentage of invalid measurements (data loss) was caused by two factors, namely the lack of compatibility of the model with the eye tracker and the influence of the eyelid shape. In the case of the normal eyelid, the percentage of data loss was 1.34%, and for the drooping eyelid, it reached the value of 7.07%. The eyelid obstructing a part of the pupil also caused a problem in measuring its diameter. The graph (Fig. 10) displays the recorded empirical values of the pupil diameter throughout the investigation. The standard deviation of these values for the drooping eyelid was 0.26 mm, while for the normal eyelid it was 0.20 mm. Furthermore, the impact of the eye's rotation angle on the measurement error is evident. Regarding the drooping eyelid, there were also documented gross errors, surpassing 2.6 mm against a pupil diameter of 5 mm.

Furthermore, the sagging eyelid affected the accuracy of the fixation point measurement. Figure 11 illustrates the evident disparity in the dispersion of measurement outcomes between normal eyelid (Fig. 11a) and the drooping eyelid (Fig. 11b), with the former exhibiting a far greater magnitude.

To provide a clear representation of the precision of fixation points measurements, confidence ellipses for a 95 % confidence level were plotted (Fig. 12) according to

$$(x-\mu_x)^2\left(\frac{1}{\sigma_x^2}\right)+2\rho\frac{(x-\mu_x)(y-\mu_y)}{\sigma_x\sigma_y}+(y-\mu_y)^2\left(\frac{1}{\sigma_y^2}\right)=\chi^2_{p,2}. \quad (1)$$

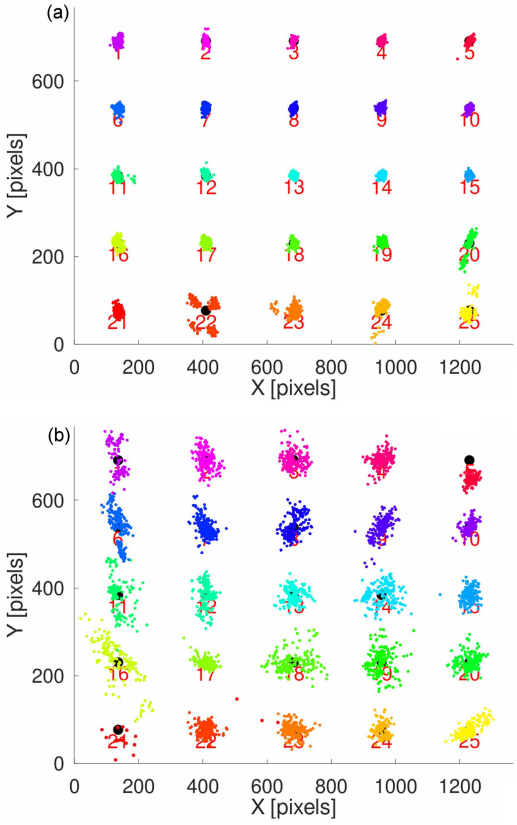

Fig. 11. The scatter of measurement results for (a) a normal and (b) a drooping eyelid.

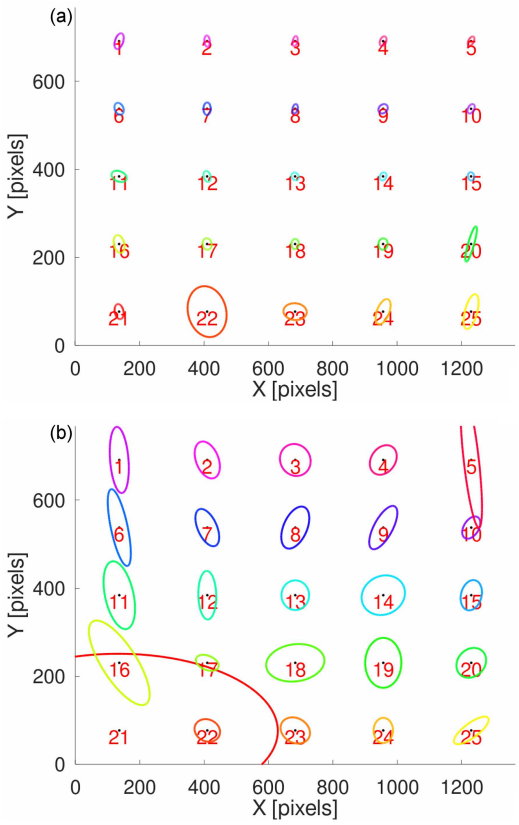

Fig. 12. The confidence ellipses for measured points of fixation, 95% confidence level, (a) a normal and (b) a drooping eyelid.

Here, μ_x and μ_y are the means of x and y; σ_x^2 and σ_y^2 are the variances of x and y; ρ is the correlation coefficient between x and y; $\chi^2_{p,2}$ is the ch-squared value for a given confidence level p with 2 degrees of freedom (here, $\chi^2_{0.95,2} \approx 5.99$ for 95% confidence).

As expected, the measurement precision is most excellent for stimuli located in the centre section of the screen and diminishes for its edges, particularly the lower one. The manifestation of this phenomenon is particularly evident in the sagging eyelid.

Concisely, the use of the artificial eye enabled the management of variables that might impact the accuracy of the measurement and confirmed the magnitude and characteristics of this impact.

4. Conclusions

The objective of this work was to effectively adapt the OEMI-7 artificial eye model to facilitate the evaluation and verification of video eye-tracking systems. By implementing a glycerin solution to regulate the refractive index and the reconfiguration of internal components, the artificial eye was able to replicate human eye properties and generate precise Purkinje reflections more effectively. Quantitative tests demonstrated that the Gazepoint GP3 HD eye tracker successfully detected the artificial eye. The data loss rate for the normal eyelid configuration was only 1.34%, while it was much higher at 7.07% for the drooping eyelid configuration.

The eyelid's conformation also influenced pupil diameter measurements' accuracy, as the drooping eyelid resulted in a standard deviation of 0.26 mm, in contrast to 0.20 mm for the normal eyelid. Significantly, gross errors of 2.6 mm or more were observed for the drooping eyelid. Concerning the accuracy of measuring fixation points, confidence ellipses produced at a 95% confidence level indicated that the drooping eyelid caused more significant variation in measurement results than the normal eyelid, especially at the screen edges.

In summary, these findings validate the modified artificial eye for evaluating different eyelid shapes and their influence on the performance of eye-tracking systems.

This work has led to the creation of a mock-up enabling testing of the influence of three factors, such as eyelid shape (normal, wide open, drooping), type of eyelashes (none, normal, long) and pupil diameter (3 mm, 5 mm, 7 mm), each factor at three levels. With a robotic station enabling programmed movement, the artificial eyes enable standardised

comparative tests and validation of different video eye tracker models. Hence, this study has established the basis for future investigations that will include more intricate arrangements, such as further eyelash type and pupil diameter diversifications.

Acknowledgments

The research was carried out within the project no. 504/04809/1142/44.000000 supporting the conduct of interdisciplinary scientific activities within the framework of funds allocated for Warsaw University of Technology from the subsidy of the Ministry of Science and Higher Education in 2023–2024.

References

[1] K. Holmqvist, M. Nyström, R. Andersson, R. Dewhurst, H. Jarodzka, J. van de Weijer, *Eye tracking: A Comprehensive Guide to Methods and Measures*, Oxford University Press, 2011.

[2] T. Kagemoto, K. Takemura, in: *Adjunct Proc. of the 36th Annual ACM Symposium on User Interface Software and Technology (UIST'23 Adjunct)*, Association for Computing Machinery, New York 2023.

[3] M. Barz, F. Daiber, A. Bulling, in: *Proc. of the 9th Biennial ACM Symp. on Eye Tracking Research & Applications (ETRA'16)*, Vol. 14, Association for Computing Machinery, New York 2016, p. 275.

[4] V. Sundstedt, *Gazing at Games: An Introduction to Eye Tracking Control*, Vol. 14, Springer, Cham 2012.

[5] O. Gordieiev, V. Kharchenko, O. Morozova, O. Illiashenko, M. Gasanov, *Int. J. Saf. Secur. Eng.* **11**, 361 (2021).

[6] J.M. Henderson, S.V. Shinkareva, J. Wang, S.G. Luke, J. Olejarczyk, *PLoS One* **8**, e64937 (2013).

[7] M.L. Mele, S. Federici, *Cogn. Process.* **13**, 261 (2012).

[8] K. Harezlak, P. Kasprowski, *Comput. Med. Imaging Graph.* **65**, 176 (2018).

[9] T.J. Anderson, M.R. MacAskill, *Nat. Rev. Neurol.* **9**, 74 (2013).

[10] G. Funke, E. Greenlee, M. Carter, A. Dukes, R. Brown, L. Menke, *Proc. Hum. Factors Ergon. Soc. Annu. Meet.* **60**, 1240 (2016).

[11] A.D. Barsingerhorn, F.N. Boonstra, J. Goossens, *Behav. Res. Methods* **50**, 2480 (2018).

[12] Tobii® Technology, Accuracy and precision test method for remote eye trackers Test Specification, 2011.

[13] J. Brand, S.G. Diamond, N. Thomas, D. Gilbert-Diamond, *Behav. Res. Methods* **53**, 1502 (2021).

[14] Gulden Ophthalmics, Silicone Human Eye Replica (Blue Iris), 2024.

[15] Rowe Technical Design, WHOLE Model Eye Rowe Technical Design, 2024.

[16] Gulden Ophthalmics, Retinoscopy Practice Eye, 2024.

[17] Ocular Instruments, Ocular Imaging Eye Model, 2021.

[18] H. Gross, F. Blechinger, B. Achtner, in: *Handbook of Optical Systems*, Vol. 4, 2008, Ch. 36.

[19] D. Wang, F.B. Mulvey, J.B. Pelz, K. Holmqvist, *Behav. Res. Methods* **49**, 947 (2017).

[20] Prusa Research, Original Prusa MK4S 3D Printer, 2024.

[21] A. Ostaszewska-Lizewska, in: *2022 IEEE 11th Int. Conf. on Intelligent Systems (IS)*, IEEE 2022.