Towards Invisible Eye Tracking with Lens-Coupled Lateral Photodetectors

Daniele M. Crafa
danielemaria.crafa@polimi.it
Politecnico di Milano
Milano, Italy

Susanna Di Giacomo
susanna.digiacomo@polimi.it
Politecnico di Milano
Milano, Italy

Dario Natali
dario.natali@polimi.it
Politecnico di Milano
Milano, Italy

Carlo Fiorini
carlo.fiorini@polimi.it
Politecnico di Milano
Milano, Italy

Marco Carminati
marco1.carminati@polimi.it
Politecnico di Milano
Milano, Italy

ABSTRACT

A novel low-power and easy to integrate sensing configuration for wearable eye tracking is presented. Within the context of infrared oculography based on individual photosensors, we proposed to couple a set of photodetectors to the lateral edges of a standard lens, acting as waveguide for the IR light, instead of directing them towards the eyeball. This allows to embed the detectors in the rim, thus being fully hidden in the eyewear, invisible to the user and robustly integrated with the glasses. A preliminary setup with four photodiodes whose signals are processed by an agile two-layer neural network was realized and characterized. Here we demonstrate both experimentally and by means of simulations the feasibility of this patent-pending approach. Detected maps of light patterns respond to different impinging light orientations. An angular resolution of about 5° is achieved with only 4 individual photodetectors coupled to a thick rectangular glass lens. A larger number of detectors would provide better resolutions. The parameters of ray-tracing simulations were first adjusted to match the experimental data from a simplified geometry. Then, simulations were used to estimate the expected signals with an eye model, paving the way to a promising outlook. The combination of hardware and software solutions here presented aims at addressing the trade-off between power consumption and angular resolution in the estimation of the direction of gaze which is crucial for pervasive eye tracking.

CCS CONCEPTS

• **Hardware** → **Sensors and actuators**; • **Human-centered computing** → *Human computer interaction (HCI)*; • **Computer systems organization** → **Embedded systems**.

KEYWORDS

Pervasive eye tracking, infrared oculography, neural networks, optoelectronic sensing.

ETRA '24, June 04–07, 2024, Glasgow, United Kingdom

ACM ISBN 979-8-4007-0607-3/24/06
https://doi.org/10.1145/3649902.3656354

ACM Reference Format:
Daniele M. Crafa, Susanna Di Giacomo, Dario Natali, Carlo Fiorini, and Marco Carminati. 2024. Towards Invisible Eye Tracking with Lens-Coupled Lateral Photodetectors. In *2024 Symposium on Eye Tracking Research and Applications (ETRA '24), June 04–07, 2024, Glasgow, United Kingdom.* ACM, New York, NY, USA, 7 pages. https://doi.org/10.1145/3649902.3656354

1 INTRODUCTION

Eye tracking (ET) is a consolidating technology in specific application areas including medical [Kasneci et al. 2017], behavioral and cognitive studies, marketing and logistics, education and training [Duchowski et al. 2000], aerospace and automotive, human-machine interfacing, etc... A major classification of ET devices is between the group of wearable trackers and non-wearable ones. Here we focus on the former group, addressing, in particular, the integration of ET in smart glasses. Currently, the state of the art of commercial wearable eye trackers is based on video-oculography, i.e. on systems including infrared (IR) illumination, typically by means of arrays of LEDs, and one or more IR cameras acquiring the light reflected by the eye whose processing, especially at high rate (above 100 Hz) is power hungry, thus requiring external bulky battery packs as in Tobii Pro 3 and Invisible by Pupil Labs. The main challenge in this field is to reduce the power consumption of the sensor [Palmero et al. 2023] and of the related processing, while retaining sufficiently good tracking performance (the best commercial wearable trackers offer accuracy and precision in the order of 1°). In order to address this challenge, we present here a novel sensing concept of wearable ET system based on a set of individual photodetectors (PD). The replacement of cameras with discrete PD has been already proposed both at research level [Li and Zhou 2018] and commercial one, for instance in the MindLink by AdHawk where a MEMS micromirror projects a laser that raster scan the eye and 5 PD collect the reflected light, estimating the direction of gaze at a rate of 500 Hz. The novelty of our solution lies in the special placement of the PD. In fact, we demonstrate here the possibility to leverage the waveguiding effect of the lens by coupling PD on the edges of a standard, unmodified, lens. As shown in Fig. 1, this allows to completely hide the PD in the glasses rim. As compared to the standard approach in which PD point towards the eye, here the aesthetics of the glasses is unaltered, at the price of a reduced amount of collected light. As a comparison with the literature, the state of the art of wearable trackers is represented by a system with 8 LED and 8 PD evenly distributed around a circular

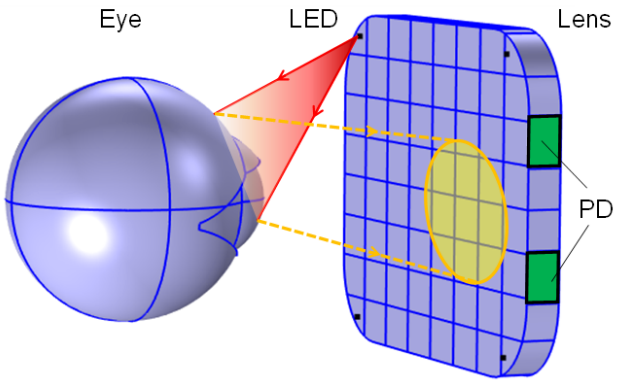

Figure 1: Proposed concept of wearable eye tracking by means of discrete IR photodetectors (PD) coupled to the edges of a standard lens acting as an optical waveguide. The domains highlighted in blue represent the geometry employed for numerical simulations.

lens and pointing towards the eye [Mastrangelo et al. 2018]. By processing 32 measurements (given by 32 different combinations of LED and PD couples) they can achieve an angular accuracy better than 1° at the price of relatively large power consumption (0.7 W if sampling at 100 Hz), bulkiness and unsightliness.

Furthermore, since the PD are now really "invisible", their number and position along the edges of the lens can be freely chosen. Here we adopt optical simulations to study the optimal placement of PD and report on the preliminary experimental validation of the approach in a simplified geometry. A simple artificial neural network was trained on a preliminary dataset acquired in this setup and used to estimate the direction of IR illumination impinging on the lens coupled to four PD.

2 PRELIMINARY SIMULATIONS AND RESULTS

Numerical simulations were performed in the COMSOL® finite-element environment, in particular using the ray tracing package. The simulated geometry is shown in Fig. 1, where the key element is clearly the eye, reflecting the IR light, placed in front of a simplified lens (a flat slab of glass of 3 mm thickness and smoothed corners) endowed with edge PD on the sides. Four LEDs emitting in the near infrared (NIR at 850 nm) with a narrow illumination cone are placed symmetrically at the four corners of the lens pointing towards the eye. A detailed optical model of the eye, including the behavior in the IR spectrum of all its tissues (of the eyeball and the surrounding skin), is extremely relevant to perform these simulations. Unfortunately, only limited information are available in the literature, where experimental data were gathered from porcine eye in the context of laser surgery [Regal et al. 2019]. We adopted the Emsley model [Emsley 1952] for the eyeball, including the eyelid, which is the most consolidated one, assuming the optical properties of the various layers as indicated in Tab. 1 where r indicates the reflectivity of the surface.

Since, in addition to the definition of boundary conditions of the materials in the simulated domain, several parameters of the simulation have to be adjusted (such as number of rays and number of reflections) we chose to begin the simulation with a simplified geometry, which was replicated in real setup physically built and used for measurements. In this way, the experimental characterization was used to simultaneously demonstrate the proof of principle of the edge readout approach and to validate simulations by tuning the parameters until the simulated values would reasonably match the measured ones.

Table 1: Optical boundary conditions for the components of the eye model

Eye Part	Eye NIR Behavior	COMSOL Surface Model
Pupil	Transparent	Pass Through
Sclera	Reflective $r \sim 0.5$	Scattering $r = 0.4$
Retina	Reflective $r \sim 0.7$	Reflective $r = 0.6$
Eyelid	Reflective $r = 0.65$	Reflective $r = 0.65$

The simplified model is pictured in Fig. 2: the lens is a rectangular slab of glass of (60 mm × 30 mm × 10 mm). Its absorbance in the IR was checked with a spectrometer (Qmini from Broadcom) and found to be negligible. An LED mounted on movable arm was used to create a directional beam of IR light impinging on the surface of the lens. Both the lens and the LED can be accurately rotated by means of goniometers. Thus, both the pitch (θ) and yaw (ϕ) angles of the beam can be controlled. The system is used to scan a grid of points on the surface of the lens. The diameter of the light spot on the surface of the lens is around 5 mm and the step size in the scan is 5°. The current signals of the four PD are conditioned by a custom-designed circuit including, for each channel, a transimpedance amplifier followed by a voltage amplifier with selectable gain and a low-power microcontroller supporting both USB and Bluetooth interfaces to stream the acquired data to a PC.

The simulations reported in Fig. 3 show the amount of light reaching the long (60 mm × 10 mm) and short (30 mm × 10 mm)

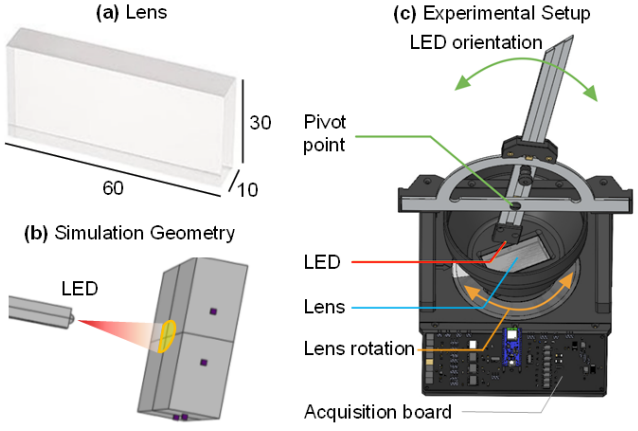

Figure 2: Test setup with a rectangular lens (a) with dimensions in mm, and a directional IR LED, simulated in COMSOL (b) and realized for the experimental validation by means of manual rotations (c) of the edge readout approach.

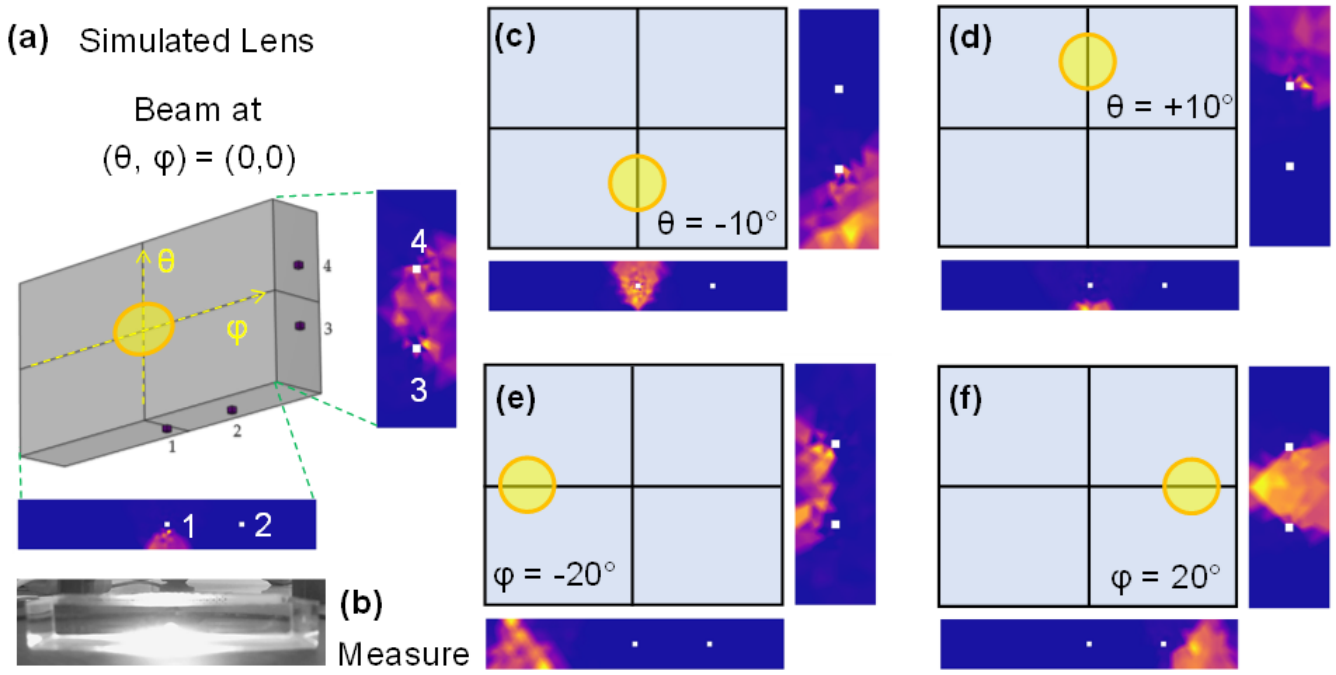

Figure 3: Simulation of the light reaching the sides of a thick rectangular lens (lateral maps) for different directions of the illumination beam: middle one (0,0) in (a) and ±10° for pitch (c,d) and ±20° for yaw (e,f).

lateral faces of the thick rectangular lens for 5 combinations of the beam angles: (0,0) which is the reference condition with the beam pointing in the middle of the lens (a), then ± 10° for the pitch (c,d) and ± 20° for the yaw (e,f). It is well evident how when changing the direction of the impinging light, the patterns of light detected on the edges change. By using an infrared camera, it was possible to image the sides of the lens observing a good qualitative matching between the simulated and the actual pattern (Fig. 3b). The parameters of the simulation were quantitatively tuned until the signals, measured by four PD (1 to 4 shown as white dots in Fig. 3) of 1 mm^2 active area and coupled to the edges by means of an optical grease confined in a 4 mm diameter, were correctly simulated for different angles of illumination. Thanks to these simulations it is also possible to find the optimal position of the PD: in our case, given the symmetry of the lens, we placed the four PD around the bottom left corner of the lens, in the points where the signal displays the largest variation during eye movements.

3 EXPERIMENTAL RESULTS

3.1 Light Maps Acquisition

By means of the experimental setup illustrated above (Fig. 2) we acquired the intensity of the light collected by the four PD while manually scanning with the beam the surface of the lens along a grid of points spaced by steps of ~5:mm. The resulting maps are shown in Fig. 4 for two conditions: (a) no modification of the lens and (b) application of a reflective coating (aluminum) along the two free sides of the lens. The light maps show sharp patterns which, as expected, are different for each PD and demonstrate the position sensitivity of the proposed approach. The color bars can be considered arbitrary units. They represent levels of the Analog to Digital Converter (ADC, with a resolution of 12 bits), with one level (least significant bit, LSB) corresponding to an input optical power of 1.8 nW.

Interestingly, it can be noted that the highest signal is recorded when the beam is pointing in the direction opposite to the location of the PD (for instance sensor 1 is placed at the bottom of the long side and the signal peak appears at the top and sensor 4 is placed on the right side while the intensity peaks on the left). This is due to the reflection paths of the light inside the glass lens. The rays reaching the detectors are mostly reflected by the opposite edge when hitting the surface of the lens at tilted angles. The addition of a reflective coating along the edges (Fig. 4b) confirms this explanation: it increases the intensity of the reflected light and introduces more regions in the patterns, as clearly visible for all PD, especially for 1 and 2. Since coating the edges of the lens with a reflective layer implies additional manufacturing, we consider in the rest of the paper the first condition, without the reflector.

3.2 Gaze Estimation Processing

In order to demonstrate the capability to reconstruct the direction of illumination from a combination of signals from only four PD, we chose to employ an artificial neural network (ANN). Among machine learning techniques, ANNs represent today a routine tool to process images and sparse data. The topology of the designed ANN is a standard fully connected one, pictured in Fig. 5a: the four inputs represent the signals acquired from the four PD. The

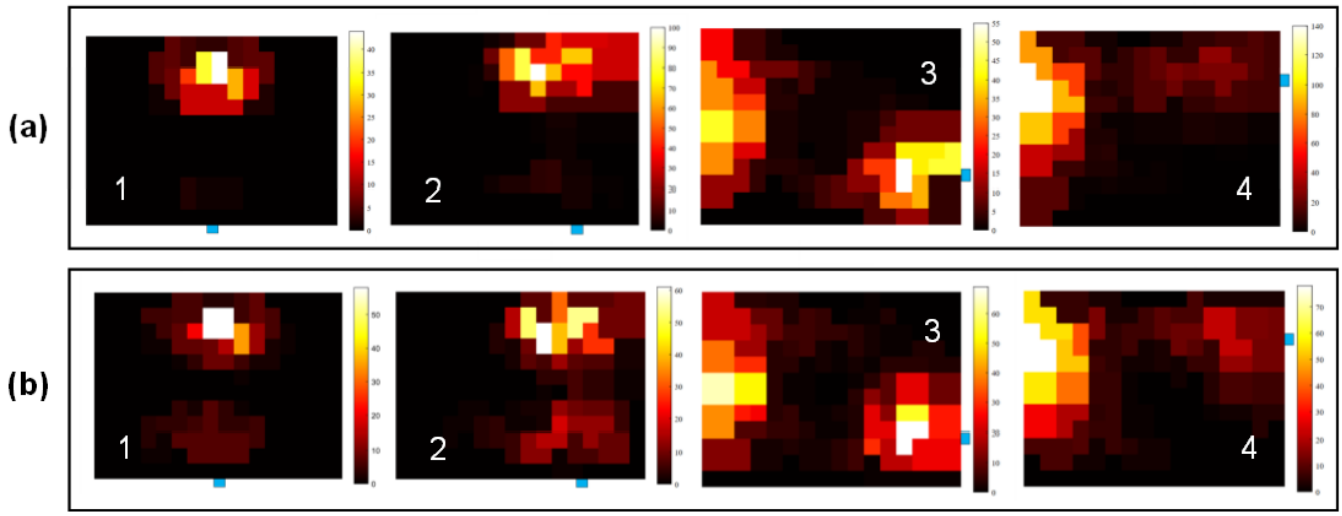

Figure 4: Planar maps showing the light intensity measured by each photodetector (1 to 4): each pixel of 5 mm size corresponds to a spot in the plane of the rectangular lens, scanned by the illumination IR beam in two conditions: (a) no modification to the lens, (b) application of a reflective layer on the free sides of the lateral edges.

hyperparameters of the ANN were optimized in Python, obtaining 2 hidden layers, with 40 and 30 neurons respectively. The output nodes represent the values of the two angles defining the direction of illumination (i.e. the direction of gaze in a bright pupil scenario). As activation functions for the hidden layers, simple rectified linear units (ReLU) were chosen, which are easily compatible with embedded implementations, for instance in low-power microcontrollers running tiny machine learning [Crafa et al. 2023; Sakr et al. 2020]. We must highlight that we opted for a very compact ANN (only 70 neurons), as compared to large convolutional neural networks (CNN) adopted to process images, also in the context of ET [Rakhmatulin and Duchowski 2020], typically running on computers with a few Watts of power dissipation, while instead we plan to embed the processing inside the smart glassed and reduce the power dissipation as much as possible (well below 1 mW, as discussed in the conclusions section).

After training the ANN with the acquired points, a fraction of the dataset was used for validation (red circles in Fig. 5). The mean absolute error is 5.7° for pitch and 13° for yaw, corresponding to root mean squared errors, normalized on the full scale range (FSR) of the angles, of 20% and 12% respectively. Figure 5b and c show two examples of dispersion of estimated points (pink stars) for two coordinates, displayed in the polar plane. Such an error is clearly larger that the accuracy offered by state-of-the-art wearable trackers (such as Pupil Labs and Tobii) typically around 1°. Two important considerations have to be made about this result. First of all, this is a preliminary result achieved with only 4 PD coupled to a very thick and flat lens and placed around only one corner. We expect that by moving to a curved and thinner lens, matching

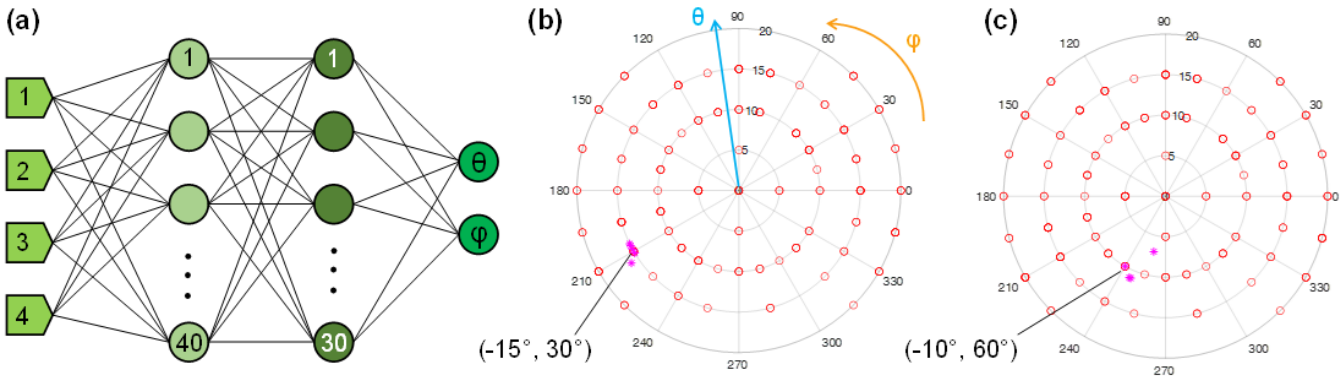

Figure 5: (a) Structure of the fully-connected neural network trained on the measured maps, (b,c) polar plots showing the validation dataset (red circles) and a few additional test points (pink stars) acquired in different experimental sessions.

the active area of the PD with its thickness, increasing the number of PD and increasing the size of a dataset used for training of the ANN, achievable, for instance, automatically by means of a motor-controlled gimbal with a synthetic eye, a better spatial resolution can be definitely achieved. Furthermore, we believe that some applications might anyway benefit from wearable ET devices offering an angular resolution of a few degrees, this enabling to target specific areas in the visual field, and which can operate in a seamless way all day long.

4 EYE RESPONSE SIMULATIONS

Thanks to the optimization of the parameters tuned with the experiments reported above, a final simulation of the full system, including the eye model, was performed in the geometry of Fig. 1. Instead of directing the IR beam on the lens as above, here the IR light illuminates the eye which reflects back towards the lens a light pattern depending of the direction of gaze. A set of simulations for a different orientation of the eye was performed. The lens is now more realistic (a flat 4 cm × 3 cm beveled slide of glass of 3 mm thickness) and 10 photodetectors (with an active area of 3 mm × 3 mm matched with the lens thickness) are placed along the edges as shown in Fig. 6 consistently with the previous tests (numbered from S1 to S10). Since here we focus on the detection capability, a single IR LED source placed in the middle of the lens was considered for simplicity. The LED optical power is 70 mW at 850 nm: given its narrow emission cone of ±13° it gives an irradiance 80 mW/cm^2. Since the LED driving will be pulsed for power saving and assuming a pulse duration of 10 μs every 1 ms (i.e. a max. sampling rate of 1 kHz), we get a max. duty cycle of 1% corresponding to an average irradiance of 0.8 mW/cm^2, compliant with the commonly assumed eye safety limit of 10 mW/cm^2. Of course, once the final shape of the lenses and glasses will be considered, further optimizations of the LED number, position, power and driving waveforms can be carried out.

A static horizontal scan of the eye was simulated with the yaw angle spanning the whole range from -40° to 40° in steps of 5°. The results are reported in the plots of Fig. 6: the horizontal axis is the yaw angle and the vertical axis the optical power detected by the ten PD. The units of the vertical scales are in W, showing power levels of tens of µW at the detector which are clearly detectable with standard photodiodes and relatively low-power readout electronics.

First of all, the symmetry of results confirms the symmetry of the geometry: couples of detectors placed in symmetrical positions with respect to the vertical symmetry axis (S3, S6 vs. S1, S8 and S4, S5 vs. S10, S9) show the same response, being horizontally mirrored as expected. Also, the middle detectors (S2, S7) show a clear symmetry with respect to positive and negative yaw values. Instead, due to the presence of the eyelid, the vertical symmetry is broken, as visible in the different signals of PD between the top and bottom halves of the lens. For each value of the angle ϕ, at least 3 different PD of the set detect different signals, suggesting that the direction of gaze can be estimated from a combination of their values, by means

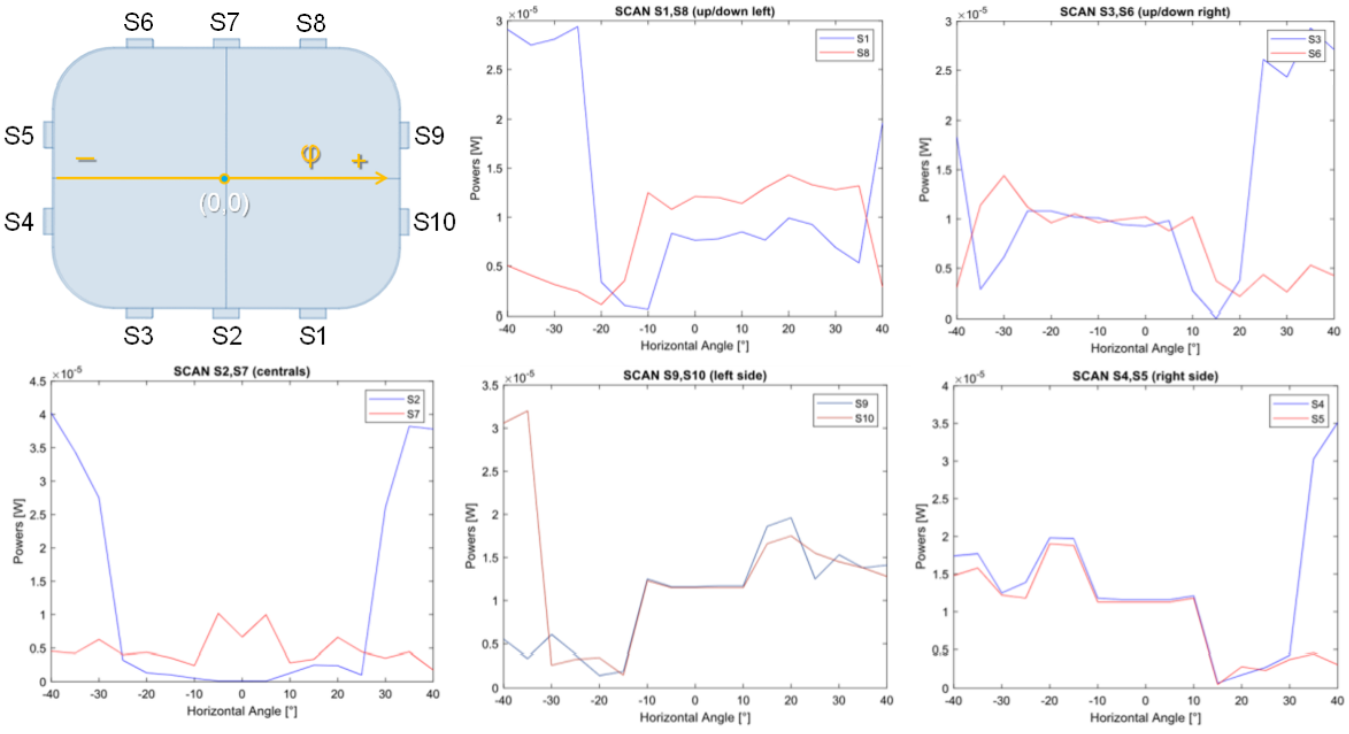

Figure 6: Simulations of the light power received during a scan of the horizontal yaw angles from -40° to +40° by 10 photodetectors symmetrically placed along the edges of the lens collecting light reflected by a model eye, illuminated by a single LED virtually placed in the middle of the lens, facing the eye. For each angle, at least 3 PD show a different signal amplitude.

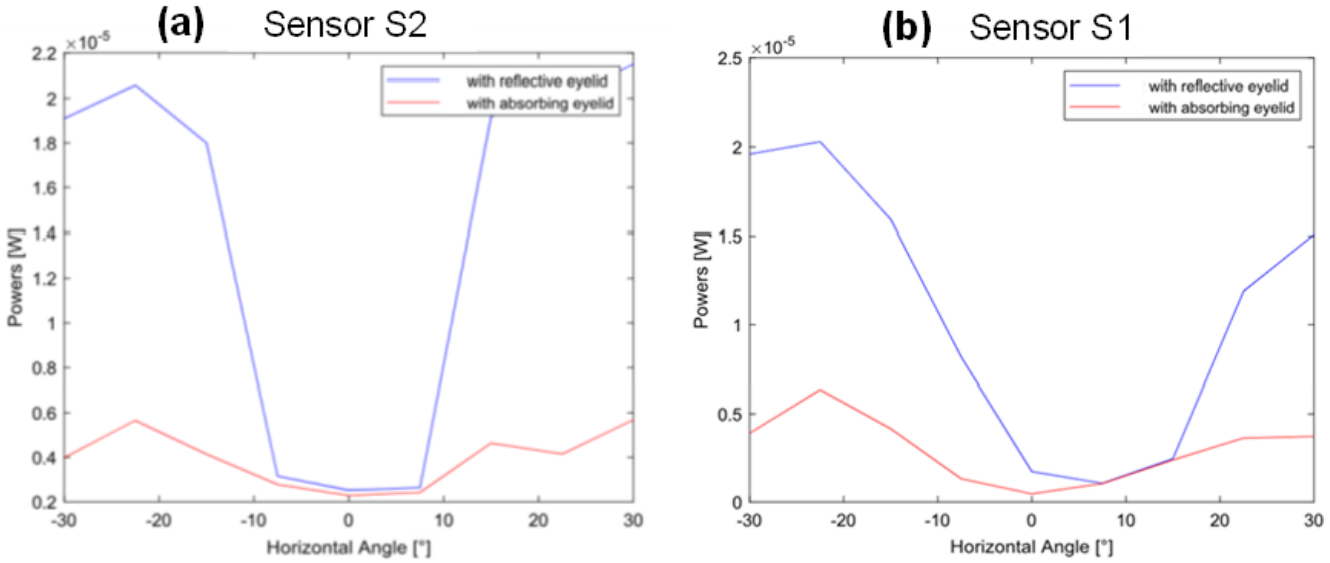

Figure 7: Impact of the eyelid reflectivity on the detected signal: comparison of simulations of reflective (blue) vs. absorbing (red) surface for two photodetectors S1 and S2 placed on the bottom side. In the case of an absorbing eyelid, the signal becomes insensitive to the eye orientation, suggesting that the reflectivity of the eyelid is crucial in our approach.

of simple processing such as look-up tables or machine learning algorithms, as the neural network illustrated above. In order to highlight the role of the eyelid, a final simulation was performed by changing the reflectivity property of the eyelid. As visible in Fig. 7, when the eyelid is changed from reflective to absorbing, the intensity measured near the extreme ends of the range (i.e. for large yaw angles) by sensors S2 and S1 drops. This confirms that, for large yaw angles, the dominant effect is due to the reflection of the eyelids.

5 CONCLUSIONS

In this work we have presented a novel optical configuration of discrete photodetectors for wearable infrared oculography. Thanks to optical coupling on the edges of the lens, leveraged as an optical waveguide operated in a reverse mode with respect to edge-lit displays, PD can be definitely hidden in the rim of the glasses, thus becoming invisible to the user. Furthermore, since the surface of the detector, attached to the lens, is no longer exposed to external environment, it results more protected against aging, dust, humidity and other sources of inaccuracy and environmental stress. The approach based on discrete PD sparsely coupled to an optical guiding medium to reconstruct its illumination by means of embedded machine learning is already successfully applied in other domains, such as the readout of scintillators with silicon photomultipliers (SiPM) for indirect detection and imaging of gamma-rays [Buonanno et al. 2020]. We have demonstrated by means of both preliminary simulations and experimental measurements (in two conditions of the edges) that this approach is also feasible for eye tracking and that the level of light detected by the PD is in the range of microWatts. This is achieved without any modification of the lens, thus paving the way to the application to any sort of lens in eyewear. Of course, it is possible to enhance the amount of detected light by machining the lens by means of additional coating layers or microlenses, further enhancing the inner light guiding effect in the infrared spectrum.

By training a standard fully-connected neural network of only 70 neurons, having the signals of only 4 PD as inputs, it was possible to reconstruct the direction of illumination on the lens with an angular accuracy of 5° and 13° along the vertical and horizontal direction respectively. Such an angular resolution would be suitable for basic functions of human/machine interfacing and gaze-activated operations in smart glasses. The compactness of the neural network allows its execution in miniaturized low-power microcontrollers: examples of wearable devices from the literature with similar computational complexity report an energy consumption of a few μJ per inference [Crafa et al. 2023], corresponding hundreds of µW with a sampling rate of hundreds of Hz, thus compatible with pervasive eye tracking. Future work includes the creation of an experimental dataset, for instance with motor-controlled ophthalmic model eyes (of increasing realism) to allow the training of a more accurate ANN followed by its validation on humans.

ACKNOWLEDGMENTS

This work was carried out in the EssilorLuxottica Smart Eyewear Lab, a Joint Research Center between EssilorLuxottica and Politecnico di Milano. We acknowledge the initial contribution of master thesis students Lodovico Lestani, Giulia Camporeale and Franscesco Guzzetta.

REFERENCES

Luca Buonanno, Davide Di Vita, Marco Carminati, and Carlo Fiorini. 2020. A Directional Gamma-Ray Spectrometer with Microcontroller-Embedded Machine Learning. *IEEE Journal on Emerging and Selected Topics in Circuits and Systems* 10, 4

(2020), 433–443.

Daniele M Crafa, Susanna Di Giacomo, Carlo Fiorini, and Marco Carminati. 2023. Neural Networks Embedded in Wearable Devices: a Preliminary Digital vs. Analog Comparison. In *2023 IEEE International Conference on Metrology for eXtended Reality, Artificial Intelligence and Neural Engineering (MetroXRAINE)*. IEEE, 587–592.

Andrew T Duchowski, Vinay Shivashankaraiah, Tim Rawls, Anand K Gramopadhye, Brian J Melloy, and Barbara Kanki. 2000. Binocular eye tracking in virtual reality for inspection training. In *Proceedings of the 2000 symposium on Eye tracking research & applications*. 89–96.

Harold H Emsley. 1952. *Visual Optics*. Hatton Press, London.

Enkelejda Kasneci, Alex A Black, Joanne M Wood, et al. 2017. Eye-tracking as a tool to evaluate functional ability in everyday tasks in glaucoma. *Journal of ophthalmology* 2017 (2017).

Tianxing Li and Xia Zhou. 2018. Battery-free eye tracker on glasses. In *Proceedings of the 24th Annual International Conference on Mobile Computing and Networking*. 67–82.

Alexander S Mastrangelo, Mohit Karkhanis, Rugved Likhite, Ashrafuzzaman Bulbul, Hanseup Kim, Carlos H Mastrangelo, Nazmul Hasan, and Tridib Ghosh. 2018. A low-profile digital eye-tracking oculometer for smart eyeglasses. In *2018 11th International Conference on Human System Interaction (HSI)*. IEEE, 506–512.

Cristina Palmero, Oleg V Komogortsev, Sergio Escalera, and Sachin S Talathi. 2023. Multi-Rate Sensor Fusion for Unconstrained Near-Eye Gaze Estimation. In *Proceedings of the 2023 Symposium on Eye Tracking Research and Applications*. 1–8.

Ildar Rakhmatulin and Andrew T Duchowski. 2020. Deep neural networks for low-cost eye tracking. *Procedia Computer Science* 176 (2020), 685–694.

Simon Regal, Denise O'Connor, Pauline Brige, Roger Delattre, Thierry Djenizian, and Marc Ramuz. 2019. Determination of optical parameters of the porcine eye and development of a simulated model. *Journal of Biophotonics* 12, 11 (2019), e201800398.

Fouad Sakr, Francesco Bellotti, Riccardo Berta, and Alessandro De Gloria. 2020. Machine learning on mainstream microcontrollers. *Sensors* 20, 9 (2020), 2638.