Pupil Tracking Under Direct Sunlight

Kamran Binaee
University of Nevada, Reno
Reno, Nevada, USA
kbinaee@unr.edu

Christian B. Sinnott
University of Nevada, Reno
Reno, Nevada, USA
csinnott@nevada.unr.edu

Kaylie J. Capurro
University of Nevada, Reno
Reno, Nevada, USA
capurro.kaylie@gmail.com

Paul R. MacNeilage
University of Nevada, Reno
Reno, Nevada, USA
pmacneilage@unr.edu

Mark D. Lescroart
University of Nevada, Reno
Reno, Nevada, USA
mlescroart@unr.edu

ABSTRACT

Pupil tracking in a bright outdoor environment is challenging due to low eye image quality and reduced pupil size in response to bright light. In this study we present research to develop robust outdoor pupil tracking without the need for shading the eyes. We first investigate the effect of camera post-processing settings in order to find values that enhance image quality for the purpose of pupil tracking under direct, oblique and overcast sunlight illuminations. We then tested the performance of the state-of-the-art pupil tracking techniques under these extreme real-world outdoor lighting conditions. Our results suggest that a key goal should be maintaining the contrast between iris and pupil to support accurate estimation of pupil position regardless of the overall eye image quality.

CCS CONCEPTS

• **Applied computing** → *Psychology*; • **Computing methodologies** → **Machine learning**; **Video segmentation**.

KEYWORDS

deep pupil tracking, outdoor eye tracking, outdoor eye tracking, eye camera settings

ACM Reference Format:
Kamran Binaee, Christian B. Sinnott, Kaylie J. Capurro, Paul R. MacNeilage, and Mark D. Lescroart. 2021. Pupil Tracking Under Direct Sunlight. In *2021 Symposium on Eye Tracking Research and Applications (ETRA '21 Adjunct), May 25–27, 2021, Virtual Event, Germany*. ACM, New York, NY, USA, 4 pages. https://doi.org/10.1145/3450341.3458490

1 INTRODUCTION

Portable eye-tracking systems provide the opportunity to investigate different aspects of human behavior ranging from low-level reflexes to high-level social behaviors. While most portable eye-trackers are designed for use in laboratories with controlled lighting condition, there has been a growing interest in data collection outside the laboratory. However, mobile eye-tracking in outdoor environments is challenging and often leads to lower-quality data due to two main reasons [Evans et al. 2012].

First, mobile eye-tracking methods rely on video oculography (VOG) to track the movements of the eye. This requires high quality images of the eyes with robust eye features such as the pupil. To this end, infrared (IR) light is typically used to illuminate the eyes, and for pupil-corneal-reflection-based gaze estimation, it allows tracking reflection(s) of IR light-emitting diodes (LEDs) across the cornea (1st Purkinje image) [Evans et al. 2012; Holmqvist et al. 2011]. The outdoor environment has an abundance of IR light that can flood the eye camera during daytime hours, leading to a drastic reduction in the eye image quality. Furthermore, environmental reflections of IR light can lead to additional Purkinje images on the eye, causing tracking to break down even when image quality is otherwise acceptable.

Second, eye-tracking in outdoor environments is challenging due to the pupillary response to light [Evans et al. 2012]. For model-based gaze estimation, robust pupil detection results in a better 3D model of the eye [Kassner et al. 2014]. In indoor environments, the pupils are often dilated, allowing the image segmentation techniques to perform well. In contrast, pupils exposed to light levels common during daylight hours constrict to a very small size. This constriction subsequently causes the image segmentation pipeline to break down, compromising overall gaze-tracking performance [Evans et al. 2012]. Furthermore, subjects tend to squint in outdoor lighting, causing the lower eyelid and eyelashes to partially or fully occlude the pupil causing an additional challenge for the eye tracking pipeline [Evans et al. 2012].

In response to some of these problems an IR blocking shade is attached or worn over the eye-tracker by the participant [Evans et al. 2012; Hausamann et al. 2020; Matthis et al. 2018; Valsecchi et al. 2020]. While suitable under certain conditions, the presence of shades may cause unnatural oculomotor or cephalomotor behaviors. For example, a subject may incline their head in order to look at an object below a shade placed over the eye-tracker. This might also cause unnatural eye movements that might not be present for instance when walking outdoors without sunglasses. Here we explore an alternative approach: optimizing the image acquisition pipeline for outdoor environments. This presents a viable solution

ETRA '21 Adjunct, May 25–27, 2021, Virtual Event, Germany

ACM ISBN 978-1-4503-8357-8/21/05…$15.00
https://doi.org/10.1145/3450341.3458490

without imposing possible demand characteristics on participants. We seek to optimize this pipeline from two perspectives. First, we report results of systematic tests to determine recommended camera settings for a variety of lighting conditions. We collected outdoor eye-tracking data across a range of camera parameters afforded by our eye tracker (Pupil Core, [Kassner et al. 2014]). Second, we characterize and report performance of three state-of-the-art pupil-tracking techniques under these conditions. These include Pupil Labs [Kassner et al. 2014], RITnet [Chaudhary et al. 2019], and DeepVOG [Yiu et al. 2019]. The last two are both deep neural network models for eye image segmentation. We then assess pupil tracking performance during an 8-minute outdoor walking session under various combinations of shade, oblique and direct sunlight.

2 RECOMMENDED CAMERA SETTINGS BASED ON PUPIL ESTIMATION ERROR

We investigate the parameters of a typical USB video device class (UVC) eye camera used in most eye tracking systems and report recommended ranges for the camera settings especially for outdoor conditions. Eye tracking data was collected under different lighting conditions including extreme cases for generating shadow and reflection artifacts. Specifically, informed by preliminary study, we recorded data under four different lighting conditions: indoors as a reference, outdoors in the shade, outdoors facing the sun directly, and outdoors facing the sun at an oblique angle. During each lighting condition, the standing participant fixated at five points on a target while every 80 ms an eye image (from left eye) was captured and stored along with its camera settings. In order to include different eye positions in the analysis, participants fixated at five points (four corners and the center of a checkerboard pattern) while the eye images were recorded for each camera setting. During a pilot study we found that gamma, brightness, and sharpness are the most important camera settings affecting the eye image quality. Therefore for each condition we considered four values for gamma, four values for brightness and three values for sharpness as shown in Fig. 1. Furthermore, we set the following fixed parameters: exposure mode = Auto, Contrast = 65, Gain = 0.

We use the residual ellipse fit error to the pupil region as the performance metric. Pupil segmentation was performed using RITnet which is reported to be robust to different illumination, eye shapes, rotation and scaling [Chaudhary et al. 2019]. We report the mean residual error for the ellipse fit to the pupil region in units of pixel. This metric serves as a proxy measure of eye image quality for each camera setting shown in Fig. 2. A successful pupil detection algorithm, typically provides a sub-pixel accuracy for the position of the pupil. For instance, As shown in Fig. 2 residual error remains below 0.3 pixels for the direct sunlight condition, when $gamma = 130$, $brightness = 0$ and $sharpness = 3$.

Our recommendations for these settings, based on the boxplot graphs in Fig. 2, give the lowest median error value with the distribution of error values most tightly clustered around it and the minimum number of outliers. The results show for different lighting conditions, different ranges of camera settings can provide the lowest pupil tracking error. For instance, the shade condition is challenging due to similar pupil and iris intensities (see supplemental figure 1-4) hence a mid-range gamma value $130 < gamma < 160$

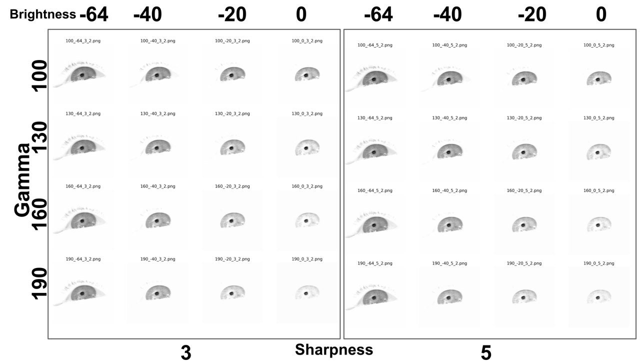

Figure 1: Sample eye images captured under direct sunlight and different camera settings for gamma and brightness. Each 4x4 sub-panel shows the eye images for a fixed sharpness value. Right panel: Pupil fit residual error for each setting and lighting condition (please zoom in for better viewing).

would stretch the pixel intensities enough to detect the pupil correctly. Whereas in oblique condition the default (minimum) value for gamma and large reduction in intensity $-60 < brightness < -40$ is required to darken the eye image for pupil detection.

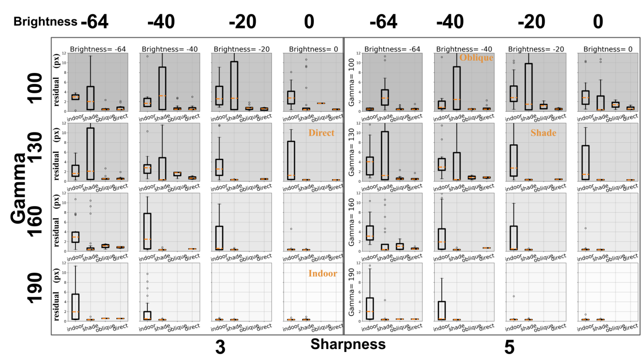

Figure 2: Pupil fit residual error for each setting and lighting condition (please zoom in for better viewing)

3 PUPIL CONFIDENCE ACROSS DIFFERENT ILLUMINATIONS & MODELS

After fine tuning the eye camera parameters for outdoor eye tracking, these settings ($gamma = 110, brightness = -30, sharpness = 3$) were used to stress-test the state-of-the-art pupil tracking pipelines. Pupil Labs [Kassner et al. 2014] uses an image processing technique that finds the location of the pupil as a disc on the surface of the eye sphere, in order to solve for the 3D position of the eye. Although this method is widely used, here we show that performance is significantly impaired for low contrast outdoor eye images. RITnet and DeepVOG, however, use deep neural network (DNN) models and they are more robust to image illumination changes since the image artifacts are usually included in their training sets [Chaudhary et al.

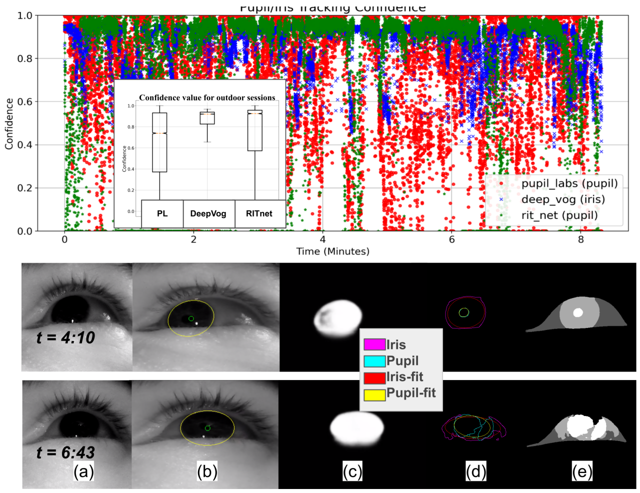

Figure 3: Top panel: confidence values in detecting the pupil or iris for an outdoor walking session: Pupil Labs (red), DeepVOG (blue), and RITnet (green). Note that for DeepVOG the confidence value is w.r.t detected iris and not the pupil. The distribution of confidence values for all four outdoor sessions is shown in the inset. Bottom panels (a) input eye image, (b) pupil detected by DeepVOG (note the iris is mistaken for the pupil due to poor contrast), (c) DeepVOG activation map, (d) RITnet output pupil and iris contour along with ellipse fits shown in color and (e) RITnet segmented eye regions (please zoom in for better viewing)

2019; Yiu et al. 2019]. It is important to note that no re-training was performed on the deepNets.

Gaze tracking data from four subjects were recorded while walking outdoors for about 10 minutes including calibration and validation at the beginning and at the end. We made sure that the subjects walked under shade, direct and oblique sunlight in order to evaluate eye image quality under realistic conditions. As shown in Fig. 3 top panel, Pupil Labs fails significantly in detecting the pupil and the DeepVOG model takes the iris as pupil for the entire duration. RITnet, however, performs better for the most part. Visual inspection of the segmented videos suggests that RITnet loses track of the pupil under extreme conditions. Pupil Labs method reports the detected pupil confidence value using a technique similar to the one explained above. However, neither RITnet and DeepVOG report the estimated pupil confidence values. In order to provide a fair comparison, we calculated the confidence values for both RITnet and DeepVOG models by Z-normalizing the residual error from ellipse fit to the detected pupil contour [Shanker et al. 1996]. Fig. 3 shows an example of the detected pupil contour and ellipse fit (which in the case of DeepVOG is the iris) for the two deepNets for one participant.

4 CONCLUSION

In this study, the space of eye camera post-processing parameters was explored in order to identify settings that improve the eye image quality for the purpose of outdoor eye tracking. Our results suggest that selection of parameters that are tailored to specific lighting conditions can preserve the eye image quality and extend the tolerance of current pupil tracking techniques. The deep networks showed strong tolerance for the change in the eye shape due to squinting that was reported in previous studies [Evans et al. 2012]. However, our findings suggest that the most critical factors seems to be the contrast between the pupil and the iris and bright reflections of the world on the iris. We could not identify a single set of parameters that works best across all illuminations, because different illuminations cause different artifacts on the eye image. For instance, the eye images under direct and oblique conditions, are visually significantly different (see supplemental figures 3 and 4), because of the difference in light source incident angle. Therefore, direct condition would require a large reduction of brightness and the oblique condition requires a large gamma to stretch the intensity difference between pupil and iris. Our results suggest the need for a technique that dynamically adapts the camera settings based on the instantaneous iris-pupil contrast in the eye image. One could use an anatomically aware technique that applies different image enhancement algorithm to different parts of the eye image. For instance, darkening the skin and eyelids while contrast-stretching the area including iris and pupil. Finally, these results suggest that retraining the existing deep networks using the image dataset presented in this study could further contribute to robust outdoor pupil tracking. This means that the deep networks will learn the similar relationship between eye parts but under extreme lighting conditions. We also note that the use of post-processing settings causes a slight decrease in frame rate, i.e. from 120 fps to 117 fps in our case.

ACKNOWLEDGMENTS

This research was supported by NSF-EPSCOR under grant number OIA-1920896 awarded to P.R.M and M.D.L.

REFERENCES

Aayush K. Chaudhary, Rakshit Kothari, Manoj Acharya, Shusil Dangi, Nitinraj Nair, Reynold Bailey, Christopher Kanan, Gabriel Diaz, and Jeff B. Pelz. 2019. RITnet: Real-time Semantic Segmentation of the Eye for Gaze Tracking. *2019 IEEE/CVF International Conference on Computer Vision Workshop (ICCVW)* (Oct 2019), 3698–3702. https://doi.org/10.1109/iccvw.2019.00568

Karen M Evans, Robert A Jacobs, John A Tarduno, and Jeff B Pelz. 2012. Collecting and analyzing eye tracking data in outdoor environments. *Journal of Eye Movement Research* 5, 2 (2012), 6.

Peter Hausamann, Christian Sinnott, and Paul R MacNeilage. 2020. Positional head-eye tracking outside the lab: an open-source solution. In *ACM Symposium on Eye Tracking Research and Applications* (Stuttgart, Germany) *(ETRA '20 Short Papers, Article 14)*. Association for Computing Machinery, New York, NY, USA, 1–5. https://doi.org/10.1145/3379156.3391365

Kenneth Holmqvist, Marcus Nyström, Richard Andersson, Richard Dewhurst, Jarodzka Halszka, and Joost van de Weijer. 2011. *Eye Tracking : A Comprehensive Guide to Methods and Measures.* Oxford University Press. http://ukcatalogue.oup.com/product/9780199697083.do

Moritz Kassner, William Patera, and Andreas Bulling. 2014. Pupil: An Open Source Platform for Pervasive Eye Tracking and Mobile Gaze-based Interaction. In *Adjunct Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing* (Seattle, Washington) *(UbiComp '14 Adjunct)*. ACM, New York, NY, USA, 1151–1160. https://doi.org/10.1145/2638728.2641695

Jonathan Samir Matthis, Jacob L Yates, and Mary M Hayhoe. 2018. Gaze and the control of foot placement when walking in natural terrain. *Current Biology* 28, 8 (2018), 1224–1233.

Murali Shanker, Michael Y Hu, and Ming S Hung. 1996. Effect of data standardization on neural network training. *Omega* 24, 4 (1996), 385–397.

Matteo Valsecchi, Arash Akbarinia, Raquel Gil-Rodriguez, and Karl R. Gegenfurtner. 2020. Pedestrians Egocentric Vision: Individual and Collective Analysis. In *ACM*

Symposium on Eye Tracking Research and Applications (Stuttgart, Germany) *(ETRA '20 Short Papers)*. Association for Computing Machinery, New York, NY, USA, Article 44, 5 pages. https://doi.org/10.1145/3379156.3391378

Yuk-Hoi Yiu, Moustafa Aboulatta, Theresa Raiser, Leoni Ophey, Virginia L. Flanagin, Peter zu Eulenburg, and Seyed-Ahmad Ahmadi. 2019. DeepVOG: Open-source pupil segmentation and gaze estimation in neuroscience using deep learning. *Journal of Neuroscience Methods* 324 (2019), 108307. https://doi.org/10.1016/j.jneumeth.2019.05.016