Precise VR Eye Tracking using Single-Shot Deflectometry

Jiazhang Wang*, Tianfu Wang**, Bingjie Xu***, Oliver Cossairt*, Florian Willomitzer****

*Department of Electrical and Computer Engineering, Northwestern University
**Department of Computer Science, ETH Zürich
***Department of Computer Science, Northwestern University
****Wyant College of Optical Sciences, University of Arizona

mailto:fwillomitzer@arizona.edu

We present a fast, accurate, and precise eye-tracking method that is based on single-shot deflectometry. Our approach estimates the gaze direction from a dense 3D reconstruction of the eye surface, including both the cornea and the sclera. The gaze estimation error is below 0.25°.

1 Introduction

A robust, fast and accurate solution to eye-tracking benefits a wide range of applications, ranging from psychology and neuroscience research, to Virtual Reality (VR). Current solutions to eye tracking can be categorized into two groups: “image-based methods” and “reflection-based methods”. The first group estimates the gaze by utilizing 2D features detected from 2D eye images, which only requires a camera and (IR) flood illumination of the eye. However, the low density of detected eye features limits the accuracy of gaze direction estimation to a few degrees. “Reflection-based methods” observe the sparse reflections of a few point light sources over the eye surface. Current so-called “glint tracking” methods sample the eye surface at $\sim$10 point source reflections. State-of-the-art reflection-based methods can reach around 1° accuracy, which makes them the preferred solution for current commercial eye trackers [1]. However, the sparsity of the captured reflection points still limits the performance of respective methods.

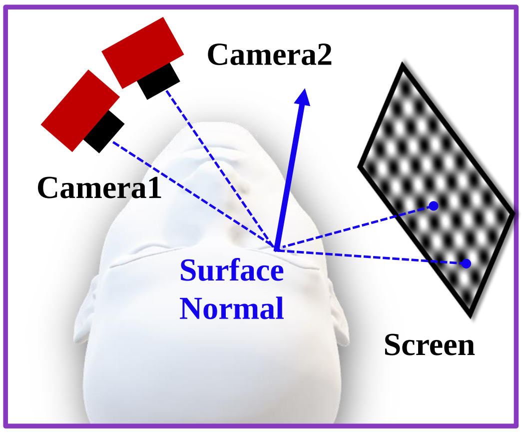

Fig. 1 *Schematic of our single-shot Deflectometry-based eye-tracking approach*

In [2, 3] we introduced a solution to eye tracking that performs a high-quality 3D measurement of the eye surface via deflectometry, and eventually estimates the gaze direction from the dense acquired surface normal and depth data. In our experiments, we can easily sample the eye surface at $3300\times$ more points than the state-of-the-art glint-tracking methods, which results in higher gaze estimation accuracy. In the following, we introduce our single-shot stereo deflectometry eye-tracking approach and our latest experimental evaluations.

2 Methods

Deflectometry is a well-established metrology principle to measure specular object surfaces with high accuracy [4]: The reflection of a screen with a known pattern is observed over the specular object surface with a camera. By analyzing the deformation of the pattern in the camera image, the corresponding pixels between screen and camera can be calculated. The inherent height-normal ambiguity problem can be solved via stereo deflectometry [4] (see Fig. 1).

Classical (stereo) deflectometry uses sinusoidal patterns paired with a four-phase-shifting method to obtain the correspondence between the screen and the camera [4]. However, this procedure requires to capture at least eight sequential camera images which makes a dynamic measurement of fast moving human eyes impossible. To enable the fastest possible acquisition, we implement a single-shot deflectometry approach by displaying a single cross sinusoidal pattern (see Fig. 1) and retrieving the correspondences (phase map) via the 2D continuous wavelet transform approach introduced in [5]. In conventional stereo deflectometry, the measurable surface area is limited by the overlap of the FoVs of both cameras. To enlarge the effective measuring area, we only use stereo deflectometry to measure a few "anchor points" in the overlap region and employ iterative surface integration methods [6] to reconstruct the whole measuring area of each camera. This delivers an accurate 3D coordinate and surface normal in each point of the measurement area (see Fig. 3).

Eventually, we estimate the gaze direction as follows: by relying on the realistic assumption that cornea and sclera of the human eye can be approximated as two spheres with different radii, we can back-trace the recovered surface normals from cornea and sclera towards the center of the eye, which delivers the centers of cornea and sclera. The gaze vector is then estimated by simply connecting the two retrieved center points (see Fig. 3). We refer to [3] for more details.

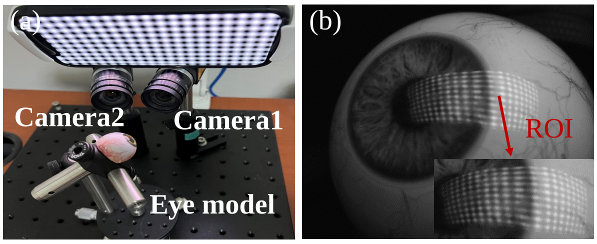

Fig. 2 *Prototype setup and sample image*

3 Experiments and Results

Our eye-tracking prototype setup consists only of off-the-shelf devices (see Fig. 2a). We use an iPhone 13 pro as display (1170 × 2532 pixels), and two machine vision cameras (FLIR fl3-u3-13s2c). The entire system is calibrated with our easy and flexible calibration method [3]. To eliminate different sources of "third-party errors" connected with the quantitative gaze estimation of real human eyes, we conduct repeatable experiments with a realistic eye model mounted on a high precision rotation stage (Fig. 2a). Figure 2b shows a sample image captured with our setup. The reflected pattern covers both the cornea and the sclera region. The evaluation result is shown in Fig. 3. The blue arrows are back-traced surface normals and the white arrow is estimated gaze. The result is overlayed with the original eye image for display purposes.

For a quantitative evaluation of our method, we rotate the eye model to 5 different agles a (-4°, -2°, 0°, 2°, and 4°) and calculate the respective gaze angle θ_a. Since the absolute gaze angle is unknown, we always calculate the *relative gaze angle* with respect to the "default" rotation angle 0° and compare the result with the “ground truth” (the angle we rotated the eye model). At each rotation position a, we perform 20 measurements, while the eye model is always moved/rotated from another position before a measurement is taken. Eventually, we evaluate the mean relative error ϵ_{0° at each rotation position a w.r.t. the 0° "default" position: $\epsilon_{0^\circ} = ||\bar{\theta_a} - \bar{\theta_{0^\circ}}| - |a - 0^\circ||$. The results are shown in Tab. 1. It can be seen that ϵ_{0° is below 0.25° for all measurements, which is a significant improvement over the current state-of-the-art. We direct the reader to [3] for more information.

Besides the discussed stereo deflectometry approach, we have also introduced an optimization-based approach that uses inverse rendering to estimate the gaze at this year’s DGaO [7, 8].

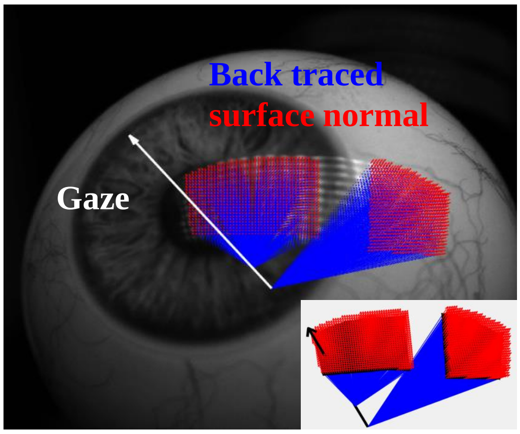

Fig. 3 *Reconstructed eye surface and estimated gaze*

Position a	-4°	-2°	0°	2°	4°
Precision σ_a	0.08°	0.06°	0.04°	0.07°	0.02°
Mean error ϵ_{0°	0.14°	0.25°	0°	0.13°	0.12°

Tab. 1 *Estimated gaze precision and mean relative error evaluation*

References

[1] I. B. Adhanom, P. MacNeilage, and E. Folmer, “Eye Tracking in Virtual Reality: a Broad Review of Applications and Challenges,” Virtual Reality pp. 1–24 (2023).

[2] J. Wang, B. Xu, T. Wang, W. J. Lee, M. Walton, N. Matsuda, O. Cossairt, and F. Willomitzer, “VR Eye-Tracking using Deflectometry,” in *Computational Optical Sensing and Imaging*, pp. CF2E–3 (2021).

[3] J. Wang, T. Wang, B. Xu, O. Cossairt, and F. Willomitzer, “Accurate Eye Tracking from Dense 3D Surface Reconstructions using Single-Shot Deflectometry,” arXiv preprint arXiv:2308.07298 (2023).

[4] M. C. Knauer, J. Kaminski, and G. Häusler, “Phase measuring deflectometry: a new approach to measure specular free-form surfaces,” in *Optical Metrology in Production Engineering*, vol. 5457, pp. 366–376 (SPIE, 2004).

[5] H. Liang, T. Sauer, and C. Faber, “Using wavelet transform to evaluate single-shot phase measuring deflectometry data,” in *Applications of Digital Image Processing XLIII*, vol. 11510, pp. 404–410 (SPIE, 2020).

[6] E. Slogsnat and K.-H. Brenner, “Non-stereoscopic method for deflectometric measurement of reflecting surfaces,” DGaO-Proceedings, pp. 1614–8436 (2009).

[7] T. Wang, J. Wang, O. Cossairt, and F. Willomitzer, “Optimization Based Eye Tracking using Deflectometric Information,” DGaO-Proceedings (2023).

[8] T. Wang, J. Wang, O. Cossairt, and F. Willomitzer, “Optimization-Based Eye Tracking using Deflectometric Information,” arXiv preprint arXiv:2303.04997 (2023).