Tracking eye position and gaze direction in near-eye volumetric displays Supplementary Materials

Marek Wernikowski*
West Pomeranian University of Technology, Szczecin
University of Cambridge

Joseph G. March
University of Cambridge

Radosław Mantiuk †
West Pomeranian University of Technology, Szczecin

Ali Özgür Yöntem
University of Cambridge

Rafał K. Mantiuk ‡
University of Cambridge

1 Eye Tracking Model

The following section describes a step-by-step procedure to compute the eye parameters, based on the eye tracking model of Guestrin and Eizenman [3]. The model computes extrinsic parameters of the eye: 3D positions of the eye rotation center $\boldsymbol{e}$, the pupil $\boldsymbol{p}$, and position of the eye nodal point $\boldsymbol{n}$ (see Figure 1). These three points define the direction of the *eye optical axis*. The *visual axis*, representing the actual viewing direction, deviates from the optical axis by specific vertical and horizontal angles, which can vary among individuals [4, 2].

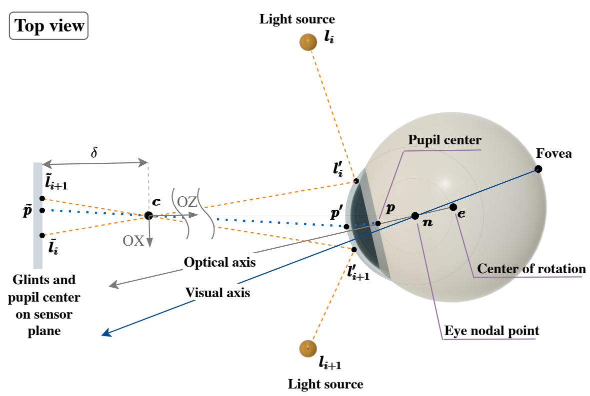

Figure 1: The geometry and symbols used in the eye tracking model.

1.1 Input parameters

Camera. The $\boldsymbol{e}$, $\boldsymbol{n}$, and $\boldsymbol{p}$ points are determined in the camera 3D coordinate system with the center located in the camera rear nodal point ($\boldsymbol{c}$). The OZ axis faces towards the eye, OX on the right, and OY upwards. The nodal point is located on the camera optical axis. Its distance from the camera sensor plane is equal to the lens's focal length if the lens is focused at infinity. When the lens is focused on a closer object (to track the eye at closer distances), the distance δ from the camera sensor plane to the nodal point can be calculated using the thin lens equation: $1/f = 1/\delta + 1/d$, where f is the focal length of the lens, and $d = \|\boldsymbol{c} - \boldsymbol{n}\|$ is the approximated distance from the camera to the eye (approximately, from the camera nodal point to the eye nodal point). In the eye tracker, positions of centers of pupil and glints are captured in the camera image coordinates and expressed in pixels. These positions are transformed to the global coordinates based on the camera intrinsic parameters. Actually, besides δ, the size of the sensor pixel (i.e., pixel pitch) Δs, and the horizontal (c_{res}) and vertical (r_{res}) resolution of the camera image must be known. Using the pinhole camera model, coordinates of a pixel (c, r) can be transformed from the camera image to the camera coordinates using the following equation:

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} \Delta s & 0 & -c_{res}/2 \\ 0 & \Delta s & -r_{res}/2 \\ 0 & 0 & -\delta \end{bmatrix} \cdot \begin{bmatrix} c \\ r \\ 1 \end{bmatrix}. \quad (1)$$

In Figure 1, the positions of pupil and glint centers in the camera space, calculated based on Equation 1, are indicated by $\tilde{\boldsymbol{p}}$, $\tilde{\boldsymbol{l}}_i$, and $\tilde{\boldsymbol{l}}_{i+1}$, respectively ($i = 1..n$).

Lights. At least two IR light sources used in our system have a dual purpose: they illuminate the eye and they introduce specular reflections (glints), which let us estimate the position of the eye and allow for free head movements. The positions of the lights $\boldsymbol{l}_i$ and $\boldsymbol{l}_{i+1}$ are expressed in camera coordinates must be known in advance.

Eye intrinsic parameters. The intrinsic parameters of the eye are individual features of a person. They can be estimated using the calibration procedure [3]. In this work, we use average values presented in Table 1.

Eye parameter	Assumed value
Distance between $\boldsymbol{p}$ and $\boldsymbol{n}$	4.2 mm [2]
Distance between $\boldsymbol{e}$ and $\boldsymbol{n}$	5.3 mm [5]
Cornea radius	7.8 mm [2, 5]
Cornea refractive index	1.3375 [1]
α angle	5.0° for the left eye, −5.0° for the right eye [4, 2]
β angle	1.5° [4, 2]

Table 1: Typical estimates of the eye parameters, taken directly from [3].

1.2 Algorithm

The algorithm for determining the position of points $\boldsymbol{n}$, $\boldsymbol{p}$, and $\boldsymbol{e}$ can be divided into steps related to the calculations of (1) direction ($\vec{u}$) of a line containing $\boldsymbol{c}$ and $\boldsymbol{n}$, (2) position of the nodal point $\boldsymbol{n}$, (3) position of the pupil center $\boldsymbol{p}$, (4) direction of the eye optical axis, and position of the eye rotation center $\boldsymbol{e}$.

*e-mail: marek.wernikowski@zut.edu.pl
†e-mail: radoslaw.mantiuk@zut.edu.pl
‡e-mail: rafal.mantiuk@cl.cam.ac.uk

(1) Direction from the camera nodal point ($\boldsymbol{c}$) to the eye nodal point ($\boldsymbol{n}$). The eye nodal point $\boldsymbol{n}$ lies on the plane, which is constrained by the nodal point of the camera $\boldsymbol{c}$, light source $\boldsymbol{l}_i$, and its corresponding glint $\tilde{\boldsymbol{l}}_i$. The planes for two light sources intersect, forming a line that intersects the points $\boldsymbol{n}$ and $\boldsymbol{c}$, as shown in Figure 2. The direction of the line can be determined by first finding the direction of the normal for each plane, then finding the direction orthogonal to the two normal vectors:

$$\vec{u} = [(\boldsymbol{l}_i - \boldsymbol{c}) \times (\tilde{\boldsymbol{l}}_i - \boldsymbol{c})] \times [(\boldsymbol{l}_{i+1} - \boldsymbol{c}) \times (\tilde{\boldsymbol{l}}_{i+1} - \boldsymbol{c})], \tag{2}$$

where $\times$ is the cross product.

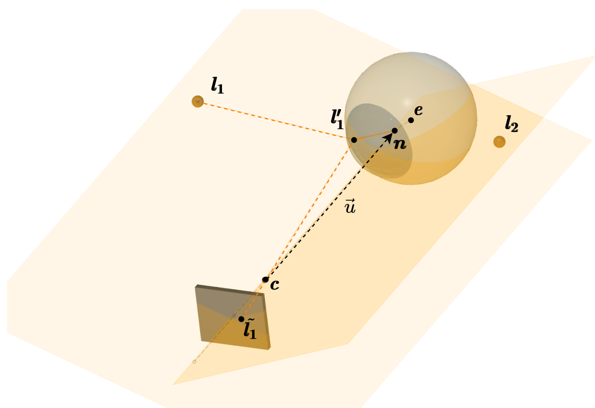

Figure 2: The plane is defined by three points: position of the light source $\boldsymbol{l}_1$, camera nodal point $\boldsymbol{c}$, and the center of the light glint in the image $\tilde{\boldsymbol{l}}_1$.

(2) Position of the eye nodal point The position of $\boldsymbol{n}$ can be determined using a set of formulas. Firstly, the $\boldsymbol{n}$ lies on the vector ($\boldsymbol{n}$-$\boldsymbol{c}$) found in the previous step:

$$\boldsymbol{n} = \boldsymbol{c} + k_0\vec{u}, \tag{3}$$

where k_0 is an unknown coefficient. Vectors $(\boldsymbol{c} - \tilde{\boldsymbol{l}}_i)$ intersect the cornea at points $\boldsymbol{l'}_i$:

$$\boldsymbol{l'}_i = \tilde{\boldsymbol{l}}_i + k_i(\boldsymbol{c} - \tilde{\boldsymbol{l}}_i), \tag{4}$$

where k_i are line coefficients and $\hat{\boldsymbol{l}}_i = normz(\boldsymbol{c} - \tilde{\boldsymbol{l}}_i)$ vector ($normz$ is the vector normalization operator). Additionally, we approximate the cornea curvature's shape to a sphere, therefore:

$$\|\boldsymbol{l'}_i - \boldsymbol{n}\| = c_r, \tag{5}$$

where c_r is the cornea radius.

The law of reflection states that the angle of incidence equals the angle of reflection. Because of that, the angle between the incidence vector $(\tilde{\boldsymbol{l}}_i - \boldsymbol{l'}_i)$ and the normal to the cornea curvature at the point of reflection $(\boldsymbol{l'}_i - \boldsymbol{n})$ must be equal to the angle between the normal and the reflection vector $(\boldsymbol{c} - \boldsymbol{l'}_i)$. Since $a \cdot b = |a||b|cos(\theta)$, and $|a| = |b| = 1$, we formulate it as:

$$\arccos\left((\boldsymbol{c} - \boldsymbol{l'}_i) \cdot (\boldsymbol{l'}_i - \boldsymbol{n})\right) = \arccos\left((\boldsymbol{l}_i - \boldsymbol{l'}_i) \cdot (\boldsymbol{l'}_i - \boldsymbol{n})\right), \tag{6}$$

All Equations (3) - (6), considered for two light sources, can be expanded to 13 scalar formulas. There are a total of 12 unknown variables (three values per $\boldsymbol{n}$, $\boldsymbol{l'}_1$, and $\boldsymbol{l'}_2$, one value per k_0, k_1, and k_2) that can be found analytically. To simplify implementation, we used the fast Nelder-Mead downhill solver method to find the best candidate for $\boldsymbol{n}$. We minimized the difference between the incidence and reflection angles (Eq. (6)) for both lights at the same time to find the best estimation.

(3) Center of the pupil. The vector $\vec{v} = \boldsymbol{c}$-$\tilde{\boldsymbol{p}}$ coming from pupil center $\tilde{\boldsymbol{p}}$ detected in the camera image plane and passing through the camera nodal point $\boldsymbol{c}$ intersects with the cornea surface at point $\boldsymbol{p'}$— see Figure 3. Then, this ray is refracted and reaches the pupil center $\boldsymbol{p}$. Determining the location of $\boldsymbol{p'}$ requires calculating the intersection of $\vec{v}$ with the surface of the cornea sphere:

$$\boldsymbol{p'} = \vec{v} \wedge (\|(x,y,z) - \boldsymbol{n}\|^2 - c_r^2). \tag{7}$$

Then, the ray $\vec{v}$ is refracted with a refractive index $\eta = 1.3375$ and the refracted ray $\boldsymbol{v'}$ is calculated. In the final step, the intersection of $\boldsymbol{v'}$ with the surface of the sphere containing the pupil center $\boldsymbol{p}$ is computed:

$$\boldsymbol{p} = \boldsymbol{v'} \wedge (\|(x,y,z) - \boldsymbol{n}\|^2 - \|\boldsymbol{n} - \boldsymbol{p}\|^2). \tag{8}$$

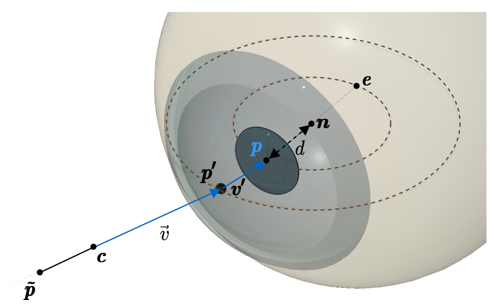

Figure 3: Computation of the pupil center position ($\boldsymbol{p}$).

(4) Direction of the optical axis and position of the eye rotation center. The line passing through the nodal point $\boldsymbol{n}$ and pupil center $\boldsymbol{p}$ determines the direction of the eye optical axis. The distances from the pupil center to the eye rotation center $\boldsymbol{e}$ is an intrinsic parameter of the eye. So we can calculate the position of the eye rotation center knowing that its center lies on the optical axis of the eye as:

$$\boldsymbol{e} = \boldsymbol{p} + \frac{\boldsymbol{n} - \boldsymbol{p}}{\|\boldsymbol{n} - \boldsymbol{p}\|} \cdot \|\boldsymbol{p} - \boldsymbol{e}\|, \tag{9}$$

where $\|\boldsymbol{p} - \boldsymbol{e}\|$ is the length of $(\boldsymbol{n} - \boldsymbol{p})$ vector and $\|\boldsymbol{p} - \boldsymbol{e}\|$ is the distance from the pupil center to the eye rotation center.

References

[1] M. Corbett, N. Maycock, E. Rosen, and D. O'Brart. Corneal topography: principles and applications. Springer, 2019. 1

[2] A. G. Gale. A note on the remote oculometer technique for recording eye movements. Vision research, 22(1):201–202, 1982. 1

[3] E. D. Guestrin and M. Eizenman. General theory of remote gaze estimation using the pupil center and corneal reflections. IEEE Transactions on biomedical engineering, 53(6):1124–1133, 2006. 1

[4] A. Slater and J. Findlay. The measurement of fixation position in the newborn baby. Journal of Experimental Child Psychology, 14(3):349–364, 1972. 1

[5] L. R. Young and D. Sheena. Survey of eye movement recording methods. Behavior research methods & instrumentation, 7(5):397–429, 1975. 1