

Covert Monitoring of the Point-of-Gaze

Moshe Eizenman

Dept. of Electrical and Computer Engineering,
Dept. of Ophthalmology and Vision Sciences, and
Institute of Biomaterials and Biomedical Engineering,
University of Toronto
Toronto, ON, Canada
eizenm@ecf.utoronto.ca

Dmitri Model

Dept. of Electrical and Computer Engineering,
University of Toronto
Toronto, ON, Canada
dmitri.model@utoronto.ca

Elias D. Guestrin

Dept. of Electrical and Computer Engineering and
Institute of Biomaterials and Biomedical Engineering,
University of Toronto
Toronto, ON, Canada
elias.guestrin@utoronto.ca

Abstract— Gaze estimation systems use calibration procedures that require active subject participation to estimate the point-of-gaze accurately. Consequently, these systems do not support covert monitoring of visual scanning patterns. This paper presents a novel gaze estimation methodology that does not use calibration procedures that require active user participation. This methodology uses multiple infrared light sources for illumination and a stereo pair of video cameras to obtain images of the eyes. Each pair of images is analyzed and the centers of the pupils and the centers of curvature of the corneas are estimated. These points, which are estimated without a personal calibration procedure, define the optical axis of each eye. To estimate the point-of-gaze, which lies along the visual axis, the angle between the optical and visual axes is estimated by a procedure that minimizes the distance between the intersections of the visual axes of the left and right eyes with the surface of a display while subjects look naturally at the display (e.g., watching a video clip). Simulation results demonstrate that for a subject sitting 75 cm in front of an 80 cm x 60 cm display (40" TV) the RMS error of the estimated point-of-gaze is 17.8 mm (1.3°).

Keywords – Covert gaze monitoring, Calibration-free gaze estimation, Remote gaze estimation.

I. INTRODUCTION

The point-of-gaze (PoG) is the point within the visual field that is imaged on the highest acuity region of the retina that is known as the fovea. Systems that estimate the PoG are used in a large variety of applications [1] such as studies of emotional and cognitive processes [2], [3], driver behavior [4], marketing and advertising [5], pilot training [6], ergonomics [7] and human-computer interfaces [8], [9].

Gaze estimation systems use calibration procedures to estimate subject-specific parameters that are needed for the accurate calculation of the PoG. In these procedures, subjects

are required to fixate specific points at specific time instances. The need for active subject participation during the calibration procedure makes it impossible to monitor the subject's visual scanning patterns covertly. Covert monitoring reduces biases to visual scanning parameters in studies of natural visual scanning behavior [2].

This paper presents a methodology that can estimate the PoG covertly. A stereo pair of video cameras is used to estimate the center of curvature of the cornea and the *optical* axis of the eye without any personal calibration [10]-[13]. Then, the angle between the *optical* and *visual* axes (human gaze is directed along the visual axis) is estimated by using the fact that when subjects look naturally at a display, the visual axes of left and right eyes intersect on the surface of the display.

A mathematical model for remote estimation of the center of curvature of the cornea and the optical axis of the eye that does not require any personal calibration is described in the next section. The estimation of the visual axis of each eye, the PoG and the subject-specific angle between the optical and visual axes is presented in Section III. Simulation results are presented in Section IV and the conclusions are presented in Section V.

II. ESTIMATION OF THE CENTER OF CURVATURE OF THE CORNEA AND THE OPTICAL AXIS OF THE EYE

Fig. 1 presents a general model of gaze estimation systems with any number of cameras and any number of light sources (all points are represented as 3-D column vectors (bold font) in a right-handed Cartesian world coordinate system). Light sources are modeled as point sources, video cameras are modeled as pinhole cameras and the front surface of the cornea is modeled as a spherical section. In this model, the line connecting the center of curvature of the cornea, \mathbf{c} , and the pupil center, \mathbf{p} , defines the optical axis of the eye. The line connecting the center of curvature of the cornea, \mathbf{c} , and the

This work was supported in part by a grant from the Natural Sciences and Engineering Research Council of Canada (NSERC), and in part by scholarships from NSERC and the Vision Science Research Program (Toronto Western Research Institute, University Health Network, Toronto, ON, Canada).

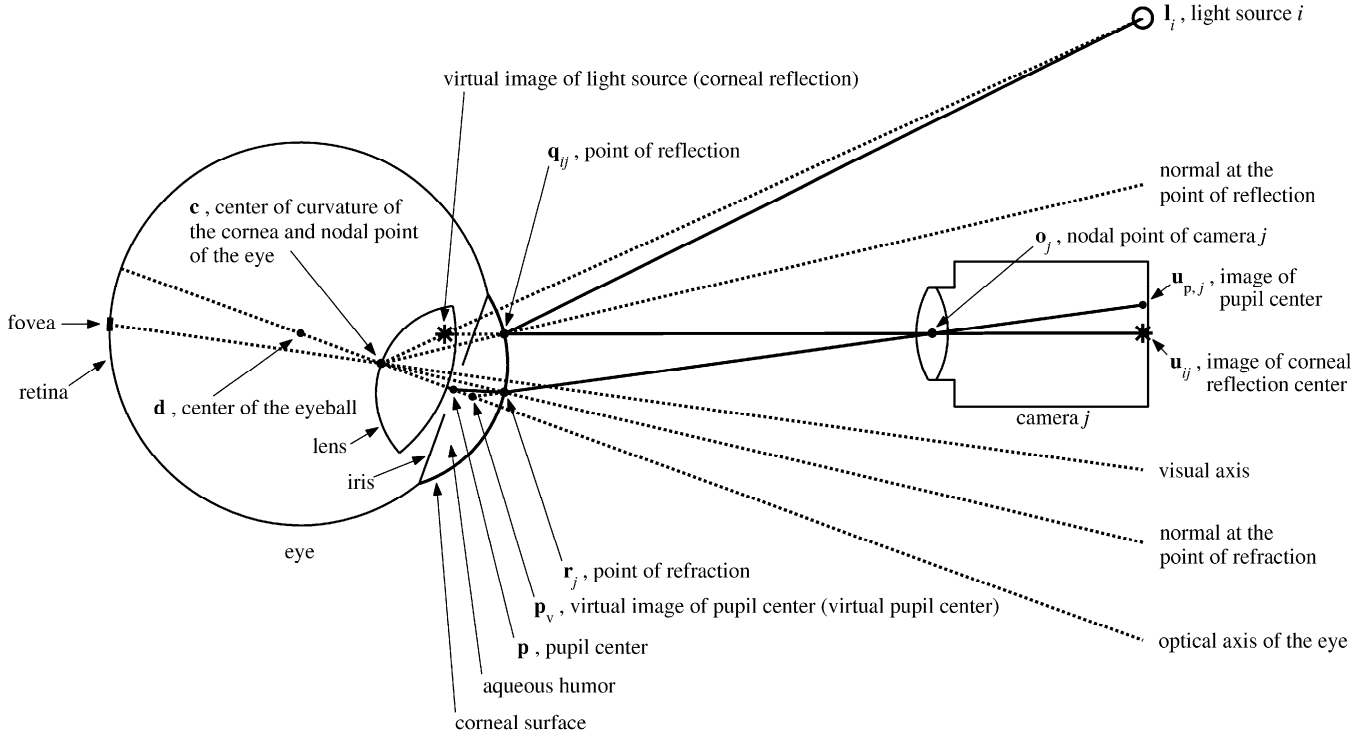


Fig. 1. Ray-tracing diagram (not to scale in order to be able to show all the elements of interest), showing schematic representations of the eye, a camera and a light source.

center of the fovea defines the visual axis. The PoG is at the intersection of the visual axis with the display. The following paragraphs present a methodology for the estimation of the optical axis of each eye that does not require a priori knowledge of subject-specific eye parameters.

First, consider a ray that comes from light source i , \mathbf{l}_i , and reflects at a point \mathbf{q}_{ij} on the corneal surface such that the reflected ray passes through the nodal point (a.k.a. camera center, center of projection) of camera j , \mathbf{o}_j , and intersects the camera image plane at a point \mathbf{u}_{ij} . According to the law of reflection, the incident ray, the reflected ray and the normal at the point of reflection are coplanar. Since any line going through the center of curvature of the cornea, \mathbf{c} , is normal to the spherical corneal surface, vector $(\mathbf{q}_{ij} - \mathbf{c})$ is normal to the corneal surface at the point of reflection \mathbf{q}_{ij} . It then follows that points \mathbf{l}_i , \mathbf{q}_{ij} , \mathbf{o}_j , \mathbf{u}_{ij} , and \mathbf{c} are coplanar. In other words, the center of curvature of the cornea, \mathbf{c} , belongs to each plane defined by the nodal point of camera j , \mathbf{o}_j , light source i , \mathbf{l}_i , and its corresponding image point, \mathbf{u}_{ij} . Noting that three coplanar vectors \mathbf{a}_1 , \mathbf{a}_2 and \mathbf{a}_3 satisfy $\mathbf{a}_1 \times \mathbf{a}_2 \cdot \mathbf{a}_3 = 0$, this condition can be formalized as

$$\underbrace{(\mathbf{l}_i - \mathbf{o}_j) \times (\mathbf{u}_{ij} - \mathbf{o}_j)}_{\text{normal to the plane defined by } \mathbf{l}_i, \mathbf{o}_j \text{ and } \mathbf{u}_{ij}} \cdot (\mathbf{c} - \mathbf{o}_j) = 0. \quad (1)$$

Notice that (1) shows that, for each camera j , all the planes defined by \mathbf{o}_j , \mathbf{l}_i and \mathbf{u}_{ij} contain the line defined by points \mathbf{c} and \mathbf{o}_j . If the light sources, \mathbf{l}_i , are positioned such that at least two of those planes are not coincident, the planes intersect at the line

defined by \mathbf{c} and \mathbf{o}_j . If \mathbf{b}_j is a vector in the direction of the line of intersection of the planes, then

$$\mathbf{c} = \mathbf{o}_j + k_{c,j} \mathbf{b}_j \text{ for some } k_{c,j}. \quad (2)$$

In particular, if two light sources are considered ($i = 1, 2$),

$$\mathbf{b}_j = \frac{[(\mathbf{l}_1 - \mathbf{o}_j) \times (\mathbf{u}_{1j} - \mathbf{o}_j)] \times [(\mathbf{l}_2 - \mathbf{o}_j) \times (\mathbf{u}_{2j} - \mathbf{o}_j)]}{\|[(\mathbf{l}_1 - \mathbf{o}_j) \times (\mathbf{u}_{1j} - \mathbf{o}_j)] \times [(\mathbf{l}_2 - \mathbf{o}_j) \times (\mathbf{u}_{2j} - \mathbf{o}_j)]\|}, \quad (3)$$

where $[(\mathbf{l}_1 - \mathbf{o}_j) \times (\mathbf{u}_{1j} - \mathbf{o}_j)]$ is the normal to the plane defined by \mathbf{o}_j , \mathbf{l}_1 and \mathbf{u}_{1j} , and $[(\mathbf{l}_2 - \mathbf{o}_j) \times (\mathbf{u}_{2j} - \mathbf{o}_j)]$ is the normal to the plane defined by \mathbf{o}_j , \mathbf{l}_2 and \mathbf{u}_{2j} .

Having two cameras, the position of the center of curvature of the cornea, \mathbf{c} , can be found as the intersection of the lines given by (2)–(3), $j = 1, 2$. Since, in practice, the estimated coordinates of the images of the corneal reflection centers, \mathbf{u}_{ij} , are corrupted by noise, those lines may not intersect. Therefore, \mathbf{c} is found as the midpoint of the shortest segment defined by a point belonging to each of those lines. It can be shown that, in such case, \mathbf{c} is given by

$$\mathbf{c} = \frac{1}{2} \begin{bmatrix} \mathbf{b}_1 & \mathbf{b}_2 \end{bmatrix} \begin{bmatrix} \mathbf{b}_1 \cdot \mathbf{b}_1 & -\mathbf{b}_1 \cdot \mathbf{b}_2 \\ -\mathbf{b}_1 \cdot \mathbf{b}_2 & \mathbf{b}_2 \cdot \mathbf{b}_2 \end{bmatrix}^{-1} \begin{bmatrix} -\mathbf{b}_1 \cdot (\mathbf{o}_1 - \mathbf{o}_2) \\ \mathbf{b}_2 \cdot (\mathbf{o}_1 - \mathbf{o}_2) \end{bmatrix} + \frac{1}{2} (\mathbf{o}_1 + \mathbf{o}_2). \quad (4)$$

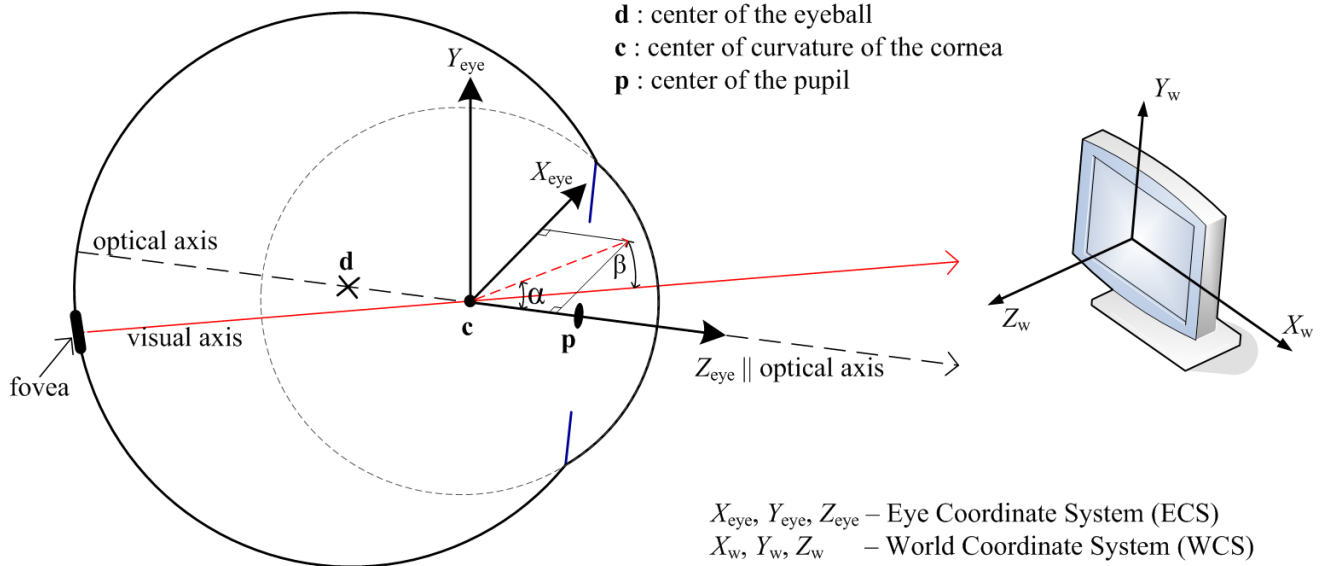


Fig. 2. Simplified schematics of the eye. The *optical axis* of the eye connects the center of the pupil with the center of curvature of the cornea. Gaze is directed along the *visual axis*, which connects the center of the region of highest acuity of the retina (fovea) with the center of curvature of the cornea.

Next, consider an imaginary ray that originates at the pupil center, \mathbf{p} , travels through the aqueous humor and cornea (effective index of refraction ≈ 1.3375) and refracts at a point \mathbf{r}_j on the corneal surface as it travels into the air (index of refraction ≈ 1), such that the refracted ray passes through the nodal point of camera j , \mathbf{o}_j , and intersects the camera image plane at a point $\mathbf{u}_{p,j}$. This refraction results in the formation of a virtual image of the pupil center (virtual pupil center), $\mathbf{p}_{v,j}$, located on the extension of the refracted ray, i.e.,

$$\mathbf{p}_{v,j} = \mathbf{o}_j + k_{p,j} \underbrace{(\mathbf{o}_j - \mathbf{u}_{p,j})}_{\mathbf{h}_j} \quad \text{for some } k_{p,j}. \quad (5)$$

In strict terms, the spatial location of $\mathbf{p}_{v,j}$ depends on the position of the nodal point of the camera, \mathbf{o}_j , relative to the eye. Therefore, in general, the spatial location of $\mathbf{p}_{v,j}$ will be slightly different for each of the two cameras. Despite this, an approximate virtual image of the pupil center, \mathbf{p}_v , can be found as the midpoint of the shortest segment defined by a point belonging to each of the lines given by (5), $j = 1, 2$, i.e.,

$$\mathbf{p}_v = \frac{1}{2} \begin{bmatrix} \mathbf{h}_1 & \mathbf{h}_2 \end{bmatrix} \begin{bmatrix} \mathbf{h}_1 \cdot \mathbf{h}_1 & -\mathbf{h}_1 \cdot \mathbf{h}_2 \\ -\mathbf{h}_1 \cdot \mathbf{h}_2 & \mathbf{h}_2 \cdot \mathbf{h}_2 \end{bmatrix}^{-1} \begin{bmatrix} -\mathbf{h}_1 \cdot (\mathbf{o}_1 - \mathbf{o}_2) \\ \mathbf{h}_2 \cdot (\mathbf{o}_1 - \mathbf{o}_2) \end{bmatrix} + \frac{1}{2} (\mathbf{o}_1 + \mathbf{o}_2). \quad (6)$$

Since \mathbf{c} is on the optic axis of the eye and assuming that \mathbf{p}_v is also on the optic axis (Fig. 1), (3)–(6) provide a closed-form solution for the reconstruction of the optic axis of the eye in 3-D space without the knowledge of any subject-specific eye parameter. In particular, the direction of the optic axis of the eye is given by the unit vector

$$\boldsymbol{\omega} = \frac{\mathbf{p}_v - \mathbf{c}}{\|\mathbf{p}_v - \mathbf{c}\|}. \quad (7)$$

The PoG, \mathbf{g} , is defined as the intersection of the visual axis, rather than the optic axis, with the scene. The visual axis is the line defined by the nodal point of the eye and the center of the fovea, and deviates from the optic axis by as much as 5° [14]. Since the nodal point is within 1 mm of the center of curvature of the cornea [14], \mathbf{c} , it can be assumed to be coincident with \mathbf{c} (Fig. 1). The estimation of the visual axis, the PoG and the angle between the optical and visual axes is described in the next section.

III. ESTIMATION OF THE VISUAL AXES, THE POINT-OF-GAZE AND THE ANGLE BETWEEN THE OPTICAL AND VISUAL AXES

Two coordinate systems are used to describe the relation between the optical and visual axes of the eye. The first is a stationary right-handed Cartesian World Coordinate System (WCS) with the origin at the center of the display, the X_w -axis in the horizontal direction, the Y_w -axis in the vertical direction and the Z_w -axis perpendicular to the display (see Fig. 2). The second is a non-stationary right-handed Cartesian Eye Coordinate System (ECS), which is attached to the eye, with the origin at the center of curvature of the cornea, the Z_{eye} axis that coincides with the optical axis of the eye and X_{eye} and Y_{eye} axes that, in the primary gaze position, are in the horizontal and vertical directions, respectively. The X_{eye} - Y_{eye} plane rotates according to Listing's law [15] around the Z_{eye} axis for different gaze directions.

In the ECS, the unknown 3-D angle between the *optical* and the *visual* axes of the eye can be expressed by the horizontal¹,

¹ The angle between the projection of the visual axis on the X_{eye} - Z_{eye} plane and the Z_{eye} axis. It is equal to 90° if the visual axis is in the $-X_{eye}$ direction.

α , and vertical², β , components of this angle (see Fig. 2). The unit vector in the direction of the *visual* axis with respect to the ECS, \mathbf{v}_{ECS} , is then expressed as

$$\mathbf{v}_{\text{ECS}}(\alpha, \beta) = \begin{bmatrix} -\sin(\alpha)\cos(\beta) \\ \sin(\beta) \\ \cos(\alpha)\cos(\beta) \end{bmatrix}. \quad (8)$$

The unit vector in the direction of the visual axis with respect to the WCS, \mathbf{v} , can be expressed as

$$\mathbf{v}(\alpha, \beta) = \mathbf{R} \mathbf{v}_{\text{ECS}}(\alpha, \beta) \quad (9)$$

where \mathbf{R} is the rotation matrix from the ECS to the WCS (independent of α and β), which can be readily calculated from the orientation of the optical axis of the eye and Listing's law [15].

Because the visual axis goes through the center of curvature of the cornea, \mathbf{c} , and the PoG is defined by the intersection of the visual axis with the display ($Z_w = 0$), the PoG in the WCS is given by

$$\mathbf{g}(\alpha, \beta) = \mathbf{c} + k(\alpha, \beta) \mathbf{v}(\alpha, \beta) = \mathbf{c} + k(\alpha, \beta) \mathbf{R} \mathbf{v}_{\text{ECS}}(\alpha, \beta) \quad (10)$$

with

$$k(\alpha, \beta) = -\frac{\mathbf{c} \bullet \mathbf{n}}{\mathbf{v}(\alpha, \beta) \bullet \mathbf{n}} \quad (11)$$

where $\mathbf{n} = [0 \ 0 \ 1]^T$ is the normal to the display surface and “ \bullet ” denotes transpose.

The estimation of $\alpha^L, \beta^L, \alpha^R$ and β^R is based on the fact that at each time instant the visual axes of both eyes intersect on the surface of the display (superscripts “L” and “R” are used to denote parameters of the left and right eyes, respectively). The unknown angles $\alpha^L, \beta^L, \alpha^R$ and β^R can be estimated by minimizing the distance between the intersections of the left and right visual axes with that surface (left and right PoGs).

The objective function to be minimized is then

$$F(\alpha^L, \beta^L, \alpha^R, \beta^R) = \sum_i \|\mathbf{g}_i^L(\alpha^L, \beta^L) - \mathbf{g}_i^R(\alpha^R, \beta^R)\|_2^2 \quad (12)$$

where the subscript i identifies the i -th gaze sample.

The above objective function is non-linear, and thus a numerical optimization procedure is required to solve for the unknown angles $\alpha^L, \beta^L, \alpha^R$ and β^R . However, since the deviations of the unknown angles $\alpha^L, \beta^L, \alpha^R$ and β^R from the

expected “average” values $\alpha_0^L, \beta_0^L, \alpha_0^R$ and β_0^R are relatively small, a linear approximation of (10) can be obtained by using the first three terms of its Taylor's series expansion:

$$\mathbf{g}(\alpha, \beta) \approx \mathbf{g}(\alpha_0, \beta_0) + \left. \frac{\partial \mathbf{g}(\alpha, \beta)}{\partial \alpha} \right|_{\alpha_0, \beta_0} (\alpha - \alpha_0) + \left. \frac{\partial \mathbf{g}(\alpha, \beta)}{\partial \beta} \right|_{\alpha_0, \beta_0} (\beta - \beta_0) \quad (13)$$

Let

$$\mathbf{g}_0 = \mathbf{g}(\alpha_0, \beta_0) = \mathbf{c} + k_0 \mathbf{v}_0 \quad (14)$$

where $\mathbf{v}_0 = \mathbf{v}(\alpha_0, \beta_0)$ and $k_0 = k(\alpha_0, \beta_0)$.

Then, using (10),

$$\mathbf{a} = \left. \frac{\partial \mathbf{g}(\alpha, \beta)}{\partial \alpha} \right|_{\alpha_0, \beta_0} = k_\alpha \mathbf{v}_0 + k_0 \mathbf{v}_\alpha \quad (15)$$

$$\mathbf{b} = \left. \frac{\partial \mathbf{g}(\alpha, \beta)}{\partial \beta} \right|_{\alpha_0, \beta_0} = k_\beta \mathbf{v}_0 + k_0 \mathbf{v}_\beta \quad (16)$$

where, from (8)-(9),

$$\mathbf{v}_\alpha = \left. \frac{\partial \mathbf{v}(\alpha, \beta)}{\partial \alpha} \right|_{\alpha_0, \beta_0} = \mathbf{R} \begin{bmatrix} -\cos(\alpha_0) \cos(\beta_0) \\ 0 \\ -\sin(\alpha_0) \cos(\beta_0) \end{bmatrix} \quad (17)$$

$$\mathbf{v}_\beta = \left. \frac{\partial \mathbf{v}(\alpha, \beta)}{\partial \beta} \right|_{\alpha_0, \beta_0} = \mathbf{R} \begin{bmatrix} \sin(\alpha_0) \sin(\beta_0) \\ \cos(\beta_0) \\ -\cos(\alpha_0) \sin(\beta_0) \end{bmatrix} \quad (18)$$

and, from (11),

$$k_\alpha = \left. \frac{\partial k(\alpha, \beta)}{\partial \alpha} \right|_{\alpha_0, \beta_0} = -k_0 \frac{\mathbf{v}_\alpha \bullet \mathbf{n}}{\mathbf{v}_0 \bullet \mathbf{n}} \quad (19)$$

$$k_\beta = \left. \frac{\partial k(\alpha, \beta)}{\partial \beta} \right|_{\alpha_0, \beta_0} = -k_0 \frac{\mathbf{v}_\beta \bullet \mathbf{n}}{\mathbf{v}_0 \bullet \mathbf{n}}. \quad (20)$$

Using the above linear approximation, the sum of the squared distances between the left and right PoGs in the objective function (12) can be expressed as

² The angle between the visual axis and its projection on the $X_{\text{eye}}\text{-}Z_{\text{eye}}$ plane. It is equal to 90° if the visual axis is in the $+Y_{\text{eye}}$ direction.

$$F(\alpha^L, \beta^L, \alpha^R, \beta^R) = \sum_i \|\mathbf{M}_i \mathbf{x} + \mathbf{y}_i\|_2^2 \quad (21)$$

where $\mathbf{M}_i = \begin{bmatrix} \mathbf{a}_i^L & \mathbf{b}_i^L & -\mathbf{a}_i^R & -\mathbf{b}_i^R \end{bmatrix}$ is a 3x4 matrix, $\mathbf{y}_i = \mathbf{g}_{0,i}^L - \mathbf{g}_{0,i}^R$ is a 3x1 vector and $\mathbf{x} = [\alpha^L - \alpha_0^L \quad \beta^L - \beta_0^L \quad \alpha^R - \alpha_0^R \quad \beta^R - \beta_0^R]^T$ is a 4x1 vector of unknown angles. The subscript “ i ” is used to explicitly indicate the correspondence to the specific time instance “ i ” or i -th gaze sample.

The solution to (21) can be obtained in a closed form using least squares as

$$\mathbf{x}_{\text{opt}} = -(\mathbf{M}^T \mathbf{M})^{-1} \mathbf{M}^T \mathbf{y} \quad (22)$$

where the optimization over several time instances is achieved by stacking the matrices on top of each other:

$$\mathbf{M} = \begin{bmatrix} \mathbf{M}_1 \\ \mathbf{M}_2 \\ \vdots \\ \mathbf{M}_N \end{bmatrix} ; \quad \mathbf{y} = \begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \\ \vdots \\ \mathbf{y}_N \end{bmatrix} . \quad (23)$$

Finally, the estimates of the subject-specific angles are given by

$$\begin{bmatrix} \hat{\alpha}^L & \hat{\beta}^L & \hat{\alpha}^R & \hat{\beta}^R \end{bmatrix}^T = \begin{bmatrix} \alpha_0^L & \beta_0^L & \alpha_0^R & \beta_0^R \end{bmatrix}^T + \mathbf{x}_{\text{opt}} . \quad (24)$$

Since the objective function (21) is a linear approximation of the objective function (12), several iterations of (14)-(24) might be needed to converge to the true minimum of the objective function (12). In the first iteration, α_0^L , β_0^L , α_0^R and β_0^R are set to zero. In subsequent iterations, α_0^L , β_0^L , α_0^R and β_0^R are set to the values of $\hat{\alpha}^L$, $\hat{\beta}^L$, $\hat{\alpha}^R$ and $\hat{\beta}^R$ from the preceding iteration.

The above methodology to estimate the angle between the *optical* and *visual* axes is suitable for “on-line” estimation as a new matrix \mathbf{M}_i is added to \mathbf{M} and a new vector \mathbf{y}_i is added to \mathbf{y} for each new estimate of the centers of curvature of the corneas and optical axes.

IV. NUMERICAL SIMULATIONS

The numerical simulations use system parameters (position of light sources and camera parameters) that are similar to the parameters of the two-camera remote PoG estimation system described in [13]. The eye parameters are as follows: interpupillary distance = 60 mm; radius of curvature of the cornea = 7.8 mm; pupil radius = 4 mm; distance between the center of the pupil and the center of curvature of the cornea = 4.5 mm.

The position of the center of curvature of the cornea of the subject’s right eye, \mathbf{c}^R , was randomly selected from a uniform distribution in a 10 cm³ volume around the point [30 0 750]^T (i.e., 750 mm from the display surface) while the center of curvature of the cornea of the subject’s left eye, \mathbf{c}^L , was positioned at $\mathbf{c}^R - [60 0 0]^T$ (i.e., 60 mm to the left of \mathbf{c}^R). The subject-specific angles between the optical and visual axes were randomly drawn for each simulation from the uniform distribution in the range of (-5, 0) for α^R , (0, 5) for α^L and (-5, 5) for β^R and β^L . The PoGs were randomly drawn from a uniform distribution over an 80 cm x 60 cm display surface, which was defined as the plane $Z_w = 0$.

For each eye position and PoG, the pupil boundary and the corneal reflections were first reconstructed in the WCS and then projected onto the image planes of the two cameras of the simulated PoG estimation system. White Gaussian noise with a standard deviation of 0.1 pixels (similar to the noise observed in the experimental system described in [13]) was then added to the coordinates of the projected features (corneal reflection images (\mathbf{u}_{ij}), pupil boundary image points). An ellipse was fitted to the noise-corrupted pupil boundary points in each camera image and its center (center of the pupil image) was used as an estimate of the image of the pupil center ($\mathbf{u}_{p,j}$). Using the resulting \mathbf{u}_{ij} ’s and $\mathbf{u}_{p,j}$ ’s, the centers of curvature of the corneas and the optical axes of the two eyes were estimated following the methodology described in Section II. The estimated centers of curvature of the corneas and the orientations of the optical axes were then used as inputs to the algorithm that estimates the angle between the optical and visual axes of each eye as described in Section III. The algorithm was used in an “on-line” mode.

For each set of eye parameters (α^L , β^L , α^R and β^R) and eye positions (\mathbf{c}^R , \mathbf{c}^L), 1000 PoGs were used to calculate the RMS error between the estimated PoGs and the actual PoGs. These calculations were repeated for 100 different sets of eye parameters and eye positions. The RMS error (\pm standard deviation) was 17.8 (\pm 9.4) mm.

V. CONCLUSIONS

A novel gaze estimation methodology that does not use calibration procedures requiring active user participation was presented. Previous gaze estimation methodologies that were presented in the past as “calibration-free” estimated the PoG as the intersection of the *optical* axis of the eye with the scene (e.g., [10]). Using this approach, the RMS error for the simulations described in Section IV is 52.9 (\pm 20.1) mm. The relatively large and unpredictable (subject-dependant) PoG estimation errors of these systems limited their use. The methodology presented in this paper reduces the PoG estimation error of previous “calibration-free” methodologies by at least a factor of three. Therefore, this novel methodology can be used in applications that require more accurate covert monitoring of visual scanning patterns.

REFERENCES

- [1] A. T. Duchowski, “A breadth-first survey of eye-tracking applications,” Behavior Research Methods, Instruments, & Computers, vol. 34, no. 4, pp. 455-470, Nov. 2002.

- [2] M. Eizenman, L. H. Yu, L. Grupp, E. Eizenman, M. Ellenbogen, M. Gemar, and R. D. Levitan, "A naturalistic visual scanning approach to assess selective attention in major depressive disorder," *Psychiatry Research*, vol. 118, no. 2, pp. 117-128, May 2003.
- [3] K. Rayner, "Eye movements in reading and information processing: 20 years of research," *Psychological Bulletin*, vol. 124, no. 3, pp. 372-422, Nov. 1998.
- [4] J. L. Harbluk, Y. I. Noy, P. L. Trbovich, and M. Eizenman, "An on-road assessment of cognitive distraction: Impacts on drivers' visual behavior and braking performance," *Accident Analysis & Prevention*, vol. 39, no. 2, pp. 372-379, Mar. 2007.
- [5] G. Loshe, "Consumer eye movement patterns of Yellow Pages advertising," *Journal of Advertising*, vol. 26, no. 1, pp. 61-73, 1997.
- [6] P. A. Wetzel, G. Krueger-Anderson, C. Poprik, and P. Bascom, "An eye tracking system for analysis of pilots' scan paths," United States Air Force Armstrong Laboratory Tech. Rep. AL/HR-TR-1996-0145, Apr. 1997.
- [7] J. H. Goldberg and X. P. Kotval, "Computer interface evaluation using eye movements: methods and constructs," *International Journal of Industrial Ergonomics*, vol. 24, no. 6, pp. 631-645, Oct. 1999.
- [8] R. Sharma, V. I. Pavlović, and T. S. Huang, "Toward multimodal human-computer interface," *Proceedings of the IEEE*, vol. 86, no. 5, pp. 853-869, May 1998.
- [9] T. E. Hutchinson, K. P. White, W. N. Martin, K. C. Reichert, and L. A. Frey, "Human-computer interaction using eye-gaze input," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 19, no. 6, pp. 1527-1534, Nov./Dec. 1989.
- [10] S.-W. Shih, Y.-T. Wu, and J. Liu, "A calibration-free gaze tracking technique," in *Proc. of the 15th International Conference on Pattern Recognition*, Sep. 2000, vol. 4, pp. 201-204.
- [11] S.-W. Shih and J. Liu, "A novel approach to 3-D gaze tracking using stereo cameras," *IEEE Transactions on Systems, Man, and Cybernetics—Part B: Cybernetics*, vol. 34, no. 1, pp. 234-245, Feb. 2004.
- [12] E. D. Guestrin and M. Eizenman, "General theory of remote gaze estimation using the pupil center and corneal reflections," *IEEE Transactions on Biomedical Engineering*, vol. 53, no. 6, pp. 1124-1133, Jun. 2006.
- [13] E. D. Guestrin and M. Eizenman, "Remote point-of-gaze estimation requiring a single-point calibration for applications with infants," in *Proc. of the 2008 Symposium on Eye Tracking Research & Applications*, Savannah, GA, USA: ACM, 2008, pp. 267-274.
- [14] L. R. Young and D. Sheena, "Survey of eye-movement recording methods," *Behavior Research Methods & Instrumentation*, vol. 7, no. 5, pp. 397-429, 1975.
- [15] H. Helmholtz, *Helmholtz's treatise on physiological optics*. Translated from the 3rd German ed. Edited by J. P. C. Southall. Rochester, NY: Optical Society of America, 1924.