Investigation of the Cross-Ratios Method for Point-of-Gaze Estimation

Jeffrey J. Kang, Moshe Eizenman*, Elias D. Guestrin, *Member, IEEE*, and Erez Eizenman

***Abstract*—The cross-ratios method for point-of-gaze (PoG) estimation uses the invariance property of cross-ratios in projective transformations. The inherent causes of the subject-dependent PoG estimation bias exhibited by this method have not been well characterized in the literature. Using a model of the eye and the components of a system (camera, light sources) that estimates PoG, a theoretical framework for the cross-ratios method is developed. The analysis of the cross-ratios method within this framework shows that the subject-dependent estimation bias is caused mainly by: 1) the angular deviation of the visual axis from the optic axis and 2) the fact that the virtual image of the pupil center is not coplanar with the virtual images of the light sources that illuminate the eye (corneal reflections). The theoretical framework provides a closed-form analytical expression that predicts the estimation bias as a function of subject-specific eye parameters. The theoretical framework also provides a clear physical interpretation for an existing empirically derived two-step procedure that compensates for the estimation bias and shows that the first step of this procedure is equivalent to moving the corneal reflections to a new plane that minimizes the distance from this plane to the virtual image of the pupil center.**

***Index Terms*—Cross-ratios, eye model, eye parameters, point-of-gaze (PoG), 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 known as the fovea. Systems that estimate the PoG are used in a large variety of applications that include studies of mood, perception, and attention disorders [1], studies of driver behavior [2], ergonomics [3], multimodal human–computer interfaces [4], and assistive devices for motor-disabled persons [5]. Most modern approaches to remote, noncontact gaze estimation are based on the analysis of eye features extracted from video images. The most commonly used features are the centers of the pupil and one or more corneal reflections. The corneal reflections (first Purkinje images, glints) are virtual images of light sources (usually infrared) that illuminate the eye and are created by the front surface of the cornea, which acts as a convex mirror.

Relatively recently, a new method for PoG estimation that exploits the invariance of cross-ratios in projective geometry [6] was presented as having several desired attributes: being simple to implement, having a simple model, using a single uncalibrated camera, tolerating natural head movements, and not requiring a subject-specific calibration procedure. It was observed later, however, that the original cross-ratios method results in very large PoG estimation bias for some subjects [7]–[9]. In order to reduce the estimation bias, error-compensation techniques using a subject-specific calibration procedure have been proposed [7]–[9]. However, despite success in reducing the estimation bias, the basic causes of the estimation bias are poorly understood. In this paper, we identify and analyze the main causes of estimation bias within the cross-ratios method for PoG estimation.

In the following section, the original cross-ratios method is described, and the main sources of error are identified. In Section III, a detailed analysis of the main causes of estimation bias is presented, and analytic expressions that quantify the subject-dependent estimation bias are provided. In Section IV, a calibration-based technique to compensate for the estimation bias is analyzed and explained. Finally, in Section V, the work is summarized and conclusions are presented.

Manuscript received April 2, 2007; revised September 15, 2007. 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 Ontario Graduate Scholarship program (OGS). *Asterisk indicates corresponding author.*

J. J. Kang was with the Department of Electrical and Computer Engineering, and the Institute of Biomaterials and Biomedical Engineering, University of Toronto, Toronto, ON, Canada. He is now with the Faculty of Law, University of Toronto, Toronto, ON, Canada (e-mail: jeff.kang@utoronto.ca).

*M. Eizenman is with the Department of Electrical and Computer Engineering, the Department of Ophthalmology and Vision Sciences, and the Institute of Biomaterials and Biomedical Engineering, University of Toronto, Rosebrugh Building, 164 College St., Toronto, ON M5S 3G9, Canada (e-mail: eizenm@ecf.utoronto.ca).

E. D. Guestrin is with the Department of Electrical and Computer Engineering, and the Institute of Biomaterials and Biomedical Engineering, University of Toronto, Toronto, ON, Canada (e-mail: elias.guestrin@utoronto.ca).

E. Eizenman was with the Department of Electrical and Computer Engineering, and the Rotman School of Management, University of Toronto, Toronto, ON, Canada (e-mail: erez.eizenman@utoronto.ca).

Digital Object Identifier 10.1109/TBME.2008.919722

II. Cross-Ratios Method and the Main Causes of Estimation Bias

A. Original Method

The cross-ratios method for PoG estimation was originally proposed by Yoo *et al.* [6], who described a novel mapping of the pupil center in an eye image to a corresponding PoG estimate on a plane. To estimate the PoG on a plane using the method described in [6], four light sources are placed in the same plane and the subject's eye is imaged using a video camera. Fig. 1(a) shows a typical system setup in which the four light sources ($\mathbf{l}_1$, $\mathbf{l}_2$, $\mathbf{l}_3$, $\mathbf{l}_4$) are placed around a computer screen (the scene plane) upon which the PoG is estimated ($\hat{\mathbf{g}}$), and the camera is placed under the screen. Fig. 1(b) shows an image of the eye as captured by the camera in which the four corneal reflections ($\mathbf{u}_1$, $\mathbf{u}_2$, $\mathbf{u}_3$, $\mathbf{u}_4$) and the pupil center, $\mathbf{u}_p$, are identified. Yoo *et al.* [6] suggested that the cross-ratios formed by features on the scene

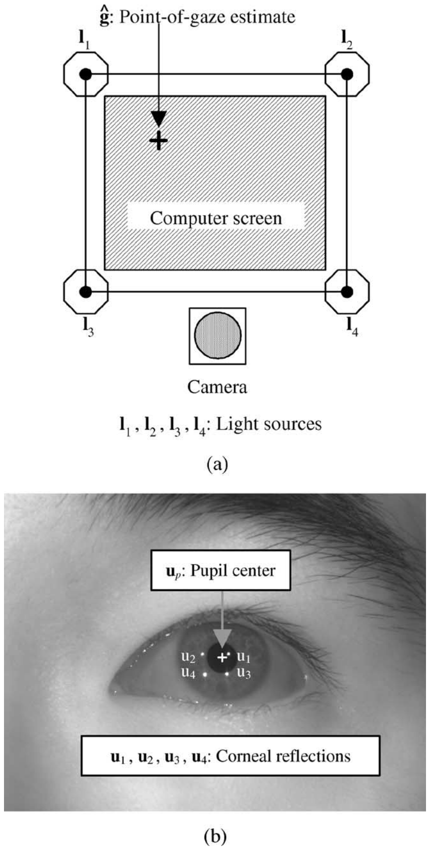

Fig. 1. (a) System setup with four light sources positioned around a computer screen, and a camera under the screen. (b) Image of the eye as captured by the camera, showing the four corneal reflections and the pupil center.

plane ($\mathbf{l}_1$, $\mathbf{l}_2$, $\mathbf{l}_3$, $\mathbf{l}_4$, $\hat{\mathbf{g}}$) were equal to the cross-ratios formed by the corresponding features on the camera's imaging plane ($\mathbf{u}_1$, $\mathbf{u}_2$, $\mathbf{u}_3$, $\mathbf{u}_4$, $\mathbf{u}_p$). Following this assumption, they derived equations to calculate the PoG estimate, $\hat{\mathbf{g}}$, from the coordinates of the light sources and the coordinates of the centers of the pupil and corneal reflections in the eye images. However, as it was shown later (e.g., [7]–[9]), the cross-ratios method can result in very large PoG estimation bias.

B. Identification of the Main Causes of Estimation Bias

The analysis of the main causes of the estimation bias is based on the optical model from Fig. 2(a), where the surface of the cornea is modeled as a convex spherical mirror with radius R and center at $\mathbf{c}$ (center of curvature of the cornea), the system's camera is modeled as a pinhole camera with its nodal point at $\mathbf{o}$, and the light sources, $\mathbf{l}_i$, are modeled as point sources. The diagram from Fig. 2(b) is used to identify the main causes of estimation bias and derive an approximate closed-form expression that predicts the PoG estimation bias as a function of the system configuration and a set of subject-specific eye parameters.

First, consider a ray of light coming from a light source, $\mathbf{l}_i$, that reflects on the corneal surface such that the reflected ray goes through the nodal point of the camera, $\mathbf{o}$, and intersects the camera's imaging plane at $\mathbf{u}_i$. This reflection results in the formation of a virtual image of the light source (corneal reflection) on the extension of the reflected ray, behind the corneal surface. Let $\mathbf{v}_i$ be defined as the intersection of the extension of the reflected ray with the line joining the center of curvature of the cornea, $\mathbf{c}$, and the corresponding light source, $\mathbf{l}_i$. Even though, in general, the corneal reflection is not exactly on the line joining the center of curvature of the cornea and the corresponding light source, since the image of $\mathbf{v}_i$ and the image of the corneal reflection are identical (i.e., $\mathbf{u}_i$), for the analysis carried out in this paper, $\mathbf{v}_i$ can be used in place of the corneal reflection without introducing any error. For this reason and for the sake of brevity, the point $\mathbf{v}_i$ will be hereafter referred to as "corneal reflection."

Next, consider an imaginary ray coming from the pupil center, $\mathbf{p}$, that travels through the aqueous humor and cornea (effective index of refraction $\approx$1.3375 [10]) and refracts at the corneal surface as it travels into the air (index of refraction $\approx$1) such that the refracted ray goes through the nodal point of the camera, $\mathbf{o}$, and intersects the camera's imaging plane at $\mathbf{u}_p$ (image of the pupil center). This refraction results in the formation of a virtual image of the pupil center (hereafter virtual pupil center, for brevity) on the extension of the refracted ray. This virtual image is located between the pupil center, $\mathbf{p}$, and the corneal surface. Let $\mathbf{p}_v$ be defined as the intersection of the extension of the refracted ray with the optic axis of the eye. Even though, in general, the virtual pupil center is not exactly on the optic axis of the eye, since the image of $\mathbf{p}_v$ and the image of the virtual pupil center are identical (i.e., $\mathbf{u}_p$), for the analysis carried out in this paper, $\mathbf{p}_v$ can be used in place of the virtual pupil center without introducing any error. For this reason and for the sake of brevity, the point $\mathbf{p}_v$ will be hereafter referred to as "virtual pupil center."

The cross-ratios method requires that: 1) the light sources on the scene plane ($\mathbf{l}_1$, $\mathbf{l}_2$, $\mathbf{l}_3$, $\mathbf{l}_4$) be related to the images of their corresponding corneal reflections on the imaging plane ($\mathbf{u}_1$, $\mathbf{u}_2$, $\mathbf{u}_3$, $\mathbf{u}_4$) by a projective transformation and 2) that the PoG estimate, $\hat{\mathbf{g}}$, be related to the image of the pupil center, $\mathbf{u}_p$, by the same projective transformation. Only if these two conditions were met, would the cross-ratios formed by features on the scene plane ($\mathbf{l}_1$, $\mathbf{l}_2$, $\mathbf{l}_3$, $\mathbf{l}_4$, $\hat{\mathbf{g}}$) be equal to the cross-ratios formed by the corresponding features on the camera's imaging plane ($\mathbf{u}_1$, $\mathbf{u}_2$, $\mathbf{u}_3$, $\mathbf{u}_4$, $\mathbf{u}_p$). For the previous two conditions to hold: 1) the light sources ($\mathbf{l}_1$, $\mathbf{l}_2$, $\mathbf{l}_3$, $\mathbf{l}_4$) and the PoG estimate, $\hat{\mathbf{g}}$, must be coplanar; 2) the corneal reflections ($\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{v}_3$, $\mathbf{v}_4$) and the virtual pupil center, $\mathbf{p}_v$, must be coplanar; 3) the images of the corneal reflections ($\mathbf{u}_1$, $\mathbf{u}_2$, $\mathbf{u}_3$, $\mathbf{u}_4$) and the image of the pupil center, $\mathbf{u}_p$, must be coplanar; and (iv) the PoG estimate, $\hat{\mathbf{g}}$, must be collinear with the virtual pupil center, $\mathbf{p}_v$, and the center of curvature of the cornea, $\mathbf{c}$. If the aforementioned requirements were satisfied, the projective transformation that forms the basis of the cross-ratios method [6] would be the composition of two perspective projections. The first perspective projection would have a projective center at the center of curvature of the cornea, $\mathbf{c}$, projecting the light sources ($\mathbf{l}_1$, $\mathbf{l}_2$, $\mathbf{l}_3$, $\mathbf{l}_4$) to their corresponding corneal reflections ($\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{v}_3$, $\mathbf{v}_4$) [see Fig. 2(b)] and the PoG estimate, $\hat{\mathbf{g}}$, to the virtual pupil center, $\mathbf{p}_v$. The second perspective projection is a camera projection, centered at the nodal

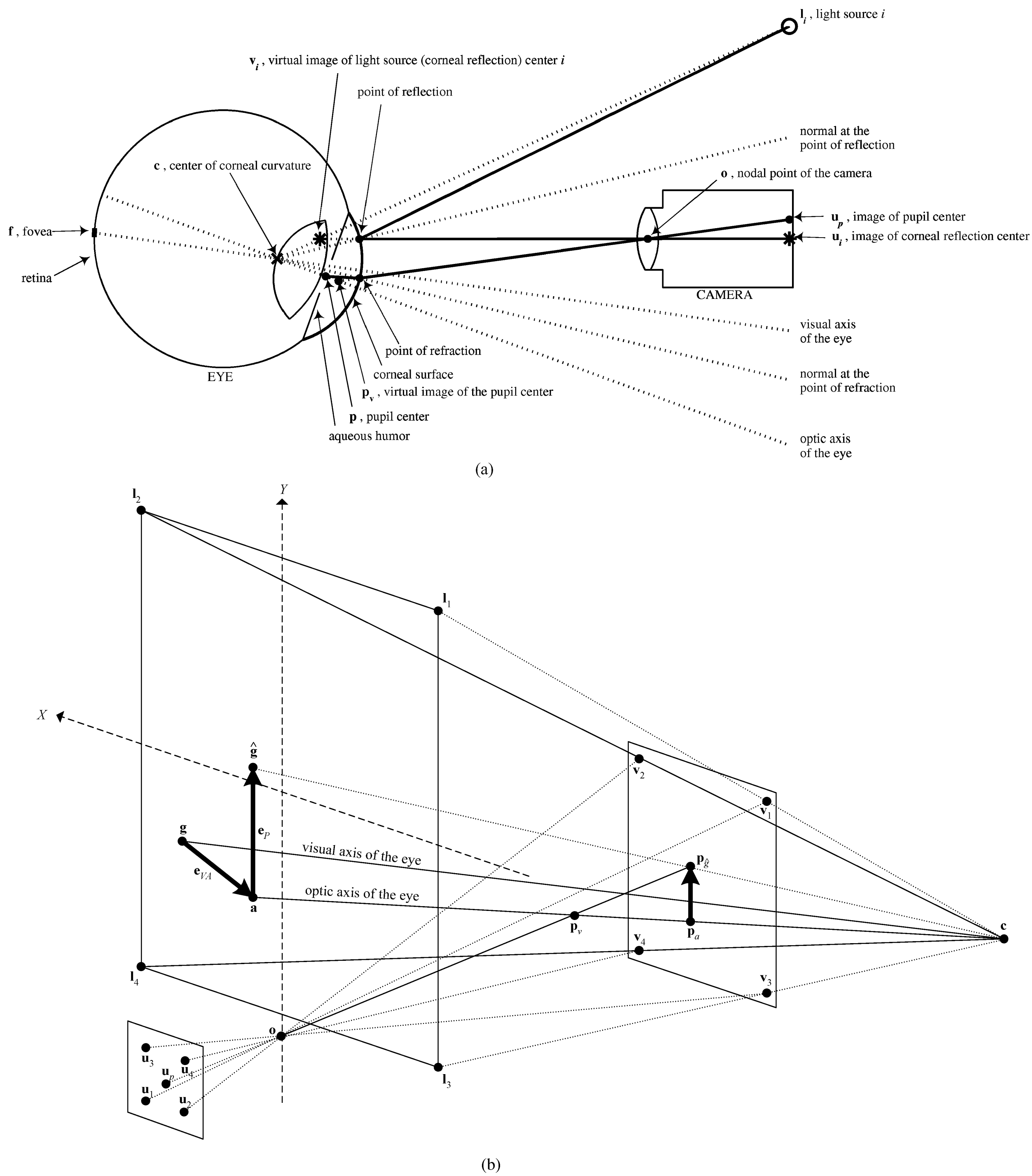

Fig. 2. (a) Ray-tracing diagram adapted from [11]. Eye parts are drawn to scale, whereas the rest of the diagram is not to scale in order to be able to show all the elements of interest. (b) A geometrical description of the two perspective projections relating points on the scene plane to points in the eye, and points in the camera's imaging plane. The four light sources ($\mathbf{l}_1, \mathbf{l}_2, \mathbf{l}_3, \mathbf{l}_4$) are placed on the scene plane, and produce the corneal reflections in the eye ($\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4$) that form the corneal reflection plane. The corneal reflections are then imaged on the camera's imaging plane ($\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \mathbf{u}_4$). The point $\mathbf{o}$ is the nodal point of the camera, $\mathbf{c}$ is the center of curvature of the cornea, $\mathbf{p}_v$ is the virtual pupil center, and $\mathbf{u}_p$ is the image of the pupil center. The point $\mathbf{a}$ is the intersection of the optic axis of the eye with the scene plane, $\mathbf{g}$ is the PoG (intersection of the visual axis of the eye with the scene plane), and $\hat{\mathbf{g}}$ is the PoG estimate.

point of the camera, $\mathbf{o}$, and projects the corneal reflections to their images on the camera's imaging plane ($\mathbf{u}_1$, $\mathbf{u}_2$, $\mathbf{u}_3$, $\mathbf{u}_4$), and the virtual pupil center to its image on the imaging plane, $\mathbf{u}_p$.

An analysis of the aforementioned four requirements reveals that since the virtual pupil center, $\mathbf{p}_v$, and the center of curvature of the cornea, $\mathbf{c}$, are on the optic axis of the eye [see Fig. 2(a)], if all four requirements were met, the PoG estimate calculated by the cross-ratios method would be coincident with the intersection of the optic axis of the eye with the scene, $\mathbf{a}$, as shown in Fig. 2(b). However, the PoG, $\mathbf{g}$, is defined as the intersection of the visual axis [line defined by the center of the fovea, $\mathbf{f}$, and the nodal point of the eye, which is assumed to be coincident with the center of curvature of the cornea, $\mathbf{c}$, as shown in Fig. 2(a)] with the scene. Therefore, the deviation of the visual axis from the optic axis results in an estimation bias represented by vector $\mathbf{e}_{VA} = \mathbf{a}–\mathbf{g}$ [see Fig. 2(b)].

The analysis of the three coplanarity requirements reveals that requirement (1) is met since the light sources and the PoG are all in the scene plane, and requirement (3) is met since the images of the corneal reflections ($\mathbf{u}_1$, $\mathbf{u}_2$, $\mathbf{u}_3$, $\mathbf{u}_4$) and the image of the pupil center, $\mathbf{u}_p$, are all on the camera's imaging plane. However, requirement (2) is not met. Even though for the system setup depicted in Fig. 1(a), the corneal reflections ($\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{v}_3$, $\mathbf{v}_4$) are exactly coplanar when the center of curvature of the cornea, $\mathbf{c}$, is in the plane of symmetry of the system (vertical plane perpendicular to the screen plane containing the center of the screen and the nodal point of the camera), the virtual pupil center, $\mathbf{p}_v$, as shown in Fig. 2(b), is not on the plane of the corneal reflections. Since $\mathbf{p}_v$ is located in front of the corneal reflection plane, when $\mathbf{p}_v$ is viewed from the perspective of the system's camera with its nodal point at $\mathbf{o}$, it appears as if it were located at $\mathbf{p}_{\hat{g}}$ on the corneal reflection plane. The PoG estimate, $\hat{\mathbf{g}}$, obtained with the cross-ratios method is then the projection of $\mathbf{p}_{\hat{g}}$ onto the scene plane. Therefore, the fact that the virtual pupil center is not on the corneal reflection plane results in an additional estimation bias represented by vector $\mathbf{e}_P = \hat{\mathbf{g}}–\mathbf{a}$ [see Fig. 2(b)]. Notice that since $\mathbf{a}$ can be thought of as the projection of $\mathbf{p}_a$ onto the scene plane, where $\mathbf{p}_a$ is defined as the intersection of the optic axis of the eye with the corneal reflection plane, it follows that $\mathbf{e}_P$ is the projection of the vector ($\mathbf{p}_{\hat{g}}–\mathbf{p}_a$) onto the scene plane.

Based on this analysis, the total PoG estimation bias, $\mathbf{e} = \hat{\mathbf{g}}–\mathbf{g}$, can be expressed as

$$\mathbf{e} = \hat{\mathbf{g}} - \mathbf{g} = \underbrace{(\hat{\mathbf{g}} - \mathbf{a})}_{\mathbf{e}_P} + \underbrace{(\mathbf{a} - \mathbf{g})}_{\mathbf{e}_{VA}}. \quad (1)$$

An analytical expression for the estimation bias is derived in the next section.

III. Analytical Expression for the Estimation Bias

An analytical expression for the estimation bias as a function of a set of system and subject-specific eye parameters is derived with the aid of Fig. 2(b). Approximations and simplifications are made in order to obtain a closed-form expression that is easy to interpret, and provides insight into the effects of the subject-specific eye parameters on each cause of estimation bias. The coordinates of all points in Fig. 2(b) are measured with respect to a right-handed 3-D Cartesian world coordinate system (WCS) that is defined such that the origin is located at the center of the screen, and the XY-plane is coincident with the plane of the screen (scene plane), with the X-axis being horizontal, the Y-axis being vertical pointing up, and the Z-axis extending outward from the scene plane. Within this coordinate system, the locations of the four light sources are expressed as $\mathbf{l}_i = [\pm w/2 \;\; \pm h/2 \;\; 0]^T$, the PoG, i.e., the intersection of the visual axis with the scene plane is expressed as $\mathbf{g} = [g_X \, g_Y \, 0]^T$, the intersection of the optic axis with the scene plane is expressed as $\mathbf{a} = [a_X \, a_Y \, 0]^T$, and the PoG estimate calculated by the cross-ratios method is expressed as $\hat{\mathbf{g}} = [\hat{g}_X \, \hat{g}_Y \, 0]^T$. To simplify the analytical expressions, it is assumed that the center of curvature of the cornea, $\mathbf{c}$, is set at a location on the Z-axis, i.e., $\mathbf{c} = [0 \; 0 \; c_Z]^T$.

The estimation bias associated with the deviation of the visual axis from the optic axis, $\mathbf{e}_{VA} = \mathbf{a}–\mathbf{g}$ [see Fig. 2(b)], can be approximated by

$$\mathbf{e}_{VA} = \begin{bmatrix} e_{VA,X} \\ e_{VA,Y} \\ e_{VA,Z} \end{bmatrix} = \mathbf{a} - \mathbf{g} \approx \begin{bmatrix} -c_Z \, \tan(\alpha_{\text{eye}}) \\ -c_Z \, \tan(\beta_{\text{eye}}) \\ 0 \end{bmatrix} \quad (2)$$

where c_Z is the distance between the center of curvature of the cornea and the scene plane, and α_{eye} and β_{eye} are, respectively, the horizontal and vertical signed angles that describe the orientation of the visual axis with respect to the optic axis.

An approximate analytical expression for the estimation bias due to the fact that the virtual pupil center is not on the corneal reflection plane can be derived from the following set of observations. Each corneal reflection $\mathbf{v}_i$ is on the line joining the corresponding light source, $\mathbf{l}_i$, and the center of curvature of the cornea, $\mathbf{c}$. Since the distance D_{cr} of each corneal reflection from the center of curvature of the cornea is approximately constant and similar for all corneal reflections, the positions of the corneal reflections can be expressed as

$$\mathbf{v}_i \approx \mathbf{c} + D_{\text{cr}} \frac{\mathbf{l}_i - \mathbf{c}}{||\mathbf{l}_i - \mathbf{c}||}. \quad (3)$$

Note that the distance between each light source and the eye is much larger than the radius of curvature of the cornea, R, and, therefore, D_{cr} is approximately equal to $R/2$. Recalling that $\mathbf{l}_i = [\pm w/2 \;\; \pm h/2 \;\; 0]^T$, it then follows that the Z-coordinate of each $\mathbf{v}_i$ is given by

$$v_{i,Z} \approx c_Z \left(1 - D_{\text{cr}} \frac{1}{\sqrt{w^2/4 + h^2/4 + c_Z^2}} \right) \quad (4)$$

which is the same for all four corneal reflections. In other words, the corneal reflections are on the plane $Z = v_{i,Z}$, which is parallel to the scene plane ($Z = 0$).

The pupil center, $\mathbf{p}$, is located on the optic axis at a distance K from the center of curvature of the cornea, $\mathbf{c}$. Due to refraction at the cornea–air interface, the camera observes the virtual pupil center, $\mathbf{p}_v$, located on the optic axis at a distance $K_v > K$ from the center of curvature of the cornea, $\mathbf{c}$. Noting that the optic axis has the direction of $\mathbf{a}$–$\mathbf{c}$ [see Fig. 2(b)], $\mathbf{p}_v$ can be expressed as

$$\mathbf{p}_v = \mathbf{c} + K_v \frac{\mathbf{a}-\mathbf{c}}{\|\mathbf{a}-\mathbf{c}\|} = \begin{bmatrix} \dfrac{K_v}{\sqrt{a_X^2+a_Y^2+c_Z^2}} a_X \\ \dfrac{K_v}{\sqrt{a_X^2+a_Y^2+c_Z^2}} a_Y \\ c_Z \left(1 - \dfrac{K_v}{\sqrt{a_X^2+a_Y^2+c_Z^2}}\right) \end{bmatrix}. \tag{5}$$

From the perspective of the camera with nodal point at $\mathbf{o}$, it appears as if the virtual pupil center were located at $\mathbf{p}_{\hat{g}}$, which is defined by the intersection of the line passing through $\mathbf{o}$ and $\mathbf{p}_v$ with the corneal reflection plane [see Fig. 2(b)]. The point $\mathbf{p}_{\hat{g}}$ can be expressed as

$$\mathbf{p}_{\hat{g}} = \mathbf{o} + k_{o\hat{g}}(\mathbf{p}_v - \mathbf{o}) \tag{6}$$

where $k_{o\hat{g}}$ is such that $p_{\hat{g},Z} = v_{i,Z}$. By substituting (5) into (6) and equating the resulting expression for $p_{\hat{g},Z}$ with the expression for $v_{i,Z}$ from (4) and setting the nodal point of the camera at the scene plane, i.e., $o = [o_X\, o_Y\, 0]^T$, it follows that $k_{o\hat{g}}$ is given by

$$k_{o\hat{g}} \approx \frac{1 - D_{\rm cr}/\sqrt{w^2/4 + h^2/4 + c_Z^2}}{1 - K_v/\sqrt{a_X^2+a_Y^2+c_Z^2}} \tag{7}$$

and the X- and Y-coordinates of $\mathbf{p}_{\hat{g}}$ are expressed as

$$\begin{bmatrix} p_{\hat{g},X} \\ p_{\hat{g},Y} \end{bmatrix} = \begin{bmatrix} o_X \\ o_Y \end{bmatrix} + k_{o\hat{g}} \left(\frac{K_v}{\sqrt{a_X^2+a_Y^2+c_Z^2}} \begin{bmatrix} a_X \\ a_Y \end{bmatrix} - \begin{bmatrix} o_X \\ o_Y \end{bmatrix} \right). \tag{8}$$

Finally, consider the mapping from the corneal reflection plane to the scene plane. Since the corneal reflection plane is parallel to the scene plane, and the center of curvature of the cornea, $\mathbf{c}$, which acts as the center of projection, is located on the Z-axis, it follows that the mapping of the X- and Y-coordinates of a point on the corneal reflection plane to the corresponding X- and Y-coordinates on the scene plane is given by the scale factor

$$\eta = \frac{c_Z}{c_Z - v_{i,Z}} \approx \frac{\sqrt{w^2/4 + h^2/4 + c_Z^2}}{D_{\rm cr}} \tag{9}$$

where c_Z is the distance from the center of projection (the center of curvature of the cornea) to the scene plane and $(c_Z - v_{i,Z})$ is the distance from the center of projection to the corneal reflection plane. Since the PoG estimate, $\hat{\mathbf{g}}$, is the projection of $\mathbf{p}_{\hat{g}}$ onto the scene plane, $\hat{\mathbf{g}}$ is given by

$$\hat{\mathbf{g}} = \begin{bmatrix} \hat{g}_X \\ \hat{g}_Y \\ \hat{g}_Z \end{bmatrix} = \begin{bmatrix} \eta\, p_{\hat{g},X} \\ \eta\, p_{\hat{g},Y} \\ 0 \end{bmatrix}. \tag{10}$$

From (7) to (10), and recalling that the nodal point of the camera was set at the scene plane, i.e., $o_Z = 0$, and that $a_Z = 0$ by definition, it follows that the estimation bias associated with the fact that the virtual pupil center is not on the corneal reflection plane is given by

$$\mathbf{e}_P = \hat{\mathbf{g}} - \mathbf{a} \approx \lambda(\mathbf{a} - \mathbf{o}) \tag{11}$$

where

$$\lambda = \frac{K_v}{D_{\rm cr}}\gamma - 1 \tag{12}$$

and

$$\gamma = \frac{\sqrt{w^2/4 + h^2/4 + c_Z^2} - D_{\rm cr}}{\sqrt{a_X^2+a_Y^2+c_Z^2} - K_v} \approx \frac{\sqrt{w^2/4 + h^2/4 + c_Z^2}}{\sqrt{a_X^2+a_Y^2+c_Z^2}}. \tag{13}$$

The approximation in (13) is based on the fact that $c_Z >> D_{\rm cr}$ and $c_Z >> K_v$.

The factor λ in (12) is a function of $D_{\rm cr}$ and K_v. In strict terms, $D_{\rm cr}$ (the distance between each corneal reflection and the center of curvature of the cornea) is a function of: 1) the radius of curvature of the cornea, R and 2) the relative position of the center of curvature of the cornea with respect to the camera and each light source. Similarly, K_v (the distance between the virtual pupil center and the center of curvature of the cornea) is a function of: 1) the distance between the pupil center and the center of curvature of the cornea, K; 2) the radius of curvature of the cornea, R; 3) the effective index of refraction of the aqueous humor and cornea combined, n_1; and 4) the relative position and orientation of the eye in space with respect to the camera. In order to keep the expressions for $D_{\rm cr}$ and K_v simple, their values are first obtained under an assumption of paraxial rays, and then, multiplied by respective correction factors that account for the fact that the rays are not paraxial. In this way, $D_{\rm cr}$ and K_v are expressed as functions of the eye parameters R, K, and n_1 as

$$D_{\rm cr} \approx \mu \frac{R}{2} \tag{14}$$

where μ is a correction factor, and

$$K_v \approx \nu \frac{n_1 R K}{R + (n_1 - 1)K} \tag{15}$$

where ν is another correction factor and $n_1 = 1.3375$ [10].

From (1), (2), and (11), the total PoG estimation bias, $\mathbf{e} = \hat{\mathbf{g}}$–$\mathbf{g}$, can be expressed in terms of the PoG, $\mathbf{g}$, as

$$\mathbf{e} = \hat{\mathbf{g}} - \mathbf{g} \approx \lambda(\mathbf{g} - \mathbf{o}) + (\lambda + 1) \begin{bmatrix} -c_Z \tan(\alpha_{\rm eye}) \\ -c_Z \tan(\beta_{\rm eye}) \\ 0 \end{bmatrix}. \tag{16}$$

By expressing $\mathbf{a}$ in terms of $\mathbf{g}$ using (2), and substituting the resulting expression for $\mathbf{a}$ into (13), it follows from (12), (14), and (15) that λ can be expressed as

$$\lambda \approx \frac{\nu}{\mu} \frac{2 n_1 K}{R + (n_1 - 1)K}\gamma - 1 \tag{17}$$

where

$$\gamma \approx \frac{\sqrt{w^2/4 + h^2/4 + c_Z^2}}{\sqrt{[g_X - c_Z \tan(\alpha_{\rm eye})]^2 + [g_Y - c_Z \tan(\beta_{\rm eye})]^2 + c_Z^2}}. \tag{18}$$

The expression for the total PoG estimation bias (16)–(18) is a function of: 1) the system configuration (the position of the light sources, $\mathbf{l}_i = [\pm w/2 \;\; \pm h/2 \;\; 0]^T$, and the position of the nodal point of the camera, $\mathbf{o}$); 2) the eye position ($\mathbf{c} = [0 \quad 0 \quad c_Z]^T$);

3) the PoG, $\mathbf{g} = [g_X\, g_Y\, 0]^T$; and 4) the subject-specific eye parameters R, K, $n_1, \alpha_{\mathrm{eye}}$, and β_{eye}. The large variation in estimation bias among subjects is caused by differences in these subject-specific eye parameters.

To evaluate the ability of (16)–(18) to predict the PoG estimation bias of the cross-ratios method, a PoG estimation system was implemented and tested with six subjects who were asked to fixate on a sequence of nine points that were presented on a computer screen. The PoG estimation system included four near-infrared (850 nm) light sources that were placed in a rectangular formation around a computer monitor and a 640 by 480 charged-coupled device (CCD) video camera (*Dragonfly,* Point Grey Research, Vancouver, Canada) with a 75 mm lens that was placed centered below the monitor to capture images of the right eye. Within the WCS that was defined at the beginning of this section, the locations of the four light sources were $\mathbf{l}_i = [\pm w/2 \;\; \pm h/2 \;\; 0]^T$, where $w = 500$ mm and $h = 360$ mm, whereas the position of the nodal point of the camera was $\mathbf{o} = [0 \;\; -210 \;\; 0]^T$ mm. A chinrest was used to minimize head movements and to keep the right eye approximately at the center of the camera's image while maintaining an approximately constant distance of 1 m between the right eye and the screen. For each of the nine fixation points, the coordinates of the centers of the pupil and corneal reflections in the eye images were estimated for 100 consecutive video frames (at 30 frames per second). Using the cross-ratios method described in [6], the PoG estimate, $\hat{\mathbf{g}}$, was calculated for each video frame.[1] The average PoG estimates for each subject and for each fixation point are presented in Fig. 3 ("+" markers). These results show that the bias in the PoG estimates varied significantly among subjects. The rms error from estimation bias varied from 17.3 mm (subject 3) to 164.5 mm (subject 1). It is important to note that the rms dispersion of individual PoG estimates around the corresponding average PoG estimates (caused by noise in the measurements of the centers of the pupil and corneal reflections and by small fixation eye movements) was much smaller and was similar for all subjects [varied from 3.7 mm (subject 1) to 4.7 mm (subject 4)].

To calculate the PoG estimation bias with (16)–(18), the subject-specific eye parameters R, K, $n_1, \alpha_{\mathrm{eye}}$, and β_{eye}, as well as the correction factors μ and ν, have to be known. After setting $n_1 = 1.3375$, the four remaining subject-specific eye parameters (R, K, α_{eye}, and β_{eye}) were obtained through the eye parameter estimation procedure described in [11].[2] The values of these four subject-specific eye parameters for the six subjects are listed in Table I. Using the model from [11], it was found that for the experimental conditions described before, the correction factor μ varies between 1.01 and 1.04, while the correction factor ν varies between 1.001 and 1.007. Furthermore, note that the value of the factor γ in (18) varies between 1.02 and 1.04. By taking the average values of γ, μ, and ν, (17) can be simplified as

$$\lambda \approx \lambda_{\mathrm{exp}} = 2.01 \frac{n_1 \hat{K}}{\hat{R} + (n_1 - 1)\hat{K}} - 1. \qquad (19)$$

[1] See note in Subsection A of the Appendix.

[2] See note in Subsection B of the Appendix.

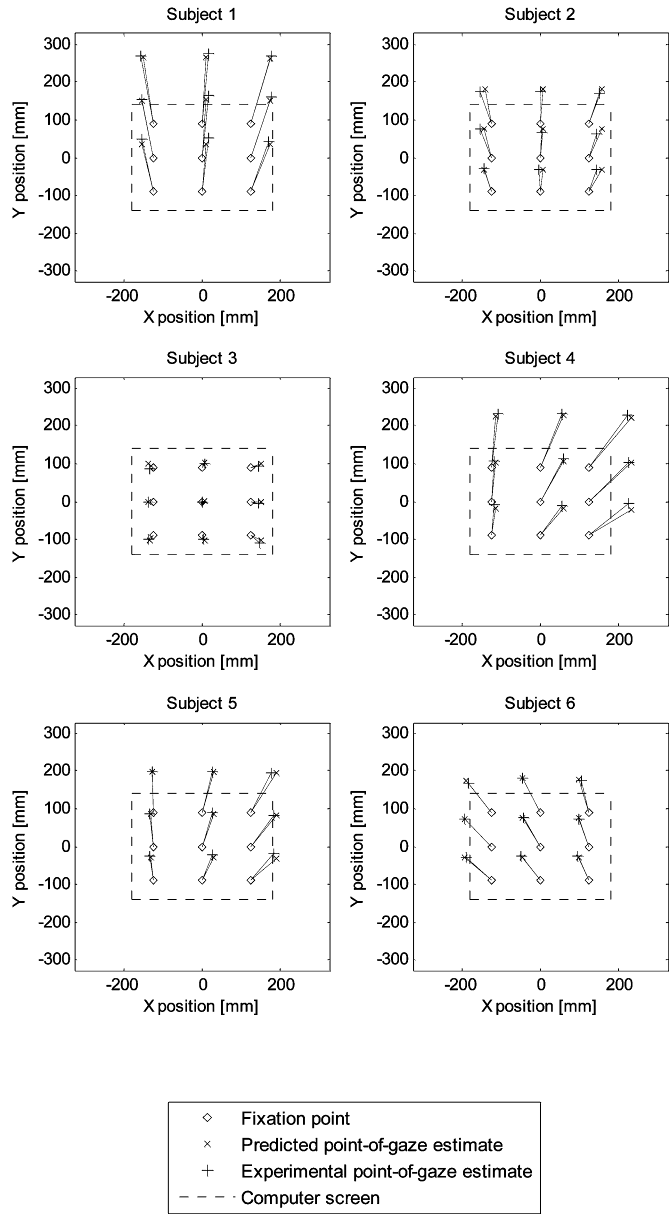

Fig. 3. PoG estimates predicted analytically as compared to the experimental estimates.

Substituting the values of the estimated subject-specific eye parameters (see Table I) and the system parameters into (16) and (19), the PoG estimation bias for each subject and each fixation point was calculated (see Fig. 3, "×" markers). The results shown in Fig. 3 suggest that the approximate analytical expression given by (16) and (19) provides a very good prediction of the highly variable experimental PoG estimation bias observed for the six subjects. The rms error between the PoG estimation bias predicted by the approximate analytical expression given by (16) and (19) and the experimental PoG estimation bias varied between 5.5 mm (subject 6) and 12.2 mm (subject 2). Note that, in the best case, the rms error between the estimation bias predicted by the analytical expression and the experimental estimation bias is similar in magnitude to the dispersion error caused by image noise and small fixation eye movements

TABLE I
SUBJECT-SPECIFIC EYE PARAMETERS

	$\hat{R}$ (mm)	$\hat{K}$ (mm)	$\hat{\alpha}_{\text{eye}}$ (degrees)	$\hat{\beta}_{\text{eye}}$ (degrees)
Subject 1	8.01	4.62	-0.46	-3.73
Subject 2	7.45	3.88	-0.36	-1.48
Subject 3	8.07	3.98	-0.38	1.66
Subject 4	8.46	5.24	-2.49	-0.84
Subject 5	7.50	4.21	-1.37	-1.06
Subject 6	7.88	3.96	2.30	-1.94

(3.7–4.7 mm). The differences between the predicted and observed estimation bias can be attributed to the approximations and assumptions that were made during the development of the analytical expressions as well as to experimental fixation inaccuracies. Nevertheless, using this analytical expression, the PoG estimation bias was predicted for all subjects to within 0.7° of visual angle.

IV. Compensation for the Estimation Bias

To minimize the PoG estimation bias of the cross-ratios method, Coutinho and Morimoto [8] suggested a compensation procedure that is based on the work of Yoo and Chung [7]. In [7], a procedure is described whereby the image coordinates of each corneal reflection are modified according to

$$\tilde{\mathbf{u}}_i' = \tilde{\mathbf{u}}_o + \alpha(\tilde{\mathbf{u}}_i - \tilde{\mathbf{u}}_o) \tag{20}$$

where $\tilde{\mathbf{u}}_i$ describes the image coordinates of the corneal reflection, $\tilde{\mathbf{u}}_i'$ describes the modified image coordinates of the corneal reflection, $\tilde{\mathbf{u}}_o$ describes the image coordinates of a fifth corneal reflection generated by a light source placed at the nodal point of the system's camera,[3] and α is a constant that governs the magnitude of the change in the image coordinates of the corneal reflections. The parameter α is determined for each subject through a calibration procedure. In this section, we explain this correction methodology using the theoretical framework established in the previous section. Specifically, we will show how the modification of the image coordinates of the corneal reflections compensates for one of the main causes of estimation bias. Furthermore, the analysis will provide a physical interpretation for α.

As was shown in the previous section, one of the main causes of estimation bias in the cross-ratios method is associated with the fact that the virtual pupil center is not in the corneal reflection plane. It then follows that a correction methodology that is based on the modification of the image coordinates of the corneal reflections should be equivalent to moving each of the corneal reflections along their lines of projection (the line joining the corresponding light source, $\mathbf{l}_i$, to the center of curvature of the cornea, $\mathbf{c}$) to a new plane that also contains the virtual pupil center. Based on the analytical formulation of Section III, the moved corneal reflections, $\mathbf{v}_i'$, can be expressed as

$$\mathbf{v}_i' = \mathbf{c} + A(\mathbf{v}_i - \mathbf{c}) \tag{21}$$

[3]In practice, this is accomplished by using a ring light source positioned around the camera's lens such that the center of the ring is coincident with the nodal point, and the plane of the ring is perpendicular to the camera's optic axis.

where A is such that $v_{i,Z}' = p_{v,Z}$ (note that the new plane is parallel to the plane of the original corneal reflections). From (4) and (21), $v_{i,Z}'$ can be expressed as

$$v_{i,Z}' \approx c_Z\left(1 - D_{\text{cr}}\frac{A}{\sqrt{w^2/4 + h^2/4 + c_Z^2}}\right). \tag{22}$$

By equating $v_{i,Z}'$ from (22) with $p_{v,Z}$ from (5), A is given by

$$A \approx \frac{K_v}{D_{\text{cr}}}\frac{\sqrt{w^2/4 + h^2/4 + c_Z^2}}{\sqrt{a_X^2 + a_Y^2 + c_Z^2}} \approx \frac{K_v}{D_{\text{cr}}}\gamma = \lambda + 1 \tag{23}$$

where the approximation with respect to γ is based on (13), and the last equality is based on (12) [the full expression for λ is given in (17)–(18)]. Recall that K_v is the distance of the virtual pupil center to the center of curvature of the cornea and D_{cr} is the distance of the corneal reflections to the center of curvature of the cornea.

Next, consider the projection of the moved corneal reflections, $\mathbf{v}_i'$, onto the camera's imaging plane. Points on the imaging plane are described in pixels with respect to a 2-D image coordinate system (ICS) having an x-axis in the direction of the rows and a y-axis in the direction of the columns of the imaging sensor. In the ICS, the image coordinates of the moved corneal reflections can be expressed as

$$\tilde{\mathbf{u}}_i' = \begin{bmatrix} \tilde{u}_{i,x}' \\ \tilde{u}_{i,y}' \end{bmatrix} = \begin{bmatrix} \dfrac{f}{s_x}\dfrac{(\mathbf{v}_i' - \mathbf{o})^T\mathbf{i}_{\text{cam}}}{(\mathbf{v}_i' - \mathbf{o})^T\mathbf{k}_{\text{cam}}} + q_x \\ \dfrac{f}{s_y}\dfrac{(\mathbf{v}_i' - \mathbf{o})^T\mathbf{j}_{\text{cam}}}{(\mathbf{v}_i' - \mathbf{o})^T\mathbf{k}_{\text{cam}}} + q_y \end{bmatrix} \tag{24}$$

where $\mathbf{o}$ is the nodal point of the camera, and $\mathbf{i}_{\text{cam}}, \mathbf{j}_{\text{cam}},$ and $\mathbf{k}_{\text{cam}}$ are the columns of the rotation matrix of the camera, $\mathbf{R} = [\mathbf{i}_{\text{cam}}\ \mathbf{j}_{\text{cam}}\ \mathbf{k}_{\text{cam}}]$, describing the orientation of the camera with respect to the previously defined WCS. The parameter f is the camera's effective focal length, s_x and s_y are, respectively, the pixel pitch in the direction of the rows and columns of the imaging sensor, and q_x and q_y define the image coordinates of the principal point [12]. By substituting (21) into (24), the image coordinate of the moved corneal reflections in the direction of the x-axis is given by

$$\begin{aligned} \tilde{u}_{i,x}' &= \frac{f}{s_x}\frac{[\mathbf{c} + A(\mathbf{v}_i - \mathbf{c}) - \mathbf{o}]^T\ \mathbf{i}_{\text{cam}}}{(\mathbf{v}_i' - \mathbf{o})^T\mathbf{k}_{\text{cam}}} + q_x \\ &= \left(\frac{f}{s_x}\frac{(\mathbf{c} - \mathbf{o})^T\mathbf{i}_{\text{cam}}}{(\mathbf{v}_i' - \mathbf{o})^T\mathbf{k}_{\text{cam}}} + q_x\right) \\ &\quad + A\left[\left(\frac{f}{s_x}\frac{(\mathbf{v}_i - \mathbf{o})^T\mathbf{i}_{\text{cam}}}{(\mathbf{v}_i' - \mathbf{o})^T\mathbf{k}_{\text{cam}}} + q_x\right)\right. \\ &\quad \left.- \left(\frac{f}{s_x}\frac{(\mathbf{c} - \mathbf{o})^T\mathbf{i}_{\text{cam}}}{(\mathbf{v}_i' - \mathbf{o})^T\mathbf{k}_{\text{cam}}} + q_x\right)\right]. \end{aligned} \tag{25}$$

Since the distance between the eye features ($\mathbf{v}_i'$, $\mathbf{v}_i$, $\mathbf{c}$) is much shorter than the distance between the eye features and the camera's imaging plane, $(\mathbf{v}_i' - \mathbf{o})^T\mathbf{k}_{\text{cam}} \approx (\mathbf{v}_i - \mathbf{o})^T\mathbf{k}_{\text{cam}} \approx$

$(\mathbf{c}-\mathbf{o})^T\mathbf{k}_{\text{cam}}$, and therefore, (25) can be approximated by

$$\tilde{u}'_{i,x} \approx \underbrace{\frac{f}{s_x}\frac{(\mathbf{c}-\mathbf{o})^T\mathbf{i}_{\text{cam}}}{(\mathbf{c}-\mathbf{o})^T\mathbf{k}_{\text{cam}}}+q_x}_{\tilde{u}_{c,x}} + A\Bigg[\underbrace{\left(\frac{f}{s_x}\frac{(\mathbf{v}_i-\mathbf{o})^T\mathbf{i}_{\text{cam}}}{(\mathbf{v}_i-\mathbf{o})^T\mathbf{k}_{\text{cam}}}+q_x\right)}_{\tilde{u}_{i,x}} - \underbrace{\left(\frac{f}{s_x}\frac{(\mathbf{c}-\mathbf{o})^T\mathbf{i}_{\text{cam}}}{(\mathbf{c}-\mathbf{o})^T\mathbf{k}_{\text{cam}}}+q_x\right)}_{\tilde{u}_{c,x}}\Bigg]$$

$$\approx \tilde{u}_{c,x} + A(\tilde{u}_{i,x}-\tilde{u}_{c,x}) \qquad (26)$$

where, in the direction of the x-axis, $\tilde{u}_{i,x}$ is the coordinate of the image of $\mathbf{v}_i$ (the corneal reflection), $\tilde{u}'_{i,x}$ is the coordinate of the image of $\mathbf{v}'_i$ (the moved corneal reflection), and $\tilde{u}_{c,x}$ is the coordinate of the image of $\mathbf{c}$ (the center of curvature of the cornea). Since an analogous expression is obtained in the direction of the y-axis, the image coordinates of the moved corneal reflection can be expressed in vector form as

$$\tilde{\mathbf{u}}'_i \approx \tilde{\mathbf{u}}_c + A(\tilde{\mathbf{u}}_i - \tilde{\mathbf{u}}_c). \qquad (27)$$

Notice that since any line passing through the center of curvature of the cornea, $\mathbf{c}$, is normal to the corneal surface, a ray of light originating at $\mathbf{o}$ and directed toward $\mathbf{c}$ will be reflected back on itself. Therefore, the corneal reflection of the light source at $\mathbf{o}$, i.e., $\mathbf{v}_o$, will be collinear with $\mathbf{c}$ and $\mathbf{o}$. It, then, follows that the image of $\mathbf{c}$ and the image of $\mathbf{v}_o$ are identical, i.e., $\tilde{\mathbf{u}}_c = \tilde{\mathbf{u}}_o$. When $\tilde{\mathbf{u}}_c$ in (27) is replaced with $\tilde{\mathbf{u}}_o$, it has the same form as (20), the original formula for modifying the image coordinates of the corneal reflections proposed in [7]. A comparison of these two equations shows that if the value of α is equal to A, the image coordinates of the corneal reflections will be modified in such a way that is equivalent, within the approximations, to moving the corneal reflections to a new plane that contains the virtual pupil center, and is parallel to the original corneal reflection plane. Moreover, the value of α, obtained through calibration in [8], can be estimated using (23).

The system described in Section III was modified to include a ring light source around the lens of the camera such that the plane of the ring is perpendicular to the optic axis of the camera and the center of the ring is coincident with the nodal point of the lens of the camera. This ring light source simulates a light source placed at the nodal point of the camera. The experiment described in Section III was repeated for the new system configuration. Following the calibration procedure described in [8], the value of α was determined for each subject (see Table II, column 1). Using the approximate expression of (23) (i.e., $A \approx \lambda + 1$) together with the approximate expression for λ from (19), the value of A was also calculated for each subject (see Table II, column 2). The results in the first two columns of Table II show that, for each subject, the value of α obtained using the procedure described in [8] is similar (within 3%) to the value of A obtained using the approximate expression of (23). Since the approximate expression for A was derived to minimize the distance between the plane of the moved corneal reflections and the virtual pupil center, this implies that the value of α, obtained by the procedure described in [8], also minimizes this distance.

TABLE II
ESTIMATION BIAS COMPENSATION

	Compensation Parameters					
	α	$A \approx \lambda+1$	$m_{\alpha,X}$ (mm)	$-e_{VA,X}$ (mm)	$m_{\alpha,Y}$ (mm)	$-e_{VA,Y}$ (mm)
Subject 1	1.30	1.30	-10.0	-8.0	-65.7	-65.3
Subject 2	1.16	1.19	0.6	6.3	-26.8	-25.9
Subject 3	1.11	1.13	-3.6	-6.6	29.9	29.0
Subject 4	1.35	1.37	-42.2	-43.4	-15.5	-14.7
Subject 5	1.24	1.27	-20.2	-23.9	-22.2	-18.5
Subject 6	1.14	1.15	39.4	40.2	-32.9	-33.9

By modifying the image coordinates of the corneal reflections using this value of α in (20), the procedure described in [8] compensates for the estimation bias associated with the fact that the virtual pupil center is not on the plane of the corneal reflections (i.e., $\mathbf{e}_P$).

Note that in the original derivation of the compensation procedure [7], the authors suggested that the corneal reflections should be moved to a plane tangent to the surface of the cornea, and therefore, α should be 2. Since the virtual pupil center is not located on the surface of the cornea, as was assumed in [7], this value of α does not, in general, reduce the distance of the corneal reflection plane to the virtual pupil center. Consequently, when α is 2, the estimation bias is not necessarily reduced, and may, in fact, be increased, as it was observed in [8].

In the compensation procedure described in [8], after the PoG estimates are calculated using the modified image coordinates of the corneal reflections, a second step is performed in which the PoG estimates are modified by adding $\mathbf{m}_\alpha = [m_{\alpha,X}\ m_{\alpha,Y}]^T$, an offset vector that accounts for the deviation of the visual axis from the optic axis. This offset, $\mathbf{m}_\alpha$, corresponds to the negative of the estimation bias $\mathbf{e}_{VA}$. For each subject, $\mathbf{m}_\alpha$ was obtained using the calibration procedure of [8], and $\mathbf{e}_{VA}$ was calculated using (2) for the subject's eye parameters α_{eye} and β_{eye} (see Table I) and $\mathbf{c} = [0\ 0\ 1000]^T$ mm. Table II shows that the values of $\mathbf{m}_\alpha$ are very similar (within 6 mm) to the values of–$\mathbf{e}_{VA}$ that were calculated using (2). This implies that by adding $\mathbf{m}_\alpha$, the procedure described in [8] compensates for the estimation bias that is associated with the deviation of the visual axis from the optic axis (i.e., $\mathbf{e}_{VA}$).

In summary, the previous analysis shows that the first stage of the bias compensation procedure proposed in [8], which depends on the parameter α, compensates for the estimation bias due to the fact that the virtual pupil center is not on the corneal reflection plane (i.e., $\mathbf{e}_P$), whereas the second step, which depends on the parameter $\mathbf{m}_\alpha$, compensates for the estimation bias due to the deviation of the visual axis from the optic axis (i.e., $\mathbf{e}_{VA}$). Although α and $\mathbf{m}_\alpha$ were presented in [8] simply as parameters to be optimized, the theoretical framework developed in this paper provides analytical expressions that establish the condition of optimality for the values of α [(17), (18), and (23)] and $\mathbf{m}_\alpha$ [negative of (2)] as functions of subject-specific eye parameters. Furthermore, the theoretical framework from this paper provides a clear physical interpretation of the two steps of the bias compensation procedure from [8].

V. Conclusion

In this paper, we examined the cross-ratios method for PoG estimation and identified the two main causes of estimation bias: 1) the deviation of the visual axis from the optic axis and 2) the fact that the virtual pupil center is not on the corneal reflection plane. The estimation bias associated with the deviation of the visual axis from the optic axis varies among subjects according to the horizontal and vertical angles of deviation, α_{eye} and β_{eye}, respectively. The estimation bias associated with the fact that the virtual pupil center is not on the corneal reflection plane was shown to vary among subjects according to a parameter that is a function of the distance of the pupil center to the center of curvature of the cornea, K, and the radius of curvature of the cornea, R. When these four subject-specific eye parameters are known, the estimation bias of the cross-ratios method can be predicted by an analytical expression (see Section III).

The empirically derived estimation bias compensation procedure proposed in [8] (refined version of the procedure from [7]) was analyzed using the theoretical framework developed in Section III. The analysis provides a clear physical interpretation of how the calibration procedure from [8] compensates for the two main causes of PoG estimation bias inherent to the cross-ratios method. Furthermore, the analysis provides analytical expressions that establish the condition of optimality for the values of the parameters of the estimation bias compensation procedure as functions of subject-specific eye parameters.

Appendix

A. Note on the Theoretical Analysis vs. the Experimental Results

The theoretical analysis carried out in this paper assumed that the PoG is estimated using the image of the pupil center. For the cross-ratios method [6], however, the PoG is estimated using the center of the pupil in the eye images [see Fig. 1(b)], i.e., the center of the pupil image. As it was pointed out in [13], due to the refraction at the cornea and the perspective projection at the camera, in general, the image of the pupil center does not fall at the center of the pupil image. For the experimental conditions described at the end of Section III, it can be shown through numerical simulations that the rms distance between PoG estimates obtained with the cross-ratios method using the center of the pupil image and the corresponding PoG estimates using the image of the pupil center can be expected to range from 0.6 to 12 mm (depends on the eye parameters, increasing with the pupil diameter—pupil diameters ranging from 2 to 8 mm were considered). In contrast, for most of the subjects tested (see Fig. 3), the rms PoG estimation bias exhibited by the cross-ratios method is much larger than 12 mm, reaching as much as 164.5 mm for subject 1. These results clearly indicate that the difference between the center of the pupil image and the image of the pupil center does not play a key role in the explanation of the large estimation bias exhibited by the cross-ratios method. In order to simplify the analysis of the main sources of PoG estimation bias in the cross-ratios method, it was assumed that the image of the pupil center was coincident with the center of the pupil image. Even though this assumption limits the ability of the analysis to fully model the PoG estimation bias in the cross-ratios method, the simplified analysis provides valuable insight into the main sources of bias in this method.

B. Note on the Estimation of the Eye Parameters for the Prediction of the Estimation Bias in the Cross-Ratios Method

The model used to estimate the eye parameters [11] assumes that the image of the pupil center is used to estimate the PoG. However, the eye parameters provided in Table I were obtained by using the center of the pupil image. As a result of using the center of the pupil image instead of the image of the pupil center to estimate the eye parameters, the estimated value of the distance between the pupil center and the center of curvature of the cornea, $\hat{K}$, can be expected to be larger than the true value, K, by 0.2–5%, depending on the actual eye parameters, including the pupil diameter.

Because the theoretical analysis from this paper and the model for the estimation of the eye parameters [11] both assume that the PoG is estimated using the image of the pupil center, whereas the actual estimation of the PoG with the cross-ratios method and the actual estimation of the eye parameters both used the center of the pupil image, using $\hat{K}$ to predict the bias of the cross-ratios method [(16) and (19)] tends to minimize the effect of the modeling assumption that the image of the pupil center coincides with the center of the pupil image.

References

[1] 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," *Psychiat. Res.*, vol. 118, no. 2, pp. 117–128, May 2003.

[2] 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," *Accid. Anal. Prev.*, vol. 39, no. 2, pp. 372–379, Mar. 2007.

[3] J. H. Goldberg and X. P. Kotval, "Computer interface evaluation using eye movements: Methods and constructs," *Int. J. Ind. Erg.*, vol. 24, no. 6, pp. 631–645, Oct. 1999.

[4] R. Sharma, V. I. Pavlović, and T. S. Huang, "Toward multimodal human–computer interface," *Proc. IEEE*, vol. 86, no. 5, pp. 853–869, May 1998.

[5] T. E. Hutchinson, K. P. White, W. N. Martin, K. C. Reichert, and L. A. Frey, "Human–computer interaction using eye-gaze input," *IEEE Trans. Syst. Man Cybern.*, vol. 19, no. 6, pp. 1527–1534, Nov./Dec. 1989.

[6] D. H. Yoo, J. H. Kim, B. R. Lee, and M. J. Chung, "Non-contact eye gaze tracking system by mapping of corneal reflections," in *Proc. 5th IEEE Int. Conf. Autom. Face Gesture Recog.*, 2002, pp. 94–99.

[7] D. H. Yoo and M. J. Chung, "A novel non-intrusive eye gaze estimation using cross-ratio under large head motion," *Comput. Vis. Image Understanding*, vol. 98, no. 1, pp. 25–51, 2005.

[8] F. L. Coutinho and C. H. Morimoto, "Free head motion eye gaze tracking using a single camera and multiple light sources," in *Proc. 19th Brazilian Symp. Comput. Graphics Image Process.*, 2006, pp. 171–178.

[9] J. J. Kang, E. D. Guestrin, W. J. MacLean, and M. Eizenman, "Simplifying the cross-ratios method of point-of-gaze estimation," presented at the *30th Canadian Med. Biol. Eng. Conf. (CMBEC30)*, Toronto, ON, Canada, Jun. 2007 (in *Proc. FICCDAT 2007*, Toronto, ON, Canada, Jun. 2007).

[10] M. C. Corbett, E. S. Rosen, and D. P. S. O'Brart, *Corneal Topography: Principles and Applications*. London, U.K.: BMJ Books, 1999, p. 6.

[11] E. D. Guestrin and M. Eizenman, "General theory of remote gaze estimation using the pupil center and corneal reflections," *IEEE Trans. Biomed. Eng.*, vol. 53, no. 6, pp. 1124–1133, Jun. 2006.

[12] E. Trucco and A. Verri, *Introductory Techniques for 3-D Computer Vision*. Upper Saddle River, NJ: Prentice-Hall, 1998.

[13] A. Villanueva and R. Cabeza, "Models for gaze tracking systems," *EURASIP J. Image Video Process.*, vol. 2007, article ID 23570, 2007, pp. 1–16.

Jeffrey J. Kang was born in Tianjin, China, in 1980. He received the B.Sc. degree in computer engineering from the University of New Brunswick, Fredericton, NB, Canada, in 2003, and the M.A.Sc. degree in electrical and computer engineering in 2006 from Edward S. Rogers Sr. Department of Electrical and Computer Engineering, University of Toronto, Toronto, ON, Canada, where he was engaged in research on point-of-gaze estimation and gaze-controlled reading assistive systems at the Institute of Biomaterials and Biomedical Engineering.

Moshe Eizenman was born in Tel-Aviv, Israel, in 1952. He received the B.A.Sc., M.A.Sc., and Ph.D. degrees in electrical engineering from the University of Toronto, Toronto, ON, Canada, in 1978, 1980, and 1984, respectively.

Since 1984, he has been with the Faculty of the University of Toronto, where he is currently an Associate Professor in the Department of Electrical and Computer Engineering, and the Department of Ophthalmology and Vision Sciences, and also at the Institute of Biomaterials and Biomedical Engineering. He is also a Research Associate at the Eye Research Institute, Canada, and the Hospital for Sick Children, Toronto. His current research interests include detection and estimation of biological phenomena, eye tracking and gaze estimation systems, visual evoked potentials, and the development of vision.

Elias D. Guestrin (S'98–M'07) was born in Paraná, Entre Ríos, Argentina, in 1974. He received the Electronics Engineer degree from the Universidad Tecnológica Nacional, Facultad Regional Paraná, Paraná, Entre Ríos, Argentina, in 2000, and the M.A.Sc. degree in electrical and computer engineering from Edward S. Rogers Sr. Department of Electrical and Computer Engineering, University of Toronto, Toronto, ON, Canada, in 2003. He is currently working toward the Ph.D. degree at the Department of Electrical and Computer Engineering, and the Institute of Biomaterials and Biomedical Engineering, University of Toronto.

His current research interests include gaze estimation, mathematical modeling of optical systems, and signal and image processing.

Erez Eizenman was born in Toronto, ON, Canada, in 1982. He received the Jeffrey Skoll combined B.A.Sc./M.B.A. degree from the University of Toronto, Toronto, ON, Canada, in 2008.