An Automatic Personal Calibration Procedure for Advanced Gaze Estimation Systems

Dmitri Model and Moshe Eizenman

Abstract— **Gaze estimation systems use calibration procedures to estimate subject-specific parameters that are needed for the calculation of the point-of-gaze. In these procedures, subjects are required to fixate on a specific point or points in space at specific time instances. Advanced remote gaze estimation systems can estimate the optical axis of the eye without any personal calibration procedure, but use a single calibration point to estimate the angle between the *optical* axis and the *visual* axis (line-of-gaze). This paper presents a novel calibration procedure that does not require active user participation. To estimate the angles between the optical and visual axes of each eye, this procedure minimizes the distance between the intersections of the visual axes of the left and right eyes with one or more observation surfaces (displays) while subjects look naturally at these displays (e.g., watching a video clip). Theoretical analysis and computer simulations show that the performance of the proposed procedure improves when the range of angles between the visual axes and vectors normal to the observation surfaces increases. Experiments with 4 subjects show that the subject-specific angles between the optical and visual axes can be estimated with a root-mean-square error of 0.5°.**

Index Terms—**Remote gaze estimation, point of gaze, personal calibration, calibration free, optical axis, visual axis.**

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 fovea. Eye gaze tracking (EGT) systems that estimate PoG are used in a large variety of applications such as studies of cognitive processes [1], driver behavior [2], psychology and psychiatry [3, 4], marketing and advertising [5], pilot training [6] and human-computer interfaces [7, 8].

D. Model is a Ph.D. student with the Department of Electrical and Computer Engineering, University of Toronto, Toronto, ON M5S 3G4, Canada (phone: 1-416-978-2255; fax: 1-416-978-4317; e-mail: dmitri.model@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, Toronto, ON M5S 3G9, Canada (e-mail: eizenm@ecf.utoronto.ca).

All EGT systems use personal calibration procedures to estimate subject-specific parameters that are needed for the calculation of the point-of-gaze. During these calibration procedures, subjects are required to fixate on a specific point or points in space at specific time instances. EGT systems that do not use physiological model of the eye, use calibration procedures with multiple calibration points [9] to determine subject-specific mappings between locations of eye features (such as pupil center, corneal reflections, etc.) to gaze locations on the observation surface [10]. Systems that employ model of the eye to estimate PoG [11-13] use calibration procedures to estimate subject-specific eye-model parameters. As was shown in [14, 15], a system that uses a stereo pair of video cameras could estimate the center of curvature of the cornea and the *optical* axis of the eye without any user calibration. However, since human gaze is not directed along the *optical* axis, but rather along the *visual* axis [16], a one-point calibration procedure is still required for the estimation of the angle between the optical and visual axes. During this calibration procedure, the subject is required to look at a calibration point so that its image will fall on the fovea. However, in some applications, such as covert monitoring or studies with young children or mentally challenged subjects, it is impossible to reliably perform even a one-point calibration routine. Therefore, one of the most important goals in the field of gaze tracking technology is to estimate human PoG without calibration procedures that require active user participation.

This paper describes a new methodology to estimate the 3D angle between the optical and visual axes of the eye. This methodology does not require subjects to fixate on any specific calibration point(s). Instead, it relies on the assumption that the visual axes of both eyes intersect on an observation surface and uses this constraint to calculate the subject-specific angles between the optical and visual axes [17, 18]. This assumption holds true for the vast majority of subjects and viewing conditions, but can be violated for certain group of patients (e.g., patients with strabismus).

The paper is organized as follows. The next section describes the algorithm for the estimation of subject-specific angles between the optical and visual axes. The properties of the proposed algorithm are discussed in Section III. Section IV presents both simulation and experimental results. Finally, conclusions are presented in Section V.

d : center of the eyeball
c : center of curvature of the cornea
p : center of the pupil

X_{eye}, Y_{eye}, Z_{eye} – Eye Coordinate System (ECS)
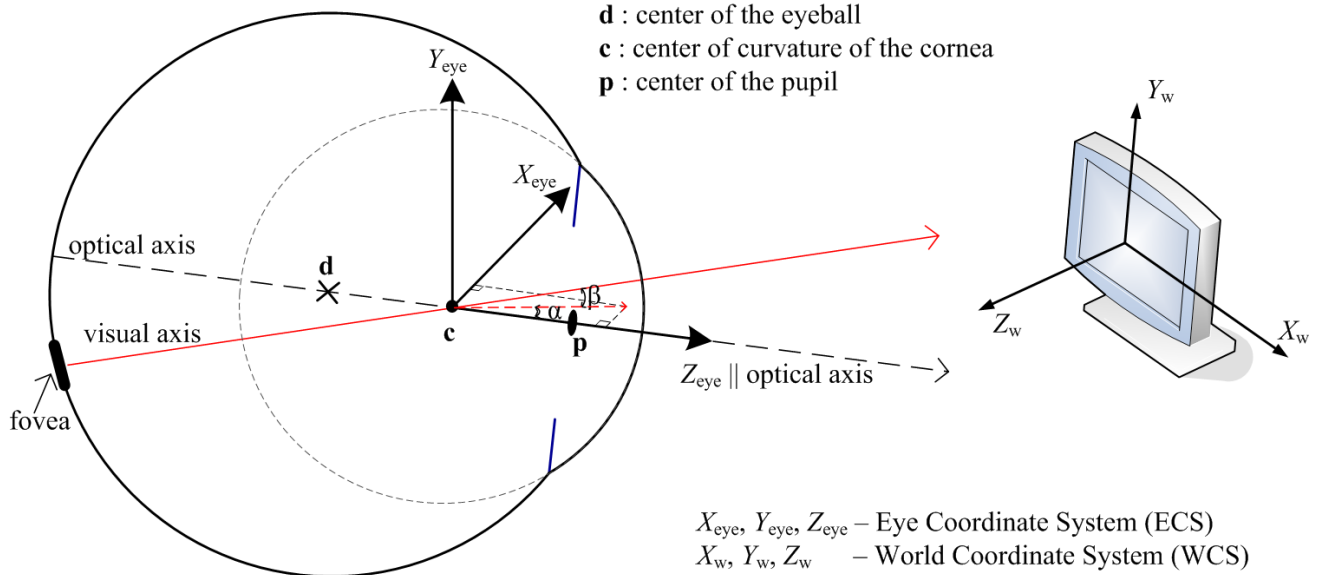X_w, Y_w, Z_w – World Coordinate System (WCS)

Fig. 1. A simplified schematic diagram of the eye (not to scale). 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.

II. AUTOMATIC CALIBRATION METHODOLOGY

A. Model

Fig. 1 presents a simplified schematic diagram of the eye [12]. The line connecting the center of curvature of the cornea with the center of the pupil defines the *optical* axis. The line connecting the center of the fovea with the center of curvature of the cornea defines the *visual* axis or the line-of-gaze. The average magnitude of the angle between the *optical* and *visual* axes is 5°. This angle has both horizontal (nasal) and vertical components, which exhibit considerable inter-personal variation [16].

To develop an algorithm that estimates the angle between the *optical* and *visual* axes, two coordinate systems are defined. The first coordinate system is a stationary right-handed Cartesian World Coordinate System (WCS) with the origin at the center of the main 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. 1). The second coordinate system 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 [19] around the Z_{eye} axis for different gaze directions.

In the ECS, the unknown 3D angle between the *optical* and the *visual* axes of the eye can be expressed by the horizontal[1],

α, and vertical[2], β, components of this angle (see Fig. 1). The unit vector in the direction of the *visual* axis with respect to the ECS, $\mathbf{v}_{ECS}$, is then expressed as:

$$\mathbf{v}_{ECS}(\alpha,\beta) = \begin{bmatrix} -\sin(\alpha)\cos(\beta) \\ \sin(\beta) \\ \cos(\alpha)\cos(\beta) \end{bmatrix} \quad (1)$$

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

$$\mathbf{v}(\alpha,\beta) = \mathbf{R}\,\mathbf{v}_{ECS}(\alpha,\beta) \quad (2)$$

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 [19].

Because the visual axis goes through the center of curvature of the cornea, $\mathbf{c}$, the PoG (ψ) in the WCS is given by:

$$\psi(\alpha,\beta) = \mathbf{c} + k(\alpha,\beta)\mathbf{v}(\alpha,\beta) = \mathbf{c} + k(\alpha,\beta)\mathbf{R}\,\mathbf{v}_{ECS}(\alpha,\beta) \quad (3)$$

where k is a line parameter defined by the intersection of the visual axis with the observation surface. For example, if the observation surface is a plane defined by $\{\mathbf{x}\,|\,\mathbf{n}\bullet\mathbf{x}+h=0\}$, then $k(\alpha,\beta)$ will be given by:

$$k(\alpha,\beta) = -\frac{h+\mathbf{n}\bullet\mathbf{c}}{\mathbf{n}\bullet\mathbf{v}(\alpha,\beta)} \quad (4)$$

where $\mathbf{n}$ is the normal to the display surface.

---

[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.

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

Note that since EGT systems that use a stereo pair of video cameras can estimate the center of curvature of the cornea and the *optical* axis of the eye without any user calibration [12, 15, 20], **c** and **R** are known.

B. Automatic Calibration Algorithm

The proposed automatic calibration algorithm for the estimation of α^L, β^L, α^R and β^R is based on the assumption that at each time instant the visual axes of both eyes intersect on the surface of an observation surface (display). The superscripts "L" and "R" are used to denote parameters of the left and right eyes, respectively. The unknown angles α^L, β^L, α^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

$$F(\alpha^L, \beta^L, \alpha^R, \beta^R) = \sum_i \left\| \boldsymbol{\psi}_i^L(\alpha^L, \beta^L) - \boldsymbol{\psi}_i^R(\alpha^R, \beta^R) \right\|_2^2 \quad (5)$$

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 α^L, β^L, α^R and β^R. However, since the deviations of the unknown angles α^L, β^L, α^R and β^R from the expected 'average' values α_0^L, β_0^L, α_0^R and β_0^R are relatively small, a linear approximation of (3) can be obtained by using the first three terms of its Taylor's series expansion:

$$\boldsymbol{\psi}(\alpha, \beta) \approx \boldsymbol{\psi}(\alpha_0, \beta_0) + \left. \frac{\partial \boldsymbol{\psi}(\alpha, \beta)}{\partial \alpha} \right|_{\alpha_0, \beta_0} (\alpha - \alpha_0) +$$
$$\left. \frac{\partial \boldsymbol{\psi}(\alpha, \beta)}{\partial \beta} \right|_{\alpha_0, \beta_0} (\beta - \beta_0) \quad (6)$$

Let

$$\boldsymbol{\psi}_0 = \boldsymbol{\psi}(\alpha_0, \beta_0) = \mathbf{c} + k_0 \mathbf{v}_0, \quad (7)$$

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

Then, using (3),

$$\mathbf{a} = \left. \frac{\partial \boldsymbol{\psi}(\alpha, \beta)}{\partial \alpha} \right|_{\alpha_0, \beta_0} = k_\alpha \mathbf{v}_0 + k_0 \mathbf{v}_\alpha \quad (8)$$

$$\mathbf{b} = \left. \frac{\partial \boldsymbol{\psi}(\alpha, \beta)}{\partial \beta} \right|_{\alpha_0, \beta_0} = k_\beta \mathbf{v}_0 + k_0 \mathbf{v}_\beta \quad (9)$$

where, from (1)-(2),

$$\mathbf{v}_\alpha = \left. \frac{\partial \mathbf{v}(\alpha, \beta)}{\partial \alpha} \right|_{\alpha_0, \beta_0} = \mathbf{R} \begin{bmatrix} -\cos(\alpha_0)\cos(\beta_0) \\ 0 \\ -\sin(\alpha_0)\cos(\beta_0) \end{bmatrix} \quad (10)$$

$$\mathbf{v}_\beta = \left. \frac{\partial \mathbf{v}(\alpha, \beta)}{\partial \beta} \right|_{\alpha_0, \beta_0} = \mathbf{R} \begin{bmatrix} \sin(\alpha_0)\sin(\beta_0) \\ \cos(\beta_0) \\ -\cos(\alpha_0)\sin(\beta_0) \end{bmatrix} \quad (11)$$

and, from (4),

$$k_\alpha = \left. \frac{\partial k(\alpha, \beta)}{\partial \alpha} \right|_{\alpha_0, \beta_0} = -k_0 \frac{\mathbf{v}_\alpha \bullet \mathbf{n}}{\mathbf{v}_0 \bullet \mathbf{n}} \quad (12)$$

$$k_\beta = \left. \frac{\partial k(\alpha, \beta)}{\partial \beta} \right|_{\alpha_0, \beta_0} = -k_0 \frac{\mathbf{v}_\beta \bullet \mathbf{n}}{\mathbf{v}_0 \bullet \mathbf{n}} \quad (13)$$

Using the above linear approximations, the sum of the squared distances between the left and right PoGs in the objective function (5) can be expressed as

$$F(\alpha^L, \beta^L, \alpha^R, \beta^R) = \sum_i \left\| \mathbf{M}_i \, \mathbf{x} + \mathbf{y}_i \right\|_2^2 \quad (14)$$

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, $\mathbf{y}_i = \boldsymbol{\psi}_{0,i}^L - \boldsymbol{\psi}_{0,i}^R$ is a 3x1 vector and $\mathbf{x} = \begin{bmatrix} (\alpha^L - \alpha_0^L) & (\beta^L - \beta_0^L) & (\alpha^R - \alpha_0^R) & (\beta^R - \beta_0^R) \end{bmatrix}^T$ is a 4x1 vector of unknown angles ("T" denotes transpose). The subscript "i" is used to explicitly indicate the correspondence to the specific time instance "i" or i-th gaze sample.

The solution to (14) can be obtained in a closed form using least squares as

$$\mathbf{x}_{opt} = -(\mathbf{M}^T \mathbf{M})^{-1} \mathbf{M}^T \mathbf{y}, \quad (15)$$

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

$$\mathbf{M} = \begin{bmatrix} \mathbf{M}_1 \\ \mathbf{M}_2 \\ \vdots \\ \mathbf{M}_N \end{bmatrix} \quad ; \quad \mathbf{y} = \begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \\ \vdots \\ \mathbf{y}_N \end{bmatrix} \quad (16)$$

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

$$\begin{bmatrix} \hat{\alpha}^L & \hat{\beta}^L & \hat{\alpha}^R & \hat{\beta}^R \end{bmatrix}^T = \begin{bmatrix} \alpha_0^L & \beta_0^L & \alpha_0^R & \beta_0^R \end{bmatrix}^T + \mathbf{x}_{opt} \quad (17)$$

Since the objective function (14) is a linear approximation of the objective function (5), several iterations of (7)-(17) might be needed to converge to the true minimum of the objective function (5). 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 the optical axes that are provided by the EGT.

III. PROPERTIES OF THE OBJECTIVE FUNCTION

In a noise-free case (i.e., **c** and **R** are noise-free**)**, the minimum of the objective function (5) is achieved when

$$\psi_i^{\mathrm{R}}\left(\alpha^{\mathrm{R}},\beta^{\mathrm{R}}\right)=\psi_i^{\mathrm{L}}\left(\alpha^{\mathrm{L}},\beta^{\mathrm{L}}\right)\quad \forall i \tag{18}$$

In this case, two PoGs on the observation surface are sufficient to estimate the four unknown angles. However, when the algorithm uses estimates of the centers of curvature of the corneas and estimates of the optical axes that include noise, more than two PoGs are required for a stable solution. Moreover, the performance of the algorithm depends on the spatial distribution of these PoGs. To simplify the analysis, eq. (18) can be interpreted as an implicit definition of $\left(\alpha^{\mathrm{R}},\beta^{\mathrm{R}}\right)$ as a function of $\left(\alpha^{\mathrm{L}},\beta^{\mathrm{L}}\right)$:

$$\alpha^{\mathrm{R}}=f_i\left(\alpha^{\mathrm{L}},\beta^{\mathrm{L}}\right) \tag{19}$$

$$\beta^{\mathrm{R}}=g_i\left(\alpha^{\mathrm{L}},\beta^{\mathrm{L}}\right) \tag{20}$$

The solution, $\left(\hat{\alpha}^{\mathrm{R}},\hat{\alpha}^{\mathrm{L}},\hat{\beta}^{\mathrm{R}},\hat{\beta}^{\mathrm{L}}\right)$, is determined by the intersection of 3D surfaces[3], f_i and g_i, for all gaze points ($i=1,\ldots,N$). In general, the range of angles between the surfaces at the point of intersection determines the stability of the solution (i.e., when the range of all the angles between the surfaces is small, all surfaces are approximately parallel to each other and the solution is susceptible to noise). The purpose of the following analysis is to define the parameters that affect the range of angles between the surfaces f_i, $i=1,\ldots,N$ and g_i, $i=1,\ldots,N$, at the solution point.

The angle between two surfaces f_i and f_j at the solution $\left(\hat{\alpha}^{\mathrm{L}},\hat{\beta}^{\mathrm{L}}\right)$ is given by:

$$\cos^{-1}\left(\frac{\nabla f_i \bullet \nabla f_j}{\left\|\nabla f_i\right\|_2\left\|\nabla f_j\right\|_2}\right), \tag{21}$$

where $\nabla f_i=\left(\dfrac{\partial f_i}{\partial \alpha^{\mathrm{L}}},\dfrac{\partial f_i}{\partial \beta^{\mathrm{L}}},-1\right)$. By substituting f with g, one can obtain a similar expression for the angle between surfaces g_i and g_j.

By using eq. (7)-(11) and employing the implicit derivation of (18) with respect to α^{L} and β^{L}, the partial derivatives $\dfrac{\partial \alpha^{\mathrm{R}}}{\partial \alpha^{\mathrm{L}}}=\dfrac{\partial f}{\partial \alpha^{\mathrm{L}}}$, $\dfrac{\partial \alpha^{\mathrm{R}}}{\partial \beta^{\mathrm{L}}}=\dfrac{\partial f}{\partial \beta^{\mathrm{L}}}$, $\dfrac{\partial \beta^{\mathrm{R}}}{\partial \alpha^{\mathrm{L}}}=\dfrac{\partial g}{\partial \alpha^{\mathrm{L}}}$ and $\dfrac{\partial \beta^{\mathrm{R}}}{\partial \beta^{\mathrm{L}}}=\dfrac{\partial g}{\partial \beta^{\mathrm{L}}}$ can be calculated by the following two equations:

$$k_\alpha^{\mathrm{L}}\mathbf{v}^{\mathrm{L}}+k^{\mathrm{L}}\mathbf{v}_\alpha^{\mathrm{L}}=$$

$$k_\alpha^{\mathrm{R}}\mathbf{v}^{\mathrm{R}}\frac{\partial\alpha^{\mathrm{R}}}{\partial\alpha^{\mathrm{L}}}+k^{\mathrm{R}}\mathbf{v}_\alpha^{\mathrm{R}}\frac{\partial\alpha^{\mathrm{R}}}{\partial\alpha^{\mathrm{L}}}+k_\beta^{\mathrm{R}}\mathbf{v}^{\mathrm{R}}\frac{\partial\beta^{\mathrm{R}}}{\partial\alpha^{\mathrm{L}}}+k^{\mathrm{R}}\mathbf{v}_\beta^{\mathrm{R}}\frac{\partial\beta^{\mathrm{R}}}{\partial\alpha^{\mathrm{L}}} \tag{22}$$

$$k_\beta^{\mathrm{L}}\mathbf{v}^{\mathrm{L}}+k^{\mathrm{L}}\mathbf{v}_\beta^{\mathrm{L}}=$$

$$k_\alpha^{\mathrm{R}}\mathbf{v}^{\mathrm{R}}\frac{\partial\alpha^{\mathrm{R}}}{\partial\beta^{\mathrm{L}}}+k^{\mathrm{R}}\mathbf{v}_\alpha^{\mathrm{R}}\frac{\partial\alpha^{\mathrm{R}}}{\partial\beta^{\mathrm{L}}}+k_\beta^{\mathrm{R}}\mathbf{v}^{\mathrm{R}}\frac{\partial\beta^{\mathrm{R}}}{\partial\beta^{\mathrm{L}}}+k^{\mathrm{R}}\mathbf{v}_\beta^{\mathrm{R}}\frac{\partial\beta^{\mathrm{R}}}{\partial\beta^{\mathrm{L}}} \tag{23}$$

The angles between two surfaces can be calculated by substituting the estimated partial derivatives into (21).

To study parameters that affect the angles between surfaces, f_i and f_j, we will make use of several examples. In the first example, a subject is sitting 75 cm in front of an observation surface (a plane at Z=0, Fig. 2(a)), while looking at two arbitrary points, say, $\mathbf{p}_1=[-10,0,0]^T$ and $\mathbf{p}_2=[10,0,0]^T$ [4]. Fig. 2(b) shows the two surfaces, f_1 and f_2, that correspond to these point and Fig. 2(c) shows a cross-section of the functions f_1 and f_2 at $\beta^{\mathrm{L}}=\hat{\beta}^{\mathrm{L}}$. Fig. 2(c) demonstrates clearly, that even though the two curves intersect at $\alpha^{\mathrm{L}}=\hat{\alpha}^{\mathrm{L}}$, the angle between the two surfaces at the solution is relatively small (1.2°). The analysis of this example in the Appendix provides a close form expressions for the partial derivatives of each surface at the solution point. The partial derivatives of f, are given by:

$$\frac{\partial\alpha^{\mathrm{R}}}{\partial\alpha^{\mathrm{L}}}=\frac{\partial f}{\partial\alpha^{\mathrm{L}}}=1-\frac{\Delta x}{k}\frac{\mathbf{n}\bullet\mathbf{v}_\alpha}{\mathbf{n}\bullet\mathbf{v}} \tag{24}$$

$$\frac{\partial\alpha^{\mathrm{R}}}{\partial\beta^{\mathrm{L}}}=\frac{\partial f}{\partial\beta^{\mathrm{L}}}=-\frac{\Delta x}{k}\frac{\mathbf{n}\bullet\mathbf{v}_\beta}{\mathbf{n}\bullet\mathbf{v}} \tag{25}$$

Equations (24) and (25) show that the partial derivatives, $\dfrac{\partial\alpha^{\mathrm{R}}}{\partial\alpha^{\mathrm{L}}}$ and $\dfrac{\partial\alpha^{\mathrm{R}}}{\partial\beta^{\mathrm{L}}}$, depend on the following parameters: 1) the distance between two eyes, $\Delta x=\left\|\mathbf{c}^{\mathrm{R}}-\mathbf{c}^{\mathrm{L}}\right\|_2$; 2) the line parameter, k (defined in (4)); 3) the angle between the normal to the observation surface at the point-of-gaze, $\mathbf{n}$, and the gaze direction, $\mathbf{v}$; and 4) the angle between the normal to the observation surface, $\mathbf{n}$, and the derivative of the gaze direction: $\mathbf{v}_\alpha$, in (24) or, $\mathbf{v}_\beta$, in (25). For the above example (configuration described in Fig. 2(a)): $\mathbf{n}=\begin{bmatrix}0&0&1\end{bmatrix}^T$, $\mathbf{v}\approx\begin{bmatrix}0&0&-1\end{bmatrix}^T$, $\mathbf{v}_\alpha\approx\begin{bmatrix}1&0&0\end{bmatrix}^T$ and $\mathbf{v}_\beta\approx\begin{bmatrix}0&1&0\end{bmatrix}^T$, so $\dfrac{\partial\alpha^{\mathrm{R}}}{\partial\alpha^{\mathrm{L}}}\approx1$ and $\dfrac{\partial\alpha^{\mathrm{R}}}{\partial\beta^{\mathrm{L}}}\approx0$ for both $\mathbf{p}_1$ and $\mathbf{p}_2$. Note that since the partial derivatives in Equations (24) and (25) are similar for the two points, the angle between the two surfaces will be relatively small.

---

[3] Equations (19) and (20) give rise to parametric surfaces $\left(\alpha^{\mathrm{L}},\beta^{\mathrm{L}},f_i(\alpha^{\mathrm{L}},\beta^{\mathrm{L}})\right)$ and $\left(\alpha^{\mathrm{L}},\beta^{\mathrm{L}},g_i(\alpha^{\mathrm{L}},\beta^{\mathrm{L}})\right)$, respectively. These surfaces are referred to as f_i and g_i.

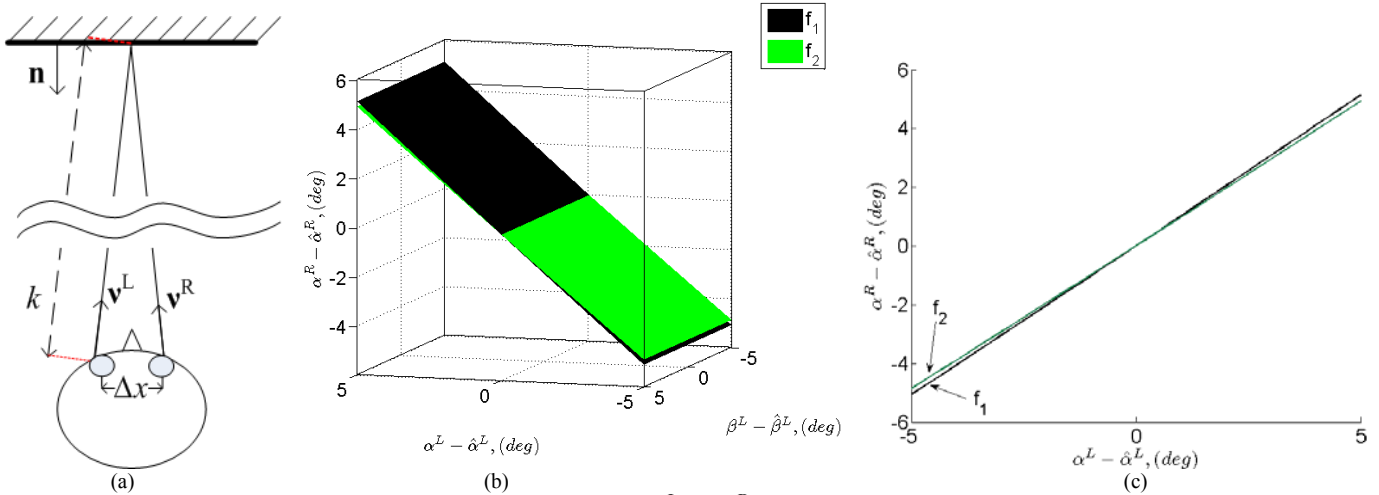[4] All coordinates in this paper are specified in centimetres.

Fig. 2. (a) Top view. Subject looks at a observation surface (a plane Z=0). $\mathbf{v}^L$ and $\mathbf{v}^R$ are vectors in the direction of the line-of-gaze of the left and right eyes respectively. $\mathbf{n}$ is the normal to the plane. Δx is the horizontal separation between to eyes; (b) Plot of $f_i\left(\alpha^L,\beta^L\right)$ for two PoGs: $\mathbf{p}_1$=[-10,0,0]^T and $\mathbf{p}_2$=[10,0,0], which lie on a plane Z=0 in the WCS. (c) Plot of two cross-sections: $f_i\left(\alpha^L,\beta^L=\hat{\beta}^L\right)$, i=1,2.

Based on Eq. (24), larger differences between $\partial f_1/\partial\alpha^L$ and $\partial f_2/\partial\alpha^L$ can be achieved if the normal to the observation surfaces, $\mathbf{n}_i$, and the derivative of the gaze direction, $\mathbf{v}_{\alpha,i}$, are at larger angles to each other and if the inner products, $\mathbf{n}_1\bullet\mathbf{v}_{\alpha,1}$ and $\mathbf{n}_2\bullet\mathbf{v}_{\alpha,2}$, have different signs. This situation can be realized if the same two points of gaze, $\mathbf{p}_1$ and $\mathbf{p}_2$, are on two different observation planes with different orientations, say, Z-X-10 = 0 and Z+X-10 = 0, respectively, (Fig. 3(a)). In this case, the normal to the observation surface, $\mathbf{n}_i$, and the derivative of the gaze direction, $\mathbf{v}_{\alpha,i}$, are at approximately 45° to each other and the inner products, $\mathbf{n}_1\bullet\mathbf{v}_{\alpha,1}$ and $\mathbf{n}_2\bullet\mathbf{v}_{\alpha,2}$, have different signs, and the angle between the surfaces f_1 and f_2 becomes 6.5° (see Fig. 3(b) and compare it to Fig.
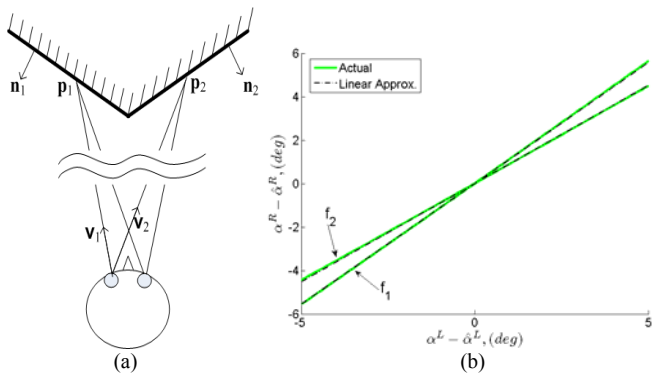

Fig. 3. (a) Top View. Subject look at a surface consisting of two planes with normals $\mathbf{n}_1$ and $\mathbf{n}_2$. $\mathbf{v}_1$ is the direction of gaze towards $\mathbf{p}_1$, while $\mathbf{v}_2$ is the direction of gaze towards $\mathbf{p}_2$. (b) Solid lines: plot of $f_i\left(\alpha^L,\beta^L=\hat{\beta}^L\right)$ from eq. (19) for two PoGs: $\mathbf{p}_1$=[-10,0,0]^T which lies on a plane Z-X-10=0 and $\mathbf{p}_2$=[10,0,0]^T which lies on a plane Z+X-10=0. Dashed lines: linear approximation based on the slope in (24).

2(c)). Fig. 3(b) shows the actual cross-section of $f_i\left(\alpha^L,\beta^L=\hat{\beta}^L\right)$ along with the linear approximation based on Eq. (24). Note the excellent agreement between the two curves.

Similarly, based on Eq. (25), the difference between the slopes of the surfaces f_i in the direction of β, $\dfrac{\partial f_i}{\partial\beta^L}$, can be increased by selecting points for which $\mathbf{n}_i$ and $\mathbf{v}_{\beta,i}$ are at relatively large angles to each other and the products $\mathbf{n}_1\bullet\mathbf{v}_{\beta,1}$ and $\mathbf{n}_2\bullet\mathbf{v}_{\beta,2}$ have different signs. For example, points $\mathbf{p}_1$=[0,-10,0]^T and $\mathbf{p}_2$=[0,10,0]^T on the planes Z-Y-10 = 0 and Z+Y-10 = 0, respectively, satisfy the above requirements.

Finally, to maximize the differences in the angles between the surfaces f_i i=1,…,N, in the directions of both α and β, one can combine the above examples and use a pyramid with a pinnacle facing the subject as an observation surface. Under the assumption that the subject will look at all four facets of the pyramid, the conditions similar to those discussed above will be achieved. Fig. 4 shows plots of $f_i\left(\alpha^L,\beta^L\right)$ and $g_i\left(\alpha^L,\beta^L\right)$ for four PoGs that lie on the four different facets of the pyramid. The intersection of $f_i\left(\alpha^L,\beta^L\right)$, i=1,…,4, provides an estimate of $\hat{\alpha}^R$ and $\left(\hat{\alpha}^L,\hat{\beta}^L\right)$, which is sufficient to determine $\hat{\beta}^R$ according to (20). As expected from the analysis in the Appendix, the surfaces $g_i\left(\alpha^L,\beta^L\right)$ i=1,…,4, are almost overlapping, as both $\dfrac{\partial g_i}{\partial\alpha^L}$ and $\dfrac{\partial g_i}{\partial\beta^L}$ are approximately constant for central points of gaze (i.e., PoGs

for which the direction of gaze is approximately along the Z_{wcs}-axis) regardless of the orientation of the observation plane.
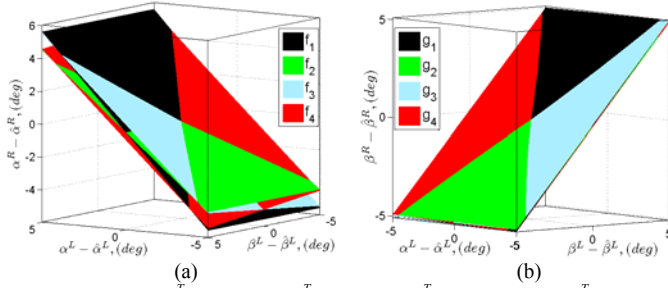

Fig. 4. $\mathbf{p}_1 = [-10,0,0]^T$, $\mathbf{p}_2 = [-10,0,0]^T$, $\mathbf{p}_3 = [0,10,0]^T$ and $\mathbf{p}_4 = [0,-10,0]^T$ that lie on the corresponding facets of the pyramid. (a) $f_i(\alpha^L, \beta^L)$, $i=1,..,4$; (b) $g_i(\alpha^L, \beta^L)$, $i=1,..,4$.

The above analysis, which was developed for the central gaze, demonstrates that the stability of the solution improves as the range of angles between $\mathbf{n}_i$ and $\mathbf{v}_i$ (as well as $\mathbf{v}_{\alpha,i}$ and $\mathbf{v}_{\beta,i}$) for the various points of gaze increases. In many practical applications, the observation surface consists of a single plane ($\mathbf{n}_i = \mathbf{n}, \forall i$). In this case, a stable solution requires that the points of gaze will span a wide range of gaze directions (e.g., by looking at larger areas of the observation plane). The performance of the algorithm with complex observation surfaces as well as with a single observation plane will be discussed in the next section.

IV. NUMERICAL SIMULATIONS AND EXPERIMENTS

A. Numerical Simulations

In the first set of numerical simulations, the performance of the algorithm was evaluated as a function of the noise in the estimation of the optical axis and the center of curvature of the cornea of each eye and as a function of the size and shape of the observation surface. In the first simulation, the head was fixed (i.e., no head movements) and the position of the center of curvature of the cornea of the subject's right eye, $\mathbf{c}^R$, was set to $[3\ 0\ 75]^T$ (i.e., approximately 75 cm from the display's surface) while the center of curvature of the cornea of the subject's left eye, $\mathbf{c}^L$, was set to $[-3\ 0\ 75]^T$ (i.e., inter-pupillary distance = 6 cm). The angles between the optical and visual axes were randomly drawn for each simulation from a uniform distribution with a 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 the observation surfaces and the optical axes of the two eyes were calculated. Noise in the estimates of the center of curvature of the cornea and the optical axis was simulated by adding independent zero-mean white Gaussian processes to the coordinates of the centers of curvature of the cornea (X, Y and Z) and to the horizontal and vertical components of the direction of the optical axis. The automatic calibration algorithm was implemented in

MATLAB[®]. One thousand PoGs were used to estimate the angle between the optical and visual axes for each set of eye parameters ($\alpha^R, \alpha^L, \beta^R$ and β^L). The initial value for the angle between the optical and visual axes was set to zero (the initial value did not affect the results as long as it was within ±20° of the actual value). The first update of the algorithm was done after 100 PoGs to prevent large fluctuations during start-up.

The simulations were performed with four different observation surfaces: a) a pyramid (30 cm x 30 cm base, 15 cm height) that was used in the theoretical analysis and in the experiments that are described in the next section, b) a plane that provides viewing angles in the range of ±14.9º horizontally and ±11.3º vertically (40 cm x 30 cm observation surface that is similar in size to a 20" monitor), c) a plane that provides viewing angles in the range of ±28º horizontally and ±21.8º vertically (80 cm x 60 cm observation surface that is similar in size to a 40" monitor), and d) a plane that provide viewing angles in the range of ±46.8º horizontally and ±38.7º (160 cm x 120 cm). For each observation surface and each level of noise in the estimation of the optical axis or in the estimation of the center of curvature of the cornea the simulations were repeated 100 times.
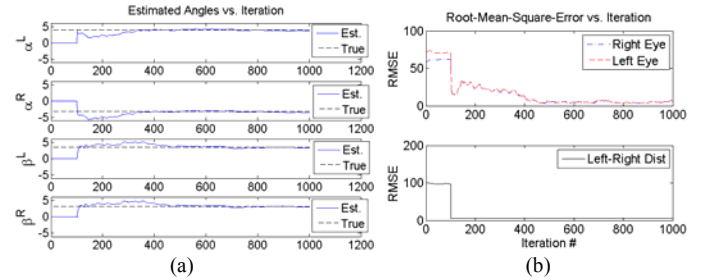

Fig. 5. Estimation results with a 40 cm x 30 cm observation surface (e.g., 20" monitor). STD of noise in the angles of the optical axis was set to 0.1° and in the coordinates of the center of curvature of the cornea to 0.5 mm. (a) Estimated subject-specific angles; (b) Root-Mean-Square Error (RMSE) of PoG estimation (top) and the distance between Left PoG - Right PoG (bottom).

Fig. 5 shows an example of one simulation with a plane that provides viewing angles in the range of ±14.9º horizontally and ±11.3º vertically (40cm x 30cm observation surface that is similar in size to a 20" monitor). The standard deviation (STD) of the noise in the angles of the optical axis was set to 0.1° and the STD of the noise in the coordinates of the center of curvature of the cornea was set to 0.5 mm. As can be seen from Fig. 5(a), after approximately 500 PoGs the solution converges to within ±0.5° of the true values of $\alpha^R, \alpha^L, \beta^R$ and β^L. Fig. 5(b) shows the root-mean-square (RMS) error in the PoG estimation (top) and the value of the objective function (bottom). The objective function was effectively minimized after the first update (i.e., with 100 PoGs) while the RMS error continues to decline until iteration 500. This is due to the fact that even though after 100 iterations the distance between the points-of-gaze of the left and right eyes was

minimized, PoGs of the both eyes had similar biases relative to the actual PoGs. From iteration 100 to 500 the PoGs of the two eyes drifted simultaneously towards their true locations.

Fig. 6 and Fig. 7 show the RMS error in the estimation of α and β as a function of noise in the optical axis and in the center of curvature of the cornea, respectively. The estimation error had no bias (i.e., zero-mean) in all the above simulations. Fig. 6 and Fig. 7 show that: a) the pyramidal observation surface provides the best performance (i.e., smaller RMS errors as compared to the planes) and b) the RMS error decreases as the range of viewing angle or equivalently the size of the observation surface increases. This finding is consistent with the theoretical analysis in the previous section that showed that by increasing the range of angles between the gaze vectors and the normal to the observation surface the solution becomes less sensitive to noise. The RMS errors for the largest observation plane (160 cm x 120 cm) and the pyramid are similar as they provide approximately the same range of angles between the normals to the observation surfaces and the gaze vectors. Fig. 6 and Fig. 7 show that for the system that was described in [14] (STD of the noise in the estimation of the optical axis is 0.4° and STD of the noise in the estimation of the center of curvature of the cornea is 1.0 mm) noise in the estimation of the optical axis contributes more significantly to the RMS error in the estimation of the angle between the optical and visual axes than noise in the estimation of the center of curvature of the cornea.
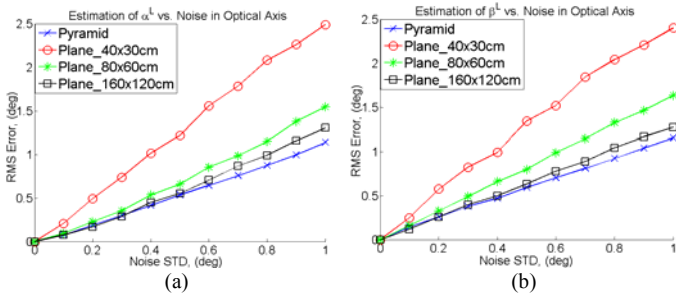

Fig. 6. RMS errors for the angle between the optical and visual axes, for four different observation surfaces, as a function of noise in the estimation of the components of the direction of the optical axis. (a) α^L ; (b) β^L .
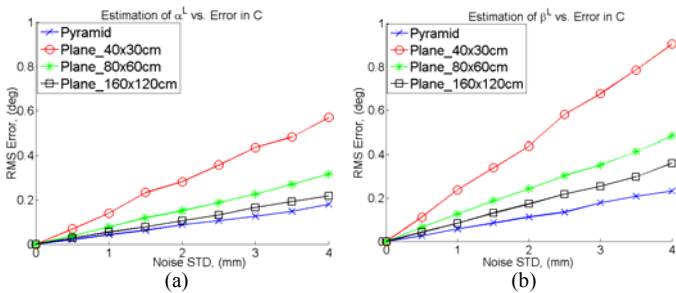

Fig. 7. RMS errors for the angle between the optical and visual axes, for four different surfaces, as a function of the noise in the estimation of the center of the curvature of the cornea. (a) α^L ; (b) β^L .

Additional simulations were performed to study the sensitivity of the algorithm to head position. For these simulations, the coordinates of the center of curvature of the cornea of the right eye, $\mathbf{c}^R$, were set at 27 different positions (all the combinations of X= -7 cm, 3 cm, 13 cm; Y= -10 cm, 0 cm, 10 cm; Z= 85 cm, 75 cm, 65 cm, e.g., $\mathbf{c}^R$ = [-7 0 75]^T, [3 0 75]^T, [13 0 75]^T, etc). For each eye position and noise level in the optical axis (the noise level in the coordinates of the center of the curvature of the cornea was set to 1 mm) the simulations were repeated 100 times. Fig. 8 shows the aggregate (all head positions) RMS error in the estimations of α and β as a function of the noise level in the optical axis. When the data in Fig. 8 are compared to the data in Fig. 6 it is clear that changes in head position introduced biases that increased the RMS errors in the estimation of the angle between the optical and visual axes. To understand these biases, it is important to note that the distance between the left and right PoGs on the observation surface is proportional to the noise level in the directions of the optical axes of the left and right eyes and to the distance between the eyes and the PoG. This implies that for a given level of the noise in the direction of the optical axes, the algorithm will attempt to minimize the distance between the location of the PoGs on the observation surface and the subject's eyes. For example, when the head is moving up relative to the location of the PoG on the observation surface, there will be a positive bias in the estimation of β^R and β^L . Therefore, in the presence of head movements (or non-symmetric distribution of gaze points relative to the head position), the minimum of the objective function represents a balance between an attempt to decrease the error between left and right PoGs due to a shift away from the actual values of α and β and an attempt to minimize the aggregate distance between the PoGs on the observation surface and the eyes.
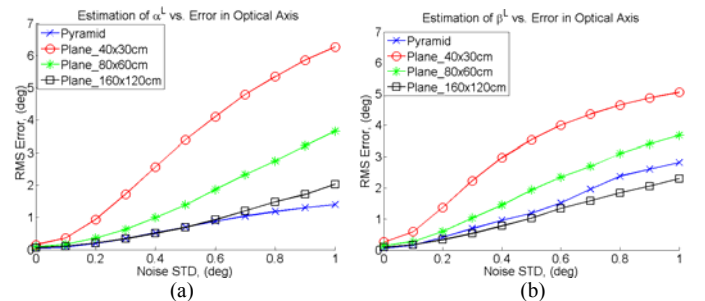
Fig. 8. RMS errors for the angle between the optical and visual axes, for four different surfaces, as a function of the noise in the estimation of the components of the direction of the optical axis. 27 head positions (all the combinations of X_{head} = -10 cm, 0 cm, 10 cm; Y_{head} = -10 cm, 0 cm, 10 cm; Z_{head} = -10 cm, 0 cm, 10 cm). (a) α^L ; (b) β^L .

Fig. 8 shows that the performance of the algorithm improves when the range of angles between the subject's gaze vectors, as he/she looks at the displays, and vectors normal to the observation surfaces increases. For a plane that provides viewing angles in the range of ±14.9° horizontally and ±11.3° vertically (40 cm x 30 cm observation surface), a noise level of

0.4º in the components of the direction of the optical axis results in a RMS error of 2.6º in the estimation of α^{L}. For the same noise level in the optical axis, the RMS error in the estimation of α^{L} is reduced to only 0.5º when a plane that provides viewing angles in the range of ±46.8º horizontally and ±38.7º vertically (160 cm x 120 cm) is used. This is similar to the performance of the algorithm with the pyramidal observation surface.

B. Experiments with real subjects and Gaze Estimation System

A two-camera Remote EGT (REGT) system [14] was used for experiments with four subjects. The STD of the noise in the estimation of the components of the direction of the optical axis in this system is 0.4° and the STD of the noise in the estimation of the coordinates of the center of curvature of the cornea is 1.0 mm. Subjects were seated at a distance of approximately 75cm from a computer monitor (plane Z=0) and head movements were not restrained. Given the limited tracking range of the system (approximately ±15º horizontally and vertically), only a relatively small observation planes (<20" monitor) or more complicated pyramid observation surfaces could be used for the experiments. Because the expected RMS errors with a 20" observation surface ($\alpha >$ 2.5º and β > 3º, see Fig. 8) are larger than the expected errors when $\alpha^{\mathrm{R}}, \alpha^{\mathrm{L}}, \beta^{\mathrm{R}}$ and β^{L} are set a-priori to -2.5º, 2.5º, 0, 0, respectively, there was no point in evaluating the performance of the algorithm with such an observation surface. Therefore, the experiments were performed with a pyramid observation surface (30 cm x 30 cm base, 15 cm height, and pinnacle towards the viewer). The base of the pyramid was mounted on the plane Z=9 (cm) and the pinnacle was located at [0,-1.5,24]^T.

During the experiment, each subject was asked to look at the four facets of the pyramid (in any order) and thousand estimates of the center of curvature of the cornea and the direction of the optical axis were obtained for each eye. The subject-specific angles between the optical and visual axes were estimated, off-line, using the automatic calibration algorithm. Next, the pyramid was removed and the subject was asked to complete a standard one-point calibration procedure [14]. The values of the estimated subject-specific angles between the optical and visual axes with the automated calibration procedure and a one-point calibration procedure are shown in Table I.

Since the estimates of the subject-specific angles between the optical and visual axes that are obtained by the one-point calibration procedure minimize the RMS error between the estimated PoG and the actual PoG, the values obtained by the one-point calibration procedure will serve as a reference ("gold standard") for the calculations of the errors of the automatic calibration procedure. The estimation errors of the automated calibration methodology are -0.08±0.59° in α (left and right combined) and -0.19±0.43° in β (left and right combined). Given the noise characteristics of the gaze estimation system, these estimation errors are in the range predicted by the numerical simulations (see Fig. 6 and Fig. 8).

TABLE I
HORIZONTAL (α) AMD VERTICAL (β) COMPONENTS OF THE ANGLE BETWEEN THE OPTICAL AND VISUAL AXES

| Subjects: | Subject-specific angles, (°) | | | | | | | |
| | AutoCalib | | | | 1-point calib | | | |
	α^{L}	α^{R}	β^{L}	β^{R}	α^{L}	α^{R}	β^{L}	β^{R}
1	1.8	-2.6	-0.1	-0.7	0.9	-2.9	-0.2	-1.0
2	1.1	-1.0	-2.8	-0.7	0.9	-1.3	-3.1	-1.5
3	0.3	-0.7	2.7	0.3	-0.1	-0.3	2.1	0.4
4	-0.2	-3.1	0.6	0.1	0.1	-2.3	0.5	0.7

'AutoCalib' – α and β estimated with the automatic calibration algorithm using a pyramidal surface. '1-point calib' – α and β estimated with a one-point calibration procedure.

Following the estimation of the angles between the optical and visual axes, the subjects looked at a grid of nine points (3x3, 8.5° apart) displayed on a computer monitor. Fifty PoGs were collected at each point. The RMS errors in the estimation of the PoG for all four subjects are summarized in Table II.

For the automatic calibration procedure the average RMS error in PoG estimation for all four subjects is 1.3°. For the one-point calibration procedure the average RMS error is 0.8°.

TABLE II
PoG ESTIMATION ERROR

| Subjects: | RMS error, (°) | | | |
| | AutoCalib | | 1-point calib | |
	Left Eye	Right Eye	Left Eye	Right Eye
1	1.2	0.8	0.6	0.5
2	1.3	1.4	1.0	0.7
3	1.4	1.0	0.8	0.8
4	1.6	1.7	0.9	1.1

'AutoCalib' – α and β estimated with the automatic calibration algorithm using a pyramidal surface. '1-point calib' – α and β estimated with a one-point calibration procedure.

V. DISCUSSION AND CONCLUSIONS

A novel calibration procedure to estimate the angles between the optical and visual axes that does not require accurate fixation on known points at specific time intervals was presented. The procedure is based on the assumption that at each time instant both eyes look at the same point in space. Using the novel procedure with a REGT system [14] that estimates the components of the direction of the optical axis with an accuracy of 0.4° show that the angles between the optical and visual axes can be estimated with an RMS error of 0.5°. These results, which were achieved with an observation surface that included four planes, are consistent with the numerical analysis in Section IV.A. Based on this numerical analysis (see Fig. 8), an eye-tracking system that estimates the components of the direction of the optical axis with accuracy that is four times better than the accuracy of the REGT system

in [14] (i.e., 0.1°) can use the novel calibration procedure with a relatively small single observation plane (30 cm x 40 cm, e.g., 20" computer monitor) to achieve a similar RMS error. By increasing the size of the observation plane to 60 cm x 80 cm, a similar RMS error can be obtained with a REGT system that can estimate the components of the direction of the optical axis with accuracy that is only two times better than the accuracy of the system described in [14] (i.e, 0.2°). The use of a single observation plane will improve considerably the utility of the novel automatic personal calibration procedure.

User-calibration-free gaze estimation systems that estimate the PoG from the intersection of the *optical* axis of one of the eyes with the display (e.g., [21]) or from the midpoint of the intersections of the *optical* axes of both eyes with the display (e.g., [22]) can exhibit large, subject dependent, gaze estimation errors. For the experiments described in Section IV.B (see Table II), calculations of the PoG by the intersection of the *optical* axis of one eye with the computer monitor results in RMS errors ranging from 1° to 4.5° (average 2.5°). By using the midpoint, the RMS errors for the four subjects are reduced to a range between 1.5° to 3.2° (average 2.0°). The automatic personal calibration procedure that was presented in this paper further reduced the RMS errors for the four subjects to a range between 0.8° to 1.7° (average 1.3°).

The experiments with a two-camera gaze estimation system [14] show that with the automatic user calibration procedure the RMS error of the PoG estimation is 1.3° (see Table II). When the same two-camera system is used with a single point calibration procedure, the RMS error is reduced to 0.8°. Similar RMS error (0.8°) is obtained when a state-of-the-art one-camera eye-tracking system [12] and multiple (>4) point calibration procedure are used. By increasing the complexity of the eye-tracking system from one camera to two cameras, the number of calibration points can be reduced (from multiple points to one point) without sacrificing accuracy. The algorithm developed in this paper eliminates the need for any calibration points, but further increases system complexity (the need to track both eyes) and reduces the accuracy of the point-of-gaze estimation. For specific applications, one should consider the trade-offs between system complexity, system accuracy and the number of calibration points.

Appendix

Approximate expressions for $\dfrac{\partial \alpha^{\mathrm{R}}}{\partial \alpha^{\mathrm{L}}}$, $\dfrac{\partial \alpha^{\mathrm{R}}}{\partial \beta^{\mathrm{L}}}$, $\dfrac{\partial \beta^{\mathrm{R}}}{\partial \alpha^{\mathrm{L}}}$ and $\dfrac{\partial \beta^{\mathrm{R}}}{\partial \beta^{\mathrm{L}}}$ (implicitly defined in equations (22) and (23)) when subjects look at centrally located points on a plane in front of them (as in Fig. 2) are derived. In this case:

$$k^{\mathrm{L}} \cong k^{\mathrm{R}} \cong k$$
$$k_\alpha^{\mathrm{L}} \cong k_\alpha^{\mathrm{R}} \cong k_\alpha \tag{A.26}$$
$$k_\beta^{\mathrm{L}} \cong k_\beta^{\mathrm{R}} \cong k_\beta$$

$$\mathbf{v}_\alpha^{\mathrm{L}} \cong \mathbf{v}_\alpha^{\mathrm{R}} \cong \begin{bmatrix} 1 & 0 & 0 \end{bmatrix}^T = \hat{\mathbf{x}}$$
$$\mathbf{v}_\beta^{\mathrm{L}} \cong \mathbf{v}_\beta^{\mathrm{R}} \cong \begin{bmatrix} 0 & 1 & 0 \end{bmatrix}^T = \hat{\mathbf{y}} \tag{A.27}$$

and

$$k_\alpha \ll k$$
$$k_\beta \ll k \tag{A.28}$$

To calculate $\dfrac{\partial \beta^{\mathrm{R}}}{\partial \alpha^{\mathrm{L}}}$, multiply both sides of eq. (22) by $\hat{\mathbf{y}}^T$ from the left. Using approximations (A.26)-(A.28), all terms except for the last one will vanish, yielding:

$$0 \cong \hat{\mathbf{y}}^T k^{\mathrm{R}} \mathbf{v}_\beta^{\mathrm{R}} \frac{\partial \beta^{\mathrm{R}}}{\partial \alpha^{\mathrm{L}}} \;\Rightarrow\; 0 \cong k^{\mathrm{R}} \frac{\partial \beta^{\mathrm{R}}}{\partial \alpha^{\mathrm{L}}} \;\Rightarrow\;$$

$$\frac{\partial \beta^{\mathrm{R}}}{\partial \alpha^{\mathrm{L}}} \cong 0 \tag{A.29}$$

To calculate $\dfrac{\partial \alpha^{\mathrm{R}}}{\partial \alpha^{\mathrm{L}}}$, multiply both sides of eq. (22) by $\hat{\mathbf{x}}^T$ from the left and use (A.29):

$$\hat{\mathbf{x}}^T k_\alpha^{\mathrm{L}} \mathbf{v}^{\mathrm{L}} + \hat{\mathbf{x}}^T k^{\mathrm{L}} \mathbf{v}_\alpha^{\mathrm{L}} = \hat{\mathbf{x}}^T k_\alpha^{\mathrm{R}} \mathbf{v}^{\mathrm{R}} \frac{\partial \alpha^{\mathrm{R}}}{\partial \alpha^{\mathrm{L}}} + \hat{\mathbf{x}}^T k^R \mathbf{v}_\alpha^{\mathrm{R}} \frac{\partial \alpha^{\mathrm{R}}}{\partial \alpha^{\mathrm{L}}} .$$

Using (A.26) :

$$\hat{\mathbf{x}}^T (k_\alpha \mathbf{v}^{\mathrm{R}} + k\mathbf{v}_\alpha) + \hat{\mathbf{x}}^T (k_\alpha \mathbf{v}^{\mathrm{L}} - k_\alpha \mathbf{v}^{\mathrm{R}}) = \hat{\mathbf{x}}^T (k_\alpha \mathbf{v}^{\mathrm{R}} + k\mathbf{v}_\alpha) \frac{\partial \alpha^{\mathrm{R}}}{\partial \alpha^{\mathrm{L}}}$$

$$\frac{\partial \alpha^{\mathrm{R}}}{\partial \alpha^{\mathrm{L}}} = 1 + \frac{\hat{\mathbf{x}}^T k_\alpha (\mathbf{v}^{\mathrm{L}} - \mathbf{v}^{\mathrm{R}})}{\hat{\mathbf{x}}^T (k_\alpha \mathbf{v}^{\mathrm{R}} + k\mathbf{v}_\alpha)} \tag{A.30}$$

Since $\quad 0 \cong \hat{\mathbf{x}}^T k_\alpha \mathbf{v}^{\mathrm{R}} \ll \hat{\mathbf{x}}^T k\mathbf{v}_\alpha \cong k \quad$ and $\mathbf{c}^{\mathrm{R}} + k^{\mathrm{R}} \mathbf{v}^{\mathrm{R}} = \boldsymbol{\psi}^{\mathrm{R}} = \boldsymbol{\psi}^{\mathrm{L}} = \mathbf{c}^{\mathrm{L}} + k^{\mathrm{L}} \mathbf{v}^{\mathrm{L}}$, using (A.26) and (A.27) : $k(\mathbf{v}^{\mathrm{R}} - \mathbf{v}^{\mathrm{L}}) = \mathbf{c}^{\mathrm{L}} - \mathbf{c}^{\mathrm{R}}$. Finally,

$$(\mathbf{v}^{\mathrm{L}} - \mathbf{v}^{\mathrm{R}}) = \frac{\mathbf{c}^{\mathrm{R}} - \mathbf{c}^{\mathrm{L}}}{k} = \frac{\Delta x}{k} \hat{\mathbf{x}} \tag{A.31}$$

(where $\Delta x = \left\| \mathbf{c}^{\mathrm{R}} - \mathbf{c}^{\mathrm{L}} \right\|_2$ is the horizontal separation between the two eyes).

Now we can further simplify (A.30):

$$\frac{\partial \alpha^{\mathrm{R}}}{\partial \alpha^{\mathrm{L}}} = 1 + \frac{\Delta x}{k^2} k_\alpha \tag{A.32}$$

Then, using eq. (12) :

$$\frac{\partial \alpha^{\mathrm{R}}}{\partial \alpha^{\mathrm{L}}} = 1 - \frac{\Delta x}{k} \frac{\mathbf{n} \bullet \mathbf{v}_\alpha}{\mathbf{n} \bullet \mathbf{v}} \tag{A.33}$$

To calculate $\dfrac{\partial \beta^{\mathrm{R}}}{\partial \beta^{\mathrm{L}}}$, multiply both sides of eq. (23) by $\hat{\mathbf{y}}^T$ from the left. Using the approximations (A.26)-(A.28), all terms except for the second one on the left-hand side and the last one on the right-hand side of (23) can be neglected, yielding:

$$\hat{\mathbf{y}}^T k^{\mathrm{L}} \mathbf{v}_\beta^{\mathrm{R}} \cong \hat{\mathbf{y}}^T k^{\mathrm{R}} \mathbf{v}_\beta^{\mathrm{R}} \frac{\partial \beta^{\mathrm{R}}}{\partial \beta^{\mathrm{L}}} \quad => \quad k^{\mathrm{L}} \cong k^{\mathrm{R}} \frac{\partial \beta^{\mathrm{R}}}{\partial \beta^{\mathrm{L}}} =>$$

$$\frac{\partial \beta^{\mathrm{R}}}{\partial \beta^{\mathrm{L}}} \cong 1 \tag{A.34}$$

To calculate $\dfrac{\partial \alpha^{\mathrm{R}}}{\partial \beta^{\mathrm{L}}}$, multiply both sides of Eq. (23) by $\hat{\mathbf{x}}^T$ from the left and use (A.34):

$$\hat{\mathbf{x}}^T k^{\mathrm{L}} \mathbf{v}^{\mathrm{L}} = \hat{\mathbf{x}}^T k_\alpha^{\mathrm{R}} \mathbf{v}^{\mathrm{R}} \frac{\partial \alpha^{\mathrm{R}}}{\partial \beta^{\mathrm{L}}} + \hat{\mathbf{x}}^T k^{\mathrm{R}} \mathbf{v}_\alpha^{\mathrm{R}} \frac{\partial \alpha^{\mathrm{R}}}{\partial \beta^{\mathrm{L}}} + \hat{\mathbf{x}}^T k_\beta^{\mathrm{R}} \mathbf{v}^{\mathrm{R}}$$

$$k_\beta \hat{\mathbf{x}}^T (\mathbf{v}^{\mathrm{L}} - \mathbf{v}^{\mathrm{R}}) = (k_\alpha \hat{\mathbf{x}}^T \mathbf{v}^{\mathrm{R}} + k \hat{\mathbf{x}}^T \mathbf{v}_\alpha) \frac{\partial \alpha^{\mathrm{R}}}{\partial \beta^{\mathrm{L}}}$$

$$\frac{\partial \alpha^{\mathrm{R}}}{\partial \beta^{\mathrm{L}}} = \frac{k_\beta \hat{\mathbf{x}}^T (\mathbf{v}^{\mathrm{L}} - \mathbf{v}^{\mathrm{R}})}{(k_\alpha \hat{\mathbf{x}}^T \mathbf{v}^{\mathrm{R}} + k \hat{\mathbf{x}}^T \mathbf{v}_\alpha)} \tag{A.35}$$

Using (A.31) and the approximation: $0 \cong \hat{\mathbf{x}}^T k_\alpha \mathbf{v}^{\mathrm{R}} \ll \hat{\mathbf{x}}^T k \mathbf{v}_\alpha \cong k$, (A.35) can be simplified to:

$$\frac{\partial \alpha^{\mathrm{R}}}{\partial \beta^{\mathrm{L}}} = \frac{\Delta x}{k^2} k_\beta \tag{A.36}$$

Then, using eq. (13) we finally get:

$$\frac{\partial \alpha^{\mathrm{R}}}{\partial \beta^{\mathrm{L}}} = -\frac{\Delta x}{k} \frac{\mathbf{n} \bullet \mathbf{v}_\beta}{\mathbf{n} \bullet \mathbf{v}} \tag{A.37}$$

REFERENCES

[1] K. Rayner, "Eye movements in reading and information processing: 20 years of research," *Psychological Bulletin*, vol. 124, no. 3, pp. 372-422, Nov. 1998.

[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," *Accident Analysis & Prevention*, vol. 39, no. 2, pp. 372-379, Mar. 2007.

[3] 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.

[4] C. Karatekin and R. F. Asarnow, "Exploratory eye movements to pictures in childhood-onset schizophrenia and attention-deficit/hyperactivity disorder (ADHD)," *Journal of Abnormal Child Psychology*, vol. 27, no. 1, pp. 35-49, Feb. 1999.

[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] 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.

[8] 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.

[9] Z. Zhu and Q. Ji, "Eye and gaze tracking for interactive graphic display," *Machine Vision and Applications*, vol. 15, no. 3, pp. 139-148, 2004.

[10] T. E. Hutchinson, K. P. White, Jr., W. N. Martin, K. C. Reichert, and L. A. Frey, "Human-computer interaction using eye-gaze input," *Systems, Man and Cybernetics, IEEE Transactions on*, vol. 19, no. 6, pp. 1527-1534, 1989.

[11] A. Villanueva and R. Cabeza, "Models for gaze tracking systems," *J. Image Video Process.*, vol. 2007, no. 4, pp. 1-16, 2007.

[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] W. Jian-Gang and E. Sung, "Study on eye gaze estimation," *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on*, vol. 32, no. 3, pp. 332-350, 2002.

[14] E. D. Guestrin and M. Eizenman, "Remote point-of-gaze estimation with free head movements requiring a single-point calibration," in *Proc. of Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2007, pp. 4556-4560.

[15] 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.

[16] R. H. S. Carpenter, *Movements of the eyes*. London, UK: Pion, 1977.

[17] D. Model, E. D. Guestrin, and M. Eizenman, "An Automatic Calibration Procedure for Remote Eye-Gaze Tracking Systems," in *31st Annual International Conference of the IEEE EMBS*, Minneapolis, MN, USA, 2009, pp. 4751-4754.

[18] M. Eizenman, D. Model, and E. D. Guestrin, "Covert Monitoring of the Point-of-Gaze," in *IEEE TIC-STH*, Toronto, ON, Canada, 2009, pp. 551-556.

[19] 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.

[20] 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.

[21] S.-W. Shih, Y.-T. Wu, and J. Liu, "A calibration-free gaze tracking technique," in *In Proceedings of 15th Int. Conference on Pattern Recognition* 2000, pp. 201-204.

[22] T. Nagamatsu, J. Kamahara, and N. Tanaka, "Calibration-free gaze tracking using a binocular 3D eye model," in *Proceedings of the 27th International Conference on Human factors in Computing Systems*, Boston, MA, USA, 2009.

Dmitri Model was born in Moscow, Russia, in 1980. He received the B.Sc. (Cum Laude) and M.Sc. degrees in electrical engineering from Technion – Israel Institute of Technology in 2002 and 2006, respectively. He is currently working toward the Ph.D. degree at the Department of Electrical and Computer Engineering, University of Toronto. His current research interests include eye-tracking and gaze estimation technology and applications, signal and image processing, and numerical optimization.

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.