REMOTE POINT-OF-GAZE ESTIMATION WITH FREE HEAD MOVEMENTS

E. D. Guestrin* and M. Eizenman**

* Department of Electrical and Computer Engineering and Institute of Biomaterials and Biomedical Engineering, University of Toronto, 4 Taddle Creek Road, Toronto, Ontario, Canada M5S 3G9.
** Departments of Electrical and Computer Engineering and Ophthalmology, and Institute of Biomaterials and Biomedical Engineering, University of Toronto, Toronto, Ontario, Canada.

elias.guestrin@utoronto.ca

Abstract: The most common approach to remote point-of-gaze estimation uses the centers of the pupil and one corneal reflection. The drawback of this methodology is that, by itself, it suffers from estimation errors when the head moves. This paper shows that it is possible to estimate the point-of-gaze with free head movements by using more than one corneal reflection. A mathematical model is discussed in detail and a specific system implementation that uses one camera and two light sources to estimate the point-of-gaze on a computer screen is described. Experimental results obtained with four subjects yielded RMS point-of-gaze estimation errors that were less than 10 mm for all experimental conditions (about 0.9º of visual angle when the eye is at a typical distance of 65 cm from the computer screen). This implies that it is possible to distinguish unambiguously about 18 x 14 points on the screen of a 19" monitor.

Introduction

The point-of-gaze is the point in space that is imaged on the center of the highest acuity region of the retina (fovea) of each eye. Estimation of the point-of-gaze is required for the determination of visual scanning patterns. Since visual scanning patterns closely follow shifts in attentional focus, they provide insight into human cognitive processes. As such, analysis of visual scanning patterns is used, for example, in the study of mood, perception, attention and learning disorders [1]-[3]. For these applications, it is desired to estimate the point-of-gaze remotely (i.e. without attaching any device to the subject) while allowing for free head movements.

Most modern remote point-of-gaze estimation systems are based on the analysis of eye features that are extracted from images captured by a video camera. The most common approach to remote point-of-gaze estimation uses the centers of the pupil and one corneal reflection [4]. The corneal reflections (first Purkinje images) are virtual images of light sources (usually infrared) created by the front surface of the cornea, which acts as a convex mirror. The drawback of this methodology is that, by itself, it does not tolerate head movements. This paper presents a system that estimates the point-of-gaze with free head movements by using more than one corneal reflection.

The following section presents a mathematical model for remote point-of-gaze estimation using the coordinates of the centers of the pupil and corneal reflections (glints) estimated from images captured by a video camera.

Mathematical Model

The point-of-gaze is formally defined as the intersection of the visual axes of both eyes with the 3D scene. The visual axis is the line connecting the center of the fovea with the nodal point[1] of the eye's optics (Figure 1). Since in the human eye the visual axis deviates from the optic axis [4], the development that follows is divided into two parts. The first part considers the problem of reconstructing the optic axis of the eye from the centers of pupil and glints in the images of the eye. The second part deals with the reconstruction of the visual axis from the optic axis, and the estimation of the point-of-gaze.

Under the assumptions that the light sources are modeled as point sources and the video camera is modeled as a pinhole camera, Figure 1 presents a ray-tracing diagram of the system and the eye, where all points are represented as 3D column vectors (bold font) in a right-handed Cartesian world coordinate system. Consider a ray that comes from light source i, $\mathbf{l}_i$, and reflects at a point $\mathbf{q}_i$ on the corneal surface such that the reflected ray passes through the nodal point[2] of the camera, $\mathbf{o}$, and intersects the camera image plane at a point $\mathbf{u}_i$. The condition that the ray coming from the point of reflection $\mathbf{q}_i$ and passing through the nodal

[1] The nodal point of an optical system is the point on the optic axis where all lines that join object points with their respective image points intersect.
[2] The nodal point of a camera is also known as center of projection, camera center, and, sometimes, lens center.

ISSN: 1727-1983

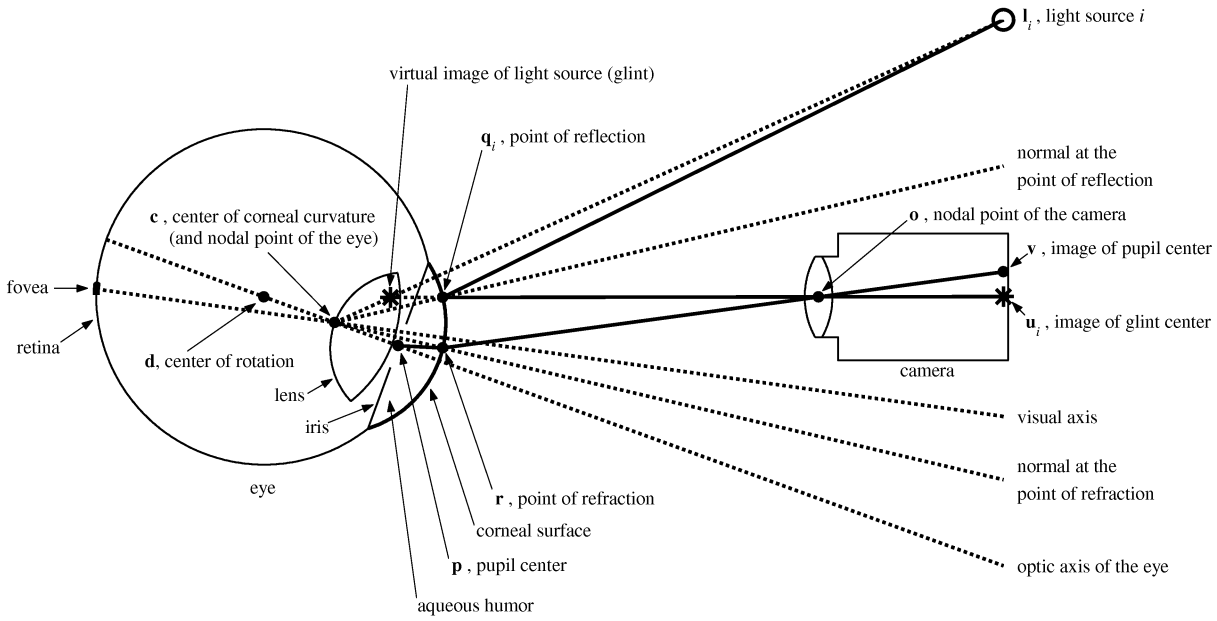

Figure 1: Ray-tracing diagram (not to scale in order to be able to show all the elements of interest), showing schematic representations of the eye, a camera and a light source.

point of the camera, **o**, intersects the camera image plane at point $\mathbf{u}_i$, can be expressed in parametric form as

$$\mathbf{q}_i = \mathbf{o} + k_{q,i}(\mathbf{o} - \mathbf{u}_i) \text{ for some } k_{q,i}\,, \tag{1}$$

whereas, if the corneal surface is modeled as a convex spherical mirror of radius R, the condition that $\mathbf{q}_i$ is on the corneal surface can be written as

$$\|\mathbf{q}_i - \mathbf{c}\| = R\ , \tag{2}$$

where **c** is the center of corneal curvature.

The law of reflection states two conditions: (a) the incident ray, the reflected ray and the normal at the point of reflection are coplanar; and (b) the angles of incidence and reflection are equal. Since vector $(\mathbf{q}_i - \mathbf{c})$ is normal to the spherical surface at the point of reflection $\mathbf{q}_i$, condition (a) implies that points $\mathbf{l}_i$, $\mathbf{q}_i$, **c** and **o** are coplanar. Noting that three coplanar vectors $\mathbf{a}_1$, $\mathbf{a}_2$ and $\mathbf{a}_3$ satisfy $\mathbf{a}_1 \times \mathbf{a}_2 \bullet \mathbf{a}_3 = 0$, condition (a) can be formalized as

$$(\mathbf{l}_i - \mathbf{o}) \times (\mathbf{q}_i - \mathbf{o}) \bullet (\mathbf{c} - \mathbf{o}) = 0\ . \tag{3}$$

Since the angle θ between two vectors **a** and **b** can be obtained from $\mathbf{a} \bullet \mathbf{b} = \|\mathbf{a}\| \cdot \|\mathbf{b}\| \cos\theta$, condition (b) can be expressed as

$$\begin{aligned}&(\mathbf{l}_i - \mathbf{q}_i) \bullet (\mathbf{q}_i - \mathbf{c}) \cdot \|\mathbf{o} - \mathbf{q}_i\| \\ &= (\mathbf{o} - \mathbf{q}_i) \bullet (\mathbf{q}_i - \mathbf{c}) \cdot \|\mathbf{l}_i - \mathbf{q}_i\|\,.\end{aligned} \tag{4}$$

Next, consider a ray that comes from the pupil center, **p**, and refracts at point **r** on the corneal surface such that the refracted ray passes through the nodal point of the camera, **o**, and intersects the camera image plane at a point **v**. The condition that the ray coming from the point of refraction **r** and passing through the nodal point of the camera, **o**, intersects the camera image plane at point **v**, can be expressed in parametric form as

$$\mathbf{r} = \mathbf{o} + k_r(\mathbf{o} - \mathbf{v}) \text{ for some } k_r\,, \tag{5}$$

whereas the condition that **r** is on the corneal surface can be written as

$$\|\mathbf{r} - \mathbf{c}\| = R\ . \tag{6}$$

The law of refraction states two conditions: (a) the incident ray, the refracted ray and the normal at the point of refraction are coplanar; and (b) the angle of incidence, θ_1, and the angle of refraction, θ_2, satisfy Snell's law (i.e. $n_1 \sin\theta_1 = n_2 \sin\theta_2$, where n_1 and n_2 are the indices of refraction of mediums 1 and 2). Since vector $(\mathbf{r} - \mathbf{c})$ is normal to the spherical surface at the point of refraction **r**, condition (a) implies that points **p**, **r**, **c** and **o** are coplanar, which can be formalized as

$$(\mathbf{r} - \mathbf{o}) \times (\mathbf{c} - \mathbf{o}) \bullet (\mathbf{p} - \mathbf{o}) = 0\ . \tag{7}$$

Since the sine of the angle θ between two vectors **a** and **b** can be obtained from $\|\mathbf{a} \times \mathbf{b}\| = \|\mathbf{a}\| \cdot \|\mathbf{b}\| \sin\theta$, condition (b) can be expressed as

$$\begin{aligned}&n_1 \cdot \|(\mathbf{r} - \mathbf{c}) \times (\mathbf{p} - \mathbf{r})\| \cdot \|\mathbf{o} - \mathbf{r}\| \\ &= n_2 \cdot \|(\mathbf{r} - \mathbf{c}) \times (\mathbf{o} - \mathbf{r})\| \cdot \|\mathbf{p} - \mathbf{r}\|\,,\end{aligned} \tag{8}$$

where n_1 is the effective index of refraction of the aqueous humor and cornea combined and n_2 is the index of refraction of air (≈ 1). In this model, the refraction at the aqueous humor-cornea interface is neglected since the difference in their indices of refraction is small relative to that of the cornea-air interface. Only the refraction at the cornea-air interface is taken into account and the aqueous humor and cornea are considered as a homogenous medium.

Finally, considering the distance K between the pupil center and the center of corneal curvature leads to

ISSN: 1727-1983

$$\|\mathbf{p}-\mathbf{c}\| = K \quad . \tag{9}$$

Since the optic axis of the eye passes through the pupil center (**p**) and the center of corneal curvature (**c**), if the above system of equations is solved for **c** and **p**, the optic axis of the eye in space can be reconstructed as the line defined by these two points. Notice that in order to solve the above system of equations, the eye parameters R, K and n_1 have to be known. These eye parameters are subject-specific and are not easily measured directly. Therefore, in general, they are obtained through a calibration procedure that is performed for each subject (an example is provided in the next section). The typical values of these eye parameters are R = 7.8 mm [4], [5], K = 4.2 mm [5], and n_1 = 1.3375 [6].

Since the point-of-gaze is defined as the intersection of the visual axis rather than the optic axis with the scene, the relation between these two axes has to be modeled. The visual axis is the line defined by the nodal point of the eye[3] and the center of the fovea (i.e. the highest acuity region of the retina corresponding to 0.6 to 1º of visual angle), and deviates from the optic axis [4] (Figure 1). In a typical adult human eye, the fovea falls about 4-5º temporally and about 1.5º below the point of intersection of the optic axis and the retina [7].

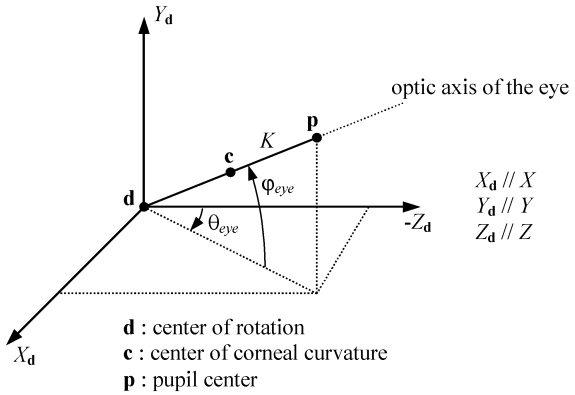

Figure 2: Orientation of the optic axis of the eye.

In order to formalize the relation between the visual and optic axes, suppose that the scene is a vertical plane (e.g. a projection screen or computer monitor) and that the world coordinate system is a right-handed 3D Cartesian coordinate system whose XY-plane is coincident with the scene plane, with the X-axis horizontal, the Y-axis vertical and the positive Z-axis coming out of the scene plane. Then, the orientation of the optic axis of the eye can be described by the pan (horizontal) angle θ_{eye} and the tilt (vertical) angle φ_{eye} defined in Figure 2, where the world coordinate system is translated to the center of rotation of the eye, **d**. As it can be derived from this figure, the angles θ_{eye} and φ_{eye} can be obtained from **c** and **p** by solving the following equation:

$$\frac{\mathbf{p}-\mathbf{c}}{\|\mathbf{p}-\mathbf{c}\|} = \begin{bmatrix} \cos\varphi_{eye}\sin\theta_{eye} \\ \sin\varphi_{eye} \\ -\cos\varphi_{eye}\cos\theta_{eye} \end{bmatrix} . \tag{10}$$

If the horizontal and vertical angles between the visual and optic axes are given by α_{eye} and β_{eye}, respectively, the orientation of the visual axis can be expressed by the pan angle ($\theta_{eye} + \alpha_{eye}$) and the tilt angle ($\varphi_{eye} + \beta_{eye}$), where all angles are signed. In particular, $\alpha_{eye} < 0$ for the right eye while $\alpha_{eye} > 0$ for the left eye, and $\beta_{eye} > 0$. The eye parameters α_{eye} and β_{eye} are also subject-specific and are usually estimated as part of a calibration procedure that is performed for each subject.

To completely define the visual axis in space, in addition to its orientation, a point through which it passes is required. The visual axis and the optic axis intersect at the nodal point of the eye. Since the nodal point remains within 1 mm of the center of corneal curvature for different degrees of eye accommodation [4], for the sake of simplicity, the nodal point is assumed to be coincident with the center of corneal curvature (**c**).

From the above discussion, it follows that the visual axis can be then described in parametric form as

$$\mathbf{g} = \mathbf{c} + k_g \begin{bmatrix} \cos(\varphi_{eye}+\beta_{eye})\sin(\theta_{eye}+\alpha_{eye}) \\ \sin(\varphi_{eye}+\beta_{eye}) \\ -\cos(\varphi_{eye}+\beta_{eye})\cos(\theta_{eye}+\alpha_{eye}) \end{bmatrix} \tag{11}$$

for all k_g. Since it was assumed that the scene plane is at Z = 0, the point-of-gaze is given by (11) for a value of k_g such that the Z-component of **g**, g_Z, equals 0, that is,

$$k_g = \frac{c_Z}{\cos(\varphi_{eye}+\beta_{eye})\cos(\theta_{eye}+\alpha_{eye})} \quad . \tag{12}$$

In the remainder of this section, it is assumed that the world coordinates of the positions of the light sources ($\mathbf{l}_i$), the nodal point of the camera (**o**) and the centers of the pupil (**v**) and glints ($\mathbf{u}_i$) in the eye images, are known. Since the centers of the pupil and glints that are estimated in each eye image are measured in pixels in an image coordinate system, they have to be transformed into world coordinates ([8], pp. 26-28, 34-39). In order to transform from image coordinates into world coordinates, all camera parameters, including the position of the nodal point (**o**), must be known. Typically, the camera parameters are estimated through a camera calibration procedure ([8], pp. 123-138), whereas the positions of the light sources are measured directly. In general, the system structure is fixed and hence these system parameters are measured/estimated only once during system set up.

The above development shows that (a) the reconstruction of the optic axis of the eye as the line defined by the center of corneal curvature (**c**) and the pupil center (**p**), using (1) through (9), *depends* on the number of light sources; and (b) once the optic axis of the eye is obtained, the reconstruction of the visual axis and the estimation of the point-of-gaze, using (10)

[3] Actually, the eye has two nodal points, 0.3 mm apart. For the sake of simplicity, a single nodal point is considered.

ISSN: 1727-1983

through (12), are *independent* of the number of light sources. For this reason, the following paragraphs concentrate on the reconstruction of the optic axis of the eye for different number of light sources.

When only a single light source is used, if the eye parameters R, K and n_1 are known, the system of equations (1) through (9) with $i = 1$, is equivalent to 13 scalar equations with 14 scalar unknowns. This means that the problem cannot be solved unless another constraint such as

$$\|\mathbf{o} - \mathbf{c}\| = known \tag{13}$$

is introduced. This constraint can be satisfied if the head is fixed relative to the system or if the distance between the eye and the camera is estimated somehow (e.g. magnetic head tracker, ultrasonic transducer, auto-focus system, etc.) [4], [9]-[13].

The use of multiple light sources allows for the solution of the system of equations (1) through (9) with $i = 1, \ldots, N$, if the eye parameters (R, K and n_1) are obtained through a calibration procedure. In this case, it is advantageous to substitute (1) into (3) to obtain

$$\underbrace{(\mathbf{l}_i - \mathbf{o}) \times (\mathbf{u}_i - \mathbf{o})}_{\substack{\text{normal to the plane defined by} \\ \mathbf{l}_i, \mathbf{o} \text{ and } \mathbf{u}_i}} \bullet (\mathbf{c} - \mathbf{o}) = 0 \,. \tag{3'}$$

This equation means that the center of corneal curvature, $\mathbf{c}$, belongs to each plane defined by the nodal point of the camera, $\mathbf{o}$, light source i, $\mathbf{l}_i$, and its corresponding image point $\mathbf{u}_i$. Moreover, all those planes intersect at the line defined by points $\mathbf{c}$ and $\mathbf{o}$. By noting that $\mathbf{a} \bullet \mathbf{b} = \mathbf{a}^T \mathbf{b}$, (3'), $i = 1, \ldots, N$ can be written in matrix form as

$$\underbrace{\begin{bmatrix} [(\mathbf{l}_1 - \mathbf{o}) \times (\mathbf{u}_1 - \mathbf{o})]^T \\ [(\mathbf{l}_2 - \mathbf{o}) \times (\mathbf{u}_2 - \mathbf{o})]^T \\ \vdots \\ [(\mathbf{l}_N - \mathbf{o}) \times (\mathbf{u}_N - \mathbf{o})]^T \end{bmatrix}}_{\mathbf{M}} (\mathbf{c} - \mathbf{o}) = \mathbf{0} \,. \tag{14}$$

From the interpretation of (3') it follows that matrix $\mathbf{M}$ has, at most, rank 2. If $\mathbf{M}$ has rank 2, the solution to (14) is given by an equation of the form

$$\mathbf{c} - \mathbf{o} = k_{c,b}\, \mathbf{b}_{norm} \,, \tag{15}$$

which defines vector $(\mathbf{c} - \mathbf{o})$ up to a scale factor. From this reasoning, it follows that (1), (2), (4), $i = 1, \ldots, N$, and (15) are equivalent to $(5N + 3)$ scalar equations with $(4N + 4)$ scalar unknowns. In particular, when $N = 2$, $\mathbf{b}_{norm}$ is a unit vector in the direction of the line of intersection of the planes whose normals are given by $[(\mathbf{l}_1 - \mathbf{o}) \times (\mathbf{u}_1 - \mathbf{o})]$ and $[(\mathbf{l}_2 - \mathbf{o}) \times (\mathbf{u}_2 - \mathbf{o})]$, thus

$$\begin{aligned} \mathbf{b}_{norm} &= \mathbf{b} / \|\mathbf{b}\| \,, \\ \mathbf{b} &= [(\mathbf{l}_1 - \mathbf{o}) \times (\mathbf{u}_1 - \mathbf{o})] \times [(\mathbf{l}_2 - \mathbf{o}) \times (\mathbf{u}_2 - \mathbf{o})] \,, \end{aligned} \tag{16}$$

and (1), (2), (4), $i = 1, 2$, and (15) are equivalent to 13 scalar equations with 12 scalar unknowns.

In the special case that $\mathbf{M}$ in (14) has rank 1 ($\mathbf{b} = \mathbf{0}$ in (16)), which means that all normals given by $(\mathbf{l}_i - \mathbf{o}) \times (\mathbf{u}_i - \mathbf{o})$ are parallel, the effective number of scalar equations decreases to $(5N + 2)$. In the case that $N = 2$, it results in the equivalent to 12 scalar equations with 12 scalar unknowns.

Consequently, if multiple light sources are used, there are enough equations to solve for the center of corneal curvature $\mathbf{c}$. Knowing $\mathbf{c}$, (5) and (6) are used to compute the point of refraction $\mathbf{r}$ (4 scalar unknowns and 4 scalar equations). Knowing $\mathbf{c}$ and $\mathbf{r}$, (7) through (9) are used to compute $\mathbf{p}$ (3 scalar unknowns and 3 scalar equations). Knowing $\mathbf{c}$ and $\mathbf{p}$, the optic axis of the eye can be reconstructed as the line defined by these two points.

The above discussion shows that *one camera and two light sources is the simplest configuration that allows for the reconstruction of the optic axis of the eye from the centers of pupil and glints while allowing for free head movements.* The above analysis also shows that after finding $\mathbf{c}$ (the center of corneal curvature), the calculation of $\mathbf{p}$ (the pupil center) is independent of the number of light sources (7 scalar equations and 7 scalar unknowns regardless of the number of light sources).

The following section provides experimental results corresponding to a system that is used to estimate the point-of-gaze on a computer screen.

Experimental Results

The point-of-gaze estimation system utilizes two near-IR (850 nm) light sources that are symmetrically positioned at the sides of a 19" computer monitor, and a video camera (640 x 480 pixels, 1/3" CCD with a 35 mm lens) that is centered under the screen. A typical image from the video camera for a subject sitting at a distance of 65 cm from the monitor (typical viewing distance), with his head approximately at the center of the region of allowed head movement, is shown in Figure 3. This specific system can tolerate only moderate head movements of about ±3 cm laterally, ±2 cm vertically, and ±4 cm backwards/forward, before the eye features are no longer in the field of view of the camera or are out of focus.

In order to be able to estimate the point-of-gaze on the screen, a set of system and subject-specific eye parameters has to be measured/estimated. Since the system components are fixed relative to the computer monitor, the system parameters (the position of the two light sources, $\mathbf{l}_1$ and $\mathbf{l}_2$, and the extrinsic and intrinsic camera parameters, which include the nodal point of the camera, $\mathbf{o}$) are measured/estimated only once during system set up. The subject-specific eye parameters (R, K, n_1, α_{eye} and β_{eye}) are obtained through a calibration procedure that is performed once for each subject. In the calibration procedure, the subject fixates on 9 evenly distributed points that are presented sequentially on the screen. For each fixation point, 100 estimates of the image coordinates of the centers of pupil and glints are obtained and the average coordinates of these features are computed. Using the average coordinates of the centers of pupil and glints, the eye parameters are optimized to minimize the sum of the square errors between the points on the screen and the estimated

ISSN: 1727-1983

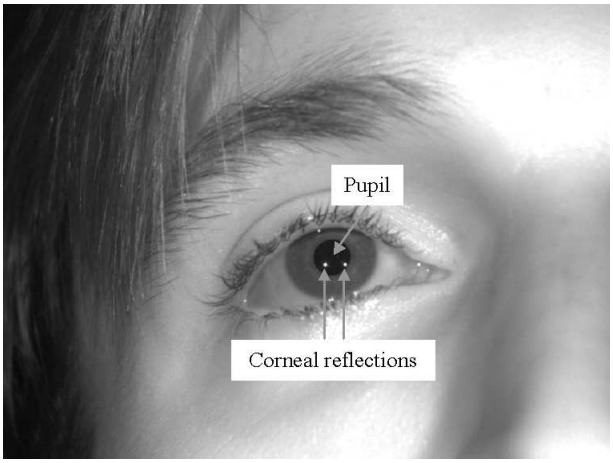

Figure 3: Sample eye image showing the pupil and the two corneal reflections (glints).

points-of-gaze [14]. The initial guess for the optimization is given by the typical values of the eye parameters provided in the previous section. During the calibration procedure the head is positioned at the center of the region of allowed head movements (central position).

To estimate the point-of-gaze, the coordinates of the centers of pupil and glints are first estimated in each image captured by the video camera [15]. These image coordinates are then transformed into world coordinates ([8], pp. 26-28, 34-39) as $\mathbf{v}$ for the pupil center, and $\mathbf{u}_1$ and $\mathbf{u}_2$ for the glints. Next, the center of corneal curvature, $\mathbf{c}$, is calculated from (1), (2), (4), $i = 1, 2$, and (15)-(16). Knowing $\mathbf{c}$, (5) and (6) are used to compute the point of refraction $\mathbf{r}$. Knowing $\mathbf{c}$ and $\mathbf{r}$, (7) through (9) are used to compute $\mathbf{p}$. Knowing $\mathbf{c}$ and $\mathbf{p}$, the optic axis of the eye in space is reconstructed as the line defined by these two points. Finally, using (10) through (12), the visual axis of the eye is obtained and the point-of-gaze on the screen is estimated.

A preliminary evaluation of the performance of this point-of-gaze estimation system was carried out through experiments with 4 subjects. In these experiments, the head of each subject was placed in the central position and 4 positions at the edges of the region of allowed head movements. The 4 edge positions correspond to lateral and backward/forward head displacements. For each head position, the subject was asked to fixate on 9 points on the computer screen and 100 estimates (≈3.3 seconds @ 30 estimates/second) of the point-of-gaze were obtained for each fixation point. The resulting RMS errors in the estimation of the point-of-gaze for the central position and the edge positions are summarized in Table 1 (RMS error). Table 1 also shows the RMS errors when the point-of-gaze was estimated for the average of the coordinates of the centers of pupil and glints (ACPG-RMS error). A typical example of point-of-gaze estimation errors for the central head position is shown in Figure 4. The ACPG-RMS errors (Table 1) correspond to the deviation of the white crosses from the centers of the dotted circles in Figure 4 and are the result of bias in the estimation of the point-of-gaze due to the deviation of the real corneal shape from the ideal spherical shape assumed in the development of the mathematical model (corneal asphericity). The dispersion of the asterisks around the white crosses is caused by noise in the estimates of the image coordinates of the centers of pupil and glints. The RMS errors shown in the last column of Table 1 correspond to the combined effects of bias and dispersion of the point-of-gaze estimates. It can be also observed that the RMS errors for the edge head positions are larger than the RMS errors for the central head position, which is also due to corneal asphericity. If the cornea were truly spherical, head displacements would have no effect on the point-of-gaze estimation error.

Table 1: Experimental RMS Point-of-Gaze Estimation Errors

Subject	Head position(s)	ACPG-RMS [a] error (mm)	RMS error (mm)
M. E.	Central	2.93	5.15
	Edges	4.25	6.12
J. K.	Central	4.88	6.41
	Edges	5.33	6.73
R. H.	Central	4.93	7.01
	Edges	5.74	7.58
E. G.	Central	4.68	6.25
	Edges	8.12	9.04

[a]RMS error when the point-of-gaze was estimated for the average coordinates of the centers of pupil and glints.

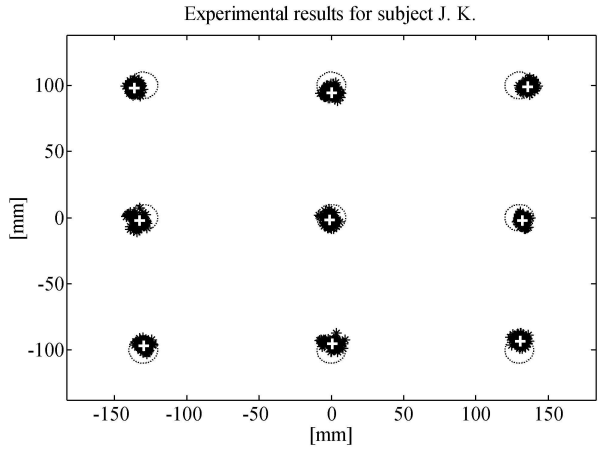

Figure 4: Experimental point-of-gaze estimation results for subject J. K. The centers of the dotted circles (10 mm radius) indicate the intended fixation points, the asterisks represent the estimates of the point-of-gaze, and the white crosses represent the estimates of the point-of-gaze for the average coordinates of the centers of pupil and glints.

In the experiments described above, the RMS error of the point-of-gaze estimation was less than 10 mm for all experimental conditions. This is equivalent to about

ISSN: 1727-1983

0.9º of visual angle when the eye is at a distance of 65 cm from the computer screen. This implies that it is possible to distinguish unambiguously about 18 x 14 points on the screen of a 19" monitor.

Conclusions

This paper presented a model for remote point-of-gaze estimation systems that use the coordinates of the centers of the pupil and corneal reflections (glints) estimated from video images.

It was shown that if only one light source is used, the point-of-gaze cannot be determined from the coordinates of the centers of the pupil and the glint, unless the head is stationary or head position information is estimated by some other means. When at least two light sources are used, it is possible to estimate the point-of-gaze while allowing for free head movements.

A specific system implementation that uses one camera and two light sources to estimate the point-of-gaze on a computer screen was described. Experimental results obtained with four subjects yielded RMS point-of-gaze estimation errors that were less than 10 mm for all experimental conditions (about 0.9º of visual angle when the eye is at a distance of 65 cm from the computer screen). This implies that it is possible to distinguish unambiguously about 18 x 14 points on the screen of a 19" monitor.

The accuracy of the point-of-gaze estimation is limited mainly by (a) the subject-specific deviation of the shape of the cornea from the spherical cornea assumed in the model; and (b) the noise in the estimation of the centers of pupil and corneal reflections.

Acknowledgement

An extended version of this work is under review for publication in the IEEE Transactions on Biomedical Engineering [16].

References

[1] EIZENMAN M., YU L. H., GRUPP L., EIZENMAN E., ELLENBOGEN M., GEMAR M., and LEVITAN R. D. (2003): 'A Naturalistic Visual Scanning Approach to Assess Selective Attention in Major Depressive Disorder', *Psychiat. Res.*, **118** (2), pp. 117-128

[2] KARATEKIN C. and ASARNOW R. F. (1999): 'Exploratory Eye Movements to Pictures in Childhood-Onset Schizophrenia and Attention-Deficit/Hyperactivity Disorder (ADHD)', *J. Abnorm. Child Psychol.*, **27** (1), pp. 35-49

[3] DE LUCA M., DI PACE E., JUDICA A., SPINELLI D., and ZOCCOLOTTI P. (1999): 'Eye Movement Patterns in Linguistic and Non-Linguistic Tasks in Developmental Surface Dyslexia', *Neuropsychologia*, **37** (12), pp. 1407-1420

[4] YOUNG L. R. and SHEENA D. (1975): 'Methods and Designs - Survey of Eye Movement Recording Methods', *Behav. Res. Meth. Instrum.*, **7** (5), pp. 397-429

[5] GALE A. G. (1982): 'A Note on the Remote Oculometer Technique for Recording Eye Movements', *Vision Research*, **22** (1), pp. 201-202

[6] CORBETT M. C., ROSEN E. S., and O'BRART D. P. S. (1999): 'Corneal Topography: Principles and Applications', (BMJ Books, London, UK), pp. 6

[7] SLATER A. M. and FINDLAY J. M. (1972): 'The Measurement of Fixation Position in the Newborn Baby', *J. Exp. Child Psychol.*, **14**, pp. 349-364

[8] TRUCCO E. and VERRI A. (1998): 'Introductory Techniques for 3-D Computer Vision', (Prentice Hall, Upper Saddle River, NJ)

[9] HUTCHINSON T. E., WHITE K. P., MARTIN W. N., REICHERT K. C., and FREY L. A. (1989): 'Human-Computer Interaction Using Eye-Gaze Input', *IEEE Trans. Syst., Man, Cybern.*, **19** (6), pp. 1527-1534

[10] MERCHANT J., MORRISSETTE R., and PORTERFIELD J. L. (1974): 'Remote Measurement of Eye Direction Allowing Subject Motion over One Cubic Foot of Space," *IEEE Trans. Biomed. Eng.*, **BME-21** (4), pp. 309-317

[11] CLEVELAND D. (1999): 'Unobtrusive Eyelid Closure and Visual Point of Regard Measurement System', Proc. Tech. Conf. on Ocular Measures of Driver Alertness, sponsored by The Federal Highway Administration – Office of Motor Carrier and Highway Safety and The National Highway Traffic Safety Administration – Office of Vehicle Safety Research, Herndon, VA, USA, 1999, pp. 57-74

[12] SUGIOKA A., EBISAWA Y., and OHTANI M. (1996): 'Noncontact Video-Based Eye-Gaze Detection Method Allowing Large Head Displacements', Proc. 18th Annual Int. Conf. IEEE Eng. Med. Biol. Soc., 1996, **2**, pp. 526-528.

[13] EBISAWA Y., OHTANI M., SUGIOKA A., and ESAKI S. (1997), "Single Mirror Tracking System for Free-Head Video-Based Eye-Gaze Detection Method', Proc. 19th Annual Int. Conf. IEEE Eng. Med. Biol. Soc., 1997, **4**, pp. 1448-1451.

[14] GUESTRIN E. D. (2003): 'A Novel Head-Free Point-of-Gaze Estimation System', M.A.Sc. Thesis, Dept. of Electrical and Computer Engineering, University of Toronto, Canada

[15] LUI B. J. M. (2003): 'A Point-Of-Gaze Estimation System for Studies of Visual Attention', M.A.Sc. Thesis, Dept. of Electrical and Computer Engineering, University of Toronto, Canada

[16] GUESTRIN E. D. and EIZENMAN M. (2005): 'General Theory of Remote Gaze Estimation Using the Pupil Center and Corneal Reflections', *IEEE Trans. Biomed. Eng.*, under review.

ISSN: 1727-1983