Non-Contact Binocular Eye-Gaze Tracking for Point-of-Gaze Estimation in Three Dimensions

Craig Hennessey, and Peter Lawrence, *Senior Member, IEEE*

***Abstract*—Binocular eye-gaze tracking can be used to estimate the point-of-gaze (POG) of a subject in real world three-dimensional (3D) space using the vergence of the eyes. In this paper, a novel non-contact, model-based technique for 3D POG estimation is presented. The non-contact system allows people to select real world objects in 3D physical space using their eyes, without the need for head-mounted equipment. Remote 3D POG estimation may be especially useful for persons with quadriplegia or advanced ALS. It would also enable a user to select 3D points in space generated by 3D volumetric displays, with potential applications to medical imaging and telesurgery. Using a model-based POG estimation algorithm allows for free head motion and a single stage of calibration. It is shown that an average accuracy of 3.93 cm was achieved over a workspace volume of 30 x 23 x 25 cm (W x H x D) with a maximum latency of 1.5 seconds due to the digital filtering employed. The users were free to naturally move and reorient their heads while operating the system, within an allowable headspace of 3 x 9 x 14 cm.**

***Index Terms*—Eye-gaze tracking, non-contact, 3D POG, binocular, high speed, human computer interface.**

I. Introduction

THE point of conscious attention of an individual can be used to provide insight into cognitive processes - information that may otherwise be difficult to obtain [1]. Eye movements, and the resulting point-of-gaze (POG) of a subject can be estimated automatically with an eye-gaze tracker. With the real-time capabilities of modern eye-gaze tracking systems the use of eye-gaze has expanded from a diagnostic tool to applications in which the POG is used for control as well [2].

Two dimensional (2D) displays are currently the standard method of visual display used with eye-gaze trackers. Considerable progress however has been made towards the development of stereoscopic, or three dimensional (3D) displays [3]. In addition to enhancing the realism of the viewing experience, 3D displays can be used to more readily view complex volumetric data sets in medical imaging (magnetic resonance and computed tomography for example), 3D computer-aided design, and telesurgery. Furthermore, autostereoscopic displays which do not require any contact with the viewers face have been developed [4], [5].

The ability to determine a user's POG in 3D space will become increasingly important as the use of 3D displays becomes more widespread. The current methods for 3D interaction typically use an electromagnetic or optically tracked stylus held by the user in 3D space against gravity [6]. Using the 3D POG for 3D interaction avoids the visual disconnect when the tracked tool cannot be physically located within the environment in which it is supposed to be acting [7]. The 3D POG also requires no physical effort other than directing the gaze to the point of interest. In addition to interaction with 3D displays, the 3D POG can be used to provide a means for interaction in real world 3D spaces using only the eyes. This could be an important advance for individuals with restricted mobility such as those with high level spinal cord injuries or advanced degenerative motor neuron diseases.

Limitations of existing eye-gaze tracking systems are application dependent. In research or clinical studies of eye movement, some inconveniences (e.g. head mounted equipment, long calibration processes) may be acceptable. For other users of a system, including the general public, the same deficiencies in usability may not be acceptable. A number of significant limitations for 2D eye-gaze tracking have been listed by Morimoto *et al* [8], difficulties which are further compounded when extending from 2D to 3D. Some of these limitations include low accuracy, low sampling rates, poor precision, complex and lengthy calibrations and uncomfortable user requirements including the need to wear the system on the users head, or to maintain a fixed head position. The usability of modern eye-gaze tracking systems may be a major reason why they are most commonly found in research based environments or specialized applications and are not widely used by the general population.

One of the areas targeted for improvement has been on increasing the usability of eye-gaze trackers with the transition from head mounted to remote eye-gaze tracking [8], which mirrors the transition to autostereoscopic displays for improving the usability of 3D displays. Head mounted systems are well suited to eye-gaze tracking applications involving user mobility such as walking or active sports [9], however, users may be averse to wearing headgear in everyday computer use. In addition, slippage of the head gear can result in increased error or require recalibration. In applications where the subject is seated, eye-gaze trackers based on remote image recording can enhance the user experience by requiring no contact with the subject's face or head.

There are two main image based techniques for estimating the POG, the Pupil-Corneal Reflection (P-CR) method [10] and methods based on models of the eye and system [11], [12], [13]. The P-CR method uses the vector formed from a reflection generated off the surface of the cornea and the center of the pupil, along with a polynomial mapping (determined through calibration [8] [14]) to determine the POG on a 2D surface such as a computer screen. The P-CR method is well suited to head mounted applications in which the distance from the eye to camera changes little, as the accuracy of the estimated POG has been shown to degrade when head motion is coupled with eye motion [8]. Model-based methods are designed to avoid the degradation in POG accuracy as head motion is implicitly compensated. With the model-based methods the image features are used to determine the position of the eyes in 3D space, the visual axis along which the user is looking, and the POG at the intersection of the visual axis and the surface of interest.

One of the first systems developed to investigate 3D POG estimation was presented by Duchowski *et al* [15] for use in a 3D virtual reality environment. A commercial Head Mounted Display (HMD) was used to provide disparity images to the left and right eyes. A commercial, binocular, head mounted eye-gaze tracker using the P-CR method for POG estimation was integrated with the HMD to determine the user's 2D POG on the left and right HMD screens. In addition to the eye-gaze tracker, an electro-magnetic tracker was attached to the head mounted apparatus to determine head position and orientation. Two stages of per-user calibration were required, the first to calibrate the eye-gaze tracker on the HMD and the second to provide estimates for the geometric parameters such as the interpupillary distance (the distance between the eyes) and the distance from the eyes to the surface of the HMD screens. Standard stereo geometry techniques [16] were then used to estimate the 3D POG based on the head pose and 2D POG estimates.

The 3D POG estimation system developed by Essig *et al* also used a binocular P-CR based head mounted eye-gaze tracker, however a neural network was used to generate the 3D POG estimates [17]. The 2D POG estimates were tracked on a remote desktop monitor and used as input to a neural network which then estimated the 3D POG. In their original work the 2D computer display used single image random dot stereograms to provide the virtual 3D display while in their later work anaglyph images were used [18]. Two stages of calibration were required, the first to calibrate the eye-gaze tracker on the desktop display and the second to train the neural network.

The system recently developed by Munn and Pelz [19] for 3D POG estimation again used a P-CR based head mounted eye-gaze tracker, however, only a single eye was used for their method. A head mounted scene camera was used to record a 2D projection of the subject's scene view, upon which the monocular 2D POG estimates were tracked. With sufficient head motion the monocular visual axis vectors over time were intersected to determine the 3D POG, provided the head mounted camera position and orientation were also accurately tracked in 3D space.

A novel binocular system by Kwon *et al* [20] estimated the 3D POG in a virtual 3D environment with a 2D parallax barrier display. The P-CR method was used to determine eye-gaze direction, along with the relative displacements of the binocular pupil centers to estimate the depth of the 3D POG. This technique required a fixed head to camera displacement which was achieved using a chin rest.

The system by Mitsugami *et al* [21] also used a fixed head position with head-mounted binocular eye-gaze tracking and the P-CR method for 2D POG estimation on a view camera scene. Binocular intersection of the view vectors was used along with the known position and orientation of the fixed head to estimate the 3D POG. A novel stochastic method was proposed for improving 3D POG estimation with increasing distance by including view lines from multiple head positions, as with Munn and Pelz. As the head position was fixed, head motion was simulated using a moving target.

The system we propose for 3D POG estimation follows the design goal of improving the usability of eye-gaze tracking with no contact required and no equipment mounted on the user's head. Our system uses a model-based method for estimating the 3D POG, which allows for head motion and does not require fixing the position or orientation of the head with a chin rest. The model-based method uses image features directly and avoids the intermediate stage of 2D POG estimation on a 2D surface, simplifying the per user calibration to a single stage. The system we propose also estimates the POG in a 3D real world volume in real-time and does not require large head motions as binocular eye-tracking is employed.

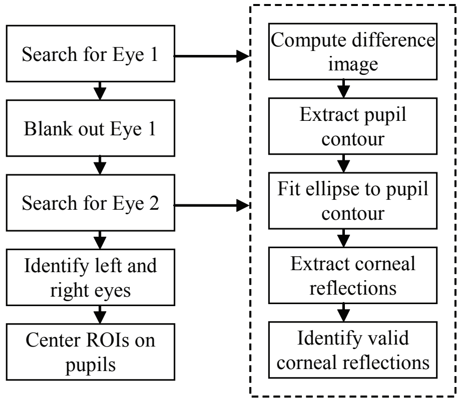
To the best of the authors' knowledge this paper has three original contributions. The first is the design of the first reported binocular system for estimating the absolute X, Y, Z coordinates of where one is looking in the real 3D world. Secondly, this is the first system that uses a model-based method for 3D POG estimation and therefore requires only a single per-user calibration stage. Finally, it is the first non-contact, head-free 3D POG eye-gaze tracking system to be reported and/or evaluated in the literature. With no attachments to the user's head or use of chin rests to fix the position of the head, the system permits eye and head motions within the field of view of the camera.

II. METHODS

The proposed system for non-contact 3D POG estimation is comprised of an image processing stage for extracting image features, a model fitting stage for computing the corneal centers and optical axes of the eyes and finally a model-based vergence algorithm for computing the 3D POG. A single per-user calibration stage is used to correct the eye models for between-subject differences.

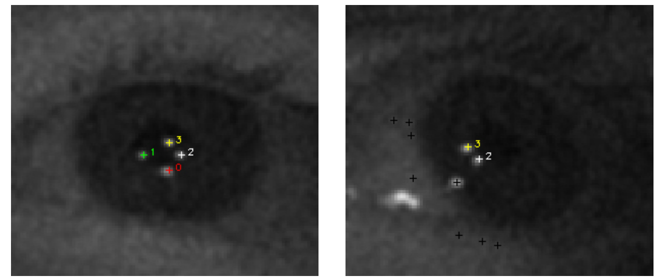
A. Image processing

The model-based POG estimation method requires accurately identified image features of both eyes from the recorded images as described in [22]. To estimate the 3D position of the cornea, the image locations of two corneal reflections are required, along with models of the system, camera and eye. To determine the direction of the optical axis, the image location of the center of the pupil is required, in addition to the previously computed 3D center of the cornea. An outline of the image processing procedure is shown in Fig. 1 in which Fig. 1(a) illustrates the overall binocular tracking and Fig. 1(b) illustrates the image processing steps for each eye. In the event that less than two eyes are detected the system will not be able



(a) Image Processing Procedure (b) 'Search for Eye N' Procedure

Fig. 1. The overall image processing loop is shown in Fig. 1(a). The search for the eyes is performed sequentially and only after both eyes have been detected are they identified as either left or right. When the ROIs are applied the image search space is greatly reduced. In Fig. 1(b) the procedure for identifying the image features required for the next stage of model fitting is presented.



(a) Reference Pattern (b) Pattern Matching

Fig. 2. An example of the four valid corneal reflections used as the reference pattern is shown in Fig. 2(a). With the large eye rotation shown in Fig. 2(b) some of the corneal reflections were distorted or lost, however, two valid corneal reflections remain, which is sufficient for POG estimation.

to estimate the 3D POG from vergence and the processing halts until the next image frame is recorded.

To aid the pupil tracking algorithm, the bright pupil and image differencing techniques are used to create a high contrast image of the pupil [23], [24]. The bright pupil image is taken using a light source located coaxially with the lens of the camera which results in a brightly illuminated pupil due to the retro-reflective property of the retina (the same phenomenon as red-eye in flash photography). The dark pupil image is formed by using off-axis lighting, which illuminates the face equivalently but does not generate a bright pupil. The difference image formed by subtracting the dark pupil image from the bright pupil image results in a high contrast pupil contour which is easily segmented. The roughly identified difference image is then used to identify the pupil contour in the bright pupil image [22]. Once the pupil contour has been identified, an ellipse is fit to the perimeter and the center of the ellipse is used as the center of the pupil [25].

The corneal reflections are found by searching the dark pupil image for high intensity image pixels located in close proximity to the identified pupil. Significant rotation of the eyes with respect to the camera, commonly occurring in 3D POG estimation, can cause the corneal reflections to appear distorted near the boundary between the cornea and scelera, or disappear completely on the rougher surface of the scelera [26]. While the location of only two corneal reflections are required for triangulation of the location of the cornea, in the system described here a set of four off-axis light sources was used to generate four corneal reflections for redundancy. Point pattern matching is used to match a reference pattern of known valid corneal reflections, shown in Fig 2(a), with the remaining visible corneal reflections, shown in Fig 2(b) [27]. The reference pattern is formed based on the relative positions of the off-axis light sources.

To achieve the desired high speed sampling rates needed for digital filtering, the amount of image information to process per system loop is significantly reduced by only processing the ROIs as opposed to the full image as described in [28]. To track the motion of the eyes within the recorded images, the left and right eye ROIs are continuously repositioned onto the left and right pupil image centers respectively. If either eye is lost, due to blinking or eye placement outside the field of view of the camera, the ROIs are resized to the full image. The full image ROIs are processed until each eye is re-acquired, after which the ROIs are reduced to encompass just the eyes, and high speed processing resumes.

B. Model Fitting

The model fitting algorithm uses the extracted image features, along with models of the physical system, camera and eye to estimate the 3D center of the cornea C, pupil P_c and ultimately the optical axis vector joining these two points as shown in Fig. 3. The 3D location of the center of the corneal sphere model is on the line of intersection between two planes. Each plane is coplanar to a ray traced from a corneal reflection image point on the surface of the camera sensor, out through the focal point of the camera lens, reflecting off the spherical surface of the cornea, and back to the originating point light source. Using the known light source positions, as well as the camera and eye models, along with the line of intersection of the two planes provides sufficient constraints for solving for the 3D coordinates of the center of the cornea. To determine the 3D position of the center of the pupil, ray-tracing is again used to trace from the pupil image center on the camera sensor, out through the camera focal point to the surface of the cornea. The pupil center ray is then refracted into the eye and intersected with a vector from the previously identified center of the cornea, at a distance r_d (from the eye-model shown in Fig. 3). The details of the 3D cornea and pupil center estimation technique have been previously described in Hennessey *et al* [22] and are an extension of earlier work by Shih and Liu [11].

The model of the physical system used here is determined through direct measurement of the locations of the camera and infrared (IR) point light sources. The camera lens is modeled as a pin-hole with the intrinsic parameters estimated using the Camera Calibration Toolbox for MATLAB [29]. The model of the eye used here is based on the schematic eye developed by Gullstrand [30] as illustrated in Fig. 3. In the simplified

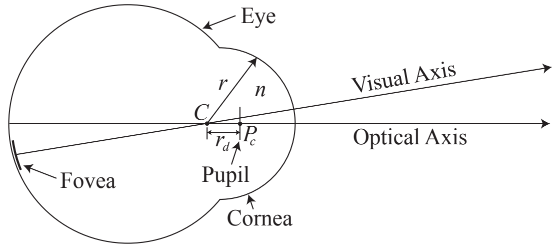


Fig. 3. The schematic eye includes three general parameters; the radius of the corneal sphere r, the distance from the center of the corneal sphere to the center of the pupil r_d and the index of refraction n of the aqueous humor fluid. The model-based method for computing the POG is based on first determining the location of the center of the cornea. With the location of the corneal center it is then possible to compute the optical axis direction. The optical axis vector is corrected through calibration to lie along the visual axis, which is offset from the optical axis due to the displacement of the fovea on the retina.

schematic eye the cornea is approximated as a uniformly spherical surface with three parameters (r, r_d, n). The three parameters vary between subjects, however, to date there has been no known method for estimating them on a per user basis based purely on remote imaging the eyes and consequently population averages are typically used. An error analysis based on the effects of these assumptions are reported in Section III-F where it is clear that system accuracy could benefit from future development of a non-contact method for estimating each of these parameters.

C. Calibration

In the model fitting procedure outlined here, the optical axis can be determined based on the simplified eye model, however, the true visual axis may lie up to 5° from the optical axis depending on the location of the fovea (high resolution portion of the retina) for an individual user [31]. The offset between the optical axis and the visual axis can be compensated with a per-user calibration.

The per user calibration procedure involves having a user observe known points in 3D space while the optical axes of the eyes are computed and the offsets required to intersect the optical axes with the test positions are determined. For 2D POG estimation using the model-based method, the test points are located on the surface of the display [11] [12]. For POG estimation in 3D, the test point can be located anywhere within the workspace volume. While a single calibration point is sufficient to determine the angular offsets, multiple calibration points located throughout the workspace display (or volume for 3D) are typically used.

For each of the N calibration test positions T_i as shown in Fig. 4(a), each optical axis OA_i is normalized and converted to spherical coordinates (1) where ϕ_i and θ_i are readily determined.

$$\left[\widehat{OA_i}\right] = \frac{[OA_i]}{\|[OA_i]\|} = \begin{bmatrix} \sin\phi_i \cos\theta_i \\ \sin\phi_i \sin\theta_i \\ \cos\phi_i \end{bmatrix} \tag{1}$$

The angular offset corrections $\Delta\phi_i$ and $\Delta\theta_i$, between the optical axis and the visual axis, can be determined using the parametric equation of a line (2) with 3 equations and 3 unknowns (t, $\Delta\phi_i$, and $\Delta\theta_i$) which can be solved for explicitly.

$$T_i = C_i + t \cdot \begin{bmatrix} \sin(\phi_i + \Delta\phi_i)\cos(\theta_i + \Delta\theta_i) \\ \sin(\phi_i + \Delta\phi_i)\sin(\theta_i + \Delta\theta_i) \\ \cos(\phi_i + \Delta\phi_i) \end{bmatrix} \tag{2}$$

All subsequent estimated optical axis vectors OA_{curr} are normalized and corrected to the visual axis VA_{curr} using proportional weighting of the calibration parameters. The similarity between the current normalized optical axis vector $\widehat{OA}_{curr}$ and each calibration optical axis vector $\widehat{OA}_i$ as determined by the Euclidean distance (3), is used to generate a list of weighting factors (4) which are then used to weight the angular offsets $\Delta\phi_i$, and $\Delta\theta_i$ determined during calibration.

$$d_i = \left\| \widehat{OA}_{curr} - \widehat{OA}_i \right\| \tag{3}$$

$$w_i = \frac{1}{d_i \cdot \sum\limits_{k=1..N} \frac{1}{d_k}} \tag{4}$$

The normalized optical axis $\widehat{OA}_{curr}$ is converted to spherical coordinates ϕ_{curr} and θ_{curr}, the weighted sum of the corrections applied to the spherical coordinates (5) and (6) and the resulting visual axis determined (7) as shown in Fig. 4(b).

$$\phi'_{curr} = \phi_{curr} + \sum_i w_i \cdot \Delta\phi_i \tag{5}$$

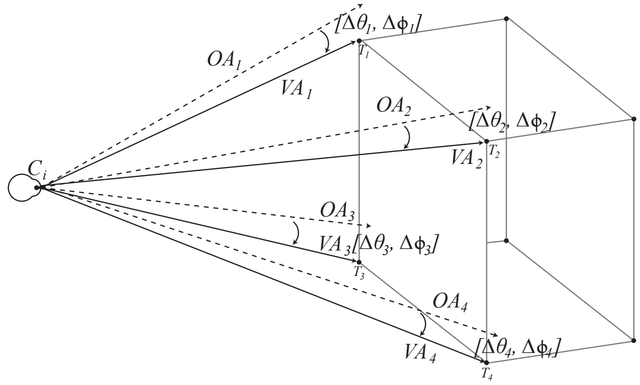
$$\theta'_{curr} = \theta_{curr} + \sum_i w_i \cdot \Delta\theta_i \tag{6}$$

$$VA_{curr} = \begin{bmatrix} \sin(\phi'_{curr})\cos(\theta'_{curr}) \\ \sin(\phi'_{curr})\sin(\theta'_{curr}) \\ \cos(\phi'_{curr}) \end{bmatrix} \tag{7}$$

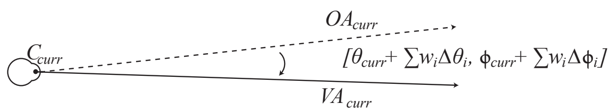
In reality the angular offsets of the eyes do not change depending on gaze direction and a single calibration point should be sufficient. However, as will be shown in the calibration experiment in Section III-D, using multiple calibration positions and the proportional weighting technique proposed here provides an improvement in overall accuracy. This is due to the additional errors introduced from using a simplified eye model and population averages for the eye model parameters as discussed in Section III-F. As the model of the eye is refined and techniques for determining the per-user eye model parameters are developed, it would be expected that the proportional weighting calibration procedure would then simplify to a single point calibration.

D. Model-Based Vergence

With 2D model-based eye-gaze tracking, the visual axis is traced from the center of the cornea into the 3D world and intersected with an object of known geometry. This object is typically the planar surface of a desktop monitor but may be



(a) Four point calibration in 3D space at a single depth plane



(b) Calibration correction of optical axis to visual axis

Fig. 4. The calibration procedure uses calibration test positions located throughout the workspace volume. In Fig. 4(a) the calibration positions T_i are shown at a single depth plane for simplicity of the figure. The calibration corrections $\Delta\phi_i$ and $\Delta\theta_i$ are used to reorient all future optical axis vectors to the visual axis as shown in Fig. 4(b).

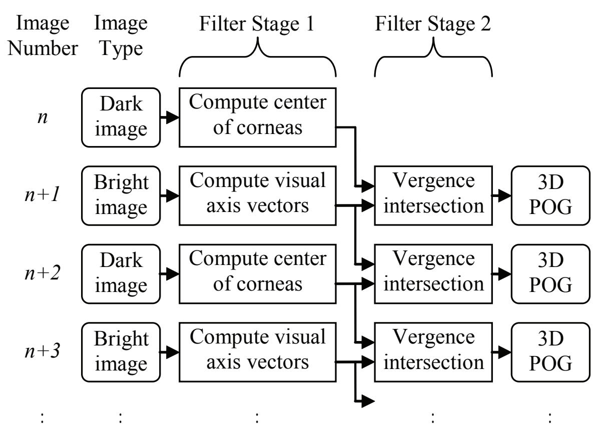


Fig. 5. The alternating bright and dark pupil images are used to generate estimates for the centers of the corneas and the visual axis vectors. The vergence of the eyes can then be used to determine the 3D POG. Note that as a result of the image differencing technique each image frame results in an update for either the corneal centers of the visual axis vectors. The 3D POG is estimated at the full camera frame rate by using the model features from the current image frame, combined with the model features from the previous image frame.

any surface, provided the location and geometry are known *a priori*. To estimate the POG in 3D space without *a priori* knowledge of the surfaces upon which the user is looking, the binocular visual axis vectors are traced from their respective corneal centers and intersected in free space. A flowchart illustrating the 3D POG estimation process is shown in Fig. 5.

The POG in 3D space is actually computed as the midpoint of the shortest distance between the two visual axis vectors, as the vectors are unlikely to exactly intersect as shown in Fig. 6 [32]. The points $P_l(s)$ and $P_r(t)$ can each be defined by a parametric equation of a line (8) and (9).

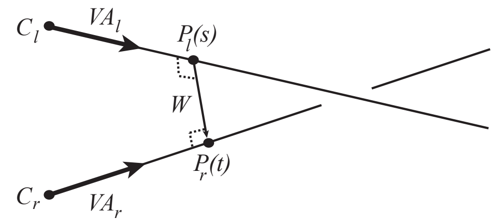


Fig. 6. The POG in 3D space is determined by computing the points $P_l(s)$ and $P_r(t)$ on each visual axis vector which result in the closest distance between the two vectors. The 3D POG is the midpoint of the vector W formed from $P_l(s)$ to $P_r(t)$, where C_l and C_r are the locations of the left and right corneal centers and VA_l and VA_r are the left and right eye visual axes respectively.

$$P_l(s) = C_l + s \cdot VA_l \tag{8}$$
$$P_r(t) = C_r + t \cdot VA_r \tag{9}$$

To minimize the distance joining the points $P_l(s)$ and $P_r(t)$, the vector W is defined from $P_l(s)$ to $P_r(t)$ and perpendicular to both VA_l and VA_r. Since W is perpendicular to both of the visual axis vectors, a system of two equations, (10) and (11), with two unknowns (the parameters s and t) can be defined and are readily determined.

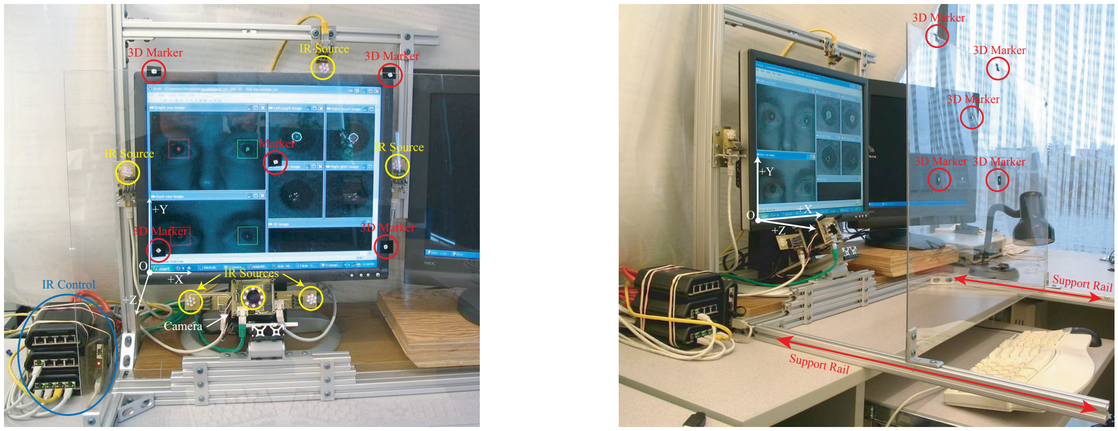
$$VA_l \cdot [P_r(t) - P_l(s)] = 0 \tag{10}$$
$$VA_r \cdot [P_r(t) - P_l(s)] = 0 \tag{11}$$

Using model-based vergence to estimate the 3D POG is only valid provided the visual axis vectors of the eyes are not parallel, *i.e.* a unique solution to (10) and (11) can be found. As the distance to the point under observation increases, the visual axis vectors of the eyes increasingly approach a parallel course. Given a constant visual axis estimation accuracy, (typically 0.5 to 1.0 degree of visual angle) this means that the spatial accuracy of the estimated 3D POG will decrease with increasing depth from the eyes.

E. Fixation filtering

The eyes are continuously in motion to keep the sensors of the eye refreshed during fixations [33]. The small motions of the eyes result in jittery visual axis vectors and can ultimately lead to poor precision in the estimated POG. With the model-based vergence technique for 3D POG estimation, increased error in the estimated visual axis vectors due to jitter can result in a much larger error in depth of the estimated 3D POG.

In the system presented here two levels of concurrent digital filtering were used to improve the precision of the 3D POG estimates as shown in Fig. 5. The first stage of lowpass filters (moving window averages) were used to stabilize the computed model features (corneal centers and visual axis vectors) while the second stage of filtering was used to stabilize the estimated 3D POG. The length of the filters can be used to tradeoff between precision and response time.



(a) Front view of experimental setup

(b) Side view of experimental setup

Fig. 7. Front and side views of the experimental setup are shown. In the front view the microcontroller and IR point light source expansion ports are located to the lower left of the screen. The off-axis IR point light sources are located around the frame and the on-axis IR ring is located in front of the camera lens. The 3D markers are located in an X grid of points on a clear Plexiglas sheet. The markers are a small cross on white paper, backed with black electrical tape for increased contrast for the subjects. In the side view the support rails are shown upon which the Plexiglas sheet can be translated in depth.

III. Experimental Design and Results

To evaluate the performance of the proposed design, the algorithms described were implemented and a set of experiments performed at the subsystem and system levels.

A. System Configuration

The system was comprised of multiple IR light sources, a high speed digital camera and a set of 3D POG markers as shown in Fig. 7. In addition to the ring of on-axis LEDs for generating the bright pupil, each off-axis light source was composed of a set of seven closely spaced LED lights to approximate a point light source. While five off-axis light sources are shown around the computer screen, only four were used in the evaluation of the system presented here. The light source located immediately to the left of the camera was left off. The selection and placement of the four light sources used were experimentally chosen such that at least two valid reflections were formed off of the surface of the cornea at all eye rotations encountered in the evaluation of the system. A microcontroller was used to synchronize the camera shutter with the alternating on-axis and off-axis LEDs. The digital camera used was a monochrome DragonFly Express from Point Grey Research, capable of recording images with a resolution of 640 x 480 pixels at a frame rate of 200 Hz. The processor of the computer used for the system was an Intel 2.66 GHz Core 2 processor with 2 GB of RAM. A C++ implementation of the 3D POG estimation algorithms achieved 200 Hz 3D POG estimation using the process flow shown in Fig. 5. Analysis of recorded data as presented in this paper was performed offline in the MATLAB environment.

The 3D test point markers were placed in an X shape on a Plexiglas sheet which was mounted on aluminum rails. The corners of the X were spaced 30 cm apart horizontally and 23 cm vertically. The rails were marked at 5 cm intervals at 6 different depths, resulting in a total workspace volume of 30 x 23 x 25 cm (width x height x depth). The total workspace volume exercised is comparable in size to modern volumetric displays [5]. An extruded aluminum structure was used to maintain the geometric positions between the camera, IR light sources and 3D position markers. The world coordinate system origin was located at an arbitrary position in 3D space. For convenience in development, it was located at the lower left corner of the monitor, with the positive X axis towards the right, the positive Y axis towards the ceiling and the positive Z axis towards the user.

B. Evaluation of filter length

The model features used to estimate the 3D POG suffer from jitter due to the natural motions of the eyes. The jittery model features can then lead to poor precision of the estimated 3D POG. To reduce the jitter and therefore increase the precision of the 3D POG, lowpass filters (moving window averages) with a user definable filter length were applied to the model features (corneal centers and visual axis vectors), as well as the final estimated 3D POG.

The accuracy and precision of the 3D POG was determined over a range of filter lengths to evaluate the effect of filtering. The experimental procedure involved a single subject, who was asked to fixate on a 3D test point located in the middle of the workspace volume while the raw image data used to compute the 3D POG was recorded.

Results

The recorded image data were then processed offline to compute the 3D POG using a variety of filter lengths. Shown in Table I are the average absolute errors, in addition to the standard deviations, over a consistent one second (200 samples) of data during the fixation. The 3D POG is listed by each coordinate (X, Y, Z) as well as the Euclidean distance error ($\sqrt{X^2+Y^2+Z^2}$). The maximum latency was determined as the time required for both filter histories to fill entirely with new fixation data. For example, at a sampling rate of 200 Hz, the 100 sample 3D POG filter requires 0.5 seconds, added to the 1 second for the 200 sample filter length for model features used in estimating the 3D POG, for a total latency of 1.5 seconds. For all further testing a filter length of 200 samples was used for the model features, and a filter length of 100 samples for the 3D POG as these produced the best results.

TABLE I
ACCURACY AND STANDARD DEVIATION OVER VARYING FILTER LENGTHS.

Model Length	3D POG Length	Latency (s)	Average Accuracy (cm) X	Y	Z	Euc.	Standard Dev. (cm) X	Y	Z	Euc.
1	1	0.005	0.34	0.43	3.30	3.41	0.26	0.30	2.57	2.50
20	10	0.15	0.17	0.43	1.65	1.79	0.09	0.14	1.29	1.19
100	50	0.75	0.12	0.40	1.07	1.20	0.03	0.05	0.61	0.53
200	100	1.5	0.15	0.43	0.44	0.70	0.02	0.01	0.40	0.27

C. Head Motion

Allowing the head to move naturally is a key goal of the proposed 3D POG estimation system. The ability to handle head motion is particularly important in 3D POG estimation as the head naturally moves and rotates while observing points in 3D space to reduce the strain on the extraocular muscles [34]. In this experiment, the allowable head space is such that both eyes remain in focus within the field of view of the camera. The experimental procedure involved a single subject, asked to observe a 3D test point located in the middle of the workspace volume. While observing the test point, the subject was asked to randomly position and rotate his/her head while exercising the full head space. A total of 24 different random locations and orientations were recorded. The first of the 24 positions was used as the calibration position. At each head position the estimated 3D POG was recorded, along with the positions of the left and right eyes (corneal centers) in 3D space.

Results

Accuracy was measured as the Euclidean distance between the estimated 3D POG and the actual 3D test point. The average error over the 23 head positions was found to be 1.96 cm with a standard deviation of 1.63 cm. From the calculated positions of the eyes the exercised head space spanned 3.2 cm horizontally, 9.2 cm vertically and 14 cm in depth.

D. Calibration Points

In the previous filter length and head motion experiments the subject observed a single test point which was calibrated at the same position. When extending the system to operate over the full workspace volume (30 x 23 x 25 cm), any number of 3D positions may be used as calibration points. While a single point is sufficient to calibrate the system, the system accuracy may be increased by ensuring the 3D POG estimation algorithm is calibrated over the entire workspace volume.

The calibration experiment procedure involved a single subject, who was asked to observe each of the 30 3D test points located throughout the workspace volume. The computed corneal center and uncalibrated optical axis vectors were recorded at each test position for offline processing. The data collection procedure was repeated twice more to generate a total of three datasets. The first data set was post processed using various combinations of calibration positions to determine the optical axis angular offsets, which were then applied to the second and third datasets and the average 3D POG accuracy computed.

The calibration positions tested used 1, 5, 10, and 30 points. The single point calibration used the same mid-volume position as in the previous filter length and head motion experiments. The 5 point calibration used the 5 test positions located on the mid-volume plane. The 10 point calibration used the 5 points located on the first and last depth planes respectively. Finally, the 30 point calibration used all the data points from the complete workspace volume.

Results

The resulting average 3D POG accuracy when each calibration set was applied to the second and third datasets are shown in Table II. An analysis of variance was performed to check for statistically significant differences in average accuracy between the calibration methods. Combining the second and third trials, a statistically significant difference was found between the techniques ($F(3,236)=7.273$, $p<0.001$). *Post hoc* analysis indicated that the average accuracy of the 1 and 5 point calibrations were worse than the 10 and 30 point calibrations, while there was no statistically significant difference between the 1 and 5 point calibrations or between the 10 and 30 point calibrations. The 10 point calibration procedure was therefore chosen for subsequent experiments as it maximized accuracy while minimizing the time required for calibration.

TABLE II
AVERAGE ACCURACY OF 3D POG ESTIMATES FOR VARIOUS CALIBRATION POSITIONS.

Calibration Points	**Dataset Number**	**Average Accuracy (cm)**	**Standard Deviation (cm)**
1 Point	2	5.47	4.04
1 Point	3	5.24	2.75
5 Point	2	4.84	4.58
5 Point	3	5.00	3.61
10 Point	2	3.19	2.83
10 Point	3	3.13	2.13
30 Point	2	3.22	2.76
30 Point	3	3.43	2.18

E. Multi-Subject Evaluation

An evaluation of the accuracy of the system was performed across a range of subjects to provide a more general indication of system performance. The experiment was evaluated over a total of 7 different subjects and exercised the full workspace (30 x 23 x 25 cm) for 3D POG estimation. The subjects were allowed freedom of head motion provided both eyes remained visible to the system camera. The subjects were all graduate students in the Electrical and Computer Engineering Department at the University of British Columbia (UBC). The subject ages ranged from 22 to 30 years old. Of the seven subjects 2 were female with 1 of 7 wearing contact lenses. The ethnicities of the subjects were 5 Caucasian and 2 Middle Eastern. The experimental procedures were certified for human experimentation by the Behavioral Research and Ethics Board of UBC under certificate H04-80920.

Each test subject was asked to observe each of the 5 points on the Plexiglas plane at the near and far depth planes to complete the 10 point calibration described in Section III-D. The calibration corrections for each subject were then used to determine the subsequent 3D POG estimates. The data collection procedure required each subject to observe each of the 5 test positions on the Plexiglas sheet while the 3D POG was recorded, then move the sheet forward 5 cm, and repeated the 5 test positions until the entire workspace was exercised. The entire workspace volume was exercised twice to generate two trials per subject.

Results

The accuracy at each depth plane of the workspace volume, averaged over the two trials for all subjects is shown in Table III, as well as the standard deviation. The accuracy reported is the average absolute error for the X, Y, and Z coordinates as well as the Euclidean distance error ($\sqrt{X^2+Y^2+Z^2}$). The depth of the planes are measured in centimeters from Z = 0 at the surface of the computer screen. The overall average accuracy and standard deviation for the entire workspace volume is also shown.

TABLE III
AVERAGE ACCURACY OF 3D POG ESTIMATION OVER THE WORKSPACE FOR ALL SUBJECTS.

Z Depth (cm)	**Average Accuracy (cm)**				**Standard**
	X	Y	Z	Euc.	**Deviation (cm)**
17.5	1.28	1.20	4.04	4.61	3.14
22.5	1.28	1.27	3.64	4.28	2.81
27.5	1.26	1.13	3.36	3.98	3.11
32.5	1.10	1.04	3.20	3.75	2.96
37.5	1.31	1.13	2.55	3.35	2.59
42.5	1.38	1.46	2.60	3.62	2.14
Overall	1.27	1.20	3.23	3.93	2.83

F. Sensitivity Analysis

The potential sources of error in the system include: 1) extracted image features errors due to limited contrast and spatial resolution of the camera, 2) the simplified model of the eye with population averages for the eye model parameters, 3) errors in the camera lens calibration, and 4) errors in the physical measurement of the system. To provide an indication of the most significant sources of error, an analysis was performed of the sensitivity of the overall average accuracy with respect to both noise in the extracted image features and variations in system parameter values.

For this experiment the pupil and corneal reflection image centers were recorded rather than the computed 3D POG. The 3D POG at each data point was then recomputed offline using the raw image data, allowing evaluation of system parameter variation on a consistent data set. A single subject was asked to perform the 10 point calibration procedure as described previously. The subject then observed each of the 30 workspace points while the image data were recorded.

Results

Random Gaussian noise with zero mean and a fixed standard deviation (SD) was added to both the X and the Y coordinates of the extracted pupil center, the 3D POG computed, and the overall system accuracy was determined. The standard deviation of the noise was then increased and the process repeated. The procedure for the addition of noise was then repeated with the random noise added to both the X and the Y coordinates of the corneal reflections. The results of the experiment are summarized in Table IV.

TABLE IV
EFFECT OF NOISE IN IMAGE FEATURE EXTRACTION ON SYSTEM ACCURACY.

Noise SD in X & Y (pixels)	0	1	2	4
	Average Accuracy (cm)			
Pupil Center	3.78	3.92	4.16	5.59
Corneal Reflection Center	3.78	4.45	7.56	24.55

To evaluate the effect of model parameter deviations, the three eye model parameters (radius of cornea r, distance from center of cornea to center of pupil r_d, and index of refraction of the aqueous humor n) and the pinhole camera parameters (focal point f and critical point c_{px} and c_{py}) obtained through camera calibration were independently varied up to ± 10 % and the average accuracy was determined as listed in Table V. The spatial coordinates of the off-axis light sources (Q) were also independently varied by up to ± 2 cm. Note that the accuracy results for the light source locations of the off-axis lights were averaged over the four lights for the X, Y, and Z coordinate variations.

TABLE V
SENSITIVITY OF AVERAGE SYSTEM ACCURACY TO PARAMETER VARIATIONS.

Variation	-10%	-5%	0%	+5%	+10%
Eye Model		Average Accuracy (cm)			
r	5.31	4.41	3.78	3.93	4.66
r_d	5.09	3.71	3.78	5.16	6.92
n	3.77	3.53	3.78	4.45	5.17
Camera Model		Average Accuracy (cm)			
f	4.34	3.56	3.78	5.10	7.20
c_{px}	4.10	3.95	3.78	3.65	3.53
c_{py}	3.92	3.85	3.78	3.75	3.72

Variation	-2 cm	-1 cm	0 cm	+1 cm	+2 cm
Light Location		Average Accuracy (cm)			
Q (X)	4.12	3.82	3.78	3.90	4.21
Q (Y)	3.95	3.81	3.78	3.84	3.96
Q (Z)	3.84	3.81	3.78	3.79	3.80

IV. DISCUSSION

With rapid and robust image processing, a high speed sampling rate was achieved. Digital filtering was employed to improve precision at the expense of increased latency. In this paper, filter lengths of 200 samples for the model features and 100 samples for the 3D POG were used. The filter lengths selected reduced the estimated POG jitter to 0.27 cm with a corresponding maximum latency of 1.5 seconds. To improve the latency of the system, fixation detection techniques may be employed to ensure that data from separate fixations are not combined in the digital filter histories, ensuring a rapid response to new fixations [31] [35].

The ability to handle head motion during 3D POG estimation is important as the head naturally reorients to reduce eye strain when observing points that require significant eye rotation. The ability to accurately estimate the 3D POG in the presence of unconstrained head motion was evaluated and an average accuracy of 1.96 cm was found over 23 different head positions and orientations. The full range of head positions spanned a head space volume of 3.2 x 9.2 x 14 cm (width x height x depth). Given the resolution of the camera sensor only a small degree of horizontal motion was possible as both eyes had to remain within the field of view of the camera. To improve the range of allowable head motion a camera with a higher resolution imaging sensor could be used to increase the field of view by decreasing the camera lens focal length without changing the effective spatial resolution.

The calibration algorithm outlined in this paper only requires a single stage for per-user calibration. Calibration is performed by having the subject observe known positions in real world 3D space while the optical-to-visual axis offsets are determined. Statistical analysis indicated that using calibration points at only a single depth (1 and 5 points) resulted in worse accuracy than using calibration points located at different depths throughout the workspace volume (10 and 30 points). Calibration with 10 point (5 on the furthest and 5 on the closest depth planes) proved the most accurate with the shortest calibration duration.

A multi-subject experiment was performed to generalize the operation of the system over a larger population sample. The subjects were allowed to move their heads naturally while observing 3D points, provided both eyes remained within the field of view of the camera. The accuracy, averaged over all subjects, improved as expected as the distance from the eye to the 3D POG was reduced. An average accuracy of 4.61 cm at Z = 17.5 cm reduced to 3.35 cm at Z = 37.5 cm. Interestingly the error increased to 3.62 cm at Z = 42.5 cm (the plane located closest to the eyes). At the closest depth plane, the 3D test points located at the corners of the plane resulted in the most extreme eye rotations of the workspace. The increase in average 3D POG error at the nearest depth plane to the eyes is a result of the distortion of the corneal reflections when the eye is rotated to significant angles with respect to the camera. Over the entire workspace volume of 30 x 23 x 25 cm (width x height x depth) an average accuracy of 3.93 cm was determined. Given the accuracy, precision and latency achieved with the system presented here, a demonstration application was developed utilizing real-time 3D POG estimation to play a 3D game of Tic-Tac-Toe on a volumetric display in Hennessey and Lawrence [36].

To evaluate robustness and help direct further research, the sources of error leading to the average accuracy achieved were investigated by determining the effect of image feature noise and system model parameter variations. The addition of noise to the extracted corneal reflection locations considerably increased the error when compared with noise added to the pupil center as shown in Table IV. To reduce the effect of error in the corneal reflections, redundant off-axis light sources were used to avoid, as much as possible, the distortion that occurs when reflections approach the boundary between the cornea and scelera. Improved eye models which account for the change in curvature of the cornea may also be investigated as a means for further improvement.

Variation of the system parameters shown in Table V indicated that average accuracy was most sensitive to the eye model and the camera lens focal length parameters. Improvement of the eye model, either through increased sophistication (*i.e.* more accurately modeling the surface of the cornea) or more accurately identifying eye parameters (rather than using population averages) may lead to improved system accuracy. For remote eye model parameter estimation, the radius of the cornea and index of refraction may potentially be determined based on externally visible reflections and refraction respectively. As the distance from the center of the cornea to the center of the pupil occurs within the eye we expect this parameter to be fairly difficult to estimate from external images. One key advantage of using model-based methods for POG estimation over the P-CR or neural network based methods is that as the models of the eye improve, the accuracy of the model-based methods for both 2D and 3D POG estimation should improve as well. The desire for a higher resolution camera previously mentioned may also improve the performance of the camera calibration. Decreasing the focal length of the camera lens to increasing the field of view will also increase the perspective projection of the camera, becoming less orthographic and increasing the needed depth information in the camera calibration images [37].

V. Conclusions

In this paper techniques for a novel non-contact, head-free eye-gaze tracking system have been developed and quantitatively evaluated for 3D POG estimation in a real world scene. The 3D POG was estimated in a real world workspace volume of 30 x 23 x 25 cm and an average accuracy of 3.93 cm was achieved over seven subjects. The completely non-contact and head-free system had an allowable head space of 3 x 9 x 14 cm with the only requirement that both eyes be visible within the field of view of the camera. Through the two stages of high speed filtering the standard deviation of the unfiltered 3D POG was lowered from 2.5 cm to 0.27 cm with a corresponding maximum latency of 1.5 seconds. Reducing the maximum latency through fixation detection remains to be investigated. The use of a model-based approach for binocular eye-gaze tracking and a model-based vergence method of visual axis vector intersection allowed for a single stage of calibration. Future work will involve integration of a higher resolution camera for improving the range of free head motion, as well as researching improved models of the eye.

Acknowledgment

The authors would like to express their appreciation for the support of the Natural Sciences and Engineering Research Council of Canada (NSERC) Chair in Design Engineering, and NSERC Discovery Grant #4924.

References

[1] E. Kowler, *Eye Movements and their Role in Visual and Cognitive Processes.* Elsevier Science, 1990, vol. 4, ch. The role of visual and cognitive processes in the control of eye movement., pp. 1–70.

[2] R. Jacob and K. Karn, *The Mind's Eye: Cognitive and Applied Aspects of Eye Movement Research*. Amsterdam: Elsevier Science, 2003, ch. Eye Tracking in Human-Computer Interaction and Usability Research: Ready to Deliver the Promises (Section Commentary), pp. 573–605.

[3] M. Halle, “Autostereoscopic displays and computer graphics,” *SIGGRAPH Comput. Graph.*, vol. 31, no. 2, pp. 58–62, 1997.

[4] N. Dodgson, “Autostereoscopic 3d displays,” *Computer*, vol. 38, no. 8, pp. 31 – 36, Aug. 2005.

[5] A. Jones, I. McDowall, H. Yamada, M. Bolas, and P. Debevec, “Rendering for an interactive 360° light field display,” in *ACM SIGGRAPH*. New York, NY, USA: ACM, 2007, p. 40.

[6] K. Meyer, H. L. Applewhite, and F. A. Biocca, “A survey of position trackers,” *Presence: Teleoper. Virtual Environ.*, vol. 1, no. 2, pp. 173–200, 1992.

[7] C. Ware, “Using hand position for virtual object placement,” *Vis. Comput.*, vol. 6, no. 5, pp. 245–253, 1990.

[8] C. H. Morimoto and M. R. M. Mimica, “Eye gaze tracking techniques for interactive applications,” *Comput. Vis. Image Underst.*, vol. 98, no. 1, pp. 4–24, 2005.

[9] D. Panchuk and J. Vickers, “Gaze behaviors of goaltenders under spatial-temporal constraints,” *Human Movement Science*, vol. 25, no. 6, pp. 733–752, Dec. 2006.

[10] T. Hutchinson, J. White, W. Martin, K. Reichert, and L. Frey, “Human-computer interaction using eye-gaze input,” *IEEE Transactions on Systems, Man and Cybernetics*, vol. 19, no. 6, pp. 1527–1534, 1989.

[11] S.-W. Shih and J. Liu, “A novel approach to 3-d gaze tracking using stereo cameras,” *IEEE Transactions on Systems, Man and Cybernetics, Part B*, vol. 34, no. 1, pp. 234–245, Feb. 2004.

[12] D. Beymer and M. Flickner, “Eye gaze tracking using an active stereo head,” in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 2, 18-20 June 2003, pp. II–451–II–458.

[13] E. Guestrin and M. Eizenman, “General theory of remote gaze estimation using the pupil center and corneal reflections,” *Biomedical Engineering, IEEE Transactions on*, vol. 53, no. 6, pp. 1124–1133, June 2006.

[14] W. J. Ryan, A. T. Duchowski, and S. T. Birchfield, “Limbus/pupil switching for wearable eye tracking under variable lighting conditions,” in *Proceedings of the 2008 symposium on Eye tracking research & applications*. New York, NY, USA: ACM, 2008, pp. 61–64.

[15] A. T. Duchowski, V. Shivashankaraiah, T. Rawls, A. K. Gramopadhye, B. J. Melloy, and B. Kanki, “Binocular eye tracking in virtual reality for inspection training,” in *Proceedings of the 2000 symposium on Eye tracking research & applications*. New York, NY, USA: ACM Press, 2000, pp. 89–96.

[16] B. K. Horn, *Robot Vision*. McGraw-Hill Higher Education, 1986.

[17] K. Essig, M. Pomplun, and H. Ritter, “Application of a novel neural approach to 3d gaze tracking: Vergence eye-movements in autostereograms,” in *Proceedings of the 26thl Meeting of the Cognitive Science Society*, K. Forbus, D. Gentner, and T. Regier, Eds., 2004, pp. 357–362.

[18] ——, “A neural network for 3d gaze recording with binocular eye-trackers,” *International Journal of Parallel, Emergent and Distributed Systems*, vol. 21, no. 2, pp. 79–95, April 2006.

[19] S. M. Munn and J. B. Pelz, “3d point-of-regard, position and head orientation from a portable monocular video-based eye tracker,” in *Proceedings of the 2008 symposium on Eye tracking research & applications*. New York, NY, USA: ACM, 2008, pp. 181–188.

[20] Y.-M. Kwon and K.-W. Jeon, “Gaze computer interaction on stereo display,” in *Proceedings of the 2006 ACM SIGCHI international conference on Advances in computer entertainment technology*. New York, NY, USA: ACM Press, 2006, p. 99.

[21] I. Mitsugami, N. Ukita, and M. Kidode, “Estimation of 3d gazed position using view lines,” in *12th International Conference on Image Analysis and Processing*, 17-19 Sept. 2003, pp. 466–471.

[22] C. Hennessey, B. Noureddin, and P. Lawrence, “A single camera eye-gaze tracking system with free head motion,” in *Proceedings of the 2006 symposium on Eye tracking research & applications*. New York, NY, USA: ACM Press, 2006, pp. 87–94.

[23] Y. Ebisawa and S. Satoh, “Effectiveness of pupil area detection technique using two light sources and image difference method,” in *Proceedings of the 15th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, Oct 28-31, 1993, pp. 1268–1269.

[24] C. H. Morimoto, D. Koons, A. Amir, and M. Flickner, “Pupil detection and tracking using multiple light sources.” *Image and Vision Computing*, vol. 18, no. 4, pp. 331–335, 2000.

[25] A. Fitzgibbon, M. Pilu, and R. Fisher, “Direct least square fitting of ellipses,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 21, no. 5, pp. 476–480, May 1999.

[26] H. Hua, C. W. Pansing, and J. P. Rolland, “Modeling of an eye-imaging system for optimizing illumination schemes in an eye-tracked head-mounted display,” *Appl. Opt.*, vol. 46, no. 31, pp. 7757–7770, 2007.

[27] G. Cox and G. de Jager, “A survey of point pattern matching techniques and a new approach to point pattern recognition,” in *Proceedings of the 1992 South African Symposium on Communications and Signal Processing*, 11 Sept. 1992, pp. 243–248.

[28] C. Hennessey, B. Noureddin, and P. Lawrence, “Fixation precision in high-speed noncontact eye-gaze tracking,” *IEEE Transactions on Systems, Man and Cybernetics, Part B*, vol. 38, no. 2, pp. 289–298, April 2008.

[29] J.-Y. Bouguet, “Camera calibration toolbox for matlab,” *www.vision.caltech.edu/bouguetj/*.

[30] D. A. Goss and R. W. West, *Introduction to the Optics of the Eye*. Butterworth Heinemann, 2001.

[31] R. Jacob, *Virtual Environments and Advanced Interface Design*. New York, NY, USA: Oxford University Press, 1995, ch. Eye tracking in advanced interface design, pp. 258–288.

[32] D. H. Eberly, *3D Game Engine Design*. Academic Press, 2001.

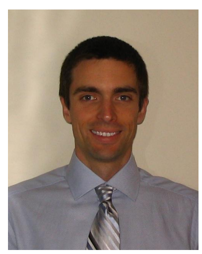
[33] R. J. K. Jacob, *Eye Movement-Based Human-Computer Interaction Techniques: Toward Non-Command Interfaces*. Norwood, N.J.: Ablex Publishing Co., 1993, vol. 4, pp. 151–190.

[34] R. S. Laramee and C. Ware, “Rivalry and interference with a head-mounted display,” *ACM Transactions on Computer Human Interaction*, vol. 9, no. 3, pp. 238–251, 2002.

[35] A. T. Duchowski, *Eye Tracking Methodology: Theory and Practice*. Springer-Verlag, 2003.

[36] C. Hennessey and P. Lawrence, “3d point-of-gaze estimation on a volumetric display,” in *Proceedings of the 2008 symposium on Eye tracking research & applications*. New York, NY, USA: ACM, 2008, pp. 59–59.

[37] X. Huang, J. Gao, and R. Yang, *Computer Vision*, ser. Lecture Notes in Computer Science. Springer Berlin / Heidelberg, 2007, vol. 4843, ch. Calibrating Pan-Tilt Cameras with Telephoto Lenses, pp. 127–137.



Craig Hennessey received the B.A.Sc in systems engineering from Simon Fraser University in 2001 and the M.A.Sc. in electrical engineering from the University of British Columbia in 2005. He is currently working towards his Ph.D. in electrical engineering, also at the University of British Columbia. His research interests include computer vision, image processing and human-computer interaction.

Peter Lawrence Peter D. Lawrence received the B.A.Sc. Degree in Electrical Engineering from the University of Toronto, Toronto, ON, Canada, in 1965, the Masters’s degree in Biomedical Engineering from the University of Saskatchewan, Saskatoon, SK, Canada, in 1967 and the Ph.D. degree in Computing and Information Science at Case Western Reserve University, Cleveland, OH, in 1970. He is also registered as a professional engineer.

He worked as Guest Researcher at Chalmers University’s Applied Electronics Department, Goteborg, Sweden between 1970 and 1972, and between 1972 and 1974 as a Research Staff Member and Lecturer in the Mechanical Engineering Department, Massachusetts Institute of Technology, Cambridge.

Since 1974, he has been with the University of British Columbia and is currently a Professor in the Department of Electrical and Computer Engineering. His main research interests include the application of real-time computing in the control interface between humans and machines, image processing, and mobile hydraulic machine modeling and control.