Behavior Research Methods, Instruments, & Computers
2004, 36 (4), 757-770

Tracking the line of primary gaze in a walking simulator: Modeling and calibration

JAMES BARABAS, ROBERT B. GOLDSTEIN, HENRY APFELBAUM,
RUSSELL L. WOODS, ROBERT G. GIORGI, and ELI PELI
Schepens Eye Research Institute, Harvard Medical School, Boston, Massachusetts

This article describes a system for tracking the line of primary gaze (LoPG) of participants as they view a large projection screen. Using a magnetic head tracker and a tracking algorithm, we find the on-screen location at which a participant is pointing a head-mounted crosshair. The algorithm presented for tracking the LoPG uses a polynomial function to correct for distortion in magnetic tracker readings, a geometric model for computing LoPG from corrected tracker measurements, and a method for finding the intersection of the LoPG with the screen. Calibration techniques for the above methods are presented. The results of two experiments validating the algorithm and calibration methods are also reported. Experiments showed an improvement in accuracy of LoPG tracking provided by each of the two presented calibration steps, yielding errors in LoPG measurements of less than 2° over a wide range of head positions. Source code for the described algorithms can be downloaded from the Psychonomic Society Web archive, http://www.psychonomic.org/archive/.

Commercial eye tracker manufacturers, such as SR Research (Osgoode, ON, Canada), ISCAN (Burlington, MA), and Applied Science Laboratories (ASL; Bedford, MA), market systems for monitoring point of regard on a display surface (combining head and eye tracking), but the manufacturers of these systems provide the devices as black-box tools. This makes it difficult for researchers to modify these trackers for special constraints, such as the large display and wide range of head movements needed for our walking simulator (Figure 1). Our simulator takes the form of a projected virtual environment (Southard, 1995) in which a participant views computer-generated images displayed on a large screen. Here, we describe a technique for addressing an important subproblem of gaze tracking—tracking the line of primary gaze (LoPG, as will be defined below). This method allows participants to use head movement to "point" at locations on our simulator screen. The techniques presented here could be extended with the addition of a head-mounted eye tracker to allow for head-movement–compensated gaze tracking (Barabas et al., 2003).

There exist a number of techniques for tracking eye movements (Young & Sheena, 1975), but relatively little has been written on calibration methods for these systems. Allison, Eizenman, and Cheung (1996) developed a combined head and eye-tracking apparatus for testing of the vestibular system, but they attributed errors to a lack of calibration. Another combined head and eye-tracking system was built for use in a projected virtual environment (Asai, Osawa, Takahashi, & Sugimoto, 2000), but few details of calibration methodology were provided. Gaze-tracking systems that combine head and eye movements have also been developed for head-mounted displays. Duchowski et al. (2002) described a virtual reality system based on a head-mounted display that incorporates a gaze-tracking system. Detailed descriptions of geometry and calibration techniques were provided, but the authors addressed a sufficiently different problem that their methods cannot be directly applied to projected virtual environments.

Toward developing a gaze-tracking system for our walking simulator, we separated the task into two parts: tracking the movement of the participant's head and tracking the rotation of the eye in its socket. In this article, we will explore the first of these two parts.

Techniques for tracking the head are useful for tasks in which a participant "points" at objects by moving his or her head. Such systems are used in tactical aircraft: With the aid of a crosshair projected on a helmet-mounted display, pilots can aim weaponry by turning their heads, aligning the crosshair with a target (King, 1995). We developed a similar system to allow participants in our walking simulator to "point" at objects on a screen. By equipping our participants with a head-mounted sight (see Figure 2) and using a magnetic head tracker and a tracking algorithm (described in the Method section), we are able to find the point on the simulator's screen that is aligned with a mark at the center of the sight. For experiments, we align this sighting mark to appear "straight ahead" of the participant's left eye. This is the primary position of gaze (Leigh & Zee, 1983). When the participant foveates a target on the projection screen through

This work was supported in part by National Institutes of Health Grant EY12890. Correspondence concerning this article should be addressed to E. Peli, Schepens Eye Research Institute, 20 Staniford St., Boston, MA 02114 (e-mail: eli@vision.eri.harvard.edu).

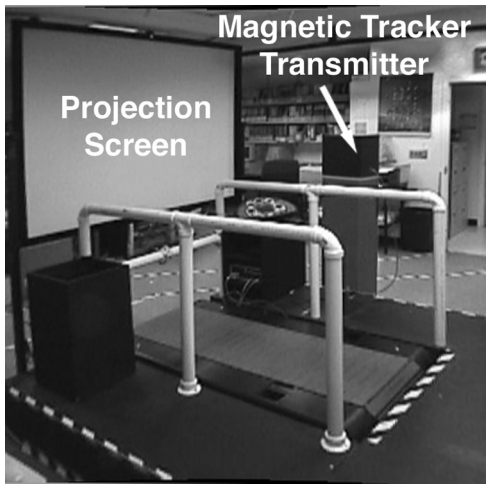

Figure 1. Experimental configuration of the walking simulator showing the locations of the walking platform, the wide projection screen, and the magnetic tracker transmitter.

the sighting mark, this locates the line of sight. The center of rotation of the eye lies approximately along the line of sight, so that the fixation axis (line between the center of rotation and the fixated target) is a close approximation of the line of sight (Atchison & Smith, 2000). When the sighting mark is placed so that the participant's eye is in the primary position of gaze while aligning the sight and the target, this is the LoPG. Thus, in our system, the LoPG connects the center of rotation of the participant's left eye and a point of regard.

Several calibration challenges must be overcome for accurate tracking of the LoPG. Our walking simulator uses a magnetic tracker (Blood, 1990) to measure location and orientation of the participant's head. This type of device, also used in biomechanics (Day, Dumas, & Murdoch, 1998), virtual reality (Livingston & State, 1997) and eye-tracking research (Caldwell, Berbaum, & Borah, 2000), is subject to environmental magnetic distortions that worsen with the distance between the magnetic transmitter and the magnetic sensor. Compensation for this distortion is required if measurements are to be taken from near the limits of the tracker's range or if the tracker is to be used in close proximity to large metallic objects (Nixon, McCallum, Fright, & Price, 1998).

Several techniques for compensating for this kind of distortion have been proposed, and a summary of these methods has been compiled by Kindratenko (2000). Most of the techniques reviewed by Kindratenko were developed for virtual reality (allowing for presentation of motion parallax effects as viewers move their heads) but are applicable to tracking the LoPG. These distortion compensation techniques generally require custom equipment and are not included in popular commercial tracking systems. For example, in a study in which the eye movements of radiologists were investigated, Caldwell et al. (2000) used an ASL eye-tracking system, but they employed a modified version of the ASL EYEPOS software and a custom calibration fixture to correct for magnetic distortion. This method allowed for more accurate gaze tracking near a metallic device, but few details of the software modifications were published.

In addition to compensating for magnetic distortion, one must also find the alignment between the physical components of the tracking system. The distances between these components and their relative orientations become parameters in the LoPG-tracking model. Although direct measurement of these parameters is possible, small errors in these measurements can lead to large errors in computed LoPG. Developers of augmented reality systems (where visual information is overlaid on an observer's view, using a see-through head-mounted display) have explored geometric models and calibration techniques like those used in eye tracking. Calibration is required for aligning overlaid images with real-world objects. Janin, Mizell, and Caudell (1993) proposed calibrating the location of information presented in a head-mounted display by aligning targets projected in the display with markers on a real-world workbench in a series of intentional gazes. We use a similar approach for LoPG calibration.

Our system allows for accurate (mean error of less than 2º) tracking of LoPG over a wide range of head locations and orientations. We compute the LoPG using a model of the geometry of gazes with a few simplifying assumptions. Doing so allows model parameters in the LoPG-tracking algorithm to correspond to actual distances and angles between parts of the tracking system, making direct verification possible. The general approach of this method can also be extended to incorporate additional or alternate tracking systems.

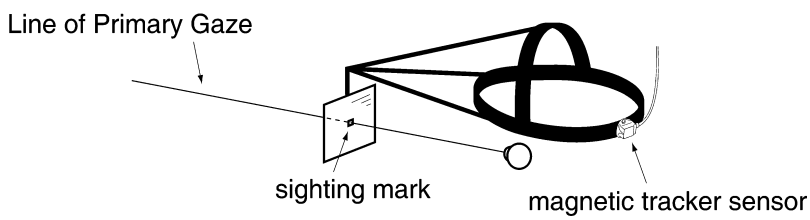

Figure 2. Head-mounted sight. The participant aligned objects of interest with a sighting mark. When viewing objects through the mark, the left eye was held in the primary position of gaze.

In this article, we first will describe our algorithms for compensating for magnetic distortion and for finding the LoPG from compensated tracker measurements. Then we will describe our calibration methodology for finding parameters for the above procedures. Finally, we will describe and present the results of two experiments used to validate our method.

METHOD

Apparatus

The LoPG tracking system we used to validate our tracking algorithm and calibration techniques was built as a part of our walking simulator. It consisted of the following.

Projection screen. A 67 × 50 in. rear projection screen was used for presentation of calibration targets. The bottom of the screen was elevated 30 in. above our walking platform (see Figure 1) to place its center near eye level for walking participants.

Magnetic tracker. An Ascension Flock-of-Birds magnetic tracker with an extended range transmitter (Ascension Technology Corporation, Burlington, VT) was used to monitor head position. This system reports location and orientation of a small magnetic sensor in relation to a transmitter. Each measurement consists of six variables: three representing location in three dimensions and three Euler angles representing orientation. The tracker transmitter was mounted on a 48-in.-tall wood stand and was placed 52 in. to the right and 36 in. in front of the projection screen center. The tracker sensor was mounted on adjustable headgear worn by the participant (Figure 2). The headgear also carried the head-mounted sight.

Head-mounted sight. A head-mounted sight was used to allow the participants to consistently direct their LoPG at projected screen targets. The sight was a 1/8-in. hollow square mark on a clear plastic panel, which was in turn mounted on a 10-in. lightweight brass boom extending from the headgear.

Computer. A Pentium computer running Microsoft Windows 2000, Matlab 6.5 (The MathWorks Inc., Natick, MA), and custom software was used to generate screen images of calibration targets, to control and record data from the magnetic tracker, and to perform calibration and gaze computations. Calibration software was written using Matlab and is available for download from the Psychonomic Society Web archive.

LoPG Tracking Algorithm

We developed an algorithm to transform raw magnetic tracker measurements into experimental variables of interest for behavior research. The algorithm consists of three steps: distortion compensation, geometric transformation, and computation of the *point of gaze*. Distortion compensation consists of correcting raw magnetic tracker readings to compensate for the spatial distortions that are characteristic of these devices. In this step, we attempt to compute the true location and orientation of the magnetic sensor from its reported location and orientation. In the second step (geometric transformation), distortion-compensated measurements are used to find the LoPG relative to the projection screen. Finally, the LoPG is used to compute the point on the screen aligned with the sighting mark. We call this on-screen location the *point of gaze*. The first and second steps of this algorithm rely on parameters that are obtained through a pair of calibration processes described below in the Calibration section.

Distortion compensation. Examination of readings from our magnetic tracker confirmed other research (Bryson, 1992) that has shown that distortion in both the location and the orientation components of measurements varied dramatically depending on location of the magnetic sensor. To account for these distortions, we use a compensation function that adds a correction factor to measurements. This correction depends on the location portion of magnetic sensor readings. Specifically, to compensate for distortion, the translational components of each raw measurement from the magnetic tracker (x_o, y_o, and z_o, representing the location in three dimensions of the magnetic sensor relative to the transmitter) are transformed with a polynomial distortion correction function (Kindratenko, 1999). The correction consists of three degree-n polynomials in three variables. Each of these three polynomials contains terms consisting of the product of x_o, y_o, and z_o, raised to powers in all combinations where the sum of the exponents is n or fewer:

$$\begin{bmatrix} x_c \\ y_c \\ z_c \end{bmatrix} = \begin{bmatrix} x_o \\ y_o \\ z_o \end{bmatrix} + \sum_{i=0}^{n} \sum_{j=0}^{i} \sum_{k=0}^{i-j} \begin{bmatrix} a_{ijk} \\ b_{ijk} \\ c_{ijk} \end{bmatrix} x_o^j \cdot y_o^k \cdot z_o^{i-j-k}, \qquad (1)$$

where a_{ijk}, b_{ijk}, and c_{ijk} are $t = (n+1)(n+2)(n+3)/6$ polynomial coefficients (obtained in the calibration process), and x_c, y_c, and z_c consist of the coordinates of the corrected measurement. Orientation of the sensor, reported by the magnetic tracker as Euler angles θ_o, ϕ_o, and ψ_o, is corrected using a similar set of polynomials:

$$\begin{bmatrix} \theta_c \\ \phi_c \\ \psi_c \end{bmatrix} = \begin{bmatrix} \theta_o \\ \phi_o \\ \psi_o \end{bmatrix} + \sum_{i=0}^{n} \sum_{j=0}^{i} \sum_{k=0}^{i-j} \begin{bmatrix} d_{ijk} \\ e_{ijk} \\ f_{ijk} \end{bmatrix} x_o^j \cdot y_o^k \cdot z_o^{i-j-k}, \qquad (2)$$

where d_{ijk}, e_{ijk}, and f_{ijk} are t polynomial coefficients, and θ_c, ϕ_c, and ψ_c are corrected orientation angles. Correcting for distortion in this way assumes that the distortion in orientation measurements depends only on the location of the magnetic sensor and not at all on orientation. Although distortion in orientation measurements does indeed vary with sensor orientation (Livingston & State, 1997), this form of distortion compensation is still effective if the sensor remains in the same relative orientation in both calibration and experimental use. Polynomial coefficients in Equations 1 and 2 correspond to the shape of the specific distortion in a given experimental

configuration and can be computed using the calibration procedure described below. In agreement with the findings of Kindratenko (2000), we found that a polynomial of degree $n = 4$ modeled the shape of the distortion well. A fourth-degree polynomial required $t = 35$ coefficients for each tracker variable (210 in total).

Geometric transformation. The second computational step of the LoPG tracking algorithm transforms the distortion-compensated magnetic tracker readings to locate LoPG in relation to the screen. This transformation employs a geometric model of the spatial relationship between the eye and the magnetic sensor and of the relationship between the screen and the magnetic transmitter (Figure 3). The set of 12 model parameters computed in the geometric calibration procedure (described below) characterizes these two spatial relationships.

The model takes each of the following locations within the tracking environment to be the origin of a three-dimensional coordinate frame: the upper left corner of the screen (coordinate frame **O**), the measurement origin of the magnetic transmitter (**B**), the measurement center of the magnetic sensor mounted on the head (**S**), and the center of rotation of the eye (**E**). The x, y, and z axes of each coordinate frame are oriented as shown in Figure 3.

We represent the spatial relationships between pairs of these coordinate frames as 4×4 homogeneous transformation matrices. Each such matrix can describe the relative locations and orientations of two coordinate frames and allows measurements in one frame to be transformed to the other by matrix multiplication. For each such matrix,

$$\underset{\text{b}\leftarrow\text{a}}{\tau} = \begin{bmatrix} r_{11} & r_{12} & r_{13} & x_{ba} \\ r_{21} & r_{22} & r_{23} & y_{ba} \\ r_{31} & r_{32} & r_{33} & z_{ba} \\ 0 & 0 & 0 & 1 \end{bmatrix}, \tag{3}$$

representing the transformation from some frame **a** to some frame **b**, x_{ba}, y_{ba}, and z_{ba} describe the three-dimensional location of the origin of **a** in the coordinate system of **b**. The r variables are the lengths of the projections of unit vectors along each of the three coordinate axes of frame **a** onto the axes of frame **b**. These projections form the rotational component of the transformation.

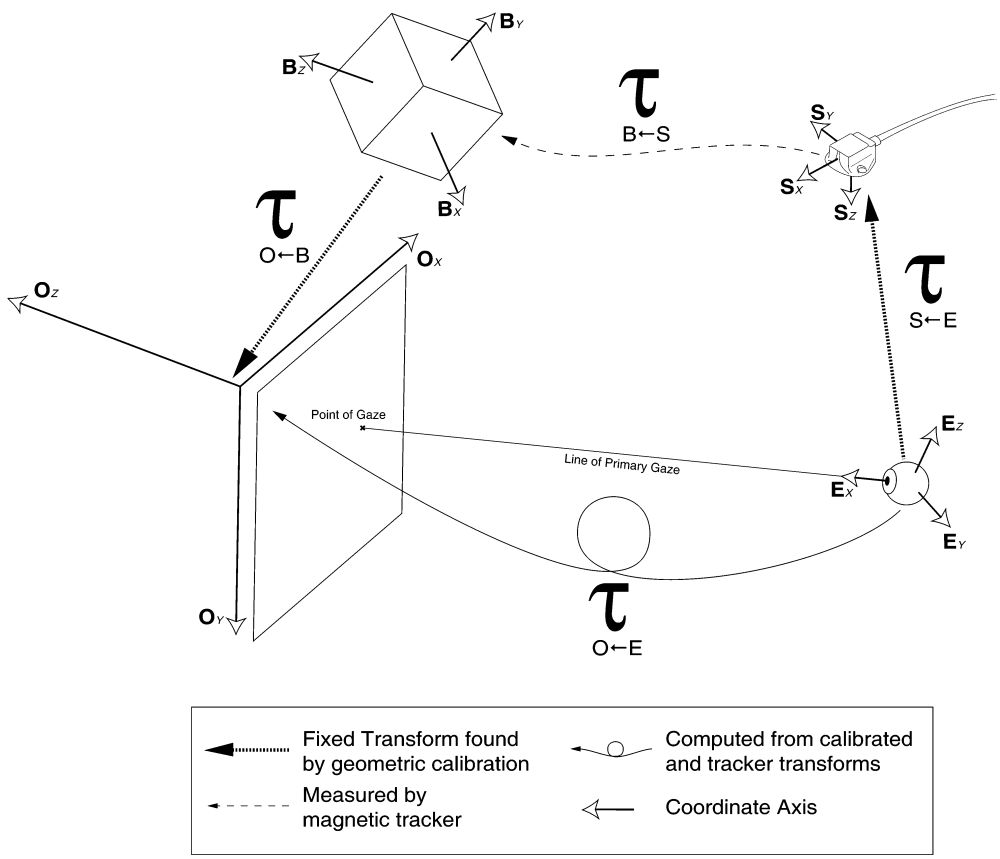

Figure 3. Geometric model for finding line of primary gaze by combining (via matrix multiplication) the spatial relationship measured by the magnetic tracker ($\tau_{B\leftarrow S}$, corrected for magnetic distortion) with those derived through geometric calibration ($\tau_{O\leftarrow B}$ and $\tau_{S\leftarrow E}$). B is the tracker transmitter, S is the tracker sensor, E is the eye, and O is the projection screen.

In addition to the matrix representation of the relationship between two coordinate frames, it is often convenient to use a translation and Euler angle representation. In this way, a transformation can be described compactly with 6 variables, instead of the 12 used in the matrix form. The three coordinates of translation and the three Euler angles that specify a transformation, $M_{b\leftarrow a} = (x_{ba}, y_{ba}, z_{ba}, \theta_{ba}, \phi_{ba}, \psi_{ba})$, can be converted to matrix form by a series of three rotations and a displacement $\tau_{b\leftarrow a}$ as shown in Equation 4 at the bottom of the page (Kuipers, 1999).

Using these conversions, the (corrected) measurements provided by the magnetic tracker can be used to derive the other spatial relationships described in Figure 3. The computations of the various transformations are described next, followed by a description of the calibration process.

Since the LoPG lies along the *x*-axis of the eye frame **E**, to find the LoPG relative to the screen, we find the location and orientation of this frame with respect to the screen frame **O**. To find this relationship, we compose the matrix representations of the spatial relationships measured by the magnetic tracker with matrices computed through calibration. The position of the magnetic transmitter with respect to the screen $\tau_{O\leftarrow B}$ (transformation from transmitter frame to screen frame) and that of the eye with respect to the sensor $\tau_{S\leftarrow E}$ remain fixed for the duration of an experiment. The transformation $\tau_{B\leftarrow S}$ is constructed using Equation 4 from the six variables $M_{B\leftarrow S} = (x_c, y_c, z_c, \theta_c, \phi_c, \psi_c)$ resulting from the distortion compensation step (Equations 1 and 2). The coordinates x_c, y_c, and z_c represent the corrected three-dimensional displacement between the origins of **B** and **S**, while θ_c, ϕ_c, and ψ_c are the corrected Euler angles describing the orientation of **S** in the coordinate system of **B**.

Since composite transformations can be represented as the product of transformation matrices, we can find the location and orientation of the eye relative to the screen $\tau_{O\leftarrow E}$ as follows:

$$\underset{O\leftarrow E}{\tau} = \underset{O\leftarrow B}{\tau}\ \underset{B\leftarrow S}{\tau}\ \underset{S\leftarrow E}{\tau}. \tag{5}$$

Once $\tau_{O\leftarrow E}$ is computed, we can find both the eye location with respect to the screen[1] (simply the last column of $\tau_{O\leftarrow E}$) and the point of intersection between the LoPG and the screen.

Point of gaze computation. The point on the screen that intersects the participant's LoPG can be found from $\tau_{O\leftarrow E}$, the result of the geometric transformation portion of the gaze-tracking algorithm. In our model, the LoPG falls along the *x*-axis of the eye coordinate frame **E**. We solve for a point that is on the *x*-axis of the eye frame and also falls in the *x*–*y* plane of the screen coordinate frame **O**—the surface of the screen. This intersection can be formulated as a linear system consisting of a point on the *x*-axis of frame **E**, $(x_o, 0, 0)$, transformed by $\tau_{O\leftarrow E}$ and set equal to a point in the *x*–*y* plane of frame **O** $(Int_x, Int_y, 0)$.

$$\underset{O\leftarrow E}{\tau}\begin{bmatrix} x_o \\ 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} Int_x \\ Int_y \\ 0 \\ 1 \end{bmatrix} \tag{6}$$

Solving this system for Int_x and Int_y gives us the intersection between the LoPG and the plane of the screen, expressed in the screen coordinate frame.

Calibration

Parameters used in the LoPG tracking algorithm are computed via two calibration steps. First, the distortion in magnetic tracker readings is measured, and coefficients for the compensation polynomials (Equations 1 and 2) are computed. This distortion compensation calibration process was carried out when our tracking system was installed. Computed coefficients can be reused, assuming that there have been no changes to the relative placement of the tracker transmitter and the distortion-causing (large metal) objects. The second geometric calibration step computes the spatial relationship between the magnetic tracker sensor and the eye, along with the relationship between the tracker transmitter and the screen image. This second step is performed once for each participant before an experimental run, since an adjustment of the headgear or of the image projected on the screen alters the spatial relationships measured in this calibration.

Distortion compensation calibration. To compute coefficients a_{ijk}, b_{ijk}, c_{ijk}, d_{ijk}, e_{ijk}, and f_{ijk} of the distortion compensation polynomials (Equations 1 and 2), we compare a set of tracker readings, taken over a grid of known points within the tracking volume, to the hand-measured locations of those points. The differences between the measured and the reported locations form a set of vectors mapping points in reported space to actual physical locations (these difference vectors were small near the transmitter and became large—see Figure 5—near the outer third of the nominal volume of our tracker's range). We then solve for the set of polynomial coefficients that, when used in Equation 1, minimizes the sum of the squared lengths of these error vectors. An identical procedure is performed on orientation data, finding coefficients that minimize the sum of the squared errors in Euler angles reported by the tracker.

$$\underset{b\leftarrow a}{\tau} = \begin{bmatrix} 1 & 0 & 0 & x_{ba} \\ 0 & 1 & 0 & y_{ba} \\ 0 & 0 & 1 & z_{ba} \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \cos(\theta_{ba}) & -\sin(\theta_{ba}) & 0 & 0 \\ \sin(\theta_{ba}) & \cos(\theta_{ba}) & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \cos(\phi_{ba}) & 0 & \sin(\phi_{ba}) & 0 \\ 0 & 1 & 0 & 0 \\ -\sin(\phi_{ba}) & 0 & \cos(\phi_{ba}) & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos(\psi_{ba}) & -\sin(\psi_{ba}) & 0 \\ 0 & \sin(\psi_{ba}) & \cos(\psi_{ba}) & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \tag{4}$$

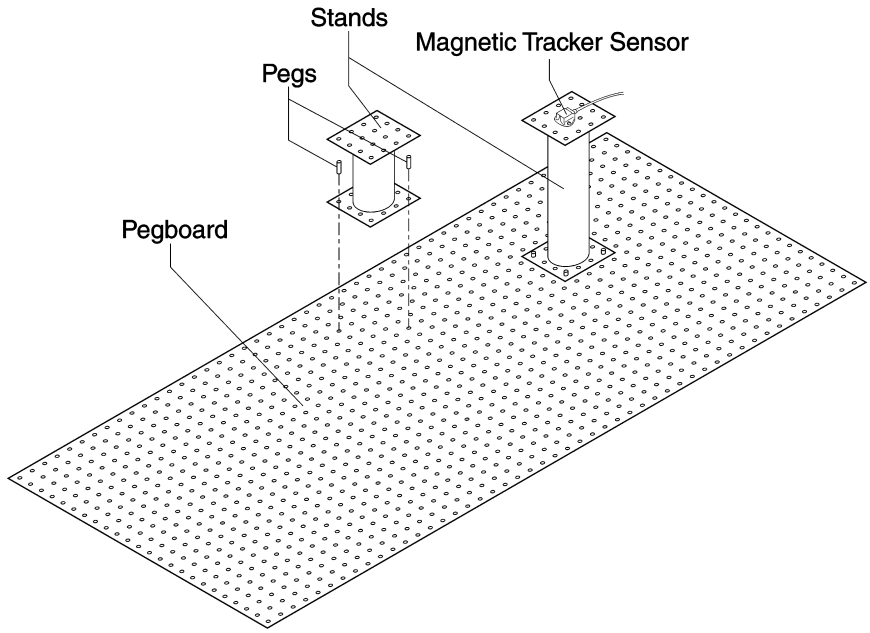

Figure 4. Pegboard apparatus for measuring distortion in magnetic tracker measurements. Two such pegboard panels were mounted on a wooden frame to calibrate a volume measuring 45 × 52 × 18 in. Stands of various heights were placed at different locations across the base pegboard to provide wide coverage of the volume.

In our walking simulator environment, we constructed a system of pegboards and stands to provide a set of known locations for placement of the magnetic sensor during this procedure. The pegboard system, illustrated in Figure 4, was constructed from 24 × 48 in. sheets of predrilled 0.25-in.-thick pegboard mounted on a wooden frame. Screws used in the construction of the frame were replaced with glue, since we found them to be causing small local distortions in magnetic tracker readings. The sensor stands were constructed from sections of 4-in.-diameter PVC pipe with flat pegboard caps on each end. Each stand held the magnetic tracker sensor rigidly on one end cap and had pegs for attachment to the pegboard on the other. The sensor was positioned so that when the stands were engaged with the pegboard, the sensor's *x*-axis (see Figure 4) pointed toward the screen and its *z*-axis pointed toward the floor. In this orientation, the magnetic tracker reported azimuth, elevation, and roll of the sensor (θ_o, ϕ_o, and ψ_o) to all be near zero (these angles deviated from zero by a few degrees, due to distortion). Several pipe lengths were used for measurements at varying heights (0, 6, 12, and 18 in.) above the pegboard plane. Other devices for placing a sensor at known locations have also been developed, including a fixture made entirely of Plexiglas (Caldwell et al., 2000) and an optical tracking system (Ikits, Brederson, Hansen, & Hollerbach, 2001).

The pegboard frame was placed in our tracking environment and was mechanically squared with the case of the magnetic transmitter. The location of the pegboard grid with respect to the magnetic transmitter was then measured.[2] Readings from the magnetic tracker were taken over both a sparse, three-dimensional grid of 64 locations spanning 45 × 52 × 18 in. (see Figure 5), and a denser, smaller grid of 128 locations covering the central 32 × 40 × 18 in. of the same space. The magnetic tracker readings from both grids were combined to form the data set of reported locations and orientations. Using Matlab, we performed linear least-squares fits on the differences between reported and pegboard-measured data sets to find the 210 coefficients for the distortion correction polynomials.[3] An illustration of the magnitude of location distortion for our tracker is provided in Figure 5.

Geometric calibration. The second calibration step computes the geometric model parameters describing spatial relationships $\tau_{O\leftarrow B}$ and $\tau_{S\leftarrow E}$. These parameters, corresponding to the alignment of the screen with the tracker transmitter ($\tau_{O\leftarrow B}$) and position of the magnetic tracker sensor relative to the eye ($\tau_{S\leftarrow E}$), are difficult to measure directly, and $\tau_{S\leftarrow E}$ changes from participant to participant or whenever the headgear is adjusted. For these reasons, we developed a calibration procedure to quickly ascertain these model parameters.

Geometric calibration of the LoPG-tracking system is performed by taking magnetic tracker recordings while participants perform a series of constrained gazes. For each gaze, the participant, using his or her left eye, aligns the head-mounted sight with a target dot projected on the screen. The participant repeats this procedure from a series of standing locations within the desired usage space.

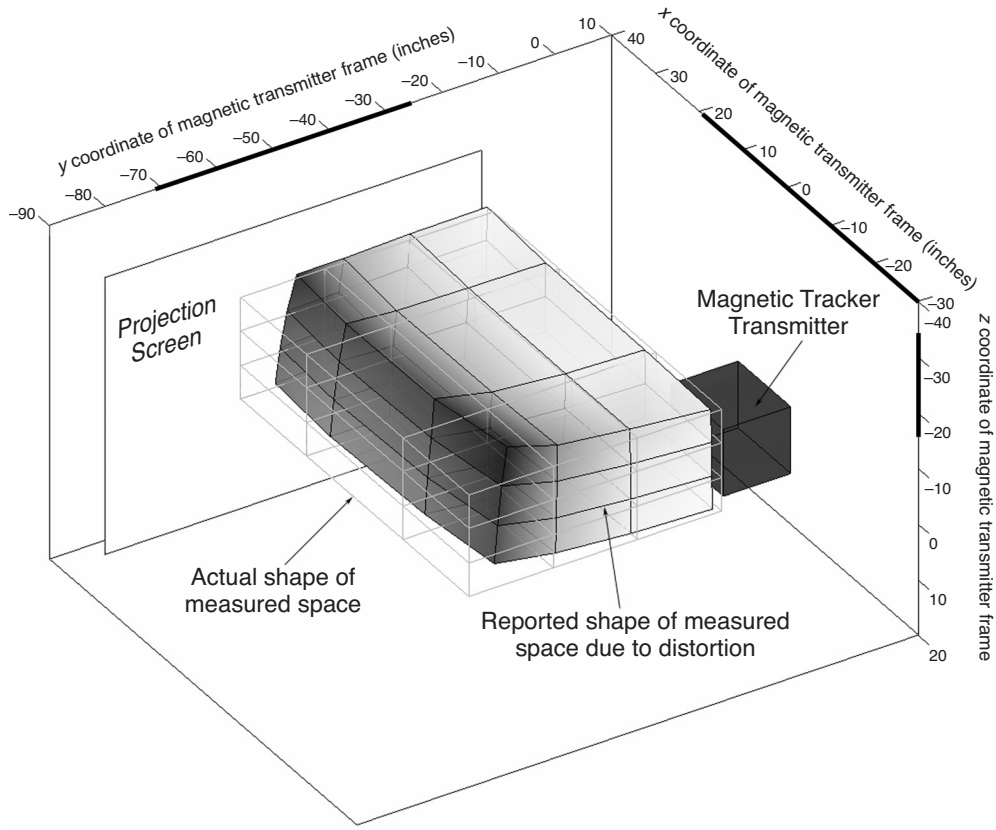

Figure 5. Distortion measured in performing distortion compensation calibration of magnetic tracker. Axis units are inches in the coordinate system of the magnetic tracker transmitter. Dark bars on axes indicate the ranges of physical locations sampled in the calibration procedure. Shade of surface indicates magnitude of distortion in location measurements from the magnetic tracker. The darkest areas indicate distortion in location measurement of over 13 in.

After a number of such gazes, a pair of $\tau_{O\leftarrow B}$ and $\tau_{S\leftarrow E}$ parameters is computed by the Gauss–Newton nonlinear least-squares fitting method, as will be described in the next section.

Geometric calibration algorithm. The fitting procedure seeks the set of model parameters that, when combined with the geometric model in our LoPG-tracking algorithm, best predict on-screen points of gaze from distortion-compensated magnetic tracker readings taken over a set of directed gazes. Our model computes *Int*, the two-dimensional on-screen point of gaze, expressed as

$$Int = \boldsymbol{F}(P, M_{B\leftarrow S}), \quad (7)$$

where $\boldsymbol{F}$ is the function that computes point of gaze from a given (corrected) magnetic tracker reading and a set of model parameters (Equations 5 and 6). P is the set of 12 model parameters characterizing, via Equation 4, the spatial relationships $\tau_{O\leftarrow B}$ and $\tau_{S\leftarrow E}$ (the 12 parameters describe, for each of the two transformations, a three-dimensional displacement between coordinate frames and three Euler angles describing relative orientations of the coordinate frames). $M_{B\leftarrow S}$ is the magnetic tracker measurement (three distances and three angles) corrected for distortion by Equations 1 and 2. We assess quality of the fit found by the geometric calibration procedure by the size of the average error across all gazes in a calibration. This *mean gaze point error* is expressed as the sum across all calibration gazes of the distances along the screen between the actual calibration targets and the model-predicted points of gaze. The fitting procedure seeks model parameters, P, that minimize this sum, given a set of magnetic tracker measurements $M_{B\leftarrow S}$, a set of corresponding screen intersections *Int*, and an initial "guess" for P. Starting with the given guess for P, the fitting procedure iteratively adjusts these parameters until a local minimum in mean gaze point error is found.[4] Minimization is performed in Matlab using the nlinfit function (from the statistics toolbox; an implementation of the Gauss–Newton fitting method). In our implementation, this technique finds a P corresponding to a given set of 20 measurements of *Int* and $M_{B\leftarrow S}$ in a few seconds on a 700-MHz PC.

Experimental Validation

To test the calibration and LoPG-tracking techniques presented in this article, we first calibrated the distortion

compensation system as described above and then performed two target-sighting experiments. In the first experiment, we performed the geometric calibration procedure, using the beam of a laser pointer to represent the line of gaze of the human participant. Doing so allowed the simulated LoPG to be precisely aligned with calibration targets projected on the screen and allowed testing without the variability of human participants' ability to maintain head position. To fix the relative placement of the laser and sensor, both were attached to a small wooden block. The sensor was placed on the block, with its *x*-axis approximately parallel to and 2 in. above the laser beam. This alignment enabled us to easily measure the actual spatial relationship between the sensor and the laser for validation of the parameters found by the geometric calibration procedure. (We found it much more difficult to measure the equivalent relationship, $\tau_{S\leftarrow E}$, on a human participant.) The wood block holding the sensor and laser was fastened to an aluminum and plastic tripod raised to the approximate height of a human participant's head. To gather data for testing the geometric calibration procedure, the tripod was placed at 10 different locations. These locations were chosen within the range of possible standing positions of our study participants (an area about 36 × 36 in.). For each of the 10 tripod locations, readings from the magnetic tracker were taken, panning and tilting the tripod head so that the laser beam fell on each calibration target. Four of the targets were just inside the corners of the screen, one was placed in the center, and five other locations were arbitrarily chosen. The targets remained in the same locations throughout each experiment. Due to the distance between the sensor and the pivot point of the tripod head, for a given tripod location the actual magnetic sensor locations varied by up to 8 in. as the tripod head was tilted at different angles to align the laser.

In a second experiment, a similar set of 100 directed gazes was performed by a human participant. The participant used the head-mounted sight (Figure 2) to align his LoPG with each of the 10 targets from each of 10 standing locations. The 5 calibration targets that were arbitrarily placed in the first experiment were in different locations for calibration of the human participant.

Data from both experiments were independently used to calibrate the LoPG-tracking system, and the benefit of distortion compensation and geometric calibration procedures were examined. For the distortion compensation step, we used polynomial coefficients computed from the 192 pegboard locations as described above. In all cases, we used as an initial guess for the geometric calibration procedure a set of hand-measured parameters that approximated the actual geometry of the system.

RESULTS

For each of the two experiments, collected data consisted of a set of 100 six-variable tracker readings (each reading taken when the laser/LoPG was aligned to fall at the center of an on-screen calibration target), as well as the 100 two-dimensional screen locations of the corresponding calibration targets (10 repetitions of 10 targets). Location and orientation data from the magnetic tracker were passed through the distortion compensation functions (Equations 1 and 2), yielding a second set of 100 corrected tracker readings for each experiment. We used subsets of this body of distortion-corrected data to derive the transformation matrices $\tau_{O\leftarrow B}$ and $\tau_{S\leftarrow E}$, using the geometric calibration procedure described above, and tested these fits against the remaining data. We compared known screen–target locations to point-of-gaze computations made with and without distortion compensation or geometric calibration. We also investigated the behavior of the error as a function of the number of included calibration points by varying the size of the calibration data sets used to compute $\tau_{O\leftarrow B}$ and $\tau_{S\leftarrow E}$.

Accuracy of Predicted Point of Gaze

Using the distortion-corrected location and orientation measurements from the entire set of 100 laser–target alignments from the first experiment and the geometric calibration procedure, the $\tau_{O\leftarrow B}$ and $\tau_{S\leftarrow E}$ matrices were computed. Equations 5 and 6 were then used to predict point of gaze for each of the 100 laser–target alignments, and mean gaze point error (mean distance along the screen between calibration target and point of gaze found with the LoPG-tracking algorithm) was computed. For the data set from the first experiment, mean error was 1.50 in. (Figure 6D). Error can also be described as the angle between the screen target and the computed point of gaze when viewed from the location of the magnetic sensor. Measured in this way, mean angular error was 1.55º. For reasons described in the Discussion section, we will report the remainder of our mean error results in inches. At our screen–eye distances, 1º corresponds approximately to 1 in. of error.

To examine the effect of distortion compensation on mean gaze point error, we also performed the geometric calibration procedure, using raw magnetic tracker measurements not processed by the distortion correction polynomials. As seen in Figure 6B, without distortion compensation, mean gaze point error was significantly larger than with compensation (Wilcoxon signed ranks test, $Z = 3.66$, $p < .001$) at 1.93 in. when calibrated using data from the same experiment (an increase of 29%). Without distortion compensation, clear outliers from the clusters of points of gaze can be seen for most of the screen targets in Figure 6B. These outlying points of gaze were all from the tripod location most distant from the magnetic transmitter, where distortion was largest. Distortion compensation provided most benefit for this tripod location.

To examine the effect of the geometric calibration procedure on accuracy of predicted LoPG, we also computed mean gaze point error across the collected data set using only a hand-measured estimate of $\tau_{O\leftarrow B}$ and $\tau_{S\leftarrow E}$: Distortion correction was included, but no geometric calibration was performed. As one might expect, without

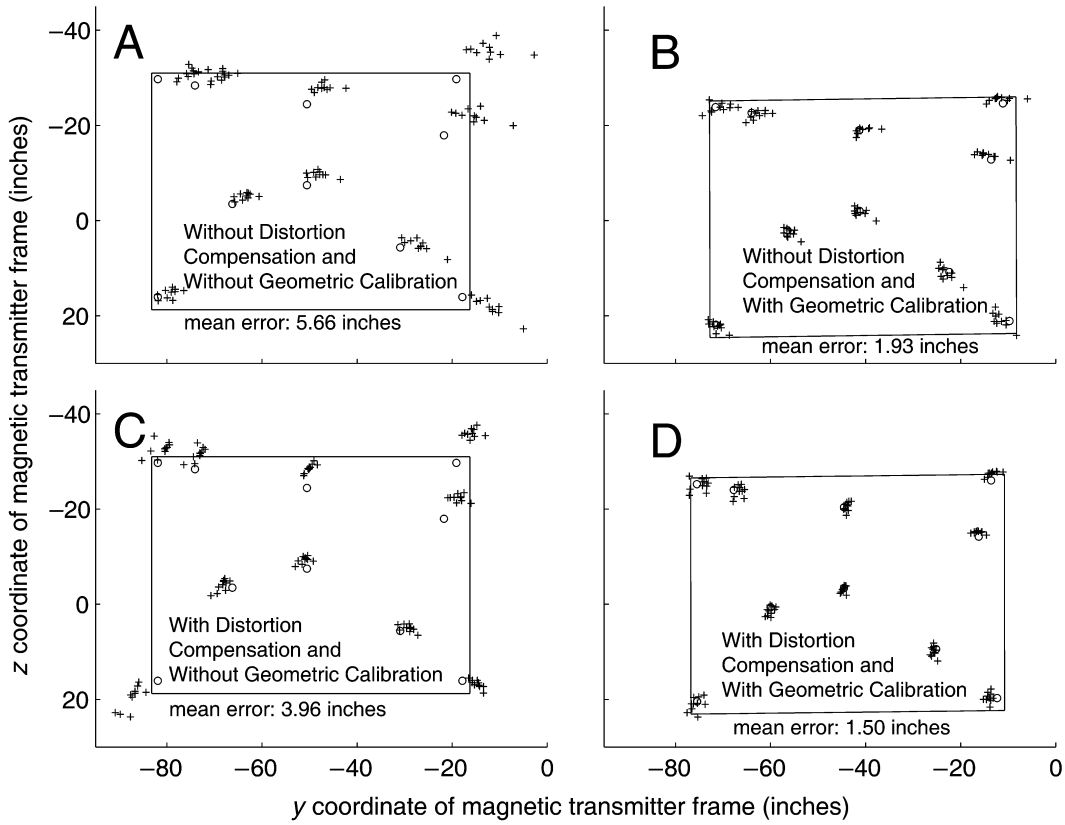

Figure 6. Effects of distortion compensation (panels C and D) and geometric calibration (panels B and D) on computation of points of gaze. Data were collected with a tripod-mounted laser pointer used to simulate line of primary gaze (LoPG). Circular marks indicate screen location of calibration targets. Crosses are laser screen intersections (points of gaze) computed by LoPG-tracking algorithm. Within each cluster, each cross represents a measurement from a different tripod location. Rectangular outline shows location of projection screen. Screen outline and targets appear at hand-measured location (panels A and C) and location found by geometric calibration (panels B and D).

calibration, errors were larger, and this difference was significant (Wilcoxon signed ranks test, $Z = 8.09$, $p < .001$). Mean gaze point error using an uncalibrated model was 3.96 in., as can be seen in Figure 6C.

We also examined the ability of the geometric model to compute points of gaze without both distortion compensation and geometric calibration. The hand-measured $\tau_{O\leftarrow B}$ and $\tau_{S\leftarrow E}$ and raw magnetic tracker readings were used in Equations 5 and 6. Mean gaze point error was 5.66 in. (Figure 6A). The addition of either geometric calibration or distortion compensation yielded a significant reduction in error (Wilcoxon signed ranks test, $Z > 6.1$, $p < .001$).

Conducting the same analysis for the data from the second experiment, we found mean gaze point error to be lower than that in the first experiment when geometric calibration was performed. When both distortion compensation and geometric calibration were used, mean error was 0.72 in. When geometric calibration was performed on data not corrected for distortion, mean error was 1.06 in. (Wilcoxon signed rank test, $Z = 5.20$, $p < .001$). Reasons for the reduced errors in the second experiment are proposed in the Discussion section. Without geometric calibration, error was much larger (Wilcoxon signed rank test, $Z > 8.1$, $p < .001$), since we were unable to measure $\tau_{S\leftarrow E}$ with the same accuracy for the second experiment. With and without distortion correction, mean errors were 39.8 and 42.6 in., respectively.

Effect of sensor location on error. Using the raw data (not compensated for distortion) from the first experiment and computing point of gaze without performing geometric calibration, the farther the sensor was from the transmitter, the worse the error (Spearman correlation, $r = .46$, $p < .001$), and that error was manifest as an increasing error in the y-dimension (Spearman correlation, $r = .50$, $p < .001$; Figure 7A). When both the distortion correction and the geometric calibration were implemented, the errors were more evenly distributed (Figure 7B), resulting in no significant correlation with distance from the transmitter (Spearman correlation, $r = .04$, $p < .71$). There was, however, a significant tendency for the errors to increase toward the screen (i.e., in the x-dimension; Spearman correlation, $r = .22$, $p = .003$).

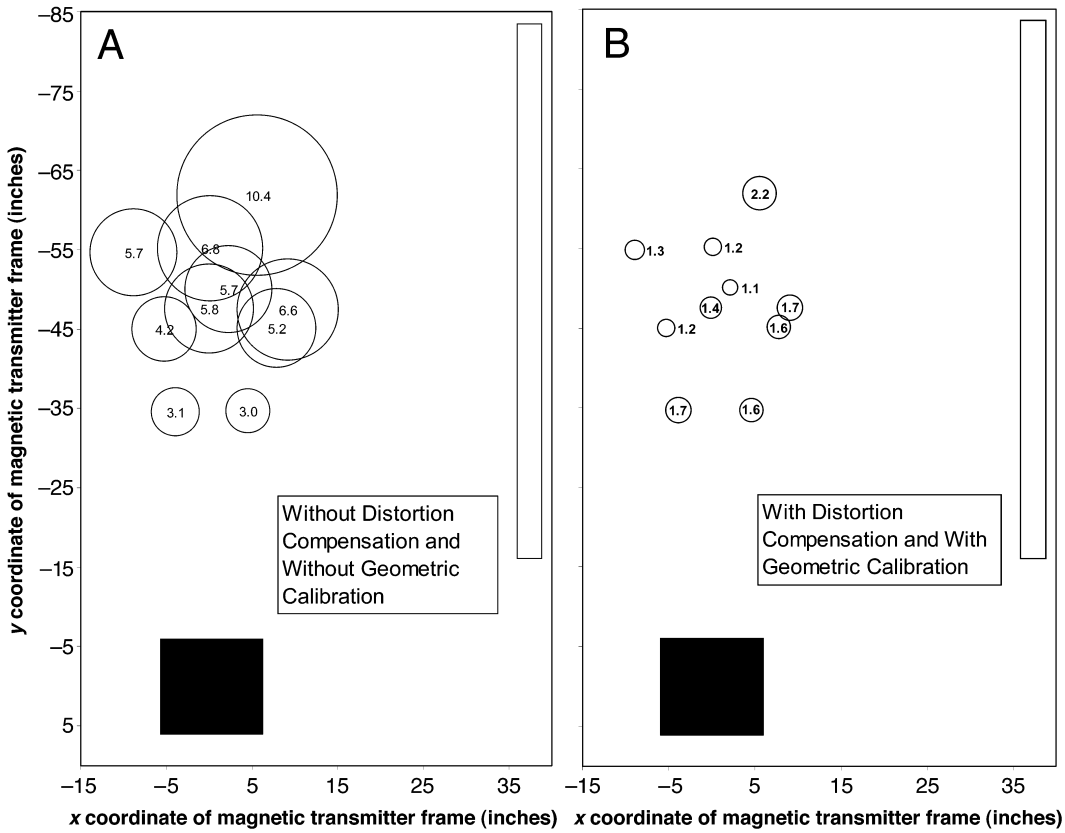

Figure 7. The mean gaze point error at each of the 10 tripod locations, on the left using the raw data from the magnetic tracker collected in the first experiment (see Figure 6A) and on the right using the same data corrected using both the distortion compensation and the geometric calibration (Figure 6D). The tripod and laser pointer were directed toward the projection screen, which is shown as the tall white rectangle on the right of each panel. The tracker transmitter is shown as the black square. Each circle represents one of the tripod locations, with radius equal to the mean gaze point error from that location. Mean gaze point error for each location is also shown with the corresponding circle.

This may be because for tripod locations closer to the screen, an angular error (due to distortion, for example) of the same size results in a larger error distance along the screen.

Short calibration. In order to use our LoPG-tracking system as a tool for gathering data from untrained study participants, it is advantageous to calibrate the system using as few data points as necessary (e.g., to reduce fatigue). To find the number of calibration data points needed to accurately predict point of gaze, we created calibration sets from subsets of the 100 measurements collected in each experiment. We used these small calibration sets to compute $\tau_{O\leftarrow B}$ and $\tau_{S\leftarrow E}$ and then used those computed parameters to calculate screen intersections across the full set of 100 measurements. We then compared these computed intersections with the locations of the actual locations of the on-screen targets. We performed this technique for calibration sets of varying sizes, across 100 random orderings of the 100 data points collected in each experiment. For the first experiment, we found that after about 20 points, the geometric calibration system converges to a point of gaze with a mean gaze point error of less than 2 in., with only small improvements attained by including additional points in the calibration set (Figure 8).

For the second experiment, overall gaze point error was lower when calibration was performed with the entire data set. For this data set, 2 in. of error could be reached with about 12 data points (Figure 9). Calibrations performed with more than 15 data points provided only a small additional benefit.

In the processing of calibration data from both experiments, geometric calibration occasionally failed completely, producing larger mean gaze point error than if geometric calibration had not been performed at all. These failures occurred with very small calibration sets when distortion-related errors provided inconsistent constraints for the model. Failures also occurred when sets included targets from only one part of the screen, insufficiently constraining the model. Fitting failures never occurred with sets of 10 points or more and never occurred when distortion compensation was performed. In all the

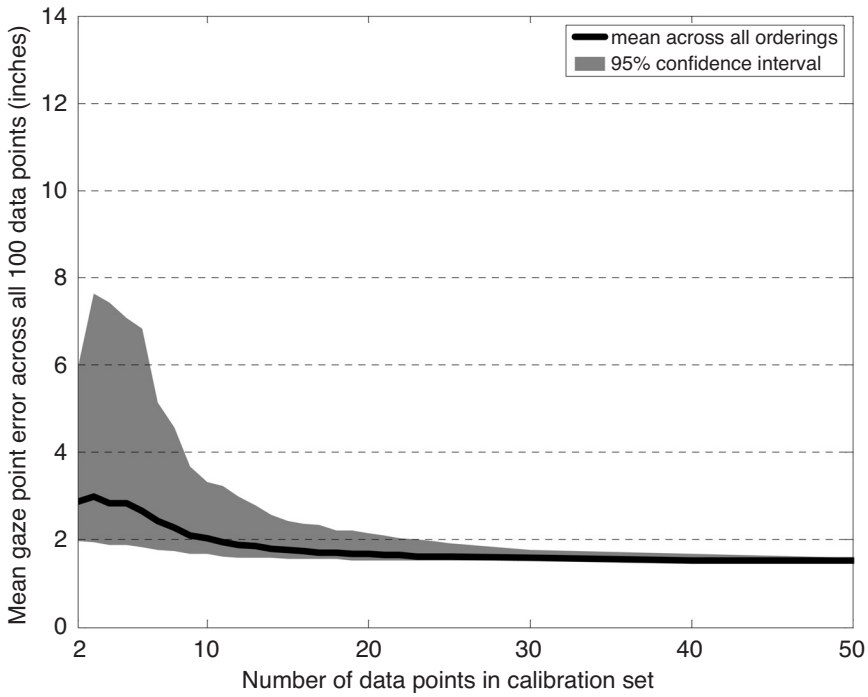

Figure 8. Error in computed point of gaze derived when model was calibrated with increasing numbers of calibration points, across 100 random orderings of the data set collected in the first experiment (using a laser pointer). Mean gaze point error was reduced to within 2 in. (about 2º) after about 23 directed gazes.

cases in which calibration failed, the addition of one or two calibration points to the set was enough to reverse the failure. In general, better fits were obtained from calibration sets with greater variety in sampled tracker orientations and screen locations.

DISCUSSION

The LoPG-tracking model and calibration techniques presented here succeed at accurately predicting the on-screen point of gaze while participants look through a head-mounted sight. The validation experiments show that, in agreement with Caldwell et al. (2000), the use of polynomial distortion compensation can improve the accuracy of gaze point calculations (Figures 6C and 6D). In addition, benefit was also seen from geometric calibration. The techniques presented for tracking LoPG provide a foundation for future gaze-tracking systems.

An interesting result of performing calibration both with and without distortion compensation was that although point of gaze was predicted more accurately with distortion compensation, the tracking system still performed relatively well in the presence of uncompensated distortion, provided that geometric calibration was performed. In the presence of measurement errors caused by distortion, the minimization procedure gives model parameters that may not correspond well to the actual spatial relationships being modeled but interact with the measurement distortion to produce reasonably accurate screen intersection predictions. If somewhat larger LoPG-tracking errors can be tolerated, one might be able to achieve reasonable results by omitting distortion compensation altogether. Omission of distortion compensation would likely result in much greater error, however, for head positions or orientations outside the range of calibration.

Possible sources of gaze point errors remaining when both geometric calibration and distortion compensation were performed include measurement noise, inaccuracies in magnetic tracker measurements not corrected by the distortion compensation system, inaccuracy of the model parameters found by the geometric calibration system, or some combination of these. Tests of stability in the tracker measurements showed that readings from the tracker taken milliseconds or hours apart agreed: both differ from reference measurements by only a few hundredths of an inch and by a few tenths of a degree. This small measurement noise is clearly not enough to explain the gaze point errors, which are orders of magnitude greater.

Our method makes the simplifying assumption that the distortions in magnetic tracker orientation readings depend on location of the magnetic sensor only, and not on the interaction between sensor location and orientation. Therefore, it is likely that our method did not completely correct for the magnetic distortion. When the magnetic sensor is kept in a similar orientation in both

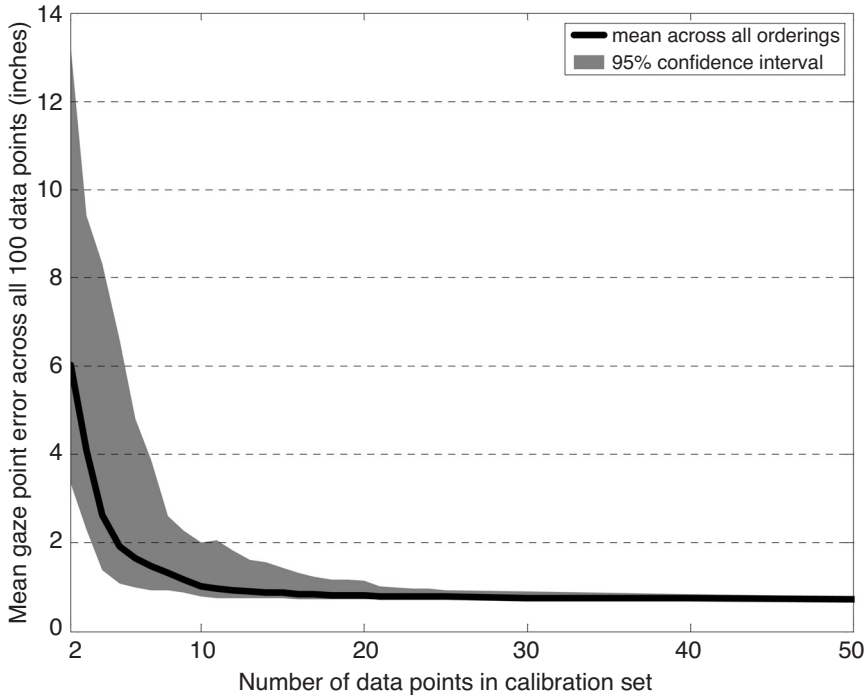

Figure 9. Error in computed point of gaze derived when model was calibrated with increasing numbers of calibration points, across 100 random orderings of the data set collected in the second experiment (with a human participant). Mean gaze point error was reduced to less than 2 in. (about 2º) in about 12 directed gazes.

distortion compensation calibration and actual experimental use, this approach is more likely to produce good results. In the first experiment, our method provided an improvement over omitting orientation distortion compensation entirely, since sensor alignment was similar in calibration and experiment. It is likely, however, that distortion still plays a role in gaze point errors. In the second experiment, we found even lower (than that reported in the Results section) mean gaze point error when magnetic tracker measurements were corrected for location distortion, but not for orientation distortion. As the magnetic tracker was rotated on its side when attached to the headgear of the head-mounted sight (Figure 2), it was no longer in the relative orientation needed to see a benefit from our compensation for orientation distortion. Expanding the distortion compensation calibration procedure to measure the sensor at multiple orientations for each grid location sampled could further reduce error due to distortion. This change would increase both the number of data points required and the time needed to perform the initial data collection procedure, but it would not impact per-participant calibration demands.

The geometric calibration algorithm seeks a $\tau_{O\leftarrow B}$ and $\tau_{S\leftarrow E}$ yielding a local minimum in gaze point error, starting from the initial guess derived from hand measurements of these parameters. Control experiments were performed to verify that no other, potentially lower, minima could be found near the minima arrived at by calibration. Although the minimization procedure failed to converge for intentionally inconsistent initial guesses, tests conducted using parameters displaced from hand-measured values by several inches in all directions all still converged on the same LoPG. Thus, uncorrected distortion in magnetic tracker measurements was most likely the leading cause of gaze point error.

Comparison between points of gaze computed with (Figure 6D) and without (Figure 6C) the geometric calibration step shows that attempts to hand measure the spatial relationships $\tau_{O\leftarrow B}$ and $\tau_{S\leftarrow E}$ (i.e., without geometric calibration) lead to relatively large gaze point errors. Application of the geometric fitting procedure substantially improves the ability of the model to compute LoPG. Although it is not surprising that better results can be obtained when calibration is performed, it is interesting to note that the parameters resulting from calibration deviated considerably from hand-measured parameters. In fact, the optimized parameters comprising $\tau_{S\leftarrow E}$ indicated spatial relationships that were clearly inconsistent with the actual configuration of the laser and magnetic sensor. For example, parameters producing the least mean gaze point error contained values for the distance between the magnetic sensor and the LoPG that were several inches greater than physical measurements indicated. Although there is some imprecision (we estimate about a quarter of an inch) in measuring straight-line distance between eye and magnetic sensor, this imprecision alone cannot ex-

plain the difference. Since calibration performed with data from the human participant seemed less prone to this inconsistency, we suspect that the constraints imposed on the movement of the laser and sensor by the tripod head may be responsible. In positioning the laser to align it with calibration targets, the sensor remained within 47° of level, and "roll" of the sensor could not be performed. These constraints lead to less variation in the calibration data and, in turn, may have underconstrained the fitting procedure used in geometric calibration.

A limitation of the geometric calibration procedure described here is the inability of the calibration procedure to locate the eye along the LoPG. Although this limitation does not impact the ability of the system to predict points of gaze on the screen, it does impact the veracity of the model parameters. Although the geometric model used specifies the eye as a coordinate frame, with an origin located some distance from the magnetic tracker sensor, the geometric calibration procedure constrains only the x-axis of this coordinate frame. This results in an infinite number of possible sets of model parameters that still predict the same point of gaze on the screen (i.e., screen intersection remains the same even if the eye were to slide forward or back along the LoPG or to rotate about its x-axis). This means that the parameters making up $\tau_{S\leftarrow E}$ found by the fitting procedure correspond to one possible location of the eye along the LoPG. Control experiments showed that the fitted location of the eye along the LoPG depended on the initial guess used for the fitting procedure, even though these alternate initial guesses still yielded the same points of gaze.[5] If the intermediate results of the model are to be used, for example, for computing the location of the eye relative to the screen for placement of a camera in a virtual environment, the inability of this system to precisely locate the eye might be overcome by performing an additional calibration. For example, a calibration similar to the one used in this article, but requiring the participant to place the eye in other, nonprimary positions of gaze, could be used to find actual eye location. This technique is currently being investigated.

REFERENCES

Allison, R. S., Eizenman, M., & Cheung, B. S. K. (1996). Combined head and eye tracking system for dynamic testing of the vestibular system. *IEEE Transactions on Biomedical Engineering*, **43**, 1073-1082.

Asai, K., Osawa, N., Takahashi, H., & Sugimoto, Y. Y. (2000, March). *Eye mark pointer in immersive projection display.* Paper presented at the IEEE Virtual Reality 2000 Conference, New Brunswick, NJ.

Atchison, D. A., & Smith, G. (2000). *Optics of the human eye*. Oxford: Butterworth-Heinemann.

Barabas, J., Giorgi, R. G., Goldstein, R. B., Apfelbaum, H., Woods, R. L., & Peli, E. (2003, May). *Wide field 3D gaze tracking system.* Paper presented at the Association for Research in Vision and Ophthalmology Meeting, Fort Lauderdale, FL.

Blood, E. (1990). *U.S. Patent No. 4,945,305.* Washington, DC: U.S. Patent & Trademark Office.

Bryson, S. (1992, February). *Measurement and calibration of static distortion of position data from 3D trackers.* Paper presented at SPIE Stereoscopic Displays & Applications III, San Jose.

Caldwell, R. T., Berbaum, K. S., & Borah, J. (2000). Correcting errors in eye-position data arising from the distortion of magnetic fields by display devices. *Behavior Research Methods, Instruments, & Computers*, **32**, 572-578.

Day, J. S., Dumas, G. A., & Murdoch, D. J. (1998). Evaluation of a long-range transmitter for use with a magnetic tracking device in motion analysis. *Journal of Biomechanics*, **31**, 957-961.

Duchowski, A., Medlin, E., Cournia, N., Murphy, H., Gramopadhye, A., Nair, S., Vorah, J., & Melloy, B. (2002). 3-D eye movement analysis. *Behavior Research Methods, Instruments, & Computers*, **34**, 573-591.

Ikits, M., Brederson, J. D., Hansen, C. D., & Hollerbach, J. M. (2001, March). *An improved calibration framework for electromagnetic tracking devices.* Paper presented at the Virtual Reality, 2001, Proceedings. IEEE, Yokohama.

Janin, A., Mizell, D., & Caudell, T. (1993, September). *Calibration of head-mounted displays for augmented reality applications.* Paper presented at the Virtual Reality Annual International Symposium, Seattle.

Kindratenko, V. (1999). Calibration of electromagnetic tracking devices. *Virtual Reality: Research, Development, & Applications*, **4**, 139-150.

Kindratenko, V. (2000). A survey of electromagnetic position tracker calibration techniques. *Virtual Reality: Research, Development, & Applications*, **5**, 169-182.

King, P. (1995). Integration of helmet-mounted displays into tactical aircraft. *Proceedings of the Society for Information Display*, **26**, 663-668.

Kuipers, J. B. (1999). *Quaternions and rotation sequences: A primer with applications to orbits, aerospace, and virtual reality*. Princeton, NJ: Princeton University Press.

Leigh, R. J., & Zee, D. S. (1983). *The neurology of eye movements*. Philadelphia: Davis.

Livingston, M. A., & State, A. (1997). Magnetic tracker calibration for improved augmented reality registration. *Presence: Teleoperators & Virtual Environments*, **6**, 532-546.

Nixon, M. A., McCallum, B. C., Fright, W. R., & Price, N. B. (1998). The effects of metals and interfering fields on electromagnetic trackers. *Presence*, **7**, 204-218.

Southard, D. A. (1995). Viewing model for virtual environment displays. *Journal of Electronic Imaging*, **4**, 413-420.

Young, L. R., & Sheena, D. (1975). Survey of eye movement recording methods. *Behavior Research Methods & Instrumentation*, **7**, 397-429.

NOTES

1. This position is used in projected virtual reality displays to determine how to position the virtual camera when generating an image of a scene.

2. Even if the tracker is not aligned exactly with the case of the transmitter, the correction polynomials (Equations 1 and 2) we used include terms to account for such an offset: These offsets are contained in the constant ($i = 0$) and linear ($i = 1$) terms of polynomials.

3. Two linear systems were constructed from 192 six-variable measurements. Each of the two systems contained $192 \times 3 = 576$ knowns and 105 unknowns.

4. It should be noted that the calibration procedure described here does not completely constrain the model parameters used. Since no measurement of orientation on the screen of the point of gaze or distance from eye to screen during a sighting is made, neither of these can be predicted by the model. As a result, the best-fitting model parameters describing $\tau_{S\leftarrow E}$ may converge on any one of a line of solutions corresponding to possible locations of the eye along the LoPG. Each of these eye locations predicts the same LoPG location on the screen. It should be possible to compute correct eye position with a calibration procedure that includes eye-in-head rotation.

5. For this reason, we have chosen to report errors in inches along the screen surface, since angular measures of error are more sensitive to location of the eye along the LoPG. Where angular error is reported, we estimated the eye–screen distance to be the distance between the magnetic sensor and the screen.

ARCHIVED MATERIALS

The following materials associated with this article are retrievable from the Psychonomic Society's Norms, Stimuli, and Data archive, http://www.psychonomic.org/archive/.

To access the above files or links, search the archive for this article using the journal (*Behavior Research Methods, Instruments, & Computers*), the first author's name (Barabas), and the publication year (2004).

FILE: Barabas-BRMIC-2004.zip.

DESCRIPTION: The compressed archive file contains seven files. The files are:

README.txt, containing a description of scripts and data included in the archive (a 3K text file). This file can renamed to readme.m and executed as a Matlab script to reproduce some of the results reported in this manuscript.

barabas2004LoPG.mat, a Matlab workspace file containing magnetic tracker data collected in the verification experiments described in this manuscript (a 32K binary file).

findcorrectioncoeffs.m, a Matlab function for finding the distortion compensation coefficients for Equations 1 and 2 (a 5K Matlab M-file).

correctdistortion.m. a Matlab function for performing distortion compensation on magnetic tracker data using Equations 1 and 2 (a 4K Matlab M-file).

findmodelparams.m, a Matlab function for finding the parameters of the LoPG model from data collected during a series of directed gazes (a 3K Matlab M-file).

intersecterror.m, a Matlab function for finding gaze point error for a set of directed gazes (a 5K Matlab M-file).

trackereulerstomtx4.m, a Matlab function that computes a transformation matrix representation of a relationship between coordinate frames via Equation 4 (a 3K Matlab M-file).

AUTHOR'S WEB SITE: http://www.eri.harvard.edu/faculty/peli/index.html.

AUTHOR'S E-MAIL ADDRESS: barabas@mit.edu.

(Manuscript received July 18, 2003;
revision accepted for publication July 23, 2004.)