Visual Alignment Precision in Optical See-Through AR Displays: Implications for Potential Accuracy

Magnus Axholt\textsuperscript{1}, Stephen D. Peterson\textsuperscript{1}, Stephen. R. Ellis\textsuperscript{2}

\textsuperscript{1}Department of Science and Technology, Linköping University, Sweden
\textsuperscript{2} Human Systems Interface Division, NASA Ames Research Center, CA, USA

Abstract

The quality of visual registration achievable with an optical see-through head mounted display (HMD) ultimately depends on the user’s targeting precision. This paper presents design guidelines for calibration procedures based on measurements of users’ head stability during visual alignment with reference targets. Targeting data was collected from 12 standing subjects who aligned a head fixed cursor presented in a see-through HMD with background targets that varied in azimuth (0°, ±30°, ±60°) and elevation (0°, ±10°). Their data showed that: 1) Both position and orientation data will need to be used to establish calibrations based on nearby reference targets since eliminating body sway effects can improve calibration precision by a factor of 16 and eliminate apparent angular anisotropies. 2) Compensation for body sway can speed the calibration by removing the need to wait for the body sway to abate, and 3) calibration precision can be less than 2 arcmin even for head directions rotated up to 60° with respect to the user’s torso provided body sway is corrected. Users of Augmented Reality (AR) applications overlooking large distances may avoid the need to correct for body sway by boresighting on markers at relatively long distances, >> 10 m. These recommendations contrast with those for heads up displays using real images as discussed in previous papers.

Keywords: boresight, line of sight, calibration, postural sway, augmented reality

1 Introduction

Augmented Reality (AR) is a technology to merge a user’s sensory impressions of the real world with additional computer generated information. Typically, computer graphics objects are superimposed on the user’s field of view (FOV) by some sort of overlaying system. To create a compelling coexistence between objects in the real world and the computer generated virtual objects, the computer graphics must be correctly positioned and oriented with respect to the world. To achieve this goal, most AR systems maintain a 3D model of at least some of the objects in the external world and a geometrical model of the perspective projection needed for the 3D rendering. The geometrical relationships are represented by transformation matrices applied to a reference coordinate system. The 3D model also incorporates a separate pinhole camera model for the rendering projection. The camera is modeled with a pyramid frustum that estimates the user’s FOV and linear perspective. The relationship governing how incident light rays from the objects in the surrounding world intersect the camera’s projection plane can be expressed as a 4-by-4 projection matrix [1].

An optical alignment technique is commonly used to establish the position and projection of points in space of points in a see-through transparent display for the purpose of determining the relationship between the presented versus intended rendered points and the eye, i.e. orientation of the view vector. These values are called extrinsic camera parameters. Once the frustum is set up, the calibration also provides values to the projection matrix determining the linear perspective, scale, translation, and skew. These values are referred to as intrinsic camera parameters. Multiplied together, the transformation and projection matrices express where on the display surface some computer graphics must be rendered in order to appear at a specific location in space as seen by the user. Incorrectly estimated values in any of the matrices can result in discrepancies between virtual and real world objects referred to as registration errors in visual direction and spatial position.

The devices for presenting AR displays can be subdivided into video see-through and optical see-through devices. In a video see-through device the view of the real world is captured and digitized on a video feed onto which the virtual objects are superpositioned. With an optical see-through device the view of the real world is never digitized but instead travels through an optical combiner on which the virtual objects are displayed for a composite view. The two types of devices differ in ease of calibration. Firstly, the hardware setup of a video see-through device already incorporates a camera and therefore better adheres to the assumptions made in a pinhole camera model. For instance, its projection plane is usually perpendicular to the view vector resulting in a symmetrical frustum. Secondly, the video feed
makes it possible to measure distances between landmarks in the resulting image on the projection plane which allows for very exact matrix estimation using computer software. In the case of an optical see-through device, the physical display surface of the optical combiners or its conjugate images seldom coincide with the projection plane complicates direct measurement of landmarks in the image. They may be misaligned intentionally as in rotation introduced for partial binocular overlap systems or unintentionally if the head-mounted optics are not normal to the principle viewing direction from the users’ eye points. Moreover, while a camera mounted at an estimated location of the user’s eye could provide measurement to perform and verify the calibration of an optical see-through device, the accuracy of that calibration when transferred to a human user, even improved with additional calibration points, can result in a registration error 20 times larger than with a camera assisted calibration. Thus, resulting quality of the calibration is ultimately dependent on the precision of human targeting as the user is the final subjective judge of a successful boresight alignment. The experiment described in this paper measures the precision with which users can maintain the required alignments to determine the correct projection matrix. This precision is the key user factor that constraints the expected registration accuracy achievable based on subjective user alignment.

In a previous experiment [6], in which subjects performed a standing boresight calibration of a large format head-up display using a real image presentation surface, it was determined that calibration precision in such displays is affected by subject’s postural sway. It was also noted that the subjects seemed to compensate for postural sway using head rotations to maintain the boresight alignment. It is the compensatory head rotations that are of particular study in this paper. An important difference between the HUD of the previous experiment and the HMD reported in this paper is the manner that head movements of the user effects calibration. In the particular HUD used, the foreground marker is located on a screen which is fixed to the surrounding world. This makes the user make alignments primarily by translational movements with need for rotational movement. In a HMD however, the foreground marker is fixed to the head so that translation movements have lesser effects on alignment, in contrast to the major effects of rotations. This paper investigates the interplay of translational and rotational alignment precision during calibration for the purpose of establishing the limits of the calibration accuracies that may be ultimately achieved.

2 Related Work

There have been several previous studies on aiming and tracking performance with HMD and helmet mounted sights (HMS). In general, in contrast to our use of standing subjects, they have involved seated subjects. Notably, Wells and Griffin [7] offer a tabular summary of 13 other reports in addition to their own study. Previously reported performance metrics have varied, but in the cases where the target’s absolute location was known, targeting accuracy is generally reported as a radial error in visual angle. When the target’s absolute location was unknown, the subjects’ variability in targeting precision has been reported as a standard deviation, or circular error probable (CEP).

Since we are primarily interested in precision, we studied three reports [8], [7], [9], with metrics similar to ours. Nicholson [9] reports an extensive study (3 subj. × 180 rep.) where subjects with a precision (standard error) of 0.13°. Verona [8] reports on a study (6 subj. × 1 rep.) with a precision (standard error) of 0.9°. Wells and Griffins [7] (12 subj. × 1 rep.) report a precision of 0.04°. The differences in reported precision clearly arise from differences in experimental tasks, hardware and instructions. For example, in the case of Nicholson [9], targeting accuracy was measured after the participant confirmed good alignment, but participants were instructed to make their bore-sight alignment as quickly as possible. Other physical differences in the display equipment also exist. For example, in contrast to our system which used a reticule defined as a graphics object on the virtual projection surface, the previous studies used physically defined reticles superimposed on the scene by optical techniques. But as we will show below, we are able to measure alignment precision somewhat unconstrained by some of the specifics of the HMD hardware used, e.g. a well-balanced, light (0.8 kg), moderately high visual resolution (3.4 arcmin/pixel).

3 The Boresight Experiment

We constructed an experimental setup which allowed us to i) present the subjects with a head fixed visual foreground cursor (henceforth referred to as the foreground marker), ii) create clearly visible background markers that did not cause visual fatigue due to the range of accommodative demand, iii) allow presentation of 15 body-referenced orientation angles while head position and orientation was recorded by a tracking equipment.

1 The projection plane and the display surface never coincide when using half-silvered mirrors as optical combiners since the user’s only view virtual mages of the display surface.
3.1 Environment

Figure 1: The setup of the experiment environment illustrating the background screen with its marker orientation angles, the imaginary location of the projection plane and its u and v coordinates, and the eye point P from which the subject is facing either initial direction -30° or 30°.

As Figure 1 illustrates, a reference eye point in space was defined at 1.65 m above the floor and 1.76 m perpendicular distance from a vertical planar background screen covered in black cloth. A retrofitted theodolite, normally used for land surveying, was placed at this eye point and used to determine position for visual targets of known direction by projecting a light point backwards through the theodolite in marker orientation angles onto the background screen with azimuth and elevation increments of 30° and 10° respectively, (henceforth notated (az, el)).

On the floor below the eye point, so called initial directions lines of -30° and 30° were indicated with tape markers. Thus when a subject was placed turned to face the initial direction of -30° and instructed to look at marker orientation angle (30,10) on the background screen subject’s body-relative rotation would be (60,10). This technique let us test a larger range of azimuth orientations in our limited space.

The subjects whose eye height was less than 1.65 m were placed on an adjustable platform to position their eyes at the reference eye point. The subjects whose eye height was more than 1.65 m were used without correction. The resulting presentation accuracy was 0.5° in elevation and 0.3° in azimuth.

3.2 Background Marker

The background marker consisted of a 5 cm diameter red LED mounted on a reflector. The emitted light was diffused with a thin layer of white paper and masked with black tape to create the shape of a crosshair with 0.3 cm (5 arcmin) thick arms stretching 1.5 cm (29 arcmin) in either direction. The blink frequency was approximately 2 Hz.

3.3 Foreground Marker

The foreground marker was a square of 3 by 3 white pixels of 10.5 arcmin visual angle, about twice the thickness of an arm of the crosshair on the background marker at (0,0). The arms of the background marker crosshair were visible behind the foreground marker for all marker orientation angles.

3.4 HMD

The FOV of the Kaiser ProView 50ST HMD was 28° vertically and 37° horizontally. The resolution uses was VGA (640 by 480 pixels). The weight including tracker sensor and 1.5 m cable was 0.8 kg. The image presented in the two optical channels was focused at 3 m (0.33 diopeters). The difference in accommodative demand between imagery in the HMD and markers on the background screen carried between 0.17-0.23 diopeters, less than the accuracy optometrists expect from subjective refraction of the eye. Consequently, visual fatigue due to accommodative changes was avoided. The luminance for a fully black screen when the device was turned on was 1.3 cd/m². The corresponding value for a fully white screen was 21.6 cd/m². The Michelson contrast for the foreground marker of the screen when the ambient florescent office lighting was turned on during the experiment was approximately 0.6.

3.5 Tracker

The sampling rate of the Polhemus FASTTRAK was 60 Hz. The positional and orientational resolutions of the tracker were specified to be 0.0005 m and 0.025° respectively within 0.76 m work volume radius according to vendor specifications. We further restricted the working volume of our tracked area to 0.6 m in front of the tracker transmitter and 0.2 m to either side to insure our subjects remained within a region of minimal distortion through the experiment.

We performed a test to check the linearity of the working volume of the experimental setup by moving the tracker sensor through a large number of known positions in a wooden lattice. The measured working volume exhibited the characteristic bending of a two hemisphere magnetic field with increasing positional error towards the extremes. Histograms showed that at close range the effects of magnetic properties on positional error along the field lines could be estimate with linear distances. Thus, moving within 0.5 m at any distance in front of the tracker transmitter showed a measured average difference of 0.0016 m (±0.0009 m) between reference and measured position. The working volume was not exhibiting a gradual change in angle reports according to the characteristics of a magnetic field but instead reported spurious angles towards the
edges of the working volume (> 0.6 m). The angular error due to nonlinearities was too small to measure with the equipment available at the time.

3.6 Participants
The 12 participants described in Table 1 were voluntary students, staff or contractors at NASA Ames Research Center and had normal or corrected vision. All signed the informed consent forms required of all human research by the NASA Ames Research Center Institutional Review Board.

<table>
<thead>
<tr>
<th>Gender (males, females)</th>
<th>9, 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dominant eye (left, right)</td>
<td>4, 8</td>
</tr>
<tr>
<td>HMD Experience (first time, many times)</td>
<td>7, 5</td>
</tr>
<tr>
<td>Age (min, max, avg, sd)</td>
<td>23, 65, 33.1, ±11.9</td>
</tr>
<tr>
<td>Height (min, max, avg, sd)</td>
<td>1.57, 1.88, 1.74, ±0.094</td>
</tr>
<tr>
<td>Eye Height (min, max, avg, sd)</td>
<td>1.45, 1.78, 1.621, ±0.092</td>
</tr>
<tr>
<td>Eye Height, compensated (min, max, avg, sd)</td>
<td>1.61, 1.78, 1.677, ±0.045</td>
</tr>
</tbody>
</table>

Table 1: Subject data

3.7 Procedure
The subject age, length and eye height was recorded. The non-dominant eye was covered with an eye patch and the subject was placed at the eye point facing an initial direction of either -30° to the subject’s left or 30° to the right. The subjects were instructed to assume a neutral relaxed posture with arms along the sides of the body and the feet three inches apart. Shorter subjects stood on a platform to achieve a standard eye height. The instruction given to the subjects emphasized that the experiment aimed at measuring postural stability and that the subject would be presented with three different viewing conditions: Eyes Closed (EC), Eyes Open (EO) and Visual Alignment/Boresighting (VA/BS). In the EC condition the subjects stood quietly with their eyes closed while wearing the eye patch and the HMD. In the EO condition the subjects stood quietly with their dominant eye open, fixing the background marker through the HMD. In the VA/BS condition the subjects were instructed to stand as quietly as possible and only make small head movements to create and maintain an alignment between a foreground marker displayed in the HMD and a background marker viewed with their dominant eye. The foreground marker was only visible in the VA/BS condition. For all conditions the experimenter moved the background marker manually on the background screen, thereby giving the subject an indication where to orient the head and also the opportunity of a 10-15 s rest between each trial. Each repetition of nine randomly presented orientation angles took about seven minutes and subjects were allowed to remove the HMD for a five minute break after every third repetition.

3.8 Experimental Design
The experiment was executed with a repeated measures within-subjects design. All participants were subjected to the three viewing conditions randomly oriented in all marker orientation angles with three replications. We made in total 972 data recordings, 81 per subject (3 viewing conditions × 9 marker orientation angles × 3 repetitions). Presentation order of the viewing conditions was systematically and exhaustively alternated for each subject (Latin square) to mange sequence effects. Left and right initial directions were alternated on each subject to mange possible asymmetric performance.

3.9 Independent Variables
Viewing Conditions: The subjects were presented with three viewing conditions: Eyes Closed (EC), Eyes Open (EO), and Visual Alignment/Boresight (VA/BS). Marker Orientation Angles: Each subject was presented with nine orientations. The angles ranged from -60° − 0°, -10° − 10°, for six subjects and from 0° − 60°, -10° −10° for the other group of six subjects.

3.10 Dependent Variables
Sway Path: The Sway Path is defined as the sum of Euclidean distances between consecutive positions sampled during a trial. The measurement provides a metric for postural instability, i.e. the length the subject’s head has travelled during the trial period.
Angular Distance from Center Direction: The Angular Distance from Center Direction (ADCD) is the sum of deflections from the average orientation divided by the number of sample points. More precisely it is the average distance from the tracker orientational readings to the average orientation. The distance refers to the magnitude of a vector where the azimuth and elevation are the two components. The average orientation is the separate arithmetic mean of the azimuth and elevation angles. ADCD used a linear estimation acceptable for small spherical angles and can be interpreted as an orientational standard deviation as a measurement of spread or precision.

4 Results
4.1 Data Preparation
The data was processed before analysis to a) move the positional readings from the tracker sensor to a common reference point so the data could be aggregated across subjects, and to b) transform the data to a coordinate system suitable for the particular analysis.

a) To minimize distortion effects the location of the tracker was moved away from the electronics in the HMD by means of a 0.14 m Plexiglas rod. While the added displacement would limit the effects of exogenous magnetic fields, it also transforms the
reports on sway path length. To correct for this effect, data points were translated from the center of the tracker sensor to the position where the skull articulates with the first vertebra of the spine [10]. The translation vector was estimated based on anthropological/medical statistics [11] and measurements of the HMD. The following displacements were used to reduce the sway path rotational component in the: x $0$, y $-0.21$, z $0.18$ m.

b) The data was thereafter transformed into one of two different coordinate systems depending on type of analysis. For the sway path analysis the raw data was transformed from initial directions to a common region in the transverse plane. For the analysis of ADCD the raw data was transformed from the marker orientation angles to a common visual orientation (0,0) towards the background screen with respect to the subjects’ eye positions.

### 4.2 Analysis of Sway Path

As illustrated in Figure 3, and similar to the experiment conducted by Axholt et al. [6], the distribution of positional data in the transverse plane in the current experiment exhibited a pronounced anterior-posterior sway. This was true for all viewing conditions, not only the VA/BS condition. This pattern is believed to occur because upright bipedal stance is more stable in the frontal plane compared to the sagittal plane [12], [13].

The sway path distributions were negatively skewed and therefore log transformed to satisfy the ANOVA criterion of homogeneity of variance. The analysis showed a statistically significant difference in sway path length between viewing conditions, $F(2,22) = 46.064$, $p = 0.000)$. The means were 0.4306, 0.3978, and 0.3938 m for EC, EO and VA/BS respectively and are shown as asterisks in the modified Tukey plot in Figure 2.

A post-hoc Scheffé test was used to find the differences in viewing conditions that exhibited a statistically significant effect. It showed that there was a statistically significant difference between the EC and the EO condition, $F(2) = 61.236$, $p <= 0.01$, but not between EO and VA/BS. While it seems obvious to conclude that the subjects maintain a better balance with their eyes open, the main interest lies in the difference between sway path lengths for the EC and the EO. The quotient between sway path length of EO and EC for ideal conditions is usually $0.73 \pm 0.09$ [15], but it seems as if factors such as limited FOV, torques and added weight from the HMD have changed this quotient to $0.92 \pm 0.13$ for the current experiment. This also means that sway path for EC, serving as a baseline comparison, changes depending on experimental setup. This also explains some of the variability of the results reported in related works. The conclusion is that equipment change can have noticeable effect on baseline values and generalization must account for both the type of task and specific equipment.

The fact that the lengths of the sway paths were about twice as long in the current experiment compared to those reported by Axholt et al. [6] prompted us to check the quality of compensation for a common reference point (as described in 4.1 a above). It can be compared by checking the variability in the resulting point cloud after compensation. The standard deviation for EC reported by Axholt et al. was 0.0131 which is comparable to 0.0102 for EC in the current experiment. A longer sway path but similar spread suggests that subjects are behaving quite differently as a result of the HMD in this experiment compared to the HUD reported by Axholt et al.

Because of the presumed symmetry around the sagittal plane, the marker orientation angles can be “folded” around 0° azimuth (az = abs(az)) to increase the power in the statistical analysis. While the combined analysis over all conditions does not show an interaction effect of azimuth and viewing condition on length of sway path, an analysis for each viewing condition reveals that rotating the upper body induces changes in postural stability relative azimuth angle within each viewing condition, see Table 2. For EC and EO there is a significant decrease in postural stability when the subjects turn further than 30°. In VA/BS the effect is smaller.
Figure 3: Head position density plots in a $-0.02 - 0.02$ m square transverse plane for Visual Alignment/Boresight viewing condition for 15 (azimuth, elevation) marker orientation angles. Each plot holds 32,400 points which are linearly quantized to a gray scale ranging from 1 to $> 150$ points/mm$^2$, except for plots at 0 azimuth that holds 64,800 points and range from 1 to $> 300$ points/mm$^2$.

Table 2: The first three columns show a compilation of ANOVA illustrating the effect of azimuth marker orientation angle on sway path length. The last three columns show pairwise post-hoc Scheffé comparison indicating between which angles the effect occurs. (ns. denotes non-significant difference)

<table>
<thead>
<tr>
<th></th>
<th>Mean 0°</th>
<th>Mean 30°</th>
<th>Mean 60°</th>
<th>0° - 30°</th>
<th>30° - 60°</th>
<th>0° - 60°</th>
</tr>
</thead>
<tbody>
<tr>
<td>EC</td>
<td>(F(2,22) = 18.511, p &lt;= 0.001)</td>
<td>0.4246</td>
<td>0.4123</td>
<td>0.4550</td>
<td>ns.</td>
<td>(F(2) = 10.754, p &lt;= 0.01)</td>
</tr>
<tr>
<td>EO</td>
<td>(F(2,22) = 9.560, p &lt;= 0.001)</td>
<td>0.3939</td>
<td>0.3871</td>
<td>0.4123</td>
<td>ns.</td>
<td>(F(2) = 9.683, p &lt;= 0.01)</td>
</tr>
<tr>
<td>VA/BS</td>
<td>(F(2,22) = 5.610, p &lt;= 0.011)</td>
<td>0.3882</td>
<td>0.3868</td>
<td>0.4063</td>
<td>ns.</td>
<td>(F(2) = 7.663, p &lt;= 0.05)</td>
</tr>
</tbody>
</table>

A principle difference between the three viewing conditions that may explain the lower effect of VA/BS in Table 2 is that in VA/BS the subject is occupied with a visual task. We hypothesize that the added visual task in VA/BS stabilizes the head reducing the length of the sway path. This seems constant over time as no effect of repetition on sway path length was found for the VA/BS condition, (F(2,22) = 0.496, ns.).

4.3 Analysis of Angular Distance from Center Direction

The ADCD distributions were negatively skewed and therefore log transformed to satisfy the ANOVA criterion of homogeneity of variance. The ANOVA showed a statistically significant difference in ADCD between viewing conditions, (F(2,22) = 222.009, p = 0.000). The means were 0.8799, 0.4739, and 0.2483° for EC, EO, and VA/BS respectively and are shown as asterisks in

Figure 4. No effect of repetition was found (F(2,22) = 0.853, ns.).
A post-hoc Scheffé test showed that there was a statistically significant difference between the EC and the EO conditions, \(F(2) = 97.895, p = 0.01\), the EC and the VA/BS conditions, \(F(2) = 443.496, p = 0.01\), and the EO and the VA/BS conditions, \(F(2) = 124.661, p = 0.01\). This means that increasing visual input, from nothing via a blank screen to a foreground marker, increases orientational precision.

Folding the azimuth angles because of symmetry, the combined analysis shows an interaction effect between viewing condition and azimuth on ADCD, \(F(4,44) = 3.869, p = 0.009\). Analyzing the azimuth angles separately yields the results presented in Table 3. It shows an important result of this paper, namely that the subjects exhibited a statistically significant deterioration in orientational precision (approximately 28%) for an off axis visual alignment greater than 30° compared to a boresight alignment straight ahead. In this context it should also be noted that there was no effect on ADCD for elevation in any viewing condition.

<table>
<thead>
<tr>
<th>ANOVA</th>
<th>Mean 0°</th>
<th>Mean 30°</th>
<th>Mean 60°</th>
<th>Post-hoc Scheffé</th>
<th>0° - 30°</th>
<th>30° - 60°</th>
<th>0° - 60°</th>
</tr>
</thead>
<tbody>
<tr>
<td>EC</td>
<td>0.8707</td>
<td>0.7720</td>
<td>0.9969</td>
<td>ns.</td>
<td>ns.</td>
<td>(F(2) = 8.030, (p = 0.05))</td>
<td>(F(2) = 5.330, (p = 0.01))</td>
</tr>
<tr>
<td>EO</td>
<td>0.4580</td>
<td>0.4627</td>
<td>0.5009</td>
<td>ns.</td>
<td>ns.</td>
<td>ns.</td>
<td>ns.</td>
</tr>
<tr>
<td>VA/BS</td>
<td>0.2101</td>
<td>0.2693</td>
<td>0.2655</td>
<td>(F(2)=29.168,(p=0.01))</td>
<td>ns.</td>
<td>(F(2) = 23.307, (p = 0.01))</td>
<td></td>
</tr>
</tbody>
</table>

Table 3: The first three columns show a compilation of ANOVA illustrating the effect of azimuth marker orientation angle on Angular Distance to Center Direction (ADCD). The last three columns show pairwise post-hoc Scheffé comparison indicating between which angles the effect occurs. (ns. denotes non-significant difference)

Studying the VA/BS viewing condition in more detail, the density plots in Figure 5 shows that the orientational precision becomes increasingly horizontally asymmetric as subject’s body-relative azimuth rotation increases. Note that the data from each subject has been transformed into a common spherical sector (approximated as a projection plane at (0,0) as shown in Figure 1). The pronounced sideways variability that seems to be occurring in a horizontal fashion is in fact along the subject’s
transmeatal axis, i.e. from ear to ear, which must not necessarily coincide with the global horizontal axis since when the head is used as a pointing device it turns around the view vector [14]. This is a likely reason why the density plots in Figure 5 seem to by slightly diagonally distributed in (-60,10), (-60,0), (60,10) and (60,0). In further detail, Table 4 shows the ADCD of Figure 5 broken down into its horizontal (u) and vertical (v) component. Statistical analysis confirms that azimuth has a significant effect on horizontal precision, (F(2) = 19.686, p = 0.000). Both azimuth (F(2,22) = 35.604, p = 0.000) and elevation (F(2,22) = 17.018, p = 0.000) has an effect on vertical precision, but there is no interaction between the two.

<table>
<thead>
<tr>
<th>0°</th>
<th>10°</th>
<th>30°</th>
<th>60°</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>u</td>
<td>v</td>
<td>u</td>
</tr>
<tr>
<td>10°</td>
<td>0.2003</td>
<td>0.1443</td>
<td>0.2552</td>
</tr>
<tr>
<td>0°</td>
<td>0.1985</td>
<td>0.1289</td>
<td>0.2991</td>
</tr>
<tr>
<td>-10°</td>
<td>0.1956</td>
<td>0.1229</td>
<td>0.2566</td>
</tr>
</tbody>
</table>

Table 4: Angular Distance from Center Direction (ADCD) in degrees along the horizontal (u) and vertical (v) direction of the projection plane for the Visual Alignment/Boresight viewing condition.

4.4 Analysis of Sway Path and ADCD in Combination
So far we have seen that a subject’s sway path length is about the same for EO as for VA/BS but that the head is rotated a lot less in the VA/BS condition compared to EO. The average precision for VA/BS when the subject is compensating for postural sway is 0.21°. Now we shall investigate the precision with which a subject can perform a visual alignment if postural sway is not present.

One may think of the orientation changes we wish to compensate for as the rotation with respect to the background marker caused by head translation. We therefore constructed an eye vector with a length equal to the screen distance for each marker orientation angle. The eye vector was slaved to subjects’ head rotation such that the end of their eye vector came to describe a point cloud on and around the background marker in space. The point cloud for each marker orientation angle was then transformed to (0,0) so that between subject comparison could be made. The point cloud was then projected onto the projection plane illustrated in Figure 1 at (0,0) 1 m distance from subject’s eye point. Lastly, the projection was transformed to represent the lines of the sight angles to each background marker. The final collection of angles show the subjects’ orientational precision when compensated for rotational effects during targeting that could be attributed to the subject’s individual head translations caused by body/head sway.

The ANOVA performed on ADCD corrected for postural sway shows a second important result in this experiment: Boresight precision increased from 0.21° to 0.01°, approximately a factor 16 improvement, compared to when not compensating for postural sway. The distributions still show a statistically significant effect of azimuth angle. The means were 0.0131, 0.0126, and 0.0148° for 0, 30, and 60° respectively, (F(2,22) = 20.831, p = 0.000). However, as visible in Figure 7, the orientational precision is now more circularly symmetric. There is no effect of elevation (F(2,22) = 0.712, ns.).

4.5 Analysis ADCD over Time
While there was no effect on ADCD over repetitions (which, had it been present, suggests effects of fatigue during the course of the experiment), there are however temporal effects within a trial. To illustrate the variability of orientational precision during a calibration session we divided the 30 s sample period into 10 shorter second time periods and calculated the ADCD for the sample points within each time period. In addition to the Tukey plots in Figure 6, averages of each time period are connected with a black line. A quadratic polynomial was fitted to the averages (not shown for clarity) and the minimum point is subsequently where precision is the highest (marked with a red X). One can conclude that when a subject needs a few seconds to stabilize before optimal precision is achieved in time period 5 or 6. This means that data for calibration should be recorded 12-18 seconds into each calibration session as opposed to an immediate reading.

![Figure 6: Angular Distance from Center Direction (ADCD) measured in degrees for 10 time periods, each three seconds long, for each of the 15 (azimuth, elevation) marker orientation angles. The average ADCD for each time period is connected with a black line which shows the progress of precision over the 30 s calibration session. The red X denotes the point of maximum precision.](image-url)
Figure 7: Angular Distance from Center Direction (ADCD) when compensated for postural sway displayed as density plots in a -0.5° - 0.5° square projection plane for Visual Alignment/Boresight viewing condition for 15 (azimuth, elevation) marker orientation angles showing orientational measurements from all subjects’ three repetitions. Each plot holds 32,400 points which are linearly quantized to a gray scale ranging from 1 to > 1.150 points/0.05°, except for plots at 0° azimuth that holds 64,800 points and range from 1 to > 2,300 points/0.05°.

Figure 8 shows plots that are constructed in the same way as Figure 6, but instead with ADCD that has been compensated for postural sway as described in section 4.4. We note that there is hardly any curvature to the polynomial trend and therefore no optimal point for recording. (Note the change in scale on the y-axis.) Consequently we conclude that if VA/BS data is compensated for postural sway there is no need to wait for an optimal point as an immediate reading is just as precise as a reading further into the calibration session.

5 Discussion
The experiment presented in this paper, investigating the relationship between postural sway and head rotation during boresight calibration, and visual alignment in general, answered several questions.

Firstly, it has been demonstrated that compensating for postural sway increases boresight calibration precision from 0.21° to 0.01°. The VGA resolution of the HMD was approx. 0.06°/pixel which means that subjects were able to make sub-pixel judgments, by centering the red background light within the foreground markers presented on the HMD.

Figure 8: Angular Distance from Center Direction (ADCD), compensated for postural sway, measured in degrees for 10 time periods, each three seconds long, for each of the 15 (azimuth, elevation) marker orientation angles. The average for each time period is connected with a black line which shows the progress of precision over the 30 s calibration session. The red X denotes the point of maximum precision.
As a sanity check we may compare to measurements of human visual acuity which is about 0.017° (=1 arcmin), sometimes better in high contrast environments.

Secondly, when compensating for postural sway the resulting precision distribution becomes more isotropic, as shown in Figure 7. This fact means that we can treat the targeting distributions as circularly symmetric. This in turn means that we do not have to consider the specific nature if the gimbal-like rotation that describes the multi-axis head rotation and which can result in anisotropies in the density plot such as those in Figure 5.

Thirdly, we also learnt that if we compensate for postural sway we do not have to wait for the user to stabilize to obtain a reading with optimal precision.

Fourthly, it has been shown that standing subjects can use their head to target objects with a precision of 2 arcmin (=0.033°) even when their heads are turned up to 60° relative to their torso, also when not compensating for postural sway.

Fifthly, the corrected targeting precision that we have measured in this experiment (~0.017°) implies that calibration bias should be correctable to approximately this precision. Notably this fact also implies that much greater calibration accuracy may be achieved in optical see-through AR displays than heretofore has been reported.

6 Acknowledgements
Magnus Axholt and Stephen D. Peterson were supported by PhD scholarships from the Innovative Research Programme at the EUROCONTROL Experimental Centre, Brétigny-sur-Orge, France. These authors were also supported through the NASA Grant NNA 06 CB28A to the San José State University Research Foundation. The experiment was conducted at NASA Ames Research Center in Mountain View, CA (corresponding author: Stephen R. Ellis, sellis@mail.arc.nasa.gov).

References