Assessing the Relative Value of Stereoscopic 3D versus Head Tracking in Large Scale Immersive Visualization

Derrick Turner
Brigham Young University

Daniel P. Ames
Brigham Young University

Follow this and additional works at: https://scholarsarchive.byu.edu/iemssconference

Part of the Civil Engineering Commons, Data Storage Systems Commons, Environmental Engineering Commons, Hydraulic Engineering Commons, and the Other Civil and Environmental Engineering Commons

https://scholarsarchive.byu.edu/iemssconference/2014/Stream-D/8

This Event is brought to you for free and open access by the Civil and Environmental Engineering at BYU ScholarsArchive. It has been accepted for inclusion in International Congress on Environmental Modelling and Software by an authorized administrator of BYU ScholarsArchive. For more information, please contact scholarsarchive@byu.edu, ellen_amatangelo@byu.edu.
Assessing the Relative Value of Stereoscopic 3D versus Head Tracking in Large Scale Immersive Visualization

Derrick Turner, Daniel P. Ames Ph.D.
Brigham Young University

Abstract: The use of large-scale stereoscopic immersive visualization environments is growing globally as a supplement to field work in environmental science and engineering. Such systems typically include large-scale stereoscopic displays and head/ hand tracking to facilitate an immersive user experience. These components can add significant cost and complication to such a system that may not be warranted if the components do not significantly improve the user experience. This paper presents a study conducted in a large-scale visualization environment called the VuePod, to help determine the relative value of head-tracking and stereoscopic technologies in terms of improved data interpretation. Forty-eight individuals were recruited for a human subjects study involving the performance of tasks in navigation and interpretation using three distinct datasets. Each user anonymously performed tasks on each dataset in one of four different system configurations. The four configurations included motion tracking with stereoscopic 3D, motion tracking with no stereoscopic 3D, no motion tracking with stereoscopic 3D, and no motion tracking with no stereoscopic 3D. Each task was timed and a score was assigned based on the accuracy of the task performance. Scores were scaled against performance time and plotted in rank order for visual assessment of task performance in each system configuration. Visual assessment, together with Student’s T tests indicate a lack of statistically significant differences between the performance metrics of each of the test groups. These results suggest additional work is needed to further determine the relative benefit of each of the target technologies for performance of common data analysis and visualization tasks.

Keywords: Stereoscopic 3D, Head tracking, 3D immersive visualization

1 INTRODUCTION

Large-scale stereoscopic immersive visualization has been demonstrated as a useful tool for viewing, interpreting, and analyzing data (Koller, Lindstrom et al. 1995, Subaran and Baker 2000, Lin, Chen et al. 2013). Two common components of large-scale stereoscopic 3D immersive environments include stereoscopic displays and head-tracking technology (Cruz-Neira, Sandin et al. 1992, Dodgson 2005). The approach assumes that stereoscopic images paired with head-tracking technology allow users to have a more immersive experience, leading to better understanding, visualization, and manipulation of data that can result in more effective use of time and resources (Dodgson 2005, Bowman and McMahan 2007). To date, there has been no specific study of the relative benefits of stereoscopic displays and head-tracking technology in terms of creating a functional immersive experience. This paper presents a human subjects study of the relative value of stereoscopic and head-tracking technologies.

2 METHODS

2.1 Location & Equipment

The human subjects study presented here was conducted using a low-cost stereoscopic immersive visualization system called the VuePod. The VuePod is comprised of twelve 55” 3D LCD televisions and paired with a custom-built high end gaming computer containing three video cards that drive four monitors each (Hayden, Ames et al. 2014). The VuePod computer provides simultaneous stereoscopic
video output to all twelve televisions. The VuePod includes a motion tracking system from ARTrack
(http://www.ar-tracking.com/products/tracking-systems/smarttrack/) that uses two cameras to track
the position of the reflecting balls attached to glasses (for head-tracking) and to a Wii remote (for hand
tracking) within a volume in front of the televisions. The VuePod computer supports both Linux and
Windows operating system based software tools. For the current study, an open source software
application, Vrui was used. The VRUI VR Toolkit (http://dav.ucdavis.edu/~okreylos/ResDev/Vrui), is a
general purpose virtual reality software that is capable of outputting stereoscopic 3D images to multiple
screens (Kreylos 2008).

2.2 Human Subjects and Training

Human subjects research study participants were recruited by email and verbal announcement. These
participants were comprised of both male and female with ages ranging from 18 to 30 years old. The
students were primarily undergraduate civil engineering students with moderate-to-high technical and
computer skills.

User suitability was achieved by a pre-participation survey completed by each user. The survey
determined whether a potential user had any problems regarding vision, depth perception, balance, fine
motor skills, or mobility. These disabilities would potentially put the users at a disadvantage, so they
were addressed before the study. Once a user was determined suitable for the study, each user signed
a consent form regarding the logistics of the study, confidentiality, risks, and compensation. After
participation in the study, each user completed a post-study survey rating the difficulty of each task.
The rating system included values of 0 through 5, with 5 being most difficult and 0 being not difficult at
all.

A short training video was shown to each user to instruct them on use of the VuePod
(www.youtube.com/watch?v=v3PkJFbJGUY). The tutorial video taught the users how to move and
rotate objects, orient a scene, measure distances, and zoom in and out of scenes. These controls were
taught by visually showing how they were achieved and with instructions explaining their purpose and
how to perform them. The video had periodic breaks to allow the user to practice using the VuePod until
they felt comfortable using the controls and tools taught in the video. The configuration used in the
training of how to use the VuePod was head-tracking with stereoscopic 3D.

2.3 Datasets & Configuration

Three different datasets were used in this study. The first dataset was a medical scan of a human foot.
This data was retrieved from the Visible Human Project at the University of California, Davis (Kreylos
2000). This dataset was used to perform two of the four tasks in the user study. The second dataset
was a LiDAR scan of a bridge overpass as part of the EarthScope Intermountain Seismic Belt LiDAR
Project (NSF, USGS et al. 2008). The last dataset used was a point cloud LiDAR scan of San Diego,
California (San Diego 2005). The point cloud was clipped to show the Crowne Plaza Hotel, San Diego.

The study required four different system configurations such that users were tested performing the
same tasks in different environments. The four configurations included head-tracking with stereoscopic
3D (TS), no head-tracking with stereoscopic 3D (NTS), head-tracking with no stereoscopic 3D (TNS),
and no head-tracking with no stereoscopic 3D (NTNS). After the video training, each user was informed
that the level of 3D immersion had been adjusted; however, they were not informed as to exactly what
changes had been made. For example, for the NTNS group, all head-tracking and stereoscopic 3D was
disabled. Whereas for the TS group, stereoscopic 3D and head-tracking were both enabled.

2.4 Tasks

Each user was asked to perform four different tasks. By random selection, each user was placed into a
configuration group and performed all four tasks using that same configuration. There were 48
participants in the study with 12 participants performing tasks in each of the four configuration groups
(TS, NTS, TNS, and NTNS).
The first task used the medical scan of the human foot. Each user was shown the human foot oriented in a vertical position that revealed the exterior of the foot. Each user was asked to rotate the foot 180° degrees horizontally to show the inside of the human foot in a vertical position. Each user was timed while performing this task and was assigned a score from 0 to 5 based on the accuracy with which the task was performed (5 indicated the greatest accuracy). Timing and accuracy score assignments were made by the study proctor. The purpose of this task was to test how well the user was able to rotate an object using the six-degrees of freedom available. Rotation is conducted using the Wii remote controller and its’ attached tracking system reflective balls.

The second task also used the medical scan of the human foot dataset. The starting view for the users was of the human foot in an upside down vertical position showing the exterior of the foot. Each user was asked to rotate the foot vertically to show the show the exterior of the foot in a vertical position. Each user was timed while performing this task and was given a score based on the accuracy at which the task was performed of 0 to 5, with 5 being the most accurate.

The third task used the dataset of the highway overpass described previously. Each user was required to measure the vertical height from the bottom of the bridge overpass to the top of the road located below the overpass. The vertical height measured was recorded, as well as the time it took to complete the task. Notes were taken on the accuracy of the measurement based on the location of the measurement in 3D space. For example, it was noted whether or not the line measured was vertical from all angles and whether the endpoints of the lines were located on the LiDAR points or in front of or behind them. The purpose of performing this task was to determine the ease and accuracy of a simple analytical tool (distance measurement).

The final task used the LiDAR scan of the Crowne Plaza Hotel in San Diego, California. Each user was given a photo of the hotel when it was first built in 1966. The LiDAR scan was created in 2005. Each user was given 3 minutes to analyze the LiDAR scan and the photo and to note any identifiable differences between the two. The number of changes and the specific changes themselves were recorded. There were approximately 10 notable changes between the original image and the LiDAR scan.

2.5 Measuring Results

The two tasks of the medical scan of the human foot were both measured on the accuracy of how well subjects were able to follow instructions and complete the task. After completing each task, the proctor observed the final position of the human foot when the user had stated that the task was completed, and gave a score from 0 to 5 with 5 being the best possible score. Once a score was determined the time it took to complete the task was recorded.

The task of measuring the height of a highway overpass was assessed by determining how well the user duplicated the actual height of 10 units. This answer is achieved when the two endpoints of the line are located on specific LiDAR points and the line is vertical from all angles. The measurement length determined by each user was recorded as well as notes regarding their accuracy. The time to complete the task was also recorded.

The change determination task of the Crowne Plaza Hotel was assessed based on the number of changes each user identified. Each identified change was written down by the proctor and was then later assessed as per its correctness. We expected the subjects to identify 10 specific and noticeable differences between the photo and LiDAR scan. The number of changes and time spent finding those changes were used to determine how well each person was able to perform the task in the different system configurations.

3 RESULTS

Three of the four tasks were assessed based on time to complete the task and a score on how well the task was completed. Scores were scaled against time of completion by dividing the score by the time and multiplying by 100. The resulting scaled scores for each user were sorted in order from lowest to
highest within each group. Creating ranked scaled scores allowed the time and actual score to be factored together to better characterize the differences in working in each configuration.

A visual assessment of the ranked and scaled scores was performed and a Student's t-test was used to determine the statistical significance of the results. We used a level of probability of 95% with 22 degrees of freedom to determine the tabulated t-value. The tabulated t-value was 2.07 and then compared with the calculated t-value. Specific results are discussed below.

3.1 Human Foot Horizontal Orientation

The results from the human foot horizontal orientation are shown in Figure 1.

![Scaled Score vs User Rank](image)

**Figure 1: Human foot horizontal orientation scaled score results**

The TS group had the highest overall score of any of the four groups. Analyzing the plot shows that among the top four users in each group (the upper tercile), TS scored the highest, followed by the TNS group. Interestingly, results from the lower four subjects in each group (the first tercile) show the opposite, with the TS group scoring lowest. One interpretation of these results is that “power users”, or people who are comfortable with technology, are more likely to benefit from the addition of head-tracking and stereoscopic 3D whereas novice technology users are more likely to find the additional technologies to be distracting or otherwise limiting. Regardless, the visual assessment of these data is inconclusive regarding the question of the relative benefit of these two technologies.

Table 1 presents the calculated t-values from the human foot horizontal orientation task. From the results, none of the calculated t-values are greater than the tabulated t-value. This means that that 95% of the tests in one group are not significantly different from any of the other groups.

**Table 1: Student’s t-test Calculated t-values from the Human Foot Horizontal Orientation Task**

<table>
<thead>
<tr>
<th></th>
<th>TS</th>
<th>NTNS</th>
<th>NTS</th>
<th>TNS</th>
</tr>
</thead>
<tbody>
<tr>
<td>TS</td>
<td>0.00</td>
<td>0.03</td>
<td>4.36</td>
<td>0.00</td>
</tr>
<tr>
<td>NTNS</td>
<td>0.03</td>
<td>0.00</td>
<td>3.35</td>
<td>0.00</td>
</tr>
<tr>
<td>NTS</td>
<td>0.36</td>
<td>3.35</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>TNS</td>
<td>0.07</td>
<td>4.09</td>
<td>0.00</td>
<td>0.00</td>
</tr>
</tbody>
</table>

3.2 Human Foot Vertical Orientation
The results from the vertical rotation of the human foot are shown in Figure 2.

![Graph showing scaled score vs. user rank with different conditions: TS, NTNS, NTS, TNS.]

Figure 2: Human foot vertical orientation scaled score results plot

The NTS group included one user with the overall highest score. However, only two of the four NTS users in the third tercile scored above the other three configuration groups. Within the third tercile TS has the highest average when the top scorer from NTS is disregarded. When looking at the first tercile TS is the top performer with the other three configurations having similar scores. The second tercile has a wide spread with NTNS having the best score followed by TS, NTS, and TNS. When considering all three terciles, TS scored best.

Table 2 presents the calculated t-values from the human foot vertical orientation task. From the results, none of the calculated t-values are greater than the tabulated t-value. This means that 95% of the tests in one group are not significantly different from any of the other groups.

Table 2: Student’s t-test Calculated t-values from the Human Foot Vertical Orientation Task

<table>
<thead>
<tr>
<th></th>
<th>TS</th>
<th>NTNS</th>
<th>NTS</th>
<th>TNS</th>
</tr>
</thead>
<tbody>
<tr>
<td>TS</td>
<td>0.92</td>
<td>0.62</td>
<td>0.40</td>
<td>0.92</td>
</tr>
<tr>
<td>NTNS</td>
<td>0.37</td>
<td>0.62</td>
<td>0.40</td>
<td>0.37</td>
</tr>
<tr>
<td>NTS</td>
<td>0.33</td>
<td>0.07</td>
<td>0.40</td>
<td>0.33</td>
</tr>
<tr>
<td>TNS</td>
<td>0.92</td>
<td>0.62</td>
<td>0.40</td>
<td>0.92</td>
</tr>
</tbody>
</table>

3.3 Highway Overpass Measurement

Assessment of the numerical results and associated user comments for the bridge overpass measurement task indicated a high level of difficulty performing the task. Most subjects were unable to correctly measure the height of the bridge overpass. The comments and proctor observations indicate that the subjects were unable to move the cursor directly on the required LiDAR points causing many of the measurements to be taken in front of the points instead of on the points. This means that the measurements were essentially taken in random space. Another issue with this task was that the users were unable to draw a vertical line. When subjects would rotate the image it was clear that the measurement lines were not exactly vertical, thereby causing measurements to be longer than expected. Only three people were able to obtain a measurement on the points with a distance close to the actual answer. These three subjects were members of groups TS, NTS, and TNS.

3.4 Change Detection of the Crowne Plaza Hotel, San Diego
The average times to identify all changes in the Crowne Plaza Hotel scene are shown in Figure 3.

![Change detection Crowne Plaza Hotel scaled score results plot](image)

*Figure 3: Change detection Crowne Plaza Hotel scaled score results plot*

Change detection results for this task were scaled against time as with the other tasks, but with little visible difference in scores since most users used all three minutes to complete the task.

The third tercile shows that each group had the same high score, but overall in this tercile, NTS group scored highest. The other three groups show highly similar results in the top tercile. The second tercile is very similar to the third with NTS showing clearly higher scores.

Table 3 shows the results from the statistical analysis of the scaled scores from the change detection of the Crowne Plaza Hotel task. None of the calculated t-values from this task exceed the tabulated t-value of 2.07. Therefore, we can conclude that there is not a significant difference between each group and their respective results.

<table>
<thead>
<tr>
<th></th>
<th>TS</th>
<th>NTNS</th>
<th>NTS</th>
<th>TNS</th>
</tr>
</thead>
<tbody>
<tr>
<td>TS</td>
<td>0.13</td>
<td>0.40</td>
<td>0.54</td>
<td>0.45</td>
</tr>
<tr>
<td>NTNS</td>
<td>0.13</td>
<td>0.40</td>
<td>0.55</td>
<td></td>
</tr>
<tr>
<td>NTS</td>
<td>0.54</td>
<td>0.45</td>
<td>0.91</td>
<td></td>
</tr>
<tr>
<td>TNS</td>
<td>0.45</td>
<td>0.55</td>
<td>0.91</td>
<td></td>
</tr>
</tbody>
</table>

4 **DISCUSSION & CONCLUSION**

Visual assessments of ranked, scaled scores for all tasks reveal inconclusive results regarding relative advantages of head-tracking and stereoscopic displays. The TS group scored well on both human foot orientations, but did not perform well on the change detection task. The NTS group performed well on the change detection task, but did not receive very high scores on the human foot orientations when compared to the other groups.
The statistical analysis using the Student's t-test approach supports the visual analysis concluded from the scaled score plots; there is not a statistically significant difference between the different configurations tested. In short, our results are at best inconclusive as to whether stereoscopic 3D with head-tracking improves one's ability to view, interpret, or manipulate data.

Further studies are needed to fully explore the question of the relative value of head-tracking technology and stereoscopic 3D visualization in terms of improved understanding, viewing, and interpreting of data. Our current results can be used to inform such studies, and indeed we are presently developing a second user study that will include more detailed training and additional tasks which we hope will more clearly answer the question of the benefits of these technologies.

REFERENCES

Hayden, S., et al. (2014). A Mobile, Low-Cost, Large-Scale, Immersive Data Visualization Environment, for Civil Engineering Applications, Brigham Young University.