

Characterizing the Effect of Level of Immersion on a 3D Marking Task

Jürgen P. Schulze, Andrew S. Forsberg, Alexander Kleppe, Robert C. Zeleznik, David H. Laidlaw

Department of Computer Science
Brown University
Providence, RI 02912
{schulze, asf, kleppe, bcz, dhl}@cs.brown.edu

Abstract

We present a user study comparing multiple immersive environments for the performance of a task. We expected them to perform differently but it was not obvious how they would differ. Our analysis aimed to discover the environment in which subjects performed best and the environment they preferred. The task, derived from developmental biology, requires a user interactively to label spheres distributed in a volume by marking the spheres in 3D. We tested three virtual reality displays: a CAVE-like environment, a single-wall display, and a desktop system (fish tank virtual reality). All displays supported head-tracked stereo viewing and a tracked wand device. The task conditions varied in field of view, apparent size of the spheres, and whether the user was standing or sitting. We also varied the number of targets that had already been marked when the user began a trial. Popular user interaction techniques allow the user to change viewpoint and do the marking.

Data we collected led to four significant findings: a) in the Cave the subjects preferred medium sized or large spheres over small spheres, b) when only a few of the targets have been marked, larger spheres were marked faster than small spheres, c) large spheres are marked most accurately, and d) our wall display is not comparable to our fish tank virtual reality environment when the spheres are small. Additionally, occlusion and larger field of view inhibited performance in the Cave more than at the fish tank when the task was dominated by visual search.

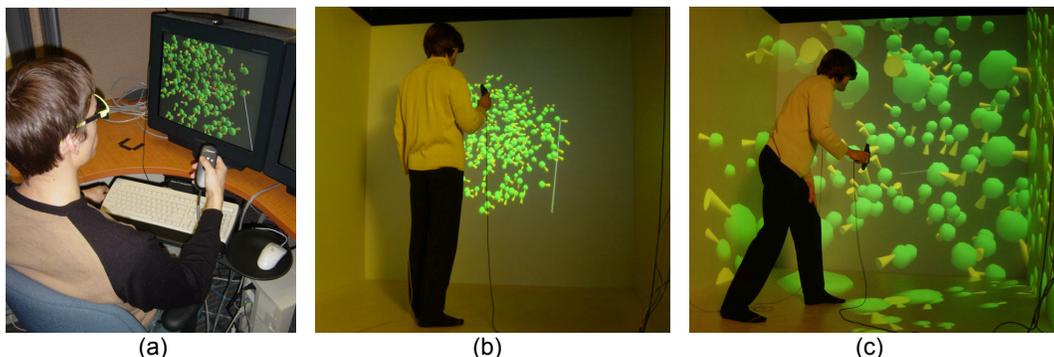


Figure 1: Marking spheres in different virtual environments: (a) at the fish tank, (b) at the single wall, and (c) in the Cave.

1 Introduction

We have implemented a software application that allows users to mark spheres in virtual reality. Our goal was to find out how this task is performed in different virtual environments and at different scales in the Cave. This work has been motivated by our collaboration with developmental biologists at Brown University with whom we built a software system for studying data sets from confocal microscopes.

Counting cells or cellular components is standard practice in studying many biological processes, such as assessing the proliferation of a tissue and determining the size of cells. The density of certain components within a particular volume is often compared in control and experimental samples. Immunohistochemical techniques that use antibodies, tagged (or labeled) with a fluorochrome or other molecule that fluoresces under particular wavelengths

of light, allow biologists to highlight (or recognize) the structures of interest within a tissue preparation. The digitized data is collected with a laser scanning confocal microscope that generates a static volume data set.

Determining cell density in a data set requires two tasks: 1) identifying an equivalent subvolume within each data set, and 2) accurately counting the cells in the subvolume. The first task requires domain knowledge to identify and register an equivalent subvolume across samples, which can be done by using anatomical features as landmarks to bound the areas in which to count cells. The second task requires uniquely identifying and tallying the cells within the volume; here the main challenges are isolating individual cells and (due to the large number of cells) avoiding double-counting. Our study reported here involved only the second task because that task did not require specialized knowledge of biology but still involved visualizing and interacting with the biological data.

Our collaborators from our Biology Department have reported verbally that viewing their data in immersive virtual reality (VR) helps them better explore and identify 3D structures. In immersive VR, the user is fully immersed in a three-dimensional world that is completely computer-generated. The user is usually given a stereo image, for instance by stereo glasses, and the user's head is tracked so that the image can be rendered for the user's viewpoint (Figure 1). Direct 3D interaction with objects is also possible. Thus, we expected immersive VR would be a good technology for the cell-counting task because head-tracked stereo viewing would ease both seeing the dense cluster of cells and annotating already counted cells. However, there are many variations of immersive VR systems and we could not predict which combination of system attributes (e.g., field of view, scale of data) would be "best". This study was aimed at exploring these issues.

For technical reasons, the direct visualization of confocal data sets (Figure 2a) with volume rendering in virtual environments does not render at interactive frame rates (15+ frames per second). The confocal data sets can be converted to isosurface data sets (Figure 2b), which typically render at higher frame rates but are not ideal for a user study because of visual artifacts, such as overlapping cells. For our user study, we simulated a simplified isosurface confocal data set by computing a number of randomly placed, non-intersecting spheres of varying radii (Figure 2c). Our collaborators from the Biology Department helped us tune parameters to make our simulated data sets as much like real data sets as possible.

Our visualization system runs in a fish tank VR environment (Figure 1a) and also in the Cave, which is a CAVE-like virtual reality system (Figure 1c). The Cave can be configured to show images only on its front wall, making it a single-wall display (Figure 1b).

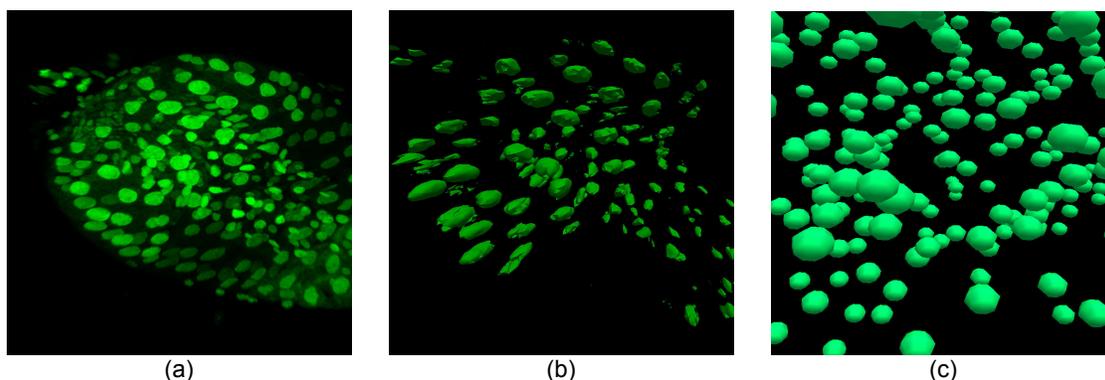


Figure 2: (a) Volume rendering of confocal images, (b) isosurface of confocal images, (c) spheres used in user study.

In our study users used cone-shaped markers to mark spheres while the computer tracked the total number of markers placed. In the Cave, we scaled the data set to three different sizes; at the single wall and the fish tank we used only the smallest size used in the Cave because only that size would fit on the smaller screen¹. Our hypothesis

¹ We found in a pilot study that users always reduced the data set scale until the whole data set fit on the screen, even though using a larger scale might have made individual cells easier to see and mark. Studying these two cases further might be interesting.

was that there was no difference in user performance between the environments because they all used the same input devices and ran the same visualization software.

Section 2 of this paper discusses previous related work. Section 3 briefly describes the software system used in the three virtual environments. Section 4 gives a detailed description of our study design and Section 5 discusses the results of the study.

2 Related Work

Most studies of the effectiveness of virtual environments at different levels of immersion have compared either extreme opposites, like a CAVE-like environment and a standard desktop computer. Arns, Cook & Cruz-Neira (1999) compared a desktop and Cave application for analysis of high-dimensional data. Gruchalla (2004) compared an application for oil well-path planning run on a stereo-enabled desktop and a Cave, but none of the hardware variations between. Pausch, Proffitt & Williams (1997) tested whether users perform measurably better when controlling the viewpoint with their head rather than their hand, but this study involved no placement task by the user. Robertson, Czerwinski & van Dantzich (1997) repeated this experiment with a desktop computer and ran a user study with a wide-screen display; neither experiment involved stereo or head-tracking. Ruddle, Payne & Jones (1999) compared a desktop and HMD for a navigation task through virtual buildings of more than 70 rooms, but the task did not involve actively interacting in 3D with the virtual environment, merely navigating in it. Slater, Usoh, Linakis & Kooper (1996) compared an HMD and TV display for reproducing on a real chessboard the state of a virtual tri-dimensional chess board after a sequence of moves was witnessed in a virtual environment. While subjects were required to touch in 3D a highlighted chess piece before it moved, the subject's ability to interact with it in 3D was not measured.

Other related work has compared different screen sizes and resolutions of displays but without stereo, head- or hand-tracking. Czerwinski et al. (2003) studied the potential benefits of display surface size for complex, multi-application 2D desktop office work. Tan, Gergle, Scupelli & Pausch (2003) compared the performance of subjects who used a large projected wall display and those who used a standard desktop monitor. The visual angle was held constant by adjusting the distance between the viewer and display, but neither display used stereo viewing.

Our study compares three virtual environments at the high end of immersion (CAVE, single wall, and fish tank) because it is not obvious which environment will be more effective for our task. In fact, a related study (Demiralp, Laidlaw, Jackson, Keefe & Zhang, 2003) found that for viewing and rotating a data set the users preferred the fish tank to the CAVE. However, in addition to viewing and rotating a data set, our task also involves selecting 3D positions. We tried to make results among the environments we used comparable by varying a single parameter between each display type (e.g., field of view and apparent data set size), and in this way our work differs from all the work described above.

The Crumbs system (Brady et al., 1995; evaluated in Swartz, Thakkar, Hix & Brady, 1999) is an early CAVE volume visualization system that has a marking technique similar to ours but uses direct volume rendering and has low frame rates. The Crumbs project has been discontinued, and to our knowledge the software has never been adapted to other virtual environments than a CAVE.

Bowman, Datey, Ryu, Farooq & Vasnaik (2002) compared a CAVE-like environment to a head-mounted display (HMD) for a navigation task. Unlike our study, their interaction involved not 3D selection but navigation with a 2D joystick. For similar reasons as Swan, Gabbard, Hix, Schulman & Kim (2003), we used a real-world task in our formal user study. However, Swan et al. (2003) did not use a fish tank VR environment.

3 Visualization System

We implemented our software system in our four-walled Cave with 8'×8' walls and for a one-monitor fish tank VR environment with a visible screen size of 15.6"×12.1". Both environments use Polhemus trackers and Wanda input devices. The Wanda is a hand-held input device (wand) with three buttons and a push-ball that reports relative force in two dimensions. We track the user's head and dominant hand. Both systems provide active stereo with shutter glasses and are equipped with Nvidia Quadro FX 3000 graphics cards. The Cave uses the 3000G version of these

cards to synchronize stereo across the Cave’s walls. The resolution is 1024×768 pixels per screen in both environments, and in both environments the frame rate was ~30fps with no markers, to ~20 with full markers.

Our software runs in the Cave and at the single wall and the fish tank with only minor adaptations to the different screen configurations. User-study subjects could manipulate position and orientation of the data set and place markers.

Manipulation: The data-set manipulation paradigm assumes that the data set is small enough that users can both naturally adjust their vantage point through head movements and move the data set relative to their position. In addition to head-tracked viewpoint adjustment, a data set can be moved and rotated by pointing at it and clicking the left button on the wand. While the button is depressed, the data set “sticks” on the pointer as if skewered by a fork. Repeated grabbing and moving allows its placement anywhere in the virtual world. Our wand device also has a trackball that rotates (or “tumbles”) the data set about its center when a force is applied.

Placing Markers: Cone-shaped yellow markers (Figures 3b and 4) are placed in the data set to mark a sphere. A sample marker (cursor) is permanently shown at a distance of 0.7’ from the wand and connected to it by a white line, so that the user knows where new markers will be placed. The marker’s position and orientation are fixed in the wand’s coordinate system and its tip points down when the wand is held button-side up. When the middle wand button is pressed, a copy of the sample marker is placed in the data set at exactly the same position and orientation as the sample marker.

4 User Study

We recruited 12 undergraduate and master’s students (10 male, 2 female) who had little or no previous experience with virtual environments and no biological domain knowledge related to the task. Subjects spent a total of about 1-1/2 hours in our lab and were paid for their efforts. Their average age was 21.1 years (standard deviation: 3.1). All but one subject were right-handed. The left-handed subject preferred to operate the wand with his right hand because he is used to using the desktop mouse with the right hand as well.

4.1 Task

To generate a data set for the study, we randomly computed positions and diameters for 250 green spheres and placed them in an invisible cube of 1’ edge length. The algorithm ensures that the spheres do not overlap. Their radii are constrained to be between 0.28 and 0.33 inches, determined empirically with our Biology Department collaborators. All spheres are placed so as to lie entirely within the 1’ cube. We calculated six such data sets to use throughout the study. The 1’ cube was scaled by a factor of 1, 4, and 7 to create the small, medium, and large data sets used in the Cave. The fish tank and single-wall environments both used the data set at scale factor 1.

Our dependent variable was the number of markers placed on the spheres within two minutes. In the virtual environments, the subjects were instructed to place markers on the spheres so that the tip of a marker was located within a sphere. Only one marker should be placed on a sphere. Once placed, markers could not be deleted or moved. The data set could be rotated and moved but not otherwise scaled or manipulated.

Table 1: Testing conditions.

Condition	CS	CM	CL	SW	FT
Environment	Cave, 4 walls	Cave, 4 walls	Cave, 4 walls	Cave, 1 wall	Monitor
Posture	standing	standing	standing	standing	sitting
Width of data set	1’	4’	7’	1’	1’
Marker length	0.7”	2.8”	4.9”	0.7”	0.7”

The study had five conditions (see Table 1), three of which were tested in the four-wall Cave, one at the single-wall display, and one at the fish tank. We randomized the order of conditions to counterbalance ordering effects. The three conditions in the Cave differed with respect to the scale of the data set. For each scale, we used a marker size about twice the size of the target spheres. Condition 1 (CS) used the data set’s original size, which was 1 ft³; the

markers were 0.7" long. Under condition 2 (CM) the cube had a 4' edge length and the markers were 2.8" long. Under condition 3 (CL) the cube was 7' on each edge and the markers were 4.9" long. Note that the markers scaled proportionally to the size of the data set, and thus to the spheres. Under all five conditions the marker tip of newly placed markers was 0.7' from the front end of the wand, and they were visually connected by a virtual white stick. At the single wall and the fish tank, data set and markers had the same physical sizes as under condition CS. We abbreviate the fish tank VR environment as FT and the single wall as SW.

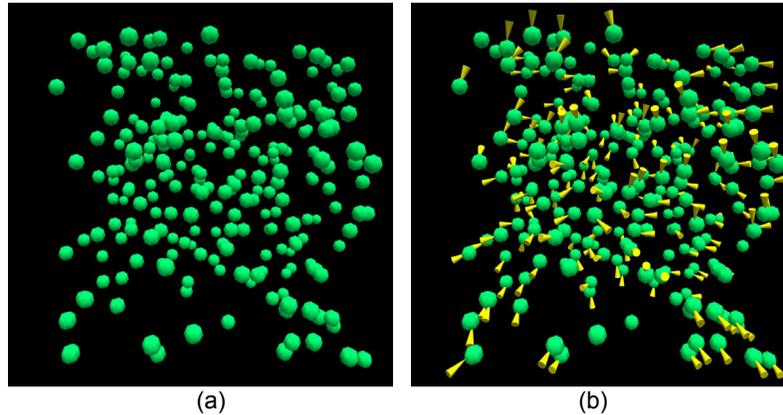


Figure 3: (a) Unmarked and (b) pre-marked data sets.

Because we were interested in the rate of marking, it was not necessary for subjects to mark the entire data set. However, we expected that marking rates might vary over time—in particular, as more cells were marked the task might become dominated by visual search for the next unmarked cell. Time constraints also limited the number and length of conditions we could use. Consequently, instead of having the subjects mark all 250 spheres, we divided the marking task in two halves: we started with an unmarked data set (UM, see Figure 3a) and gave the subjects two minutes to mark as many spheres as they could. Then we interrupted the experiment, removed the markers the subject had placed, and loaded spheres in the same configuration, but with 210 of them pre-marked (PM, see Figure 3b). The subject had another two minutes to mark the remaining 40 spheres.

For the user study we computed six data sets with randomly placed spheres and manually pre-marked each data set. The specific data set used in a given condition was selected randomly.

4.2 Protocol

At first the subjects filled out an online questionnaire asking for statistical information like age, gender, and previous experience with 2D and 3D software. Then they went to the appropriate environment and started the first practice session.

In the practice session, the subjects worked with one of the six data sets; this data set was subsequently removed from the trial pool. The experimenters walked the subject through the trial conditions for the testing condition and made sure they understood the user interface and tried out different ways to use the devices and to position themselves and the data set. For instance, at the FT the subjects were encouraged to adjust the seat, try different ways of holding the wand, and rest the elbow on the armrests or on the table. We asked the subjects to tell us when they felt comfortable using the system and were ready to start the trials.

The trial runs were timed to precisely two minutes. During the trials, the experimenters interrupted the subjects only if they made a system error, such as confusing the buttons for navigation and marker placement.

After the last trial of the last condition, the subjects filled out an online questionnaire concerning their preference of conditions and environments and their ideas on how to improve the systems. The session ended with a short debriefing in which we summarized the study goals.

5 Results

Our user study produced four types of results: task performance (rate and accuracy with which markers were placed), questionnaire answers, interaction logs, and observations we made during the tests. We present these results below.

5.1 Task Performance

In the statistical analysis we distinguish three cases of placed markers. A “hit” is a marker placed inside a sphere. A “miss” is a marker placed outside a sphere, but no more than 1.5 sphere radii away. We found empirically that beyond this distance it is not clear to which sphere a marker belongs. We considered such markers “outliers”, probably placed when users confused the navigation and marker placement buttons or accidentally pressed the marker button, and ignored them in our statistical analysis.

Figure 4 shows the number of markers the subjects placed accurately on spheres (hits). In all cases the subjects placed about 3-4 times more markers in the UM trials than in PM.

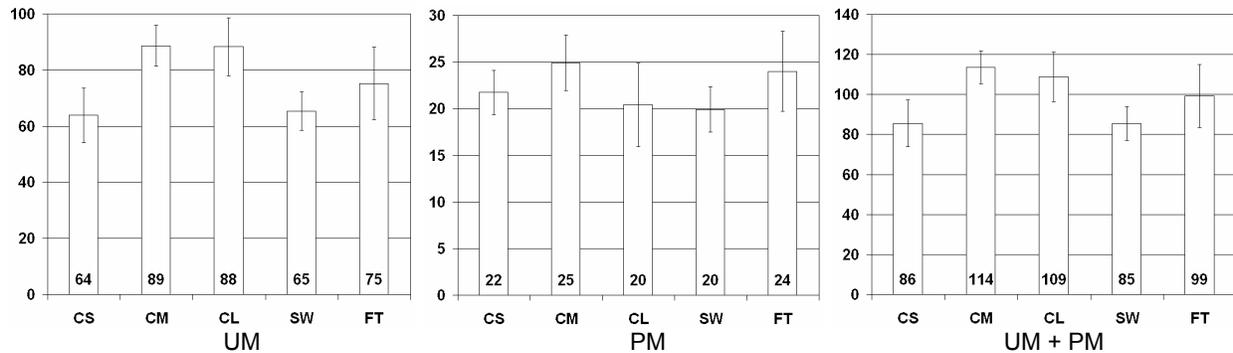


Figure 4: Averages of accurately placed markers. The error bars shown are 95% confidence intervals.

Our null hypothesis was that there is no difference in how many spheres were marked in the various environments. We analyzed the results of our study with SPSS for Windows 11.5.0 and did a within-subjects ANOVA analysis with Bonferroni correction for multiple pairwise comparisons, looking for statistical significance at the 5% level.

For the number of accurately markers placed markers in UM and PM combined, our analysis finds statistically significant differences between CS and CM ($p < 0.001$), CS and CL ($p = 0.013$), CM and SW ($p < 0.001$), and CL and SW ($p = 0.012$). There is a trend that marking performances in CS and FT are different, but not in CS or SW. The latter was not surprising because most users did not use the additional walls in CS but used almost exclusively the Cave’s front wall, so that the CS environment was nearly identical to SW.

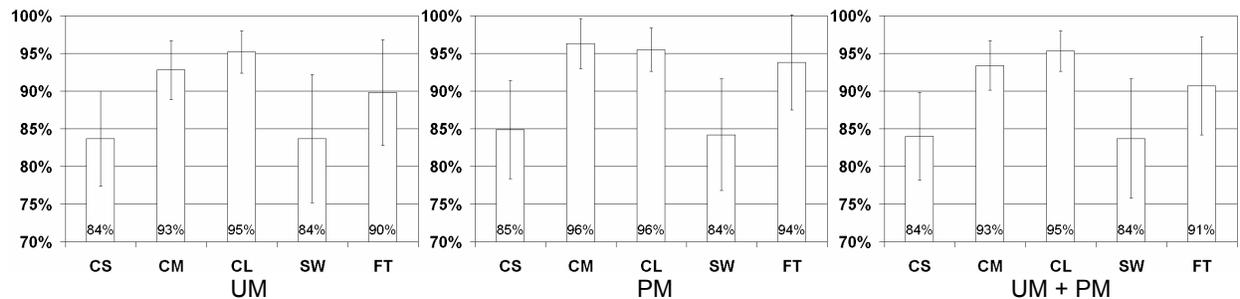


Figure 5: Average placement accuracies. The error bars shown are 95% confidence intervals.

Figure 5 shows the accuracy with which users marked spheres in the environments. Overall (UM+PM) we get the following statistically significant differences: CS and CM ($p=0.002$), CS and CL ($p=0.004$), CM and SW ($p=0.011$), and CL and SW ($p=0.017$). There is a trend that accuracy in SW and CS were lower than in any other condition, and that among the Cave conditions, the accuracy in CM and CL was significantly higher than in CS.

5.2 Questionnaire

After the study, the subjects filled out an online questionnaire concerning their thoughts on the virtual environments and the placement of the markers, and answered some more general questions such as how to improve the marking task and whether they had experienced discomfort during the study. Here we summarize the more interesting answers.

Table 2: Participants' estimate of marking rate. For each condition the table lists how many subjects listed it, how many of these correlated with the actual task performance (#correct), and how many did not (#incorrect).

	Fastest rate				Slowest rate			
	Cond.	#mentions	#correct	#incorrect	Cond.	#mentions	#correct	#incorrect
Unmarked spheres	CM	8	5	3	FT	5	4	1
	CL	2	1	1	SW	4	3	1
	CS	1	0	1	CS	3	2	1
					CL	1	0	1
Pre-marked spheres	CM	6	0	6	FT	4	0	4
	CS	3	1	2	CS	4	0	4
	CL	1	0	1	SW	3	1	2
	FT	1	0	1	CL	2	0	2

- Question: "Please rank the sizes of data sets in the Cave according to your preference."
For unmarked spheres, six subjects ranked the large scale highest and five preferred the medium scale. For pre-marked spheres, only three subjects preferred the large scale but seven preferred the medium scale. All but one subject ranked the small scale lowest for the UM trial; for PM nine subjects ranked the small scale lowest.
- Question: "Please rank the conditions (CS, CM, CL, SW, FT) according to how fast you think you marked."
Table 2 compares what subjects thought were the fastest and the slowest conditions to their actual performance. If subjects ranked two conditions evenly, we counted both conditions. There were more correct answers for unmarked than for pre-marked spheres.

Table 3: Count of subjects regarding preferred stick lengths for CS, CM, CL, SW, and FT.

	CS	CM	CL	SW	FT
Length was fine	7	5	4	7	5
Longer preferred	1	3	4	1	4
Shorter preferred	4	4	4	4	3

The following two questions concern the virtual stick connecting the wand to the tip of the marker. Its length determines the distance between the user's hand and where new markers were placed. We wanted to know what the subjects thought about this length.

- Question: "Did the stick length work for you? If not, what length would you have preferred?"
Table 3 summarizes the answers to this question. Overall, about half the subjects thought the stick length was fine. One third would have preferred a shorter stick for CS and SW. For CM, CL, and FT, the majority would have preferred a stick of a different length, but they were almost evenly split as to whether it should have been longer or shorter.
- Question: "Would you have preferred a stick whose length you could change continuously?"

Four subjects would have preferred an adjustable stick, three would not, and five subjects were worried that it would have been too complicated and/or time-consuming to use.

5.3 Interaction Logs

We logged data from the sensors during our trials. One trend we observed was that the distance subjects moved their head was directly related to the scale of the data set (Figure 6). That is, they moved more as the size of the data set increased.

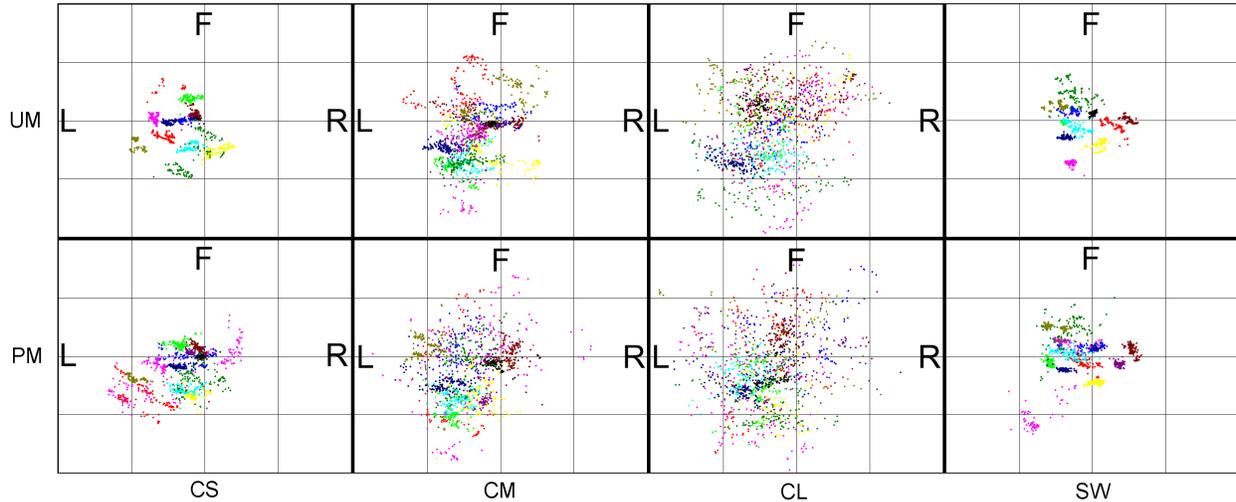


Figure 6: Top-down views of head positions of all subjects, distinguished by color. Temporal resolution: one sample per second. Columns from left to right: CS, CM, CL, SW. Rows from top to bottom: UM, PM. All plots cover an 8'x8' area (the small boxes are 2'x2'). The letters L, F and R indicate the Cave walls left, front, and right, respectively. Condition FT is not shown because the subjects sat on a chair and moved much less than in the Cave.

Another interesting trend was that subjects spent more time tumbling (with the trackball) and moving (with the left wand button) the data set in the PM than in the UM trials, probably because tumbling helped find spheres that were unmarked, and moving helped reach those spheres. Figure 7 shows the percentage of time subjects spent tumbling and moving the data in CS, CM, CL, SW, and FT. The difference in percentages for UM and PM was statistically significant for tumbling, but not for moving. We think that the large confidence intervals for moving indicate that subjects used different moving strategies. For example, rotating the data set was possible by tumbling or repeated moving.

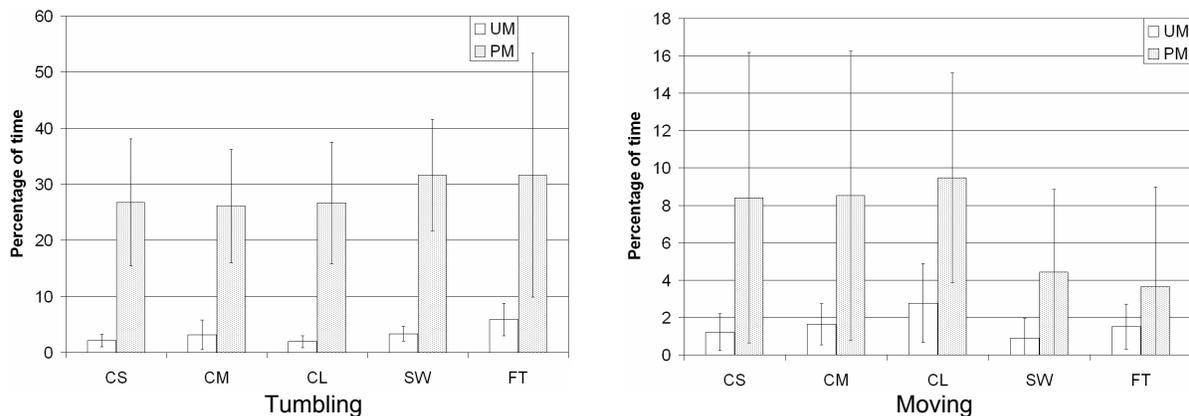


Figure 7: Percentage of time tumbling and moving. The error bars shown are 95% confidence intervals.

5.4 Discussion

We made several observations about user strategy during the task. First, “tumbling” the data set with the trackball wand control played a central role in the PM trials (see Section 5.3). Tumbling the data set more quickly gave users many views of the spheres from different angles and let them identify unmarked spheres more easily than if they used just a static viewpoint or even the varying viewpoints allowed by head-tracking alone.

Second, even though during training subjects were encouraged to walk around the 1' data sets in the Cave, they chose to stand in an essentially fixed position so that the data set projected only onto the front wall. Consequently, it appeared that the CS condition was effectively equivalent to the single wall in that subjects got no benefit from the additional display space on the left, right, and floor displays of the Cave. This hypothesis is supported by the number of markers subjects placed in CS and SW (see Section 5.1).

At the fish tank users could rest their elbows on the chair's armrests or on their thighs. However, even though both these positions were practiced in the training, only three subjects used this in the trials. One subject rested elbow on thigh during both trials; another subject rested the elbow on the armrest during PM. One subject used both hands to hold the wand throughout UM and PM, and rested both elbows on the thighs during PM. One subject used both hands to hold the wand during UM but only one hand during PM, and did not rest the arms throughout the FT trial.

Another point during the training was to make subjects aware that the wand buttons and the trackball could be operated either with the thumb or, when turned 90 degrees, with the forefinger and optionally the middle finger. All subjects chose to operate the buttons and the trackball with the thumb.

Among the Cave trials and the single wall, two conditions maximized quantity and quality in marker placement. Users placed more markers in CM and CL than in CS and SW, and they were also more accurate in CM and CL than in CS and SW.

Subjects marked at lower rates when the data set was scaled large versus small very possibly because of occlusion. In particular, markers could be occluded by spheres when they were on the far side of a sphere, so that subjects had to view all sides of a cell to be confident that it hadn't already been marked. Tumbling a small dataset about its origin could quickly reveal an unmarked cell. When the data set was scaled up (to 4' or 7' in each dimension), the subject had to look in more directions due to the larger field of view to discover an unmarked cell.

The data collected also showed that, contrary to our expectations, subjects did not perform equally well in the SW and FT environments. In particular, subjects marked more spheres in FT than SW. Possible factors in this difference include standing versus sitting; the brightness/contrast of the CRT projector versus CRT monitor; and the distance to the display (i.e., the difference between where subjects focused and the convergence of their eyes on the data set appearing in front of the display).

6 Conclusion and Future Work

Our objective in this work was to determine how Cave, single wall, and fish tank environments differ with respect to user performance in a marking task. Our results had four significant findings: a) in the Cave the subjects preferred medium sized or large spheres over small spheres, b) when only a few of the targets have been marked, larger spheres were marked faster than small spheres, c) large spheres are marked most accurately, and d) our wall display is not comparable to our fish tank virtual reality environment when the spheres are small. Additionally, occlusion and larger field of view inhibited performance in the Cave more than at the fish tank when the task was dominated by visual search.

Possible future work includes developing and user-testing techniques to handle occlusion for scaled-up data sets, for instance fixing the orientation of the markers relative to the viewer so that they are never occluded by the spheres, or using differently shaped markers. Additionally, a user study could focus on identifying the factors contributing to the performance differences between fish tank and single wall. A future study should also compare the results from this study with one that uses actual confocal microscope data.

7 Acknowledgments

We thank Professor Mel Slater of UCL for early discussions about this paper and design and analysis of the user study. We also thank Professor Kristi Wharton and Ph.D. candidates Lorena Soares and Erdem Bangi from the Department of Molecular Biology, Cell Biology and Biochemistry at Brown University for their collaboration on this work. Our software uses the VRPN library, which is supported by the NIH National Research Resource in Molecular Graphics and Microscopy at the University of North Carolina at Chapel Hill. This work was partially supported by DOE award DE-FG02-03ER25582, LLNL Research Subcontract No. B527302, and NSF (CCS-0086065). The user study reported here has been approved by the Institutional Review Board (IRB) of Brown University.

REFERENCES

- Arns, L., Cook, D., & Cruz-Neira, C. (1999). The Benefits of Statistical Visualization in an Immersive Environment. *Proceedings of the IEEE Virtual Reality Conference*, Houston, TX, pp. 88-95.
- Bowman, D.A., Datey, A., Ryu, Y.S., Farooq, U., & Vasnaik, O. (2002). Empirical Comparison of Human Behavior and Performance with Different Display Devices for Virtual Environments. *Proceedings of Human Factors and Ergonomics Society Annual Meeting*, Baltimore, pp. 2134–2138.
- Brady, R., Pixton, J., Baxter, G., Moran, P., Potter, C.S., Carragher, B., & Belmont, A. (1995). Crumbs: A Virtual Environment Tracking Tool for Biological Imaging. *IEEE Symposium on Frontiers in Biomedical Visualization Proceedings*, pp. 18–25.
- Czerwinski, M., Smith, G., Regan, T., Meyers, B., Robertston, G., & Starkweather, G. (2003). Toward Characterizing the Productivity Benefits of Very Large Displays. *Proceedings of International Conference on Human Computer Interaction (INTERACT)*, IFIP Press, pp. 252–255.
- Demiralp, C., Laidlaw, D.H., Jackson, C., Keefe, D., & Zhang, S. (2003). Subjective Usefulness of CAVE and Fish Tank VR Display Systems for a Scientific Visualization Application. *IEEE Visualization '03 Poster Abstract*, Seattle, WA, 2003.
- Gruchalla, K. (2004). Immersive Well-Path Editing: Investigating the Added Value of Immersion. *Proceedings of the IEEE Virtual Reality Conference*, Chicago, IL, pp. 157-164.
- Pausch, R., Proffitt, D., & Williams, G. (1997). Quantifying Immersion in Virtual Reality. *ACM SIGGRAPH 97 Proceedings*, pp. 13–18.
- Robertson, G., Czerwinski, M., & van Dantzich, M. (1997). Immersion in Desktop Virtual Reality. *Proceedings of the Symposium on User Interface Software and Technology (UIST)*, Banff, Alberta, Canada, pp. 11-19.
- Ruddle, R.A., Payne, S.J., & Jones, D.A. (1999). Navigating Large-Scale Virtual Environments: What Differences Occur Between Helmet-Mounted and Desk-Top Displays? *Presence: Teleoperators and Virtual Environments*, 8, pp. 157–168.
- Slater, M., Usoh, M., Linakis, V., & Kooper, R. (1996). Immersion, Presence and Performance in Virtual Environments: An Experiment with Tri-Dimensional Chess. *Proceedings of ACM VRST*, pp. 163–172.
- Swan II, J.E., Gabbard, J.L., Hix, D., Schulman, R.S., & Kim, K.P. (2003). A Comparative Study of User Performance in a Map-Based Virtual Environment. *IEEE Virtual Reality '03 Proceedings*, Los Angeles, CA, pp. 259–266.
- Swartz, K., Thakkar, U., Hix, D., & Brady, R. (1999). Evaluating the Usability of Crumbs: a Case Study of VE Usability Engineering. *Proceedings of the 3rd International Immersive Projection Technologies Workshop*, May '99, Springer-Verlag, pp. 243–252.
- Tan, D.S., Gergle, D., Scupelli, P.G., & Pausch, R. (2003). With Similar Visual Angles Larger Displays Improve Spatial Performance. *Proceedings of ACM CHI'03*, pp. 217–224.