Evaluating Depth Perception of Volumetric Data in Semi-Immersive VR

Isaac Cho
UNC Charlotte, USA
icho1@uncc.edu

Wenwen Dou
UNC Charlotte, USA
wdou1@uncc.edu

Zachary Wartell
UNC Charlotte, USA
zwartell@uncc.edu

William Ribarsky
UNC Charlotte, USA
ribarsky@uncc.edu

Xiaoyu Wang
UNC Charlotte, USA
xwang25@uncc.edu

ABSTRACT
Displays supporting stereoscopy and head-coupled motion parallax can enhance human perception of complex 3D datasets. This has been studied extensively for datasets containing 3D surfaces and 3D networks but less for so volumetric data. Volumetric data is characterized by a heavy presence of transparency, occlusion and highly ambiguous spatial structure. There are many different rendering and visualization algorithms and interactive techniques that enhance perception of volume data and these techniques' effectiveness have been evaluated. However, the effort of VR display technologies on perception of volume data is less well studied. Therefore, we conduct two experiments on how various display conditions affect a participant's depth perception accuracy of a volumetric dataset. A demographic pre-questionnaire also allows us to separate the accuracy differences between participants with more and less experience with 3D games and VR technologies. Our results show an overall benefit for stereo with head-tracking for enhancing perception of depth in volumetric data. Our study also suggests that familiarity with 3D games and VR type technology affects the users'ability to perceive such data and affects the accuracy boost due to VR displays.

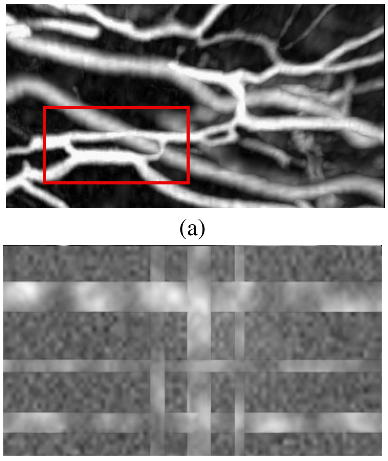
Keywords
Depth Perception, Stereoscopic, Head-Tracking, Volumetric dataset.

Categories and Subject Descriptors
I.4.8 [**Scene Analysis**]: Depth cues, Motion, Stereo, Tracking; I.3.7 [**Three-Dimensional Graphics and Realism**]: Virtual reality; I.2.10 [**Vision and Scene Understanding**]: 3D/stereo scene analysis

General Terms
Experimentation

1. INTRODUCTION
Previous research has demonstrated the utility of computer displays that provide stereoscopy and head-coupled motion parallax for enhancing human perception of complex three-dimensional datasets.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. AVI '12, May 21-25, 2012, Capri Island, Italy
Copyright ©2012 ACM 978-1-4503-1287-5/12/05... $10.00



(a)

(b)

Figure 1: Similarity comparison between our artificial dataset and actual MRI blood vessel scan. (a) Maximum Intensity Projection rendering of blood vessels [9]. (b) Our artificial dataset

This includes fully immersive displays such as CAVE's and HMD's and semi-immersive displays such as desktop VR and the virtual workbench [2]. For example, studies by Ware et al. examine the effect of the stereoscopic and kinetic depth for understanding 3D networks which are represented by tubes or lines [15]. Their results demonstrate improved user performance at finding paths in a complex 3D network when using stereoscopy and structure-from-motion.

One would expect similar results for volumetric data. Further, the addition of VR display technology could be especially important with time-varying volumetric datasets that are viewed in real-time where extensive preprocessing for optimizing transfer functions and volume rendering parameters is not possible. An example would be real-time, streaming doppler weather radar data [16]. With the increasing affordability of semi-immersive VR displays and GPUs capable of advance volume rendering, there is a pressing need to quantify the effectiveness of stereopsis and structure-from-motion on volumetric data and also to quantify how these display parameters interact with other volumetric rendering conditions. The large number of potential display hardware and rendering variables make such evaluations particularly challenging. In this paper, we take a first step towards this by using a fixed set of volumetric rendering parameters–chosen through pilot studies–and then by varying the

VR display hardware.

We present two experiments on the effects of stereoscopy and head-tracking on the perception of depth ordering of volumetric objects. The experiment is motivated by datasets such as the MRI scan of blood vessels shown in Figure 1a reproduced from [8]. As is typical of volumetric data, this dataset is has a heavy presence of transparency, occlusion and a highly ambiguous spatial structure. In Figure 1a, it is particularly challenging to determine the depth order of the blood vessels inside the red square. As discussed in [8], the volume rendering technique used makes it appear that the square-shaped vessel is in front of the diagonal one. However, in fact the diagonal one is in front of the square-shaped one.

We mimic this type of ambiguity by generating controlled experimental datasets such as Figure1b where the user's task is to determine the depth ordering of various occluding, transparent cylinders. The subjects view the datasets under a variety of display conditions including combinations of stereoscopic display, head-tracking, and small object rotations. We present a set of cylinders of various size and opacities and depth orderings to mimic datasets such as Figure1a but in a experimentally controlled manner. The two experiments include a depth ordering task, in which participants must understand the full depth ordering between 6 volumetric cylinders, and a depth discrimination task, in which participants must distinguish the relative order of just 2 cylinders within a limited exposure time (2 sec). In addition, the results also differentiate between experienced and less-experienced users with respect to use of 3D games and VR related technology. Results from both groups show an overall benefit for stereoscopy with head-tracking in enhancing depth perception of volumetric data. More interestingly, our study also suggests that familiarity with 3D games and VR related technology affects the user's depth perception accuracy.

2. PREVIOUS WORK

Displays supporting stereoscopy and head-coupled motion parallax can enhance human perception and understanding of complex 3D datasets. Structure-from-motion is sub-classified into motion parallax due to observer motion and the kinetic-depth-effect due to object rotation [15]. Prior experiments with surface and 3D network datasets show that stereoscopic display can aid depth perception when either the visual stimuli lacks other depth cues, as can occur in teleoperator environments or in simplistic computer graphic rendering, or when the visual stimuli contains a high depth complexity as measured by many occlusions [4]. For example, Ware et al. examine the effect of stereoscopy and the kinetic depth effect on a person's understanding of a 3D network which is represented by 3D tubes or lines [15]. They demonstrate great benefit for stereoscopic and kinetic depth when a user must find paths in between nodes in a complex 3D network.

The most detailed evaluations of how to improve a user's perception and spatial understanding of volumetric data focus on comparing different rendering techniques and/or different transfer functions [12, 13, 3, 1]. A few authors have studied the effectiveness of stereoscopic display for volumetric data [6, 5, 9, 8]. Our experiment complements these works by focusing more on low-level occlusion and depth ordering perception under a wider variety of display conditions.

3. ENVIRONMENT

We test the effectiveness of a semi-immersive VR system on depth perception of volumetric data. We examine the effects of display environment on two tasks: a depth ordering task and a depth discrimination task. Both tasks use a fish-tank VR setup which consists of a stereo display (Samsung Sync Master 2233RZ) and a tracked pair of Nvidia stereo glasses. The tracker is a Polhemus Fastrak. The display refresh rate is 120 Hz. The participant sits roughly 60 cm in front of the screen. The independent experimental variable is the display condition with a combination of stereoscopy, head-tracking and/or a small object rotation for a kinectic depth effect. Our synthetic volumetric dataset contains six overlapping cylinders of varying diameters and transparency. Three cylinders are vertical and three are horizontal. The voxel resolution is 512 x 256 x 256. Perlin noise is used for the internal texture of the cylinders and the texture of a large background polygon. The background polygon approximates the visual effect of having the cylinders embedded in a more complex volumetric environment.

We chose a high quality GPU-based ray-casting rendering available as a OpenSceneGraph plugin. Compared to other rendering techniques, such as per pixel lighting and Maximum Intensity Projection (MIP) rendering, the GPU based ray casting technique yields more accurate depth cues [11].

We added a black polygon with a square hole in front of the volumetric data to act as a window to hide the edges of the cylinders. This was necessary to block the view of the cylinder edges. Being able to see the cylinder ends made the depth ordering task trivial. Alternatively, simply scaling up the voxels' rendered size (to extend the cylinders off screen) was not helpful because aliasing artifacts become too visible. (Increasing the volumetric resolution exceeded the renderer's capabilities).

Following a previous study [9], we fix the data parameters, such as the Alpha gradient and transparency, to represent a reasonable clear outline of each semi-transparent cylinder (alpha = 0.9, transparency = 0.2, density = 0.025). Note that, to isolate the effects of stereo and structure-from-motion, we do not allow users to interactively adjust the transfer function in our study even though many previous studies demonstrated its utility in depth perception [7, 10].

4. EXPERIMENT DESIGN

Our two experiments examine the effect of stereoscopy and/or head-tracking on the perception of volumetric data. Experiment 1 has four display conditions and Experiment 2 has six; the details of each condition are described below. We use a within-subject design with repeated measures. Each subject is randomly assigned a sequence of conditions using latin squares. The measures in our experiment are response time and accuracy. Before each experiment, the participant signs an IRB consent form and provides demographic information such as gender, academic major and degree being sought and reports their familiarity with stereoscopic display, VR technology and gaming. After the experiment, each participant fills out a post-questionnaire regarding the their confidence with their answers to the task's spatial questions and their opinions on various aspects of the volumetric dataset. We use 7-point Likert scales where applicable.

We recruited 28 participants (16 for Experiment 1 and 12 for Experiment 2). 16 of them are undergraduate students and 12 are graduate students. 14 participants major in Computer Science and 14 participants are of other majors including psychology, nursing history and fashion design. All participants have 20/20 or corrected 20/20 vision. We provide a tutorial to familiarize the participants with the stereo display and head-tracking hardware.

5. EXPERIMENT 1: DEPTH DISCRIMINATION

The first experiment evaluates how stereoscopy and structure-from-motion affect users when performing a depth discrimination task. The participants must determine which of two cylinders, one horizontal and the other vertical, is in front of the other. Although traditionally in psychological studies, the exposure time of is usually in the range of a few hundred of milliseconds ([14]), Experiment 1 requires subjects to first locate an intersection of a pair of cylinders based on a provided cue, and then report on the depth relation of the cylinders. Therefore 2 seconds is more appropriate this task [17]. The volumetric dataset actually contains 6 cylinders, but in each trial a pair of cylinders is designated as the target pair for the depth discrimination task. On each trial, the first screen displays a 2D picture which designates which of the 9 intersections of the 6 cylinders is the target pair. Next, the screen displays the volumetric dataset for 2 seconds. Finally the screen displays a menu with three choices: "the horizontal cylinder is in front", "the vertical cylinder is in front", or "I don't know". Figure 2 shows the task procedure. The 2D picture screen is shown on the left, the volumetric data screen in the middle and the question screen on the right. Each participant performs 216 trials.

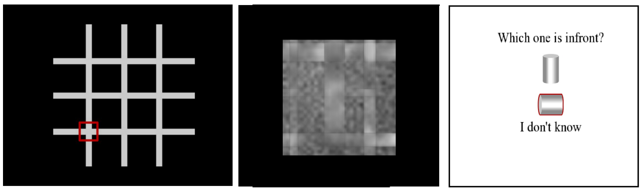


Figure 2: The depth discrimination task's 3 screen images.

Experiment 1 has six conditions. The conditions are: non-stereo without head-tracking (NSNH), stereo without head-tracking (SNH), non-stereo with head-tracking (NSH), stereo with head-tracking (SH), non-stereo with head-tracking simulation (NSHS) and stereo with head-tracking simulation (SHS). The last two conditions were added because in pilot tests some users did not utilize the head-tracking when limited to the 2 second exposure time. In the non-head-tracked conditions, a participant uses a chin rest.

We analyzed the results using ANOVA followed by Fishers̀ least significant difference (LSD) for pairwise comparisons. The factor of this experiment is the condition (NSNH, SNH, NSH, SH, NSHS, SHS) and the dependent variable is accuracy, measured by the percentage of correct depth judgements. A one-way ANOVA did not show a significant effect of display condition on the accuracy. We did expect, however, that the SH condition would perform better than the no stereo, no head-tracking condition. A post-hoc LSD comparison shows that the mean accuracy for condition of SH (M = 0.75, SD = 0.16) is significantly better than the condition of NSNH (M = 0.67, SD = 0.2) with $p = 0.01$.

Our participants were recruited from two pools: students from several computer science laboratories and students from the Psychology department's participant pool. The latter includes students from various majors taking psychology courses. In our case, the majors include history, nursing, and psychology. The participants from each pool were randomly assigned to different experiments and conditions and they participated over the same course of time. However, we found various differences between the two groups. We will refer to these two pools as CS majors and non-CS majors.

We observed that CS major participants appeared more confident with their task performance and more comfortable with using our semi-immersive VR environment. Further, a one-way ANOVA showed that if we ignore display conditions, CS majors average accuracy was 73% versus 64% for non-CS majors ($p = 0.001$). Different from the findings from overall population, the mean accuracy for the condition of NSH (M = 0.79, SD = 0.1) $p = 0.042$ is significantly better than the NSNH condition. Plausibly, these differential effects on CS majors is due to their greater experience with 3D games and VR type technologies which we found when analysing the the questionnaires.

6. EXPERIMENT 2: DEPTH ORDERING

In Experiment 2, participants perform a depth ordering task on the 6 volumetric cylinders. The four display conditions are: no stereo with no head-tracking (NSNH), stereo with no head-tracking (SNH), no stereo with head-tracking (NSH), and stereo with head-tracking (SH). The participants is asked to designated which of 6 cylinders is at a particular position, either the front, the middle, or the back. Each participant undergoes 36 trials per display condition which means 216 trials total. During each trial, the cylinders are rendered with random depth ordering such as in the middle figure of 2. Then one out of the three questions is given. Each cylinder is labelled with a number (1-6) and the participant designates which cylinder is at the queried position by pressing the corresponding number key on the keyboard. In non-head-tracking conditions participants use a chin rest. There was no time limit.

We analyzed the effect of the display condition on the response time and accuracy using ANOVA followed by Fishers̀ least significant difference (LSD) for pairwise comparisons. A one-way ANOVA shows differences in accuracy among the four display conditions. Accuracy is computed as the number of correct answers divided by total number of questions in each trial (36 questions per condition). There is a significant main effect of the condition on accuracy ($F(3, 60) = 6.242, p = 0.001$). Post hoc comparisons using the LSD test indicate that the mean accuracy for condition SH (stereoscopy with head tracking) (M = 0.63, SD = 0.19) is significantly better than the condition of no stereo with head-tracking (NSH) (M = 0.47, SD = 0.11) $p = 0.003$, and no stereo with no head-tracking (NSNH) (M = 0.42, SD = 12) $p = 0.000$. In addition, mean accuracy for condition SNH (M = 0.54, SD = 0.14) is significantly better than condition NSNH with $p = 0.02$. In other words, stereo with head-tracking led to the best accuracy in the depth-ordering task followed by stereo with no head-tracking. Unexpectedly, head-tracking alone (NSH) does not lead to significant improvement in accuracy over the no stereo no head-tracking (NSNH) condition. Also unexpectedly, adding head-tracking to stereoscopy does not improve accuracy compared to stereoscopy alone. Contrary to our expectation, the display condition does not significantly affect response time.

Our observations in Experiment 1 led us to test for any effect of game playing experience. A one-way ANOVA test shows that gaming experience is significantly different between the CS majors (M = 4.8, SD = 0.75) and non-CS majors (M = 2.3, SD = 1.5) $p = 0.005$. We use a 4x2 factorial analysis of variance to evaluate the effects of the display condition and CS major/non-major pool on depth-ordering accuracy. Results indicate a significant main effect for the student pool factor, F(1, 56) = 9.86, p = 0.003. As hypothesized, the accuracy of CS major observers is better across all conditions than the non-CS major subjects. The interaction between these two factors was not significant.

Based on this finding, we performed separate one-way ANOVA tests on the two participant groups. The effect of display condition on accuracy and response time among non-CS major observers is not statistically significant. However, there is a significant effect of the condition on accuracy ($F(3, 28) = 10.115, p = 0.000$) among the CS major participants. LSD post-hoc comparisons indicate that the mean accuracy for condition SH (stereo with head- tracking) (M = 0.72, SD = 0.12) is significantly better than all other three conditions, namely stereoscopy with no head-tracking (SNH) (M = 0.57, SD = 0.1) $p = 0.005$, no stereo with head-tracking (NSH) (M = 0.52, SD = 0.08) $p = 0.000$, and no stereo with no head-tracking (NSNH) (M = 0.46, SD = 0.08) $p = 0.000$. In addition, mean accuracy for condition SNH is significantly better than condition NSNH with $p = 0.03$. The findings share similarities with the overall analysis in that stereo with head-tracking led to the best accuracy on the depth-ordering task followed by stereo with no head-tracking. However, the findings differ in that for CS majors, combining head-tracking with stereoscopy did significantly improve accuracy compared to stereoscopy alone.

7. DISCUSSION

Our results indicate that different semi-immersive VR display conditions affect how well a user perceives depth within a volumetric dataset. More specifically, our results support the hypothesis that stereo with head-tracking significantly improves depth perception of our volumetric dataset. The SH conditions generally outperform all other combinations. Results of the depth ordering experiment echoes previous research results that stereoscopic display enhances the depth perception.

Plausibly, the CS major vs non-CS major differences in Experiment 1 are due to CS major participants' greater gaming experience. Such experience could train them to make better depth judgements by using structure-from-motion cues. For example, in Experiment 1, the CS major participants, head-tracking is better than the no stereo, no head-tracking condition even given the limited exposure time. Further, CS major participants always score higher compared to the non-CS major pool.

In Experiment 1 (depth discrimination) stereo with head-tracking shows the smallest error rate. However, stereo without head-tracking and no stereo with head-tracking are not significantly different. Also, participants choose the 'I don't know' answer the most under the no stereo, no head-tracking condition.

In Experiment 2 (depth ordering) stereo with head-tracking and stereo without head-tracking are more effective than no stereo without head-tracking. However, head-tracking with no stereo is not effective. These results are in line with prior work that showed stereoscopy improved spatial understanding. Interestingly, display condition did not affect response time.

8. CONCLUSION

Our study examined the effect of stereoscopy and head-tracking on depth discrimination and depth ordering tasks for volumetric datasets. Stereoscopy and head-tracking conditions enhance depth perception in both experiments. And we found that stereoscopy by itself improves depth perception in a depth ordering task. However, stereoscopy alone does not aid depth discrimination and head-tracking alone does not benefit participants overall in either experiment. Differential affects were found based on participant major which correlated with gaming experience.

9. ACKNOWLEDGEMENTS

This study was supported in part by a grant from the Renaissance Computing Institute of North Carolina.

10. REFERENCES

[1] C. Boucheny, G.-P. Bonneau, J. Droulez, G. Thibault, and S. Ploix. A perceptive evaluation of volume rendering techniques. *ACM Trans. Appl. Percept.*, 5(4):1–24, 2009.

[2] D. A. Bowman, E. Kruijff, J. J. LaViola, and I. Poupyrev. *3D User Interfaces: Theory and Practice*. Addison Wesley, 2005.

[3] S. Bruckner and E. Gröller. Enhancing depth-perception with flexible volumetric halos. *IEEE TVCG*, 13(6):1344–1351, 2007.

[4] E. T. Davis and L. F. Hodges. *Virtual Environments and Advanced Interface Design*, chapter Human Stereopsis, Fusion, and Stereoscopic Virtual Environments. Oxford University Press, 1995.

[5] D. J. Hancock. *Distributed volume rendering and stereoscopy display for radiotherapy threatment planning*. PhD thesis, The University of Manchester, 2001.

[6] R. J. Hubbold and D. J. Hancock. Stereo display of nested 3d volume data using automatic tunnelling. 1999.

[7] J. Kniss, G. Kindlmann, and C. Hansen. Multidimensional transfer functions for interactive volume rendering. *IEEE TVCG*, 8(3):270–285, 2002.

[8] R. Maciejewski, S. Choi, D. S. Ebert, and H. Z. Tan. Multi-modal perceptualization of volumetric data and its application to molecular docking. In *WHC '05: Proceedings of the First Joint Eurohaptics Conference and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems*, pages 511–514, Washington, DC, USA, 2005.

[9] B. Mora and D. S. Ebert. Instant volumetric understanding with order-independent volume rendering. *Computer Graphics Forum*, 23(3):489–497, September 2004.

[10] H. Pfister, B. Lorensen, C. Bajaj, G. Kindlmann, W. Schroeder, L. S. Avila, K. Martin, R. Machiraju, and J. Lee. The transfer function bake-off. *IEEE Comput. Graph. Appl.*, 21(3):16–22, 2001.

[11] S. Roettger, S. Guthe, D. Weiskopf, T. Ertl, and W. Strasser. Smart hardware-accelerated volume rendering. In *VISSYM '03: Proceedings of the symposium on Data visualisation 2003*, pages 231–238, Aire-la-Ville, Switzerland, Switzerland, 2003. Eurographics Association.

[12] A. J. Stewart. Vicinity shading for enhanced perception of volumetric data. In *Visualization, 2003. VIS 2003. IEEE*, pages 355–362, 2003.

[13] N. A. Svakhine, D. S. Ebert, and W. M. Andrews. Illustration-inspired depth enhanced volumetric medical visualization. *IEEE TVCG*, 15(1):77–86, 2009.

[14] C. Ware. *Information Visualization: Perception for Design; 2nd Edition*. Morgan Kaufmann, 2004.

[15] C. Ware and P. Mitchell. Visualizing graphs in three dimensions. *ACM Trans. Appl. Percept.*, 5(1):1–15, 2008.

[16] Z. Wartell, E. Houtgast, O. Pfeiffer, C. D. Shaw, W. Ribarsky, and F. H. Post. Interaction volume management in a multi-scale virtual environment. In *Advances in Information and Intelligent Systems*, volume 251, pages 327–349. 2009.

[17] Y.-Y. Yeh and L. D. Silverstein. Limits of fusion and depth judgment in stereoscopic color displays. *Hum. Factors*, 32:45–60, January 1990.