P-3: Evaluation of Rendering Algorithms for Presenting Layered Information on Holographic Displays

James Barabas, Quinn Y. J. Smithwick, and V. Michael Bove Jr.
Object-Based Media Group, MIT Media Laboratory, Cambridge, MA, USA

Abstract

This paper proposes evaluation of rendering methods for holographic display allowing a viewer to use monocular information to selectively focus on parts of the display volume. We describe an experiment for comparatively evaluating human performance across different holographic rendering methods, by requiring the identification of overlapping letters separated in depth.

1. Introduction

We propose an experiment that seeks to quantify the impact of one often-cited benefit of volumetric and holographic content – the ability to present monocular depth information. 3D displays projecting scene points that focus at varying depths (so that a viewer can selectively bring some points into focus at the expense of others) provide this monocular information. Monocular depth information cannot necessarily be used by the observer to judge absolute, or even relative depth of objects, but does provide a cue to depth separation [6]. These cues are present when viewing true 3D scenes, and are not available with traditional two-view stereo displays.

This paper will present an overview of display requirements for displaying monocular depth information and proposes an experiment for testing human and display performance in a layered letter identification task. Although this task examines one form of layered information presentation, we anticipate the results will be useful in informing other types of information and data visualization design, and for identifying applications where holographic displays improve task performance.

2. Motivation and Related Work

The entertainment and electronics industries have again become interested in 3D. Although most new commercial 3D display products present only two "flat" views, new research prototypes of holographic and dense light field displays are appearing as well. A number of constraints shape the commercial development of 3D displays, but in particular, geometry and data streams for stereo displays are similar enough to conventional televisions that they represent a lower-risk step than an entirely new technology. We want holograms on our screens, but do the first commercial displays have the most potential for entertainment? Computer-aided design? Medical diagnosis? To justify the investment in holographic or other dense 3D content and technology, it will be important to understand exactly which tasks and experiences are facilitated or improved by post-stereo 3D displays.

The proposed method allows for testing of human performance at discriminating layered information presented in a way that emulates a dense multi-view display, and compares it to a rendering method that exploits the full capabilities of electro holography. By comparing performance using these two methods, the task-performance benefits of holographic displays can be critically evaluated.

By performing a combined overlapping letter identification task and a depth discrimination task, we will emulate a situation likely to appear, for example, when visualizing layered information with text labels. For such visualizations to be useful, viewers should be able to read text on multiple layers, and will benefit further by being able to discriminate the depth order of layers. Understanding the critical depth separations at which identification performance improves will be useful in designing visualizations that seek to exploit overlapping information. Likewise, observing performance over a range of depths will inform designers of the practical limits for displaying text on an astigmatic display. By comparing performance at viewing stereograms and two-layer holograms, we hope to evaluate the potential benefit of monocular depth cues.

Beyond the display of text, understanding of the impact of monocular cues on discerning details in 3D displays is useful in other aspects of information design. For example, the use of layered information is integral to geographic information system visualization where many pieces of information are available for a single location on a map. The US census dataset alone contains more layers of information than can be practically visualized simultaneously. By optically layering geographic information layers, further gains in data density over state-of-the-art visualizations could be achieved. By understanding the practical limits in ability to discern layers, such displays could be further optimized.

Augmented reality applications also stand to benefit from understanding monocular cue presentation in 3D displays. In These applications, an electronic display adds information "on top of" the physical reality around the user. By placing this augmented reality in depth, these displays could make electronic annotations on physical objects easier to read. Since accommodative adjustments of the human visual system take time, labels or other virtual markers could be more quickly interpreted if they appeared in the same focal plane as the objects to which they relate. Additionally, augmentations that appear out of plane might be less likely to obscure other labels, or parts of the real scene seen through a display.

More generally, all 3D display applications that require close work (approximately at arms length or less), or seek to convey objects that appear close to the viewer, are likely to benefit most from display technology that can accurately portray monocular depth information, since these tasks are most likely to induce accommodation-vergence conflicts [3] when traditional stereo display techniques are used.

Related experiments have been documented for evaluating deep display technologies. Kim et. al [4] describe an evaluation of three 3D displays used in their lab. By photographing the light output of the display through a lens and pupil approximating that of a human observer's eye, they compare three displays' ability to provide monocular depth cues. As expected, they find no monocular cues in stereo images, but are able to show defocus-like blur in photographs from their Multi-Focus Display (MFD).

Single directional views from the MFD do not exhibit blur when the simulated eye is refocused, but four adjacent views fall on the pupil of the camera, and the relative alignment of these views changes, creating discretely duplicated images that closely approximate the appearance of blur.

Towards exploiting focusing ability to read text, designs have been presented for exploiting multi-layer LCD screens for text editing [5]. Although no experiments examining performance at reading overlapping letters have been documented, Experiments have been performed examining subject's ability to read text when overlaid on textured backgrounds. Scharff and Ahumada [9] examined human performance at searching for specific words hidden in passages of text placed on different backgrounds, emulating the appearance of web pages that use background images. They found a decrease in performance when text contrast was low, and when backgrounds contained high spatial frequencies. In an unpublished study using a two-layer LCD display, Wong et al [13] tested performance at a similar task while placing text and background at different layers. The authors suggest that performance at reading text over depth-separated backgrounds is different from when both are presented at the same depth. These studies, indicate that monocular cues to depth can increase viewer's ability to identify depth-overlapped letters on a screen and that the combination of monocular and binocular cues allow viewers to discriminate their relative depth ordering of these letters.

3. Methods

3.1. Display

The Mark II Holographic display [11] was designed to be a reconfigurable platform for experiments in horizontal-parallax computational holography. The display uses two cross-fired 18-channel tellurium dioxide acousto-optic modulators, relayed and de-scanned to project stacks of 144 fine-pitched nearly arbitrary 1-D patterns. In its current configuration, it creates 144 horizontal scanlines of up to 150 mm long, each with 2^{18} samples (about 1747 samples per mm, with a window of about 1.2 mm coherently illuminated at a time using a HeNe laser). The image of the acousto-optic modulators is viewed on a vertical diffusing screen, allowing the 30-degree-wide image volume to be seen from a range of vertical positions.

3.2. Rendering

Computing the scanlines to display an arbitrary horizontal-parallax-only computer-graphics 3D scene can be computationally intensive. Full optical simulations of 3D scenes are currently impractical for realtime applications, so approximations are used to render holograms for the display. Approximate rendering methods currently in use [8] on the Mark II display subdivide the output of the display into discrete angular bins and for each bin, sample the luminance of the 3D scene using traditional computer graphics rendering techniques. The scene is then displayed as a cloud of discrete light-emitting points, whose brightness, is modulated in a discrete set of directions according to the 2D renderings. Each light emitter, or "hogel" is created by computing a hologram of a single point, and the holograms for points covering the desired scene are added together to form a single full-screen hologram. Most published photographs of Mark II display output have placed these light emitters in a single plane, making the images more similar to high-density stereograms in that all light-emitting points appeared in focus in the same plane, which sits at the display's surface.

Scene appearance was previously demonstrated [8] for 383 emitters over 140 different viewing directions (spanning 30 degrees), which places emitters 0.19 mm apart, with each view spanning about 0.21 degrees. This geometry was chosen to place views about one pupil-width wide at a viewing distance of 0.6 meters. Even with this view spacing, the full fan of rays from a single view entering the viewer's pupil creates a strong monocular depth cue at the location of the emitter. This fan of rays is in focus only when the viewer is accommodating at the plane of the emitter, and will appear blurred otherwise. It may be possible for such a configuration to also provide other bundles of rays that come into focus at other distances. When rays from separate emitters whose view zones cross before entering the pupil, focusing the eye at the depth-plane of intersection causes these rays to overlap on the retina. The following experiment is used to comparatively evaluate the legibility of letters rendered as a single plane of emitters whose appearance changes with angle, as compared with a condition where light emitters are placed at two separate planes.

3.3. Experiment

In order to test subjects' performance at reading layered using the two rendering methods, we propose an experiment where subjects are asked to identify and compare the relative depth positions of two letters presented overlaid on each other, but separated in depth. To prevent subjects from using occlusion as a cue to relative depth, the two letters will be presented as transparent, self-illuminated surfaces. Each will be visible "through" the other irrespective of their depth order.

3.3.1. Stimuli

Letters for the task are rendered using two methods, using capitals from the SLOAN typeface [7], approximately 70 mm tall, displayed as luminous characters on a dark background. The first method uses the Reconfigurable Image Projection algorithm [8] to render a single set of 64 parallax views of the two overlaid transparent letters. The views are then used to modulate a single plane of emitters that lies 4 mm in front of the display's diffuser plane. This method will be referred to as "Stereogram" rendering. The second rendering method to be evaluated creates two layers of isotropic emitters, each in the shape of one letter. For this "Two-Layer" method, the image is effectively that of two planar holograms added together. For both rendering methods, emitters are placed on a grid, 128 emitters wide, with 144 lines of emitters.

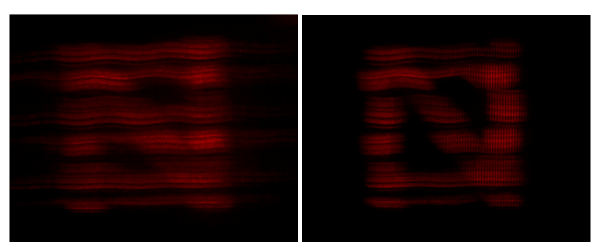

Figure 1. Example stimuli. Left image (N overlaid on S) generated using layered planar holograms. Right image (N overlaid on D) by stereogram rendering.

3.3.2. Task

Subjects are seated in front of the Mark II Holographic display, in a chair adjusted to place their eyes 600 mm from the diffuser screen. Subjects are presented with a series of letter pairs and are asked to discern the letters' depth order. Letter pairs consist of an N 75 mm behind the diffuser plane, and a second letter placed either in front or behind the first. Subjects respond with key presses, first to discriminate whether the non-N letter is in front, or behind the N, and second to identify the second letter presented. One hundred letter pairs, ranging in depth separation from –75 to 75 mm, and rendered with the two methods (Stereogram and Two-Layer), are presented in random order. After viewing a dark screen, a pair of letters appears, and remains visible for 750 ms. and then the screen goes blank. Subjects are instructed to respond to the depth order of the non-N letter by using the 8 and 2 keys on the numeric keypad (8 indicating behind and 2 indicating in front). Following this, subjects are asked to press the key on the keyboard corresponding to the non-N letter that appeared. Experimental software controls timing and display, and records responses, and response latencies.

4. Results

The experiment described above was performed on three subjects. For both rendering methods, performance at judging relative depth of the two letters was poorest with small depth separations. Errors in letter recognition were more uniform across the depth range. Performance in depth perception was significantly better with the Two-Layer rendering method (81% correct) compared to the Stereogram method (61% correct) (Wilcoxon signed-rank test: z=-2.832, p=0.005). Performance in letter recognition was not significantly better with the Two-Layer rendering method (88% correct) compared to rendering method #1 (83% correct) (Wilcoxon signed-rank test: z=-1.219, p=0.223).

5. Discussion

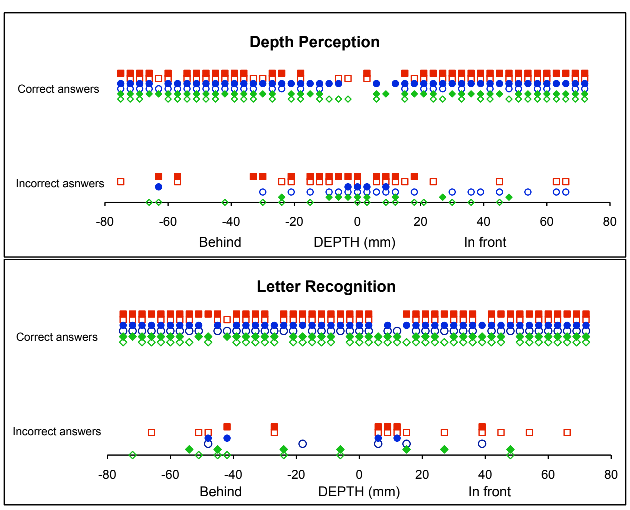

Figure 2. Results of depth discrimination and letter recognition for three subjects. Filled symbols are for Two-Layer rendering method, open symbols for Stereogram method.

Subjective evaluations showed that subjects had little or no ability to independently bring into focus layered presentations rendered using the Stereogram method although this method presents dense-parallax views conveying strong binocular depth cues. Combined with significantly higher error rates for depth perception, this method appears to be less suited for displaying layered information that the Two-Layer method. When viewing images rendered by the Two-Layer rendering method, some subjects suggested that they could focus independently on discrete depth layers although further evaluation using an optometer could verify this. Although the Two-Layer method supported better depth discrimination performance, the stereogram method may still be preferable for rendering small details, as it allows objects at all depths to be simultaneously in-focus.

Since the Mark II holographic display is only capable of displaying horizontal parallax, larger depth separations result in more difference between horizontal and vertical focus, and may impede letter identification. Additional experiments using larger depth separations could investigate this.

Future work could also include experiments with other rendering methods including the phase-added stereogram [14] or wavefront element rendering [10].

6. Acknowledgements

The authors would like to thank Fuensanta Vera-Diaz for her insightful advice on and assistance with experiment design and data analysis. This work has been supported by the Digital Life, Things That Think, and CELab consortia and the Center for Future Storytelling at the MIT Media Laboratory.

7. References

[1] BRAINARD, D. H., PELLI, D. G., AND ROBSON, T. Display characterization. In *Encyclopedia of Imaging Science and Technology*. Wiley-Interscience, 2002, pp. 172–188.

[2] HARRISON, B. L., ISHII, H., VICENTE, K. J., AND BUXTON, W. A. S. Transparent layered user interfaces: an evaluation of a display design to enhance focused and divided attention. In *CHI '95: Proceedings of the SIGCHI conference on Human factors in computing systems* (New York, NY, USA, 1995), ACM Press/Addison-Wesley Publishing Co., pp. 317–324.

[3] HOFFMAN, D. M., GIRSHICK, A. R., AKELEY, K., AND BANKS, M. S. Vergence–accommodation conflicts hinder visual performance and cause visual fatigue. *J. Vis. 8*, 3 (3 2008), 1–30.

[4] KIM, S.-K., KIM, D.-W., KWON, Y. M., AND SON, J.-Y. Evaluation of the monocular depth cue in 3d displays. *Opt. Express 16*, 26 (2008), 21415–21422.

[5] MASOODIAN, M., MCKOY, S., ROGERS, B., AND WARE, D. Deepdocument: use of a multi-layered display to provide context awareness in text editing. In *AVI '04: Proceedings of the working conference on Advanced visual interfaces* (New York, NY, USA, 2004), ACM, pp. 235–239.

[6] MON-WILLIAMS, M., AND TRESILIAN, J. R. Ordinal depth information from accommodation? *Ergonomics 43*, 3 (2000), 391 – 404.

[7] PELLI, D. G., ROBSON, J. G., AND WILKINS, A. J. The design of a new letter chart for measuring contrast sensitivity. *Clinical Vision Sciences*, 2 (1988), 187–199.

[8] PLESNIAK, W., HALLE, M., V. MICHAEL BOVE, J., BARABAS,

J., AND PAPPU, R. Reconfigurable image projection holograms. *Optical Engineering 45*, 11 (2006), 115801.

[9] SCHARFF, L. F. V., AND AHUMADA, ALBERT J., J. Predicting the readability of transparent text. *J. Vis. 2*, 9 (12 2002), 653–666.

[10] SMITHWICK, Q. Y. J., BARABAS, J., SMALLEY, D. E., AND V. MICHAEL BOVE, J. Interactive holographic stereograms with accommodation cues. H. I. Bjelkhagen and R. K. Kostuk, Eds., vol. 7619, SPIE, p. 761903.

[11] ST-HILAIRE, P. *Scalable optical architectures for electronic holography*. PhD thesis, Massachusetts Institute of Technology, Program in Media Arts and Sciences, 1994.

[12] ST-HILAIRE, P. Scalable optical architecture for electronic holography. *Optical Engineering 34*, 10 (1995), 2900–2911.

[13] WONG, B. L. W., VILLANUEVA, R., AND NEES, A. Colour and transparency on the multi-layer display (mld™). Tech. rep., Discussion Paper 2003/05. Department of Information Science, University of Otago, Dunedin, New Zealand, 2003.

[14] YAMAGUCHI, M., HOSHINO, H., HONDA, T., AND OHYAMA, N. Phase-added stereogram: calculation of hologram using computer graphics technique. S. A. Benton, Ed., vol. 1914, SPIE, pp. 25–31.