An Evaluation of Depth Perception on Volumetric Displays

Tovi Grossman, Ravin Balakrishnan
Department of Computer Science
University of Toronto
www.dgp.toronto.edu
tovi | ravin @ dgp.toronto.edu

ABSTRACT

We present an experiment that compares volumetric displays to existing 3D display techniques in three tasks that require users to perceive depth in 3D scenes. Because they generate imagery in true 3D space, volumetric displays allow viewers to use their natural physiological mechanisms for depth perception, without requiring special hardware such as head trackers or shutter glasses. However, it is unclear from the literature as to whether these displays are actually better than the status-quo for enabling the perception of 3D scenes, thus motivating the present study. Our results show that volumetric displays enable significantly better user performance in a simple depth judgment task, and better performance in a collision judgment task, but in its current form does not enhance user comprehension of more complex 3D scenes.

Categories and Subject Descriptors

H.5.2 [**User Interfaces**]: Interaction styles.

General Terms

Experimentation, Human Factors.

Keywords

Volumetric display, depth perception, evaluation

1. INTRODUCTION

Three-dimensional (3D) display technologies such as immersive VR systems [4], or non-immersive fish-tank VR systems using LCD shutter stereo-glasses [24], have significantly improved in display quality in recent years. Experimental evaluations have also shown that these displays can improve the user's ability to perceive virtual 3D scenes [1, 19, 25]. However, these technologies have some drawbacks. For example, immersive VR systems require users to wear cumbersome head mounted displays, while the less intrusive fish-tank VR displays still require users to wear special glasses. Some systems utilize head tracking to provide motion parallax, but this requires special hardware and limits the display to a single user. A more fundamental problem is that these displays create a conflict between the two mechanisms that enable human stereoscopic vision: accommodation and convergence. The resulting ambiguous depth cues can cause users to experience symptoms of asthenopia, such as nausea, dizziness and eye fatigue [13, 17].

Volumetric displays (Figure 1) [2, 7, 12] address many of these problems by actually illuminating points in 3D space to generate true 3D volumetric images. As a result, viewing images on such displays is akin to viewing real 3D objects in the physical world. Natural physiological mechanisms, such as true motion parallax and stereopsis through convergence and accommodation, can be used by the viewer for depth perception. Since images are produced in true 3D space, the convergence and accommodation cues are consistent, eliminating the symptoms of asthenopia often experienced when viewing stereoscopic displays. Special hardware, such as shutter glasses or head trackers, is not required, and the display can be viewed from almost any direction by multiple users simultaneously.

While volumetric displays are promising, an unanswered question is whether these displays improve the user's perception of a 3D scene, in comparison to the other 3D display technologies. In particular, do volumetric displays provide better 3D depth perception? To investigate this question, we conducted a formal experiment comparing user performance in three different 3D depth perception tasks on a volumetric display, a standard perspective display, and a stereoscopic display with and without head tracking. The results of this work will provide us with a better understanding of the perceptual benefits and drawbacks of volumetric displays, which is valuable in helping guide the design of future user interfaces for them.
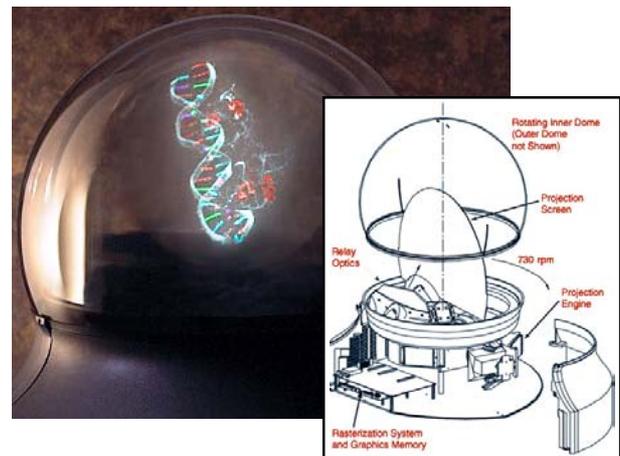
Figure 1. 3D volumetric display. Voxels illuminated in true 3D space facilitate 3D viewing from anywhere around the display without wearing special glasses. Inset: schematic of the display we used, from Actuality Systems (www.actuality-systems.com)

2. RELATED WORK

2.1 3D Display Techniques and Depth Cues

One of the biggest challenges to creating realistic depictions of 3D scenes on 2D displays is to provide sufficient depth cues to enable the user to garner appropriate 3D information about the scene. To meet this challenge, various techniques have been utilized:

One of the simplest and most effective depth cues is a perspective projection, which inversely scales the image coordinates with depth. Perspective is particularly effective when the scene has parallel lines [3]. Generally, perspective projections are provided from a single viewpoint, and users must vary this viewpoint in order to get a sufficiently rich understanding of the displayed 3D scene.

Stereopsis, resulting from binocular disparity, is another depth cue that can be exploited by presenting different images for each eye. A fused 3D image appears at the point of convergence between these two images.

Another important depth cue is motion parallax, obtained when the user's viewpoint of a 3D scene changes. Several systems have explored motion parallax cues [6, 8, 18, 19], where the user's head position is tracked, and the location of the user's eyes is estimated by offsetting them by a constant distance from the user's head [1]. The viewpoint of the virtual scene is then updated appropriately

Systems which display 3D imagery can provide some or all of the above depth cues. For example, a fish-tank VR environment [24] is a head-coupled stereo display which provides all of these cues, with a viewing volume roughly equivalent to the inside of the monitor. We include a fish-tank VR display in our study as it is one of the more viable and least intrusive forms of 3D display currently available.

More immersive forms of 3D display technology also exist, usually in the form of head mounted displays [4]. However, we will not include these forms of displays in the present study, since they have the drawback of separating the user from the real world [1, 4]. While this may be acceptable for some tasks, we wish to compare the volumetric display to other displays which could be appropriately integrated into a regular workplace environment.

2.2 Evaluation of Display Techniques

A vast body of research exists which evaluates the relative value of different types of displays for 3D viewing. A thorough review of such research is beyond the scope of this paper, and we summarize those studies which specifically focus on depth perception.

It has been shown that stereo glasses can improve user performance when detecting paths in a tree structure [1, 18, 19, 25]. Sollenberger and Milgram [18, 19] also found that scene rotation further reduced errors in such a path tracing task. In their implementation, scene rotation was controlled by the system. In a follow up study, Ware [24] found similar results when the motion was controlled by the user, with a head coupled perspective view.

McKenna [14] found head coupled perspective to also be beneficial in a 3D positioning task. Users were required to position a cursor to match the location of a cube in three dimensions. Results showed that a head coupled perspective gave the best results, while a mouse-controlled viewpoint condition decreased performance.

Arthur et al. [1] showed that users' subjective impressions of a 3D scene were more positive when a head coupled without stereo display was used, over a stereo display without head coupling.

In a more recent study, Ware and Franck [25] evaluated nine different types of viewing modes for a path tracing task. It was found that the stereo viewing mode without motion was significantly worse than all three tested stereo modes with motion, including system controlled rotation, hand coupled rotation, and head coupled rotation. Of the three viewing modes which combined stereo and motion, there were no significant differences, showing that the motion parallax cues were important, but it did not matter how they were provided.

2.3 Asthenopia in Stereo Displays

While the previous work clearly shows that stereo displays can be beneficial for 3D tasks, a drawback of these displays is that asthenopia has been associated with their use [13, 17]. This is a factor that has prevented their widespread adoption [17]. Asthenopia is caused by weakness or fatigue in the eyes, and can be accompanied by nausea, dizziness, headaches, or dimming of vision. It is believed that the main cause of asthenopia in stereo displays is due to the discrepancy in the visual representation between accommodation and convergence. The discrepancy arises because accommodation is fixed on the depth of the display surface, while the convergence distance depends on the perceived depth of the virtual object. In natural 3D vision, accommodation and convergence always produce identical depth information.

2.4 Volumetric Displays

Several volumetric display technologies exist [7, 12], with the advantage that imagery is presented in true 3D space. As a result, stereo and motion cues come without requiring special hardware such as glasses or head trackers. Further, consistent depth information is provided as there is no discrepancy between accommodation and convergence.

Balakrishnan et al. [2] discussed advantages of volumetric displays, and demonstrated interaction scenarios via wizard-of-oz prototypes. More recently, a research prototype [10] allowed users to directly interact with volumetric displays using hand and finger gestures on and above the display surface. Volumetric displays have also been used as a platform for experimental evaluation [9]. While this previous work provides fairly compelling usage demonstrations of volumetric displays, users' ability to perceive depth information in volumetric displays has not been rigorously evaluated. To date, only a few studies have been reported, and none provide conclusive evidence as to how volumetric displays compare to other existing 3D display technologies for depth perception tasks.

Rosen et al. [16] found that users could identify deformations in three dimensional objects with more accuracy on a volumetric display than on a 2D display. This is an unsurprising result, given that the 2D display used did not provide any stereo or motion cues, and more relevant, 3D displays, were not included in the study. A study of air traffic control tasks [21] did not provide conclusive results about how volumetric displays compared to other 3D displays. In another study comparing volumetric displays to 2D displays, participants navigated a simple maze using a joystick [20]. However, the maze and task were 2D, providing little insight into the value of volumetric displays for understanding the 3D scenes that they were designed to display.

In summary, our review of the literature revealed little conclusive data on the performance of volumetric displays in comparison to other 3D display techniques for 3D perception tasks. Our present study seeks to provide some data in this regard.

3. EXPERIMENT

3.1 Goals

The purpose of our study is to compare volumetric displays to other 3D display techniques with regard to user ability to perceive depth information when viewing 3D imagery. Our intention is not to run an exhaustive comparison of all 3D displays, but rather to provide data for one interesting class of display – swept volumetric – in comparison to established ones. This display is particularly interesting as it provides consistent depth information, and a 360 degree viewing angle, unlike most other autostereoscopic technologies. Empirical data as to the capabilities of this class of display will be useful in guiding future user interface designs as the technology continues to improve.

Two factors are particularly important in designing this study: the type of display techniques to compare the volumetric display against, and the tasks used for the comparison.

3.2 Manipulation of Display Techniques

The number of different 3D display techniques that can be used for comparison is quite numerous. In a study by Ware and Franck [25], for example, nine different display techniques were evaluated. To keep the size of our study manageable, we chose three display techniques to compare to the volumetric display, all of which provided significantly different results in the Ware and Franck study [25]. The first is a perspective projection on a 2D display, the second is a static stereoscopic display, and the third, is a stereoscopic display with head tracking. The third technique is closest to viewing the volumetric display, as both stereo and motion cues are provided, and the literature shows that stereo with head tracking outperforms stereo only, which in turn outperforms perspective only [25].

Including these three display techniques will allow us to determine where the volumetric display lies in this continuum of baselines. In the event that the volumetric display does not outperform the leading candidate – stereo with head tracking – we will be able to determine if the volumetric display is better than *any* of these three baselines.

We omitted hand coupled techniques from the study as previous research did not find them to result in significantly different performance from head coupled views. In all cases user input was through a keyboard, positioned in front of the display.

The display parameters were set such that apparent sizes were controlled to be equal across all displays. Other factors such as color, resolution, and brightness were set at optimal viewing values for each individual display. This allowed us to determine how the volumetric display compares to the baselines in a "best case" scenario for all displays. We now discuss each hardware setup.

3.2.1 Perspective Projection on 2D Display

A 19-inch Dell Trinitron CRT monitor was used, with the 3D scenes projected onto the 2D plane using a perspective projection. The refresh rate of the monitor was 120Hz.

3.2.2 Stereoscopic Display

The same monitor as in the 2D perspective setup was used. A StereoGraphics CrystalEyes 3D LCD shutter glasses provided stereo viewing. Left and right eye images were provided by the top and bottom half of the frame buffer respectively. The monitor ran at 120Hz with each eye receiving a 60Hz update rate, coordinated with the shutter glasses. The glasses were synchronized with the display by an infrared transmitter positioned on top of the monitor.

3.2.3 Stereoscopic Display with Head Tracking

This setup was the same as the stereoscopic display, except that, in addition, the user's head position was tracked in real-time and used to generate the correct perspective view for each eye position. Head position was measured using an Ascension Flock-of-Birds electromagnetic 6-dof tracker. The tracker had a positional accuracy of less than 2mm, and operated at 100Hz. The tracker's electromagnetic receiver was positioned on the user's forehead using a headband. The position of each eye was derived from this reported head position, and the perspective view was continuously updated accordingly. Figure 2 illustrates this setup.

3.2.4 Volumetric Display

We used a volumetric display (Figure 1) from Actuality Systems (www.actuality-systems.com). It generates a 10" diameter spherical 3D volumetric image by sweeping a semi-transparent 2D image plane around the Y-axis. Each slice consists of 768x768 pixels, and a total of 198 2D images (slices) are uniformly displayed around the Y-axis, resulting in a total of 116 million voxels. The display's refresh rate is 24Hz. The user sat in front of the display, in the same chair used for the other three display setups (Figure 3). To ensure the same viewing angles were used for all displays, a calibration program was used to ensure that the user's viewpoint was aligned with the center of the display. In this and the previous stereo with head tracking technique, we limited the amount of head movement by fading out the 3D scene if the head position deviated from a calibrated starting position by more than 10 inches. Limiting the head movements prevented users from taking on viewpoints which would trivialize the depth perception tasks. Other than fading out the scene, head movements had no effect on what was being drawn in the volumetric display. Tracking head movements in the volumetric setup also allowed us to compare the amount of head movement used with what occurred in the stereo head tracking setup. Room lights were off to enable optimal viewing.

Figure 2. Stereoscopic display with head tracking setup. Shutter glasses provide stereo viewing, and an electromagnetic tracker placed on the user's forehead provides head tracking.

Figure 3. Volumetric display setup. The electromagnetic tracker measures head movement to allow for imagery to be displayed only when head position is within a given range.

3.3 Tasks

Instead of using a single task to evaluate depth perception within a 3D scene, we chose a set of three tasks, all of which have been previously used to study human depth perception. Doing so allowed us to determine how the volumetric display compared to the other display techniques for a variety of task scenarios. The first task required participants to judge the depth of a single object in a 3D scene. The second required them to view and comprehend a complex graph which could be used for information visualization. In the third task, participants studied the position and heading of two moving objects and judged if they were on a collision course or not. In each of these tasks, the error rate was the main dependant variable, as previous studies have shown this to be highly sensitive to display technique [1, 18, 25].

Participants were not required to perform any virtual object manipulations to complete any of the tasks, ensuring that the task was purely perceptual in nature. In a pilot study, we included a 6-dof docking task that had both perceptual and manipulation components. We found that the manipulation component added an overhead to task learning, completion times, and accuracy, while providing little additional insight into our primary goal of evaluating depth perception.

In another pilot study, we allowed participants to take as much time as they wanted to complete a task. However, some would take their time, while others seemed to rush through the experiment. Participants also took different amounts of time depending on the display technique. To alleviate this problem, in all of the tasks, scenes were viewed for a controlled period of time. By enforcing an equal task completion time across all participants and conditions, we avoid the complex tradeoffs between completion time and accuracy. This design also prevented participants from "racing through the experiment", and, as such, they had no reason not to provide their best answers.

We now discuss in detail the three tasks and their procedures and designs used during the experiment.

3.3.1 Task 1: Depth ranking

In this task, participants were required to rank the depth of a sphere which was floating above the floor in a 3D graphical scene. This is similar to tasks previously used to evaluate the effect of shadows on perception of depth and spatial relationships [11, 22, 23, 26]. In our implementation, square outlines of the floor and back wall of the scene were drawn. Along both sides of the floor, nine uniformly distributed tick marks were drawn, the first just after the beginning of the floor, and the last just before the back wall. A wireframe sphere was drawn floating above the floor. The center of the sphere was exactly aligned with one of the nine tick-marks. The task of the participant was to determine the depth of the sphere, by indicating which tick mark they thought it was aligned with. The numbers 1, 3, 5, 7, and 9 were drawn beside the corresponding tick-marks, but participants were told that 2, 4, 6, 8 were also valid answers. The even numbers were omitted from the display to reduce visual clutter. The horizontal and vertical viewing angles of the scene were both 0°, parallel with the floor of the scene, and the participant's viewpoint was centered with the middle of the back wall. This ensured that participants would have to rely on their depth perception to determine their answer. Figure 4 illustrates this task. For each trial, the sphere could take on any of the nine depth values, and the size, height, and lateral position of the sphere were systematically varied during the experiment.
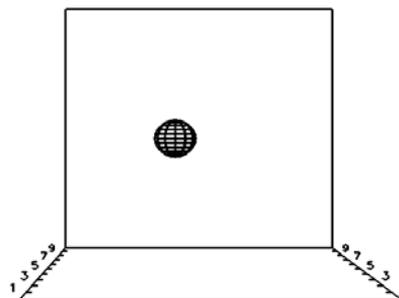
Figure 4. Depth ranking task. Participant specified the depth (1-9) of the sphere's location in 3D space.

The procedure for this task was as follows: participants hit the space bar to begin a trial. The scene was then drawn, and a timer started. Participants had exactly 3 seconds to view the scene. We determined through pilot studies that 3 seconds was a reasonable duration for participants to comprehend the scene and make an assessment of the object's location in depth. After 3 seconds, the scene disappeared, and participants were prompted to enter their answer by hitting the relevant numeric key on the keyboard. To control viewing time, participants could not enter their answer until the scene disappeared.

Participants completed 54 trials, with the sphere appearing at each depth value exactly 6 times. The height, lateral position, and size of the sphere took on random values which were generated before the experiments began, so that each participant saw the exact same 54 trials, appearing in random order. Before starting, four practice trials were given, so that the participants were familiar with the task.

3.3.2 Task 2: Path Tracing

The purpose of this task was to evaluate the effectiveness of the various displays for understanding a complex 3D information network represented by a graph. The ability to trace paths in such graphs is an important aspect to understanding such information networks. Such path tracing tasks have previously been used to evaluate various 3D viewing modes, such as perspective, stereoscopic, and head coupled stereoscopic [1, 25] Thus, our use in this context is consistent with the literature. In particular, our implementation was similar to that used in Ware et al. [25]. The computer generated a random 3D graph consisting of 36 nodes and 48 edges. The nodes were randomly placed in a 5" radius spherical volume. The nodes were divided into three groups of 12. Two of these groups were considered leaf nodes, while the third group was considered to be intermediate nodes. Each leaf node was connected to exactly two randomly chosen intermediate nodes, resulting in the 48 edges. In each graph, two leaf nodes were highlighted by in the use of a different color than the other nodes. The task of the user was to determine whether or not there was a path of length two from one highlighted node to the other. Because the highlighted nodes were both leaf nodes, they could never be connected directly. Figure 5 illustrates this task.

With this task, the participant had 8 seconds to view the graph before the scene disappeared. The 8 second parameter was determined via pilot studies as being a reasonable duration for comprehending the graph. Once the scene disappeared the participant would either hit 'y' or 'n' on the keyboard, 'y' for "yes there is a path of length two", or 'n' for "no there isn't a path of length two". The highlighted nodes were drawn before the participant hit the space bar, so that there would be no difficulty finding them once the entire graph was displayed.
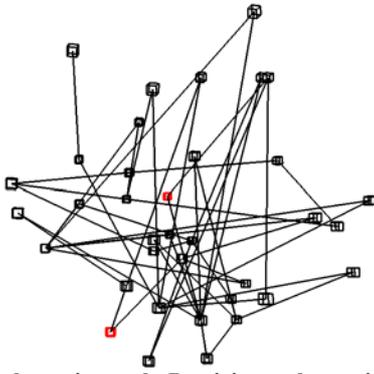
Figure 5. Path tracing task. Participant determined if there was a path of length two connecting the two highlighted nodes.

Twelve graphs with predetermined highlighted nodes were randomly generated before the experiments began. In six of these graphs, a path of length two existed, and in the other six, a path did not exist. Each graph was presented at 3 different orientations, rotated by 120° along the up-down axis, resulting in 36 trials. Four practice trials were given: in two of these practice trials, a path existed; in the other two, a path did not exist. This ensured that participants thoroughly understood the task.

The volumetric display we used had certain regions where data was slightly harder to perceive, due to the physical boundaries of the spinning screen within the display. This could potentially make it hard to see important nodes and edges of a graph, depending on its orientation. By presenting each graph in three orientations, our intention was to mitigate potential perception difficulties due to the vagaries of our display. It is important to note that this is merely a safeguard against a weakness in the particular display we used, rather than a concern with volumetric displays in general. A detailed discussion of our display's viewing artifacts is presented later.

3.3.3 Task 3: Potential Collision Judgment

In this task, participants were asked to make spatial judgments about a *dynamic* 3D scene. We chose a task that had previously been used to evaluate display techniques and visual enhancements for 3D scenes [5, 15, 21]: two objects flew together and disappeared before reaching a potential point of impact, and participants determined whether or not the objects were going to collide (Figure 6a).

The objects were drawn as wireframe cubes, and the flight paths were always at the same altitude. The trajectory of each cube varied from trial to trial, but the depth component of the trajectory was always positive (moving away from the user), and the angle between the two 3D flight paths was always 90 degrees. The horizontal and depth coordinates of the two objects were varied to create three different scenarios. In one case, the cubes were on a true collision course (Figure 6b). In the other two cases, the cubes were on a near-collision course, such that one cube would eventually pass behind the other if the animation was allowed to continue (Figure 6c, d). The cubes disappeared before they collided or passed by each other when the horizontal distance between their positions reached a threshold value of approximately 2.25cm. The size and speed of the cubes were always the same. The user's viewpoint position was equal to the altitude of the objects. This made the task particularly difficult, for if the animation continued, the objects would collide on the 2D projection of the scene, regardless of whether or not they were actually colliding in 3D space. For frame of reference, a ground plane grid was drawn below the objects flight path.
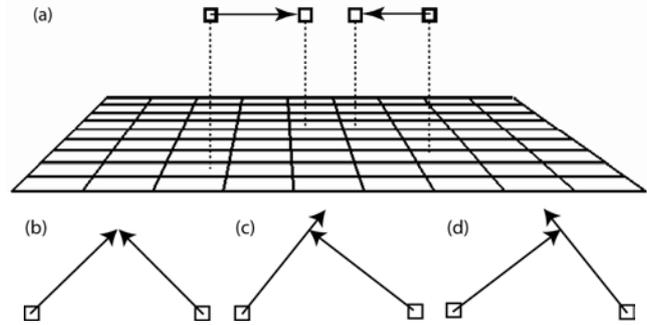
Figure 6 (a) User's perspective of the collision task (dashed lines and arrows are for illustration only). (b-d) Top view of the task, illustrating the three possible scenarios.

As in the previous two tasks, a trial began when the participant hit the space bar. The animation would then play, and when the horizontal distance between the two cubes reached the threshold value, the scene disappeared. At this point, users would either hit 'y' or 'n' on the keyboard, 'y' for "yes they were going to collide", or 'n' for "no they were not going to collide".

For this task, there were 60 trials. Trajectories were randomly generated before the experiment started, so that all participants saw the exact same 60 trials. In 30 of the trials, the objects were on a true collision course. Trials in which the left object would pass behind the right object, and vice-versa, appeared 15 times each. The 60 trials appeared in random order. Four warm-up trials were given. In two of the trials, the objects were on a true collision course. In the other two, each miss scenario occurred once.

3.4 Participants

Five female and seven male volunteers participated in the experiment. Participants were screened for adequate stereo vision using the Stereo Optical RANDDOT stereopsis test. Participants ranged in ages from 18 to 25.

3.5 Overall Experiment Design

The experiment was performed in one sitting, where participants completed all trials for one display technique before moving on to the next. The 12 participants were randomly split into 4 groups of 3. The presentation order of the four display techniques was counterbalanced using a Latin square design. For each display, participants completed the three tasks. The three tasks always appeared in the same order, with each task lasting about 5-7 minutes. The entire experiment took approximately 75 minutes.

3.6 Results

3.6.1 Task 1: Depth ranking

The main performance measure for this task was the error magnitude, defined as the average difference between the participant's answer and the correct expected answer.

Analysis of variance showed that the display type had a significant effect on error magnitude ($F_{3,33} = 252.31$, $p<.0001$). The volumetric display had the lowest error magnitude of 0.48, which was a 25% improvement on the stereoscopic display with head tracking setup, which had the next lowest error magnitude of 0.64. Pair-wise means comparisons showed this difference to be significant ($p<.05$). The perspective display was significantly worse than the other three techniques ($p<.0001$) while the difference between the stereoscopic display with and without head tracking was not significant (Figure 7).
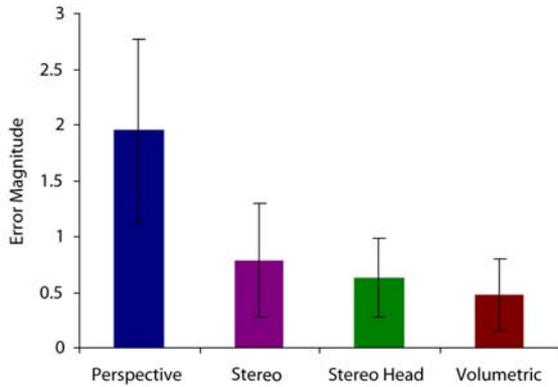
Figure 7. Error magnitudes for the depth ranking task.
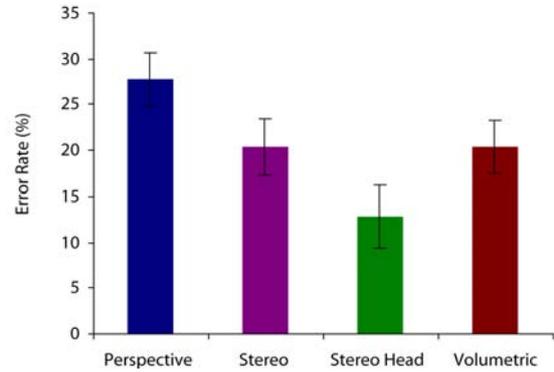
Figure 9. Error rates for the path tracing task.
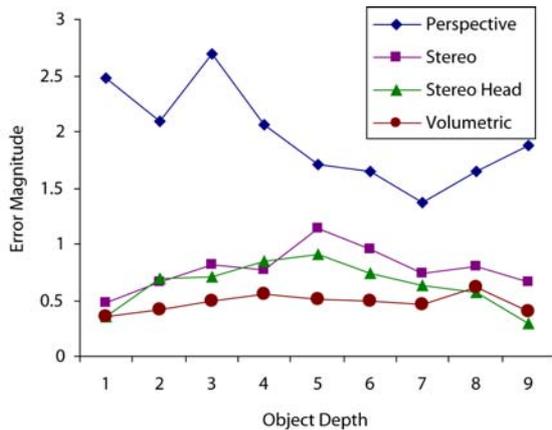
Figure 8. Error magnitudes for the depth ranking task, for each object depth.
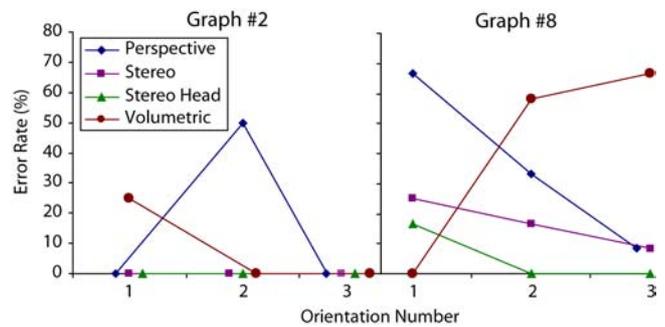
Figure 10. Interaction between display technique and graph orientation. Poor results for the perspective display are likely due to occlusions in the graph's 2D projection. Poor results for the volumetric display are likely due to display artifacts.

The depth of the object also had a significant effect ($F_{8,376} = 3.93$, p<.0001) on error magnitude (Figure 8). The two stereoscopic display techniques did best at the extreme values. Pair-wise means comparisons showed that for the stereoscopic display with head tracking, depth values 5 and 6 were significantly worse than depth values 1 and 9, and for the stereoscopic only display, depth value 1 was significantly different from depth value 5. This could be due to the fact that at the extreme values, the user could use the room (either the front of the floor, or the back wall), as a frame of reference, improving the results. Contrary to this, the results for the volumetric display seem to be fairly uniform throughout, with pair-wise comparisons showing no two values to be significantly different. The values for the perspective display are quite sporadic, likely due to users simply guessing the answer, as depth perception on these displays is quite poor.

3.6.2 Task 2: Path Tracing

The main performance measure for this task was the error rate, defined as the average number of errors per trial. Analysis of variance showed that the display type had a significant effect on average error rate grouped by participant ($F_{3,33} = 11.94$, p<.0001). In this task, the stereoscopic display with head tracking resulted in the best performance, with an average error rate of 12.7%. Error rates for the stereoscopic only, volumetric, and perspective displays were 20.4%, 20.4% and 27.7% respectively. Pair-wise means comparisons showed that the stereoscopic display with head tracking was significantly better than the volumetric display, which in turn outperformed the perspective display (p<.05) (Figure 9).

Our analysis of graph orientation found that in 6 graphs, there was a significant interaction between orientation and display (p<.05 in all 6 cases). Further analysis revealed two causes for this interaction. Firstly, error rates for the perspective display could be quite high for some graph orientations, whereas, on the other displays, the error rates were uniformly low for those particular orientations. By looking at the conditions where this occurred, it was clear that in the orientations where the perspective display performed poorly, occlusions were present in the 2D projection of the 3D graph, making the task especially difficult. In the other displays, users could either use the stereo cues or head tracking to have a better understanding of the graph in the occluded locations. Secondly, error rates for the volumetric display were high in some graph orientations, while the other displays had uniformly low errors rates for these orientations. Examining these cases, our earlier concern about some regions of the volumetric display being harder to perceive was confirmed. In particular, high error rates occurred when important graph nodes or edges were centered at the back of the display furthest away from the user's position. Figure 10 shows the error rates by orientation for two of the graphs which had significant interaction between orientation and display.

When we reanalyzed the data with the conditions where goal nodes were in a dead spot of the volumetric display, the average error rate for the volumetric display was reduced to 16.4%. This was still slightly higher than the error rate for the stereoscopic display with head tracking, but the difference was not statistically significant.

3.6.3 Task 3: Potential Collision Judgment

As with task 2, the main performance measure for this task was the error rate, defined as the average number of errors per trial. Analysis of variance showed that display type had a significant effect on average error rate grouped by condition ($F_{3,33} = 39.50$, p<.0001). As in the depth ranking task, the volumetric display provided best result, with an error rate of 19.3%. The error rates for the stereoscopic display with head tracking, stereoscopic only display, and perspective display were 22.4%, 27.9% and 48.3% respectively (Figure 11). Pair-wise means comparisons showed that the difference between the volumetric display and the stereoscopic display with head tracking was not significant, but the volumetric display had a significantly lower error rate than the stereoscopic only display (p <.05). The high error rate for the perspective display was significantly different from all other displays (p<.0001) and its value, which was close to 50%, shows that the task was basically impossible without any stereo or motion cues.
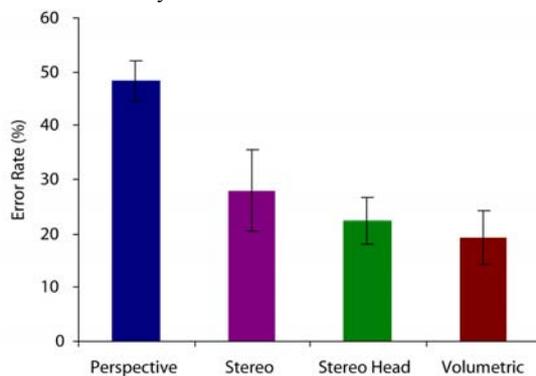
Figure 11. Error rates for potential collision judgment task.

3.6.4 Head Movement Analysis

We analyzed the head movement data for the stereoscopic display with head tracking and volumetric display conditions. Because viewing durations were different for the three tasks, we measured the amount of head movement in inches per second, allowing us to compare head movements across tasks. The head movement measure was thus obtained by dividing the total head movement by the viewing duration, for each trial.

For the depth ranking task, the volumetric display resulted in significantly more head movement than the stereoscopic display with head tracking ($F_{1,11} = 93.39$, p<.0001). This was also the case for the path tracing task ($F_{1,11} = 26.70$, p<.0001). For the collision judgment task, the rank ordering was reversed, with slightly more head movement for the stereoscopic display with head tracking (F= 4.69 p < .05). Figure 12 illustrates these results.

As can be seen in Figure 12, very similar head movements were used during the path tracing and depth ranking tasks. There are a couple of possible reasons why more head movement was used on the volumetric display. Firstly, participants may more naturally use head movements when viewing 3D scenes in true 3D space. Secondly, head movements may have increased for the volumetric display because image quality can depend on the viewpoint.

As for the collision judgment task, both displays resulted in fewer head movements, and the difference between the two displays was much smaller than in the other two tasks. Head movements may have been used less because it was a dynamic task. Our observations indicated that head movements were mostly vertical in this task, to provide a higher viewing elevation.
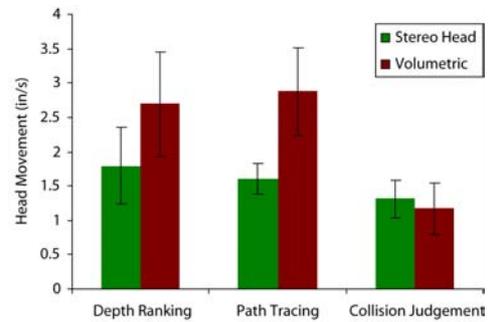
Figure 12. Head movement data (in/s) for the stereoscopic display with head tracking, and volumetric display conditions.

4. Volumetric Display Artifacts

Volumetric displays offer a number of promising properties which clearly warrant a systematic experimental evaluation of their capabilities. Because the technology is still very new, the display quality is relatively low when compared to the far more mature current generation of 2D and stereoscopic display systems used in this study. This means that there is room for improvements in the quality of the volumetric display hardware, and consequently in the user performance results which we have reported. For example, as the display's quality improves in future generations of the hardware, the significantly better performance for the volumetric display found in tasks 1 and 3 could further increase in magnitude, and for task 2, the volumetric display could outperform the other techniques.

The following is a list of problems in display quality with the volumetric display we used. These may have resulted in less than ideal performances in our reported results:

Brightness: In the display used, the projector brightness is quite low. To be able to see the interior contents of the display, the room lighting must be very low.

Image Stability: The three dimensional images in the display used are generated by a flat projection screen sweeping out the display volume. Slices which are 180° apart are therefore perfect mirror images of each other. These "front" and "back" projections are not perfectly aligned, and as a result, images appear to be shaky in certain areas of the display. The magnitude of the instability is location dependant, and ranges from non-existent to quite severe.

Refresh Rate: At 198 slices per revolution, and 24 revolutions per second, the projector is updating at an enormous rate. However because the entire viewing volume is only being updated at 24 Hz, there is a noticeable flicker in the displayed image.

Faint Spots: With low ambient lighting, most of the viewing volume can be seen quite clearly. However, objects which are inline with the user's eye position and the center of the display are much harder to discern, as the slices which these images are projected onto are parallel with the user's line of vision.

5. Discussion and Conclusions

Before conducting our study, volumetric displays were known to have a number of beneficial properties unique to the technology:

Consistent depth information: accommodation and convergence cues are consistent, so users do not suffer from asthenopia.

Minimal hardware requirements: special glasses, head mounted devices, and head tracking technology are unnecessary.

360° viewing angle: Imagery can be viewed from any angle, allowing simultaneous viewing by multiple users.

One factor which was unclear from the literature, which we explored in our study, was whether or not volumetric displays provided superior depth perception. Results of the depth ranking task show that volumetric displays do indeed provide more improved depth perception in comparison to stereoscopic displays, even when the stereo view is coupled with head tracking hardware. In addition to their other beneficial properties, this result makes volumetric displays an excellent hardware platform candidate for 3D tasks.

Despite the improved depth perception, the volumetric display did not perform as well in the path tracing task, most likely due to the viewing artifacts associated with the current generation of display which we used. This is an important result, as it shows that the quality of the display must improve before it can be adequately used for real world tasks with dense 3D environments. We discussed the viewing artifacts which were present, and if these are addressed, users' viewing experiences will likely improve. It is important to note that even with these artifacts present, users still performed best with the volumetric display in two of the three tasks.

Because of these artifacts in the volumetric display, we could not ensure uniformity in factors such as brightness, contrast, and resolution across all the display types in the experiment without reducing the viewing quality of the other displays, which would be unfair to those displays. In essence, our experiment biased against the volumetric display, and even so the results are mostly in favor of it, which means it can only do better as the technology improves.

We also found that more head movements were used when viewing the volumetric display than when using the stereoscopic display with head tracking. Although this was partially a result of users needing alternative views due to volumetric display artifacts, it also implies that users might be more comfortable moving their heads when they are viewing true 3D imagery. It would be interesting to further explore this issue and compare results with head movements used when inspecting physical 3D objects.

Our results should also be interpreted in light of the body of existing work. We intentionally chose three baseline display techniques for comparison which have been found to provide statistically different results in depth perception tasks. The results which we obtained for the baseline conditions match these previous results, with the same continuum being found, in all three tasks. This further validates our experimental method and results, providing strong evidence as to where the volumetric display lies in this continuum.

In summary, we believe that volumetric displays have potential as a 3D display platform, due to the many beneficial features including enhanced depth perception as demonstrated in the present study. That said, these displays are likely unsuitable for widespread use until better performance is observed in complex situations such as our path tracing task. This will undoubtedly occur as the display quality improves over time and the artifacts are reduced.

6. REFERENCES

[1] Arthur, K., et al. (1993). Evaluating 3D task performance for fish-tank virtual worlds. *ACM Transactions on Information Systems*, *11(3)*. p. 239-265.

[2] Balakrishnan, R. et al. (2001). User interfaces for volumetric displays. *IEEE Computer*, *March 2001*. p. 37-45.

[3] Brooks, F. (1988). Grasping reality through illusion - Interactive graphics serving science. *ACM CHI* p. 1-11.

[4] Buxton, W. et al. (1998). HMD's, Caves, & Chameleon: A human-centric analysis of interaction in virtual space. 64-68.

[5] DeLucia, P. (1995). Effects of pictorial relative size and ground-intercept information on judgments about potential collisions in perspective displays. *Human Factors*, *37*. p. 528-538.

[6] Diamond, R., Wynn, A., Thomsen, K., and Turner, J. (1982). *Three-dimensional perception for one-eyed guys*, in *Computational Crystallography*, Oxford, Clarendon Press.

[7] Ebert, D. et al. (1999). Realizing 3D visualization using crossed-beam volumetric displays. *Communications of the ACM*, *42(8)*. p. 101-107.

[8] Fisher, S. (1982). Viewpoint dependent imaging: An interactive stereoscopic display. *Processing and Display of Three-Dimensional Data II*. p. 41-45.

[9] Grossman, T. and Balakrishnan, R. (2004). Pointing at trivariate targets in 3D environments. *ACM CHI*.

[10] Grossman, T. et al. (2004). Multi finger gestural interaction with 3D volumetric displays. *ACM UIST* p. 61-70.

[11] Hubona, G. et al. (1999). The relative contributions of stereo, lighting, and background scenes in promoting 3D depth visualization. *ACM Transactions on Computer-Human Interaction*, *6(3)*. p. 214-242.

[12] Lucente, M. (1997). Interactive three-dimensional holographic displays: seeing the future in depth. *Computer Graphics issue on Current, New, and Emerging Display Systems*, *May 1997*.

[13] McCauley, M. et al. (1992). Cybersickness: Perception of self-motion in virtual environments. *Presence*, *1(3)*. p. 311-318.

[14] McKenna, M. (1992). Interactive viewpoint control and three-dimensional operations. *Symp. on I3D Graphics*. p. 53-56.

[15] Pfautz, J. (2000). Depth perception in computer graphics. *Doctoral Dissertation*, University of Cambridge.

[16] Rosen, P. et al. (2004). Perception of 3D spatial relations for 3D displays. *Stereoscopic Displays XI*. p. 9-16.

[17] Shibata, T. et al. (2005). Stereoscopic 3D display with dynamic optical correction for recovering from asthenopia. *Stereoscopic Displays and Virtual Reality Systems XII*. p. 1-9.

[18] Sollenberger, R. and Milgram, P. (1991). A comparative study of rotational and stereoscopic computer graphic depth cues. *Human Factors Society Annual Meeting*. p. 1452-1456.

[19] Sollenberger, R. and Milgram, P. (1993). Effects of stereoscopic and rotational displays in a three-dimensional path-tracing task. *Human Factors*, *35(3)*. p. 483-499.

[20] Tyler, T.,et al. (2005). The utility of perspecta 3D volumetric display for completion of tasks. *Stereoscopic Displays and Virtual Reality Systems XII*. p. 268-279.

[21] van Orden, K. and Broyles, J. (2000). Visuospatial task performance as a function of two- and three-dimensional display presentation techniques. *Displays*, *21(1)*. p. 17-24.

[22] Wanger, L. (1992). The effect of shadow quality on the perception of spatial relationships in computer generated imagery. *ACM Symp. on Interactive 3D Graphics*. p. 39-42.

[23] Wanger, L. et al. (1992). Perceiving spatial relationships in computer-generated images. *IEEE Computer Graphics and Applications.*, *12(3)*. p. 44-51, 54-58.

[24] Ware, C., Arthur, K., and Booth, K. (1993). Fish tank virtual reality. *ACM CHI Conference on Human Factors in Computing Systems*. p. 37-42.

[25] Ware, C. and Franck, G. (1996). Evaluating stereo and motion cues for visualizing information nets in three dimensions. *ACM Trans. on Graphics*, *15(2)*. p. 121-140.

[26] Yonas, A. (1979). *Attached and cast shadows*, in *Perception and pictorial representation*, p. 100-109.