An Assessment of Differences in Human Depth Understanding in Cube Displays Utilizing Light-Field Displays

Raymond Swannack and Oky Dicky Ardiansyah Prima
Graduate School of Software and Information Science
Iwate Prefectural University
152-52 Takizawa, Iwate, Japan
e-mail: g231s501@s.iwate-pu.ac.jp, prima@iwate-pu.ac.jp

***Abstract*—The proliferation of study into three-dimensional (3D) digital content affirms its popularity, not only in academics but in the business and private sectors as well. Virtual and augmented realities have been an attractive means of viewing and interacting with three-dimensional (3D) content. A Light-Field Display (LFD) allows users to experience stereoscopic images, using a similar approach to Virtual Reality but without the need for a headset. This study aims to expand upon the preliminary experiment, in which we evaluated the benefits of stereoscopic depth cues in human visual understanding through users viewing a 3D scene on a cube display. Our task scenario involves users interacting with the 3D cube display to judge the distance between objects and completing a questionnaire about the experience. For each test, using the LFD "Lume Pad" produced by Leia Inc., 3D contents were presented with and without stereoscopic depth cues. Our results show that users can judge the distance between objects with more certainty and with fewer errors with stereoscopic depth cues.**

Keywords-Light-Field Display; Fish Tank VR; 3D human perception; stereoscopic vision.

I. INTRODUCTION

Two-dimensional (2D) screens allow for almost anything imaginable to be displayed. Complicated data is transformed into graphs, charts, lists, and other visuals designed for easy accessibility. Users can see places, objects, cultures, and creatures that they have only ever heard of. However, there are limits to what can be conveyed, as a 2D screen is incapable of showing true depth because it has no real depth. This paper is an extension of our previous work, evaluating the benefits stereoscopic depth cues bring to perceived depth understanding when using a Fish Tank VR cube display [1].

Since a 2D screen has no depth, the illusion of depth must be created using depth cues. Some of these cues are stereoscopy, convergence, occlusion, relative size, height in the visual field, relative density, aerial perspective, accommodation, and motion parallax [2]. These cues are utilized to a greater or lesser extent to recreate depth on a flat surface, such as a 2D screen, and the developing technology that manipulates these cues is what 3D screens are built around.

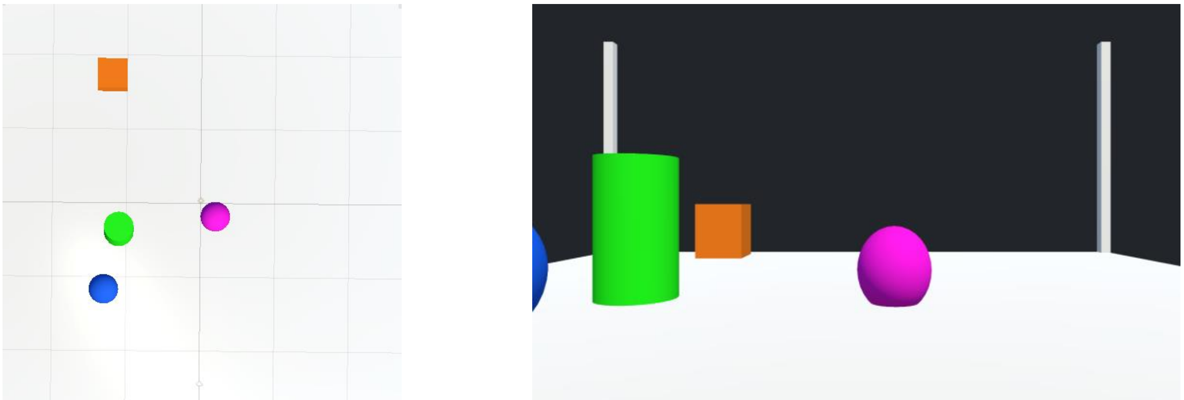
With the rise of virtual and augmented reality headsets, 3D digital content has become more commonplace. This content allows users to interact with 3D media in a fashion that was unimaginable a decade ago. A key component to these devices is that they create stereoscopic depth cues, giving the user the sensation of seeing a true 3D scene.

A lesser-known category of this type of display is handheld displays, such as smartphones and tablets. Further advancements have been made with 3D content in hand-held displays, with the vast majority of these advancements coming in the form of augmented reality applications that use the device's camera and add digital content to what the camera can see. Conversely, some handheld devices are being designed with hardware to be used primarily as AR or VR devices.

One example avenue of these hardware advancements is handheld LFDs. LFDs use curved lenses, known as lenticular lenses, to bend the images displayed behind them in such a way as to cause each of the viewer's eyes to see a different image. This creates an experience similar to VR and AR headsets but without the need for a headset. Most people have come into contact with this technology, albeit in a less technologically advanced form. Lenticular lenses placed over static images create pictures that appear to be 3D, move, or change into another image entirely. This is a popular technique used to make unique postcards and stickers. An LFD brings this technology to a digital screen. Many LFDs are very large and complicated, using multiple projectors aligned with complex lenticular lens arrays to create 3D images that can feel like someone is looking at a hologram from a science fiction movie. One such example of this is the 120-degree viewing angle LFD designed by Liu et al. [3].

Another form of 3D display is what is known as Fish Tank VR displays. These displays have a 3D shape and thus can make use of actual depth. In theory, any shape that can be created with physical objects can be made into a display, allowing for custom displays to be built specifically for individual tasks. One example is the SpheriCul [4] designed by Prima Lab. This sphere-shaped display has many different uses, such as allowing users to walk around the sphere display and see any content that is globe-shaped or to use a tracker to have the display alter its image to fit what the user should be seeing from the angle they are standing.

The purpose of this study is to improve our preliminary findings and to present the benefits to a user's understanding of perceived distance in a cube display by adding stereoscopy. To this end, three new tests were created. These tests were designed to improve upon the previous experiment and to expand upon the ideas proposed in that paper. The first test takes the lessons learned from our previous work to improve



(a) Top-Down view from Unity

(b) View from the cube display

Figure 1. A sample scene for Test One: Perceived Distance Test

upon the previous experiment, the second test focuses on the benefits of stereoscopy when other depth cues are removed, and the final test is a further expansion of our previous experiment.

The remainder of the paper is structured as follows. Section II presents the methodology of the experiment to evaluate the subject's understanding of the portrayed distance between objects on a cube display with and without stereoscopy. Section III describes the details of the hardware and software used in the experiment. The preliminary experiment is discussed in Section IV, covering both the findings and what improvements were considered for this follow-up experiment. Section V details the main experiment as well as the results and questionnaire. In Section VI the findings and implications are discussed. Section VII presents the conclusion and discusses future work.

II. Methods

The focus of this study was to measure a subject's understanding of 3D distance portrayed on a 2D screen, given different visual cues. The primary objective was to show the benefits of adding stereoscopic depth cues to a cube-type fishbowl VR device. To demonstrate this, three tests were devised. Test One asked the subjects to identify which objects were closest and farthest from a specified object. Test Two was similar to the first but with many customary depth cues removed. Finally, Test Three was a tree-tracing test where subjects were asked to find a node on a branching tree and asked to trace it back to the base of the tree.

A. Visual Cues

The primary depth cues observed throughout this experiment, to a greater or lesser extent, are motion parallax, occlusion, relative size, height in the visual field, and stereoscopy. The visual cue being studied in the greatest depth in this work is stereoscopy.

In nature, each human eye naturally sees a slightly different angle of the same scene. The brain receives these two images and creates the 3D view that humans see when observing the world around them. How the eyes see different images is called stereoscopy. When humans look at a flat object, stereoscopy is not achieved because each eye is seeing effectively the same image.

While moving or observing movement, humans perceive objects closer to themselves as moving faster than objects farther away. This is more prevalent in a vehicle moving at speed and is known as motion parallax. Occlusion occurs when an object is placed in front of another object. When this happens, one object will block vision to some of or all of the other object. This blocking tells our brain which object is nearer to us. Objects whose size is known can be used to compare the size of similar objects. This is the depth cue known as relative size. Finally, height in the visual field is how when a viewer is looking down on a scene, objects farther away appear taller than those closer to the viewer.

B. Test One: Perceived Distance Test

To assess the subject's understanding of the simulated depth within the cube display, a 3D scene was created with four objects positioned within it. This scene includes a green cylinder, a blue sphere, a pink sphere, and an orange cube. Subjects were asked to locate the green cylinder and to identify which objects they felt were closest to the green cylinder and which object they felt was the farthest from the green cylinder.

The objects were set on a white floor and beams were added to mark the corners of this floor. This was done to give the subjects more visual cues, specifically to enhance motion parallax and occlusion. Figure 1(a) shows an example layout for one scene. Subjects did not see this view. Figure 1(b) shows the same scene as it is displayed on one tablet in the cube display.

C. Test Two: Limited Depth Cue Test

Similar to Test One, a 3D scene was created with four spheres positioned within it. The colored spheres were green, blue, pink, and orange. Subjects were asked to locate the green sphere and identify which object they felt was closest to and farthest from that sphere.

(a) Top-Down view from Unity

(b) View from the cube display

Figure 2. A sample scene for Test Two: Limited Depth Cue Test

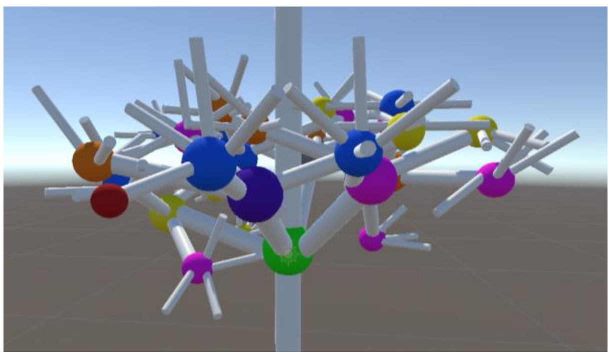


Figure 3. A sample scene for Test Three: Tree Tracing Test

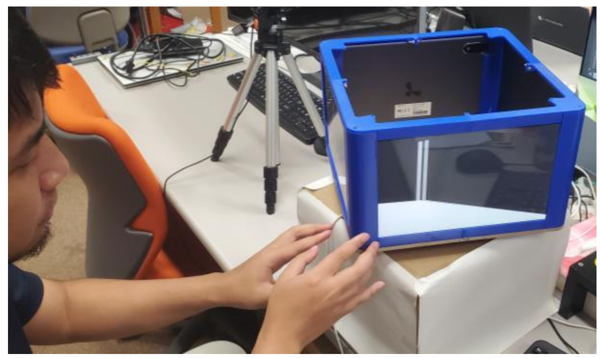


Figure 4. A subject interacting with the cube display

The difference from Test Two is that some depth cues were removed. Specifically, occlusion, height in the visual field, and relative size were either removed or steps were taken to lessen their effectiveness. The objective was to put more emphasis on the remaining depth cues and highlight the benefits of stereoscopy. To achieve this, the floor was removed, and all of the objects were made transparent, removing occlusion. The camera was lowered to no longer look down onto the scene but instead look directly into the scene. Finally, some object's sizes changed between scenes, hindering the effects of relative size. Figure 2(a) shows an example layout for one scene while Figure 2(b) shows the same scene as it is displayed on one tablet in the cube display.

D. *Test Three: Tree Tracing Test*

A branching tree was placed at the center of the scene, as can be seen in Figure 3. On the trunk of the tree, a green sphere was placed. At the end of one of the outermost branches, a red sphere was placed. These spheres marked the endpoints for the test. Subjects were asked to traverse the tree from one of these points to the other. Colored spheres were placed at each branching point to give the subject a path to follow when traversing the tree.

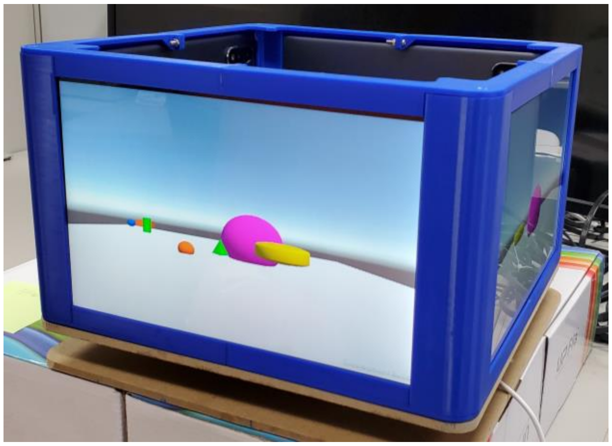
E. *Experimental Procedure*

The procedure for the experiment was as follows. Firstly, the subjects were seated at the desk and were told what was expected of them within the experiment. This verbal preparation was followed by a short demonstration of how to interact with the cube display and how to perform the first experiment. Subjects were told to take as much time as they needed and to move as much as they felt was necessary to see the cube from any angle, they felt would help them.

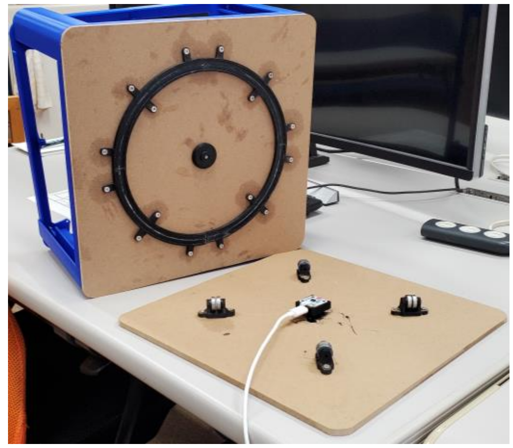
No limit was set for the distance a subject could be from the display. Although the display works best at 45-55 cm in front of the display and within a 40-degree viewing angle in front of the display, the priority was to allow the user to interact with the display in what they felt was a natural way. The optimal viewing distance and viewing area were explained and subjects were told that if they moved beyond these positions, the display may not operate optimally and their immersion might be hindered, but subjects were free to interact with the display as they saw fit. An example of a subject interacting with the cube can be seen in Figure 4.

Each test consists of two parts, with each part comprising half of the scenes designed for the test. The first part of each test was performed with the LFD turned off. The second part was performed with the LFD on. Test One consisted of ten scenes, while Test Two and Test Three consisted of twelve scenes.

When each part of the test was completed, the subject was asked how confident they felt in their choices and why they felt that way. After both parts of the test were finished, the



(a) The cube display was constructed with 3D printed parts



(b) The Encoder and the roller track

Figure 5. The Cube Display setup for the experiment

subject was asked if they felt more confident with the LFD or without the LFD and why they felt that way.

III. Hardware and Software used in Experiment

A. Lume Pad

The following research was performed using Leia Inc's Lume Pad, an LFD tablet. As discussed above, the LFD gives stereoscopic depth cues to the user. The tablet possesses a 10.1-inch screen with a resolution of 2560x1600 pixels. To create the light field effect, the tablet displays four images simultaneously and uses lenticular lenses to allow the user to see two of these images at a time. As the user moves horizontally or rotates the cube, each eye will see a new image, thus creating a different view. In this way, three different views are created when the cube is stationary, and stereoscopic depth cues are added to all the other depth cues the subject sees when interacting with the cube.

To generate each of these different views, the Lume Pad uses four images that are divided into a 2x2 grid which is saved as a single image file. This image is then split and placed into the correct locations for the lenticular lenses. Due to the need to put four images into one screen, each image can only make use of one-quarter of the total resolution, so each view has a resolution of 640x400 pixels [5].

A common method of measuring how much detail a screen can show is through pixel density. This is calculated as follows,

$$PPI = \frac{Diagonal\ in\ Pixels}{Diagonal\ in\ Inches} \quad (1)$$

$$Diagonal\ in\ Pixels = \sqrt{Width^2 + Height^2}\ . \quad (2)$$

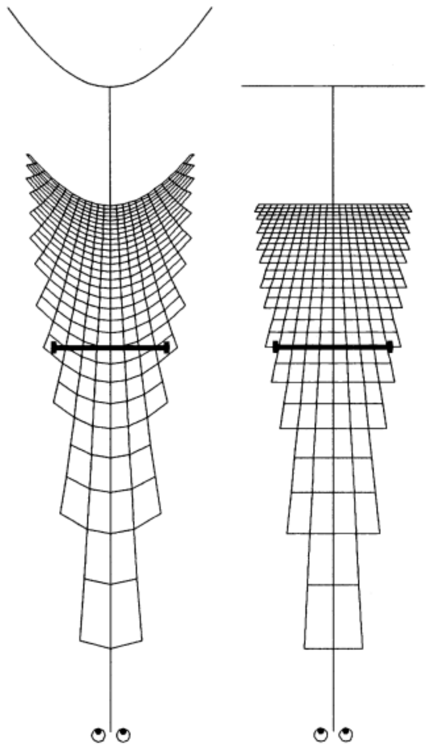
The width and height pertain to the pixel dimensions of the tablet's screen. In this case, the Lume Pad has a pixel width of 2560 pixels and a pixel height of 1600 pixels. This gives the Lume Pad a natural pixel density of 290 pixels per inch (ppi). When the LFD is used, the width and height are reduced to 640 and 400 pixels respectively. Given these values, the pixel density is about 75 ppi. This comparison appears to be poor because the LFD has 25% of the ppi that the standard tablet, but this is a misnomer due to the tablet being small and having a relatively high resolution. For a more realistic comparison, the LFD can be compared to a standard computer monitor. A popular computer screen size is 24 inches with a resolution of 1920x1080 pixels, giving this screen a pixel density of 92 ppi. This is a much more favorable comparison for the Lume Pad with the LFD turned on, as it has 81.5% of the pixel density of a standard computer screen.

To cope with this reduction in pixel density, the Lume Pad uses some techniques to cause the image to appear clearer than one might expect. The smaller screen size, when compared to a standard PC monitor is one of these techniques as well as the orientation of the lenticular lenses. The lenses in the Lume Pad are not perfectly vertical but instead are slanted slightly. This allows for smoother transitions between views. It has also been shown as an effective method of blending views together to cause the user to believe they are seeing more views than are actually being displayed [6].

B. Cube Display

Using four Lume Pad tablets, a cube display was constructed, as seen in Figure 5(a). It was decided to only utilize the sides of the cube and not the top. The reason for this is that the Lume Pad is designed to be viewed from the front. It is not designed to be placed flat on the table in front of the user and rotated in a circle. If it was used in this manner, the lenses would not work optimally and could display very confusing stereoscopic images. For these reasons, it was decided to limit the displays to the sides of the cube.

Bearings were placed on a track to allow the cube to rotate freely, as can be seen in Figure 5(b). The tablet supports, as well as the track for the roller bearings, were created using a 3D printer. A magnetic encoder was used to track the rotation of the display. The magnet was placed in the bottom of the tray holding the Lume Pads, with the encoder placed on the stationary base and then connected to a PC.



(a) Toed-in camera (b) Parallel camera

Figure 6. 3D maps of (a) toed-in cameras and (b) parallel cameras versus object distance [8].

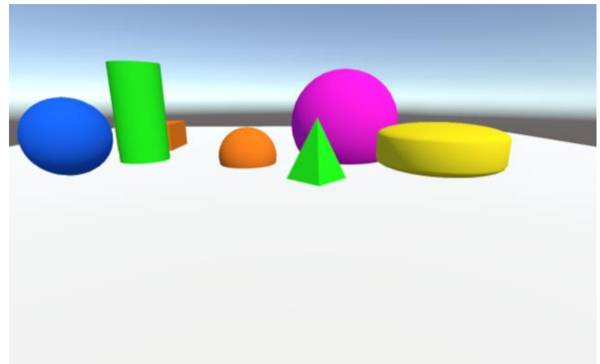


Figure 7. A sample scene from the Preliminary Experiment as seen from one tablet in the cube display

TABLE 1. PRELIMINARY EXPERIMENT ERROR RATE (%)

Subject	Non-LFD	LFD
1	10.0	0.0
2	20.0	10.0
3	20.0	10.0
Mean	16.67	6.67
St. Dev	5.774	5.774
Min	10.0	0.0
Max	20.0	10.0
Median	20.0	10.0
Mode	20.0	10.0

C. Unity

The tests performed in this experiment were developed using Unity with the Leia Unity Software Development Kit (SDK) [7]. This SDK is produced by Leia Inc to help developers create applications for their devices. The SDK allows for the utilization of the Lume Pad's features, such as the special Leia camera which creates stereoscopic depth cues, as well as the ability to change the state of the LFD from on to off or vice versa within the test.

The Leia camera consists of four cameras, aligned in parallel with each other. Aligning the cameras in parallel is important to avoid the depth place curvature known as the keystone distortion, which can be seen in Figure 6. This can occur if the cameras are not aligned properly, such as if the cameras are angled inward (toed-in cameras) [8]. This distortion is not as visible when looking at objects close to the viewer, but when looking far away from the user, the image is distorted in a way that does not align with how human eyes naturally see the world and can break the viewer's immersion in the scene.

By using Unity, this allowed a PC to perform most of the necessary computations. This in turn reduced the workload that each tablet needed to perform within the tests, limiting the lag caused by the tablet hardware. Data from the encoder would be sent to the PC, which would do the calculations for camera positions, and then send this information to the tablets via a Wi-Fi connection.

IV. PRELIMINARY EXPERIMENT

In the preliminary experiment, three subjects, all males with an average age of 23 years and normal or corrected-to-normal vision, were asked to perform an experiment similar to the one described above in Test One. Seven objects were placed in each scene and the size difference between each object was much larger, as can be seen in Figure 7. Subjects were first asked to find the green pyramid and choose which object they felt was closest to it. Then they were asked to locate the green cylinder and choose the object they felt was closest to it. Then the scene was changed, the objects were placed in new locations, and this procedure was repeated four more times. After five scenes had been completed, the LFD was turned on and the second five scenes were completed. The data collected was the error rate of each subject in choosing the correct object.

A. Results

After the experiment, subjects were asked to fill out a questionnaire to ascertain how well they believed they had understood the test as well as if the test had any adverse effects. The questions were as follows:

TABLE 2. MAIN EXPERIMENT ERROR RATE (%)

Subject	Test One		Test Two		Test Three	
	Non-LFD	LFD	Non-LFD	LFD	Non-LFD	LFD
1	10.0	20.0	41.7	41.7	0	0
2	20.0	20.0	40.0	30.0	0	0
3	10.0	10.0	50.0	20.0	0	0
4	0.0	0.0	16.7	0.0	0	0
5	0.0	40.0	33.3	41.7	0	0
6	10.0	10.0	41.7	8.3	0	0
7	20.0	10.0	33.3	33.3	0	0
8	30.0	20.0	41.7	16.7	0	0
9	20.0	20.0	50.0	30.0	0	0
10	10.0	0.0	50.0	25.0	0	0
11	20.0	10.0	41.7	16.7	0	0
12	10.0	30.0	50.0	16.7	0	0
13	30.0	30.0	50.0	75.0	0	0
14	10.0	30.0	50.0	30.0	0	0
15	20.0	10.0	33.3	25.0	0	0
16	20.0	20.0	50.0	50.0	0	0
Mean	15.00	17.50	42.08	28.75	0	0
St. Dev	8.944	11.255	9.379	17.769	0	0
Min	0.0	0.0	16.7	0.0	0	0
Max	30.0	40.0	50.0	75.0	0	0
Median	15.0	20.0	41.7	27.5	0	0
Mode	10.0	20.0	50.0	30.0	0	0

a. With the LFD turned off, how well do you feel that you understood the scene? Did you know where everything was?
b. With the LFD turned on, how well do you feel that you understood the scene? Did you know where everything was?
c. How confident were you in your understanding?
d. How much discomfort did you feel? Did your eyes hurt? Did you feel sick?

Subjects reported that they felt the scenes were difficult to understand both with and without the LFD. While the LFD helped them to feel more confident in their understanding of the scenes, they were still often uncertain about which of the objects were closest to the green target objects. Most said that they believed that the objects were either too far away or too small to feel confident about their choices. None of the subjects reported any ill effects from the experiment. After the experiment, the subjects were interviewed to try to understand their choices in the tests.

Despite the small sample size, an interesting trend was observed. The number of errors with the LFD cube display was smaller than with the standard cube display, as can be seen in Table 1. A mean error rate of 17% was recorded for the standard cube display while 7% was recorded for the LFD cube display. The standard deviations are the same, and the remaining stats are close to equal as well.

B. Improvements for the Main Experiment

Creating more comprehensible scenes was the most important issue that needed to be addressed in further experiments. More research was done concerning the depth cues that should be the most important for this style experiment. It was decided that occlusion and relative size had not been given enough consideration and had potentially been used incorrectly in the preliminary experiment.

To this end, the size of the objects within the scene was somewhat standardized, making objects roughly three meters tall and three meters wide, using Unity's internal measurements. Previously, the objects had ranged in size from as small as one meter to as large as twelve meters. By standardizing the sizes, subjects could feel more confident in understanding the distance between objects by using the size of objects within the scene as a form of visual measurement.

In the preliminary experiment, it was believed that by placing objects in the scene in such a way as to block some viewpoints, the subjects would need to interact more with the display to understand the distance between the objects. Instead, this caused the subjects to choose answers that they were not confident in. Subjects did this because there were not enough good viewing angles due to occlusion. One method used to address this was to reduce the number of objects in each scene from seven to four. With fewer objects in each scene, more angles were useful to the subject.

Further consideration was also given to how to create a test that would put more emphasis on the stereoscopic depth cues that the LFD can produce. It was believed that by removing other depth cues, more emphasis would be placed on the remaining cues. To this end, Test Two was designed, where occlusion and depth in the visual field were removed as well as relative size being reduced. Given the feedback from the

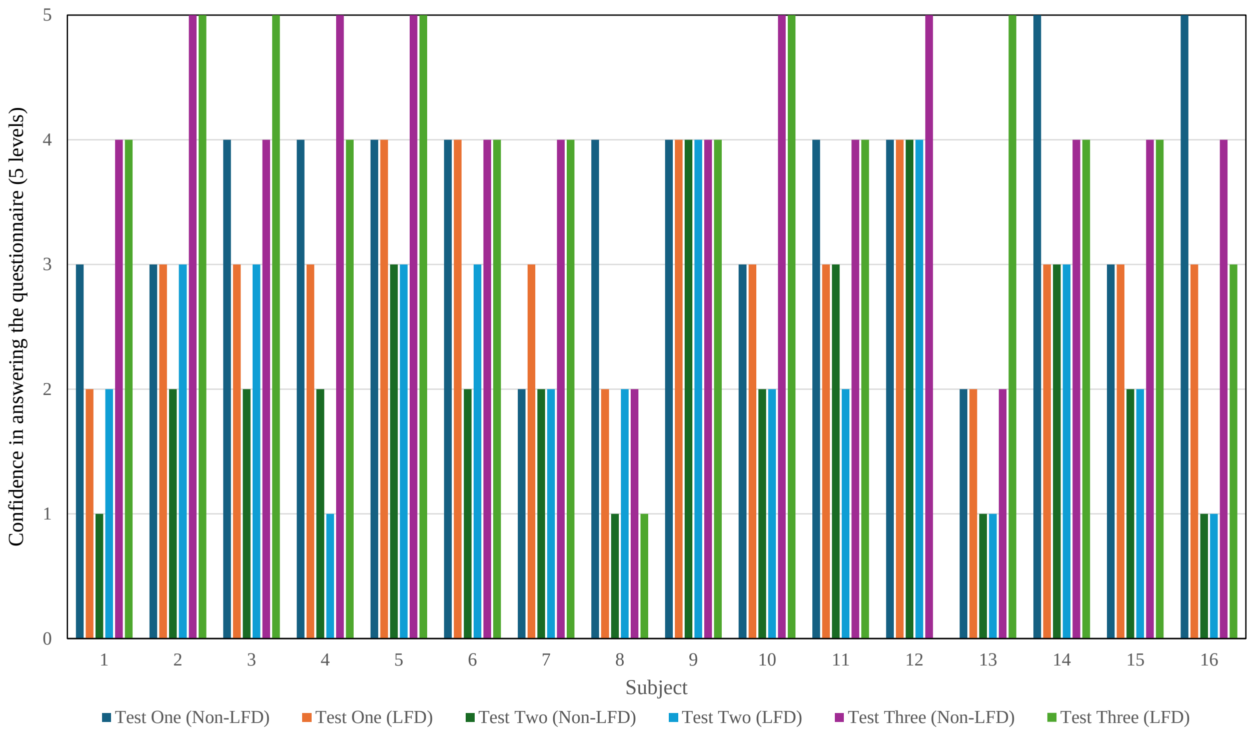


Figure 9. Confidence in answering the questionnaire (5 levels)

preliminary experiment, a lot of care was given to this experiment to make sure it would produce useful data.

V. MAIN EXPERIMENT AND RESULTS

Sixteen subjects (six females and ten males) participated in the experiment. They ranged in age from 20 to 31 with an average age of 24 years. All subjects had normal or corrected-to-normal vision. The experiment was carried out over two weeks, and the subjects were asked not to discuss the tests with other subjects.

The subjects performed all of the tests in one session, with the tests performed in the same order. The experiment started with Test One, progressed through Test Two, and concluded with Test Three. In each test, the non-LFD part was performed first and the LFD part was performed second. Every thirty minutes a short break was taken, though most subjects finished before thirty minutes had elapsed. Participants were encouraged to ask questions if they did not fully understand what was expected of them.

A. Scores

The results for Test One can be seen in Table 2. From this table, it can be observed that the non-LFD results were stronger in almost every metric. From this table, it can be seen that the addition of stereoscopic depth cues increased the error rate, standard deviation, maximum, median, and mode. Only the minimum score was unchanged.

A paired samples t-test was performed to compare the error rate between a cube display with stereoscopic depth cues via the LFD and without stereoscopic depth cues. There was not a significant difference in error rate between the cube display without stereoscopic depth cues ($M=15$, $SD=8.944$) and the cube display without stereoscopic depth cues ($M=17.5$, $SD=11.255$); $t(15) = -0.719$, $p=0.483$.

The results of Test Two were much more encouraging. The results can be seen in Table 2. The error rate decreased when the LFD was active, while the maximum score, median, and mode increased. These results are positive for the LFD, but the standard deviation increased, meaning that the scores have more variance between them, and this is backed up by the minimum score decreasing as well. Neither of these results are positive for the LFD.

A paired samples t-test was performed to compare the error rate between a cube display with stereoscopic depth cues via the LFD and without stereoscopic depth cues when other depth cues are removed. There was a significant difference in error rate between the cube display without stereoscopic depth cues ($M=42.1$, $SD=9.379$) and the cube display without stereoscopic depth cues ($M=28.8$, $SD=17.769$); $t(15)=3.228$, $p=0.006$.

The results of Test Three were not as promising. Every subject achieved a perfect score on this test which can be seen in Table 2. The reasons for this are addressed in the discussion section of this paper. Since every subject achieved a 0% error rate, no t-test or any other test was performed on these data.

B. Questionnaire

After completing each section of the test, subjects were asked how confident they felt in their answers. These responses were recorded as a score from one to five with five being extremely confident and one being not confident at all. These results can be seen in Figure 9. Subjects were then asked if they felt more confident with or without the LFD, or neither. Finally, once all the tests were completed, subjects were asked if at any point the display cube made them feel sick or uncomfortable and whether or not they felt that the LFD helped their understanding of distance on the cube display. In Test One, ten subjects stated feeling more confident with the LFD off, three stated feeling more confident with the LFD on, and three stated feeling the same level of confidence with and without the LFD. In Test Two, six subjects stated feeling more confident with the LFD off, seven stated feeling more confident with the LFD on, and three stated feeling the same level of confidence with and without the LFD. In Test Three, four subjects stated feeling more confident with the LFD off, eleven stated feeling more confident with the LFD on, and one stated feeling the same level of confidence with and without the LFD.

VI. Discussion

Test One and Test Three did not produce the desired results for reasons particular to each of the tests, while Test Two produced unexpected but useful results. In Test One, many subjects stated that the LFD confused them. Only objects that were at a medium distance were in focus. Objects that were near or far away from the viewer were blurry or sometimes they appeared to be two objects. This caused the subjects to lose confidence in their understanding of the scene. The reason for this became obvious during the interviews and highlights a limitation of LFDs.

A. Test One

The results for Test One were surprising. This test was built upon the preliminary experiment, so it was assumed that the results would follow a similar trend. Instead, the results were the opposite. While these results were not what had been desired, they still were valuable as a learning tool.

When humans look at something, their eyes angle toward that thing, this is called eye convergence. This causes whatever the person is looking at to be in focus while everything else in their field of vision is less in focus. This is essential to how stereoscopy works. LFDs simulate eye convergence with a convergence plane. This sets a focus plane, which can be seen in Figure 10, where objects on or near this plane will be well-defined while objects farther from this plane are in less focus. Objects that are not near the convergence plane will appear to be in a different location for each eye while objects near the convergence plane will appear to be near the same location. Objects that are near the same location appear to be 3D while when each eye sees an object in a different location it causes that object to be blurry or in extreme examples there appear to be two objects.

In tests performed while redesigning the scenes for the main experiment, the convergence plane's position was tested at 5, 10, 20, 25, and 30 meters, but subjects stated that 10 meters felt the most realistic to them. The test area was composed of a square that was 50 meters in length and width, so setting the plane at 25 meters was the original plan, but this caused many subjects to state that they felt dizzy or nauseous. The same feedback was received when the convergence plane was set at both 5 and 30 meters.

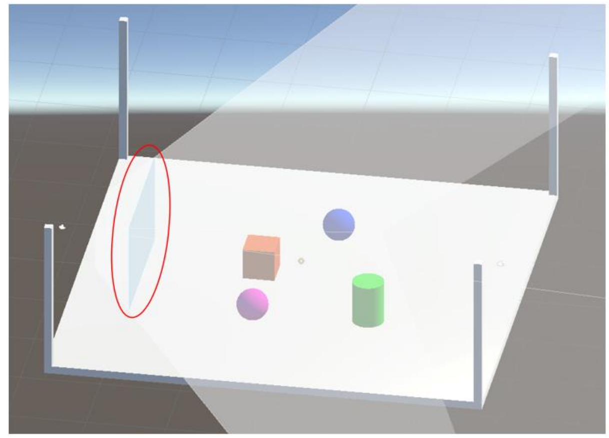


Figure 10. Convergence Plane (red circled area) as shown in Unity

In a 3D scene where objects are placed at any x and z coordinates, it is not possible to set a convergence plane that covers all objects in the scene. Subjects needed to shift their focus from one object to another to judge the distance between objects, but the LFD could not do that naturally. To combat this, the objects could have been clustered more closely together, which would have helped but not completely alleviated the problem.

Another option would be to allow the subjects to control the position of the convergence plane, but this carries some inherent risk. Based on a small trial, performed after the main experiment's results had been gathered, some individuals feel a sensation of moving when looking at a scene where no objects were in motion but the convergence plane was changing. If the convergence plane changes too quickly it can cause motion sickness, and trying to judge where the plane is located was difficult. Another option would be to allow subjects to choose objects to focus on and allow the software to move the plane accordingly. However, both of these options still cause objects to be out of focus and blurry, so neither is a perfect solution.

B. Test Two

In designing Test Two, a lot of care and consideration was given to the layout of each scene. There was a fear that by removing so many depth cues the scenes would be too confusing for subjects to understand the distance between any of the objects and this would cause the results to be inconclusive.

While the test suffers from the same problems as Test One, subjects stated that with there being fewer other cues to aid them, the stereoscopic depth cues were beneficial to their confidence. At first glance, the results are promising but questionable. When the LFD was on the error rate, median, mode, and maximum scores all increased, the standard deviation went up and the minimum score went down, meaning that the scores were less tightly grouped together but the average score was better. Overall, the results were good.

C. Test Three

The original concept for Test Three was to compare both completion time and error rate. After some preliminary testing, the time comparison was abandoned. This was largely due to three concerns. Firstly, the display is large, heavy, and expensive. Turning the cube quickly risked it falling off the roller track and getting damaged. Another worry is that the encoder might not detect the magnet's rotation and the display would not show the correct images. Also, because of the weight, quick rotations would cause a decent amount of momentum and would be difficult to stop. Again, this increases the risk of damage to the display. Secondly, the Lume Pad's hardware sometimes causes lag or dropped frames. This tablet is designed for portability, not for quickly calculating positions in 3D and rendering images based on those computations. The PC handled most of these calculations, but some work still needed to be done on the tablet. The final constraint was the Wi-Fi connection that was used to transmit the data. While the connection was usually strong, it would occasionally drop a few packets, or in one case, disconnect from the network entirely. These variables can add unintended seconds to each test and would make the results unrepresentative of the actual time taken for each task. For these reasons, the subjects were not timed.

To compensate for the removal of the time comparison from the test, attempts were made to make the test more difficult. The position of the red sphere was often moved to locations that were difficult to see from an accessible angle. It was often placed behind other objects, causing subjects to need to view it from multiple angles to understand where it was, and the path needed to find the base of the tree. However, the test was not difficult enough, as can be seen by the perfect scores achieved by each test subject.

D. Subject feedback

On the questionnaire, four subjects stated that they felt sick or uncomfortable at some point during the test, one subject asked to take a short break during the test before continuing. Many subjects stated that Test Two caused them to feel a little dizzy. Other feedback was very enlightening. One subject stated that the seams of the display, the corners of the cube that hold the tablets in place, were very thick and broke their immersion to some level. Another subject stated that when the LFD was active the edges of the spheres were harder to determine than edges of the cube or pyramid. This suggests that objects with sharp corners make for better objects for stereoscopic depth cues, especially when these objects are not near the convergence plane.

VII. Conclusion and Future Work

In this study, we examined a user's understanding, given some constraints, of a distance displayed on a display cube utilizing LFDs in an attempt to show that the human brain understands distance better with the inclusion of stereoscopic depth cues than without them. It was shown that using an LFD to display stereoscopic depth cues increased subjects' understanding of the simulated distance over when the cube did not display stereoscopic depth cues when other depth cues were removed or suppressed. The t-test results from our second test allow us to state this with confidence.

A comparison of our results can be made to the findings of Reising and Mazur [9]. In their research, pilots performed an airspace disambiguation task, in which the subjects were shown a battle situation display with and without stereoscopic depth cues. Subjects were asked to identify the number of aircraft in a given area in a target group. For example, a pilot might be asked to tell the number of friendly aircraft in the quadrant that is in front of and above the pilot's aircraft. In their study, it was shown that the addition of stereoscopic depth cues did not provide a benefit, except for when other depth cues were limited. These results align with the results from our study.

Another study that has comparable results to ours is that of Ntuen et al. [10] as it resembles Test One. In their study, a cone and three colored spheres were added to a scene from Google Maps and users were asked to identify which sphere was closest to the cone. This test was performed with and without stereoscopic depth cues and then the results were compared. The results showed that adding stereoscopic depth cues did not produce a significant improvement, which is what our results in Test One showed as well.

An improvement that can be applied to the entire experiment is to improve the refresh rate of the tablet screens used in the display. The displays generally would operate at thirty frames per second but, as discussed in the previous section, sometimes the screens would lag. Further optimizations can be made to the tablets when rendering the scenes. If this lag can be brought low enough, recording the time it takes for each subject to complete each task would be very beneficial. Response time comparisons could be made not just between with and without the LFD, but between Test One and Test Two as well. These results could show us which depth cues are the most beneficial in a subject's understanding of the 3D scene.

To further this experiment, firstly modifying Test One is a high priority. Enlarging the objects and bringing them closer together would fix some of the problems that the test suffered from. It will not fix all of the problems but based on the results and feedback from Test Two, it should fix enough of the problems that we will obtain better results.

An idea for enhancing Test Two is to create tests where only one of the depth cues is removed, such as only removing occlusion or height in the visual field. This would help to isolate the benefits of the LFD and maybe can highlight if other depth cues are more important.

Finding a method to alter Test Three is more difficult. Seeing as many of the problems encountered are related to the hardware, creating a dedicated device would be the obvious response. An LFD display cube designed from scratch could be made lighter and with better hardware. Hopefully, the lag will be overcome. The device could also be made to be lighter, allowing for a user to carefully pick it up and interact with it instead of it being attached to a track and only rotating on one axis. The drawbacks of this idea come down to the cost and complexity of LFDs. Even if the cube is easier to interact with, a great amount of care would need to be taken when handling the cube.

Comparisons can be made between our work and that of Stavness et al. [11] with their pCubee device. This is a cube display from which we drew a great deal of inspiration. One of the key differences is that pCubee uses a head tracker to create a realistic 3D experience by tracking where the user's head is and adjusting the screens accordingly. Without any form of head tracking our display works similarly, but only to a point. The current LFD works with horizontal movements but does not work with vertical movements. Without adding a form of user or more advanced device tracking this cannot be added. This is something that should be strongly considered going forward. With a redesigned device more complex device tracking can be achieved. If the device can track how high it is lifted or tilted, then the displays can compensate for this and reproduce these changes within the display. This would allow our cube to deliver a stereoscopic experience without the need for any form of user tracking.

If a form of user tracing was to be added, eye tracking would be the most beneficial, though the implementation would be very complicated. More advanced LFDs use eye tracking to adjust the images used in the LFD to create a smoother 3D experience for a single user [12]. Based on feedback, one problem with the cube display was that if the subject moved while the display was stationary, the 3D image would not shift seamlessly. While moving from the area where one image is displayed and into where another is displayed, the view would occasionally become confusing. Moving a short distance to either side solved this issue, but creating a smoother experience would still be beneficial.

Another area that has not been explored is the potential for multiple users to interact with the LFD cube display. Since an LFD has an optimal viewing area and is not tethered to a single individual, multiple users can experience the stereoscopic depth cues that the LFD produces. A test that can ascertain two or more users' understanding of the depth cues visible on the display cube would be compelling. Either a test that requires multiple users to interact with the cube, or a test that requires users to cooperate to succeed would be an idea for these tests. More research needs to be done to determine the best type of test to achieve satisfactory results. Also, this cannot be used natively along with eye tracking. If the images on the display are changing for one individual, any other individuals will see a moving scene that they are not interacting with. An option to address this would be the addition of a way to isolate the view that each subject can see such as shutter glasses.

A final area that should be explored is to look at other methods of interacting with the display cube. pCubee could be interacted with such as a stylus and a PC mouse. In a 3D tree-tracing task, the researchers found that the fastest mean user response times were achieved with a combination of pCubee and a PC mouse while using only pCubee resulted in the slowest mean response times. A comparison between the findings of pCubee and this cube display would be enlightening.

References

[1] R. Swannack and O. D. A. Prima, "Assessment of Differences in Human Depth Understanding in Cube Displays Using Light-Field Displays," The Seventeenth International Conference on Advances in Computer-Human Interactions, ACHI 2024, pp. 210-213, 2024.

[2] J. E. Cutting and P. M.Vishton, "Perceiving layout and knowing distances: The integration, relative potency, and contextual use of different information about depth," In Handbook of Perception and Cognition, Perception of Space and Motion, *Academic Press*, pp. 69-117, 1995.

[3] B. Liu et al., "Time-multiplexed light field display with 120-degree wide viewing angle," *Optics Express*, *27*(24), pp. 35728-35739, 2019.

[4] SpheriCul: A New Perspective on 3D Displays, https://www.sphericul.com/ [Retrieved on 20 August 2024]

[5] 3D Lightfield Experience Platform, https://www.leiainc.com/ [Retrieved at May 2022]

[6] C. Van Berkel and J. A. Clarke, "Characterization and optimization of 3D-LCD module design," Proceedings of the SPIE, 3012, pp. 179-186, 1997.

[7] SDK and Developer Resources, www.leiainc.com/sdk. [Retrieved at 10 June 2022]

[8] A. J. Woods, T. Docherty, and R. Koch, "Image distortions in stereoscopic video systems," Proceedings of the SPIE 1915, Stereoscopic Displays and Applications IV, 1993.

[9] J. Reising and K. Mazur, "3-D displays for cockpits: where they pay off," In *Stereoscopic Displays and Applications*, 1256, pp. 35-43, 1990.

[10] C. Ntuen, M. Goings, M. Reddin, and K. Holmes, "Comparison between 2-D & 3-D using an autostereoscopic display: the effects of viewing field and illumination on performance and visual fatigue," in *International Journal of Industrial Ergonomics*, 39(2), pp. 388-395, 2009.

[11] I. Stavness, B. Lam, and S. Fels, "pCubee: a perspective-corrected handheld cubic display," In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 1381-1390, 2010.

[12] G. Zhang et al., "Directional and Eye-Tracking Light Field Display with Efficient Rendering and Illumination," *Micromachines*, 14(7), pp. 1465, 2023.