The Destiny-class CyberCANOE – a surround screen, stereoscopic, cyber-enabled collaboration analysis navigation and observation environment

Noel Kawano, Ryan Theriot, Jack Lam, Eric Wu, Andrew Guagliardo, Dylan Kobayashi, Alberto Gonzalez, Ken Uchida, Jason Leigh; Laboratory for Advanced Visualization and Applications (LAVA), University of Hawai i at Mānoa; Honolulu, HI/USA

Abstract

The Destiny-class CyberCANOE is a hybrid-reality environment that provides 20/20 visual acuity in a 13-foot-wide, 320-degree cylindrical structure comprised of tiled passive stereo-capable organic light emitting diode (OLED) displays. Hybrid-reality systems such as Destiny, CAVE2, WAVE and the TourCAVE combine surround-screen virtual reality environments with ultra-high-resolution digital project-rooms. They are intended as collaborative environments that enable multiple users to work minimally encumbered, and hence comfortably, for long periods of time in rooms surrounded by data in the form of visualizations that benefit from being displayed at resolutions matching visual acuity and in stereoscopic 3D.

Destiny is unique in that: it is the first hybrid-reality system to use OLED displays; it uses a real-time software-based approach rather than a physical optical approach for minimizing stereoscopic crosstalk when images are viewed severely off-axis on polarized stereoscopic displays; and it used Microsoft's HoloLens augmented reality display to prototype its design and aid in its construction.

This paper will describe Destiny's design and implementation - in particular the technique for software-based crosstalk mitigation. Lastly it will describe how the HoloLens helped validate Destiny's design as well as train the construction team in its assembly.

Introduction

The CyberCANOE project seeks to construct display rich environments to support collaborative science, engineering, and education in the Hawaiian Islands. The Polynesian voyaging canoe was the historical vessel of exploration, discovery, and communication for Pacific Islanders. The CyberCANOE is the Cyber-enabled "vessel" of Collaboration, Analysis, Navigation, and Observation for University of Hawaiʻi researchers and students in the era of data-intensive science. Four CyberCANOEs have been deployed at UH campuses in West Oʻahu, Mānoa, and Hilo.

Table 1: CyberCANOE Class Identification Table

Class	Resolution	3D	Tracking	Location
Pathfinder	11520x2160	Y	N	W. Oʻahu
Voyager	11520x2160	N	N	Mānoa
Innovator	8160x1152	Y	Y	Mānoa
Explorer	11520x2160	Y	Y	Hilo
Destiny	34560x7680	Y	Y	Mānoa

CyberCANOEs are ultra-high resolution 2D & 3D visualization systems that enable teams of researchers and their students in Science, Engineering, Medicine, and the Arts, to make sense of large scale data in their research. This allows researchers to arrive at conclusions with greater speed, accuracy, comprehensiveness, and confidence. CyberCANOEs come in a variety of configurations (or "classes") which vary in total pixel resolution as well as capabilities such as stereoscopic 3D and 6 degrees of freedom (DOF) tracking.

This paper describes the Destiny-class CyberCANOE which is a Hybrid Reality environment consisting of a 320 degree tiling of displays such as CAVE2 [1] (youtu.be/LwlAI4pQnFI). However, Destiny differs from CAVE2 and other similar systems [2], such as the TourCAVE [3] and WAVE [3] in a number of important ways: Destiny has higher total resolution, uses off-the-shelf OLED 4K displays which provide a superior image brightness and contrast compared to liquid crystal displays (LCD), and is constructed at 1/3 the cost of CAVE2. Destiny is intended as a single-user system that can provide close to 20/20 acuity that overcomes the resolution limitations of current low-cost head mounted displays, such as the HTC Vive and Oculus Rift. Most importantly, while CAVE2 uses specially designed LCD panels which shifted the interleaving polarizers of the top and bottom rows of the displays, Destiny uses a software-based approach to mitigate off-axis stereo crosstalk thereby significantly reducing overall cost of the system.

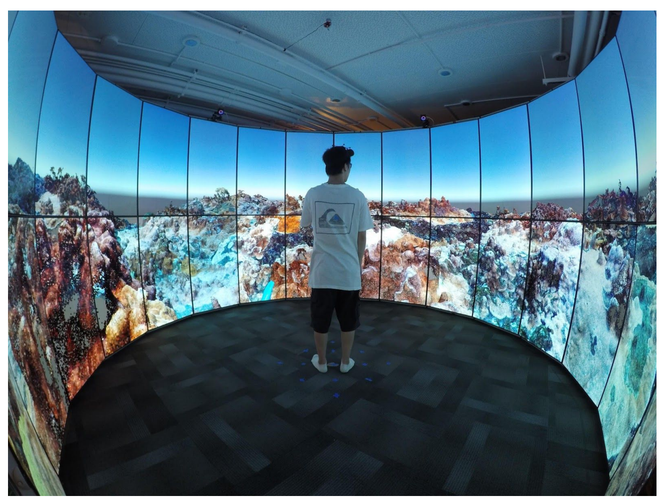

Figure 1. The Destiny-class CyberCANOE showing a visualization of Coral Reef data from the Hawai i Institute of Marine Biology (data courtesy of John Burns).

The primary contributions of this research are: the use of newly emerging OLED displays rather than LCD displays, the use of an augmented reality (AR) device as part of the design process and the use of a software-based approach to mitigate stereoscopic crosstalk.

Kawano, N., Theriot, R., Lam, J., Wu, E., Guagliardo, A., Kobayashi, D., Gonzalez, A., Uchida, K., Leigh, J., "The Destiny-class CyberCANOE – a surround screen, stereoscopic, cyber-enabled collaboration analysis navigation and observation environment." IS&T Electronic Imaging 2017, Engineering Reality of Virtual Reality, 2017.

Figure 2. External view of the Destiny-class CyberCANOE.

This paper will describe Destiny's end-user requirements, its design and construction, and the impact of using HoloLens [4] in prototyping and assisting in its construction. It will also describe the software-based crosstalk mitigation technique developed for this system and a preliminary evaluation of its efficacy.

Motivation

The goal was to build a system capable of providing nearly 20/20 acuity, so the viewer would not be able to see the individual pixels as on the HTC Vive or Oculus Rift. Consequently, emphasis on image quality (brightness, contrast, acuity) over other existing VR systems was paramount.

Current Hybrid Reality Systems such as CAVE2 are constructed with tiled LCDs. All pixels within a LCD are illuminated by a backlight, while pixels within an OLED are self-illuminated. This allows the off state of an OLED pixel to produce perfect image blacks resulting in an infinite contrast ratio and an infinite dynamic range [5].

Destiny is intended for use in scientific, engineering, medical, and art visualizations that can benefit from high resolution 2D and 3D environments. The system should provide an image that can fully envelop a user's field of view. Lastly, and perhaps most importantly, the system must be comfortable to use for long periods of time and should be minimally encumbering.

Implementation of the Physical Structure

Destiny's physical implementation was influenced by the technology that was chosen, the off axis field of view (FOV) limitations, and targeted acuity. OLED active stereo displays were not commercially available resulting in the decision to use passive stereo. While a 65" display was available, the 55" provided greater pixel density which was able to provide near 20/20 acuity. LG was used over others due to ease of tiling and their minimal bezel. Thus, the Destiny system was built using the passive stereo 55" LG OLED 4K Smart TV (model number 55EF9500).

The OLEDs (in portrait orientation) have a height and width of 48.25" x 27.625". Passive stereo-enabled LCDs have a vertical 3D FOV of 20 degrees when placed in the landscape orientation [1,2]. The OLEDs in Destiny were found to have approximately the same FOV. The minimum viewing distance of the entire display can be calculated at the intersection of the leftmost and rightmost pixel FOV lines, as seen in equation (1).

$$Minimum\ 3D\ Viewing\ Distance = \frac{Display\ Width}{2 \times tan(10°)} \quad (1)$$

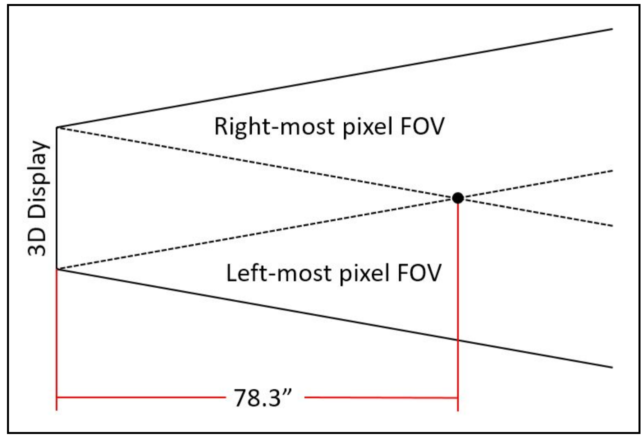

Figure 3. Top down view of a portrait-oriented display with 3D 20 degree FOV lines for the left-most and right-most pixel. Minimum 3D viewing distance occurs at the intersection of the two dotted lines.

Any given display in Destiny has a minimum viewing distance of 78.335" (Figure 3). Stereoscopic acuity can be calculated using the minimum viewing distance and the display's technical specifications (Table 2).

Table 2: Acuity Calculator for a LG model 55EF9500

	Vertical	Horizontal
Screen Dimension	48.250"	27.625"
Stereo Resolution	3840	1080
Angular Resolution	0.009	0.019
Arc Minutes	0.540	1.123
Snellen Notation	20/11	20/22

Stereo Resolution: the resolution seen by a single eye.

Angular Resolution: the angular difference between two neighbor pixels from the minimum viewing distance.

Arc Minutes: one 1/60 of one angular resolution degree.

Snellen Notation: a visual acuity measurement relative to 20 feet for every arc minute.

The number of necessary display columns can be calculated using equation (2), yielding 18 sides:

$$number\ of\ columns = \pi / tan^{-1}\left(\frac{display\ width}{2 \cdot minimum\ viewing\ distance}\right) \quad (2)$$

Two columns are absent to provide a gap for entry and exit, with a width of 55.2", making Destiny wheelchair accessible.

Destiny was constructed with 8020 [6] parts and designed in AutoCAD [7] using the 8020 plugin (Figure 4). Each column's base matches the width dimension of the displays (27.625"). The vertical structure was designed as a rectangular prism which provided stability and easy linkage to adjacent columns. 45-degree angle extrusions secure the vertical pillars and are used to support the video and power cables. In addition to providing stability, the extra column space is used as a shelving for the computer cluster.

The 8020 build design required no custom parts, therefore reducing the cost and assembly time required for the system.

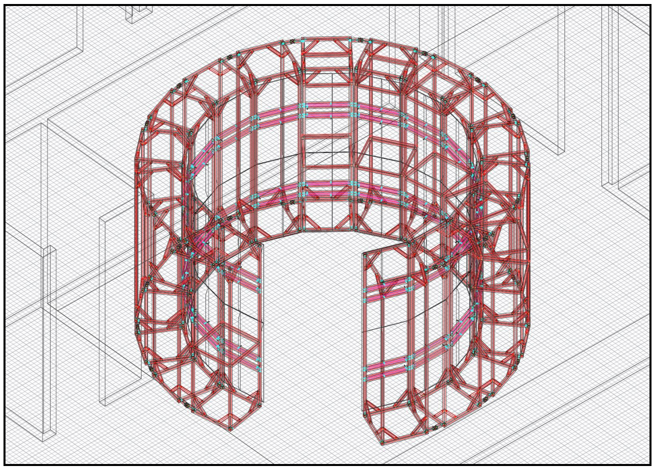

Figure 4. Wireframe design of Destiny in AutoCAD.

Use of the HoloLens

The HoloLens was used in the initial design and construction of Destiny. After the 8020 plan was completed in AutoCAD, the model was imported into Unity3D [8] for use in the HoloLens. Utilizing the HoloLens, it was possible to superimpose a one-to-one scale model of Destiny within the physical space designated for its installation (Figure 5). This allowed for the determination of proper placement, orientation, and whether the design could clear physical constraints such as ventilation ducts and electrical conduits that existed in the room.

Figure 5. Hologram of the Destiny model as seen through a HoloLens.

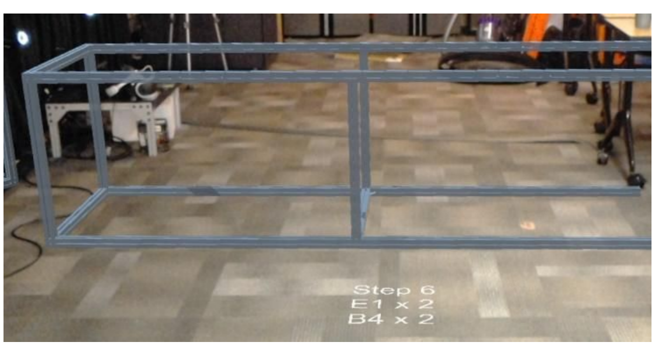

Figure 6. Step 6 of the Destiny HoloLens construction manual.

Destiny was constructed entirely by students, who were unfamiliar with the build process and therefore needed training. A preliminary user-study conducted with 6 subjects compared two training approaches. Three subjects used the HoloLens, which presented a one-to-one AR representation of the parts and assembly steps that could be superimposed on top of the physical parts for the same column construction process (Figure 6). The second group of three subjects were given a traditional paper instruction manual consisting of 2.5D orthographic diagrams for the construction of a single column of displays for Destiny (Figure 7).

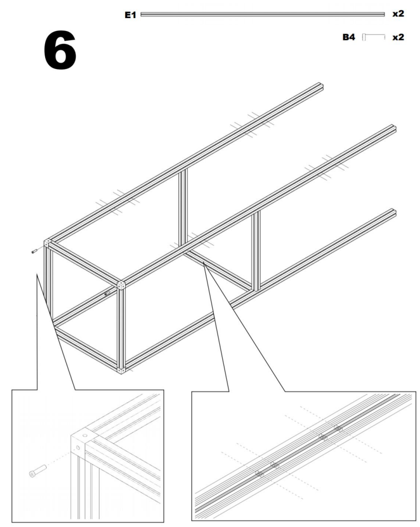

Figure 7. Step six of the Destiny orthographic construction manual.

HoloLens users reported a variety of issues during the assembly process. The small FOV requires head movement, as opposed to eye movement, in order to see a full scene. The HoloLens also had a tendency to enter sleep mode when not interacted with for a brief period of time. This resulted in the frequent relaunch and recalibration of the Destiny construction application. Users trained with the HoloLens took twice as long to complete their construction than those using the paper manual. However, by comparison to those given the paper manual, the HoloLens group reported greater confidence that they would not require any reference material for further construction tasks. More rigorous studies are needed to determine whether this greater confidence is a result of users having spent more time with the HoloLens or whether the augmented reality presentation was the primary contributor.

Despite the challenges in using the HoloLens for the assembly process, it proved very useful during the design phase.

Computing Hardware Implementation

Destiny's display capabilities require a large amount of graphics rendering power. The classic approach to driving a high resolution multi-display system is via a cluster of computers [1,2,9]. The key determinant for the size of the cluster needed is how many pixels can be rendered in at least 15 frames per second by each individual compute node. With Destiny, empirical tests were conducted using multiple 4K displays and a compute node of the following specifications: Intel Core i7 5960X 3.0GHz, 64GB DDR4 2400MHz RAM, 1TB Samsung 850 Evo SSD, 4 Nvidia GTX 980, and Windows 10 64-bit.

Table 3: Omegalib tests at different resolutions

Number of Displays	Number of Omegalib instances	Resolution per Omegalib instance	Frames per second
16	1	61440 x 2160	Error
4	1	15360 x 2160	Error
8	4	7680 x 2160	< 15
8	8	3840 x 2160	< 15
4	**4**	**3840 x 2160**	**30**

Omegalib [10], a hybrid-reality visualization framework, was used for the system validation. A point cloud of the island of O‘ahu in Hawai‘i, several asteroids, and a 3D model of the International Space Station, served as the test scene. Varied numbers of Omegalib instances at differing resolutions were used. Table 3 summarizes the results.

Based on these results it was determined that 8 computers were the minimum needed to drive all 32 of Destiny's displays, each compute node would drive four 4K displays.

In addition to the compute nodes, Destiny uses four OptiTrack Prime 13W infrared cameras mounted above the displays. Retroreflective markers attached to 3D printed controllers provide 6DOF interaction for the user. Tracking information is delivered to the computer by utilizing OptiTrack's Motive software.

Comparison to Other Hybrid-Reality Systems

Table 4 summarizes Destiny's capabilities using similar comparison metrics from Ponto [2], while also comparing to other Hybrid-Reality systems [1,9]. In addition to providing significantly better contrast and color gamut, due to its use of OLED displays, Destiny also has 3.6x the resolution of CAVE2 the former best-in-class hybrid-reality system. Furthermore, while CAVE2 cost over $900K to build, Destiny only cost $250K.

Table 4: The Destiny System Comparison

System	CAVE	CAVE2	DSCVR	Destiny
Stereo Resolution (MP)	22.1	36.2	20.7	**132.7**
Viewable 3D Resolution (MP)	10.8	19.7	20.1	**72.6**
FOV Horizontal Coverage (%)	100	100	100	**100**
FOV Vertical Coverage (%)	100	27	62	46.7
Immersive Resolution (MP)	10.8	5.4	12.5	**33.9**
Refresh Per Eye (Hz)	35	60	60	30
Immersive Bandwidth (MP/s)	378	319	750	**1016**

Stereo Resolution: the total number of megapixels of the system that can be viewed while wearing 3D glasses.

Viewable 3D Resolution: although the systems have a stereo resolution, it may not be possible to view them all at the same time due to shape or size. Viewable 3D resolution is the calculated resolution seen by one eye when standing in the center of a given system.

FOV Horizontal Coverage (%): the percentage of the view which the display surface covers, using the average human's estimated horizontal field of view [2].

FOV Vertical Coverage (%): the percentage of the view which the display surface covers, using the average human's estimated vertical field of view [2].

Immersive Resolution (MP): the product of the viewable 3D resolution and vertical and horizontal coverage values. This attempts to balance how much the display surrounds the user, while also accounting for display resolution [2].

Refresh Per Eye (Hz): a system specification describing the refresh rate per image seen by a single eye [2].

Immersive Bandwidth (MP/s): the product of the immersive resolution and the refresh per eye values. This number accounts for frame interleaving by attempting to provide a fixed-viewpoint measure of immersion [2].

Mitigating Stereoscopic Crosstalk

Stereoscopic crosstalk is defined as the incomplete isolation of the left and right image channels so that one image "leaks" or "bleeds" into the other" [11]. In passive stereo displays this occurs when a viewer moves off-axis, potentially looking at a pixel through a polarizer intended for the opposite eye (Figure 8). As the displays in Destiny are oriented in portrait mode, they have excellent vertical off-axis performance but inferior horizontal off-axis performance. CAVE2, using displays in landscape format, mitigates crosstalk via an expensive custom polarizing filter. The top and bottom rows of the filter are slightly shifted to accommodate a greater vertical off-axis field of view [1]. As Destiny is intended primarily as a single viewer system, it was possible instead to develop an equivalent solution entirely in software, hence significantly reducing cost. A Graphics Processing Unit (GPU) shader was developed to use the location of the viewer's head in real-time to adjust the pixel color which mitigated the perception of crosstalk.

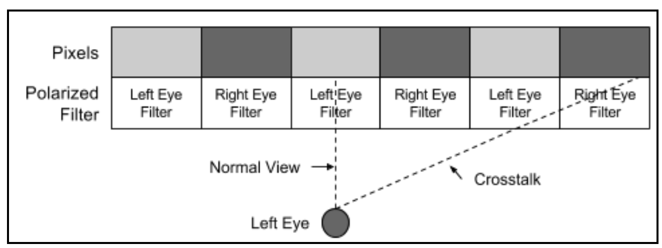

Figure 8. Off-axis crosstalk.

GPU-based Implementation

The GPU shader solution is based on the assumption that crosstalk starts at the intersection of the 3D horizontal FOV and the display column. At this intersection the user is viewing the intended pixel color, but also the unintended color of the neighboring pixel. Increasing the RGB value of intended pixel color and subtracting the RGB value of the unintended color can potentially counteract the crosstalk effect [11]. However, limitations exist due to using color subtraction. For example, given these pixels:

Intended Pixel Color in RGB	=	(255,	0,	0)
Unintended Pixel Color in RGB	=	(0,	255,	0)

This technique would not be able to counteract crosstalk as it cannot subtract green from the intended pixel color due to its green channel value being 0.

The following describes the procedure for one display column within Destiny. There are sixteen display columns and each column is calculated separately.

Perpendicular distance (PD) is calculated between the display column and the user's head. The off-center distance (OCD) is calculated from the display column's center and the user's head (Figure 9).

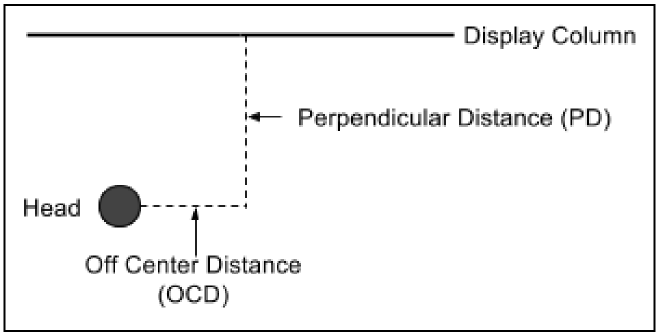

Figure 9. Perpendicular and Off Center Distance.

Given the 3D horizontal FOV is 20 degrees for each display column, it is necessary to calculate which column of pixels the 3D horizontal FOV intersects on the display column and label these left crosstalk starting pixels (LCS) and right crosstalk starting pixels (RCS) (Figure 10). These two values are passed into a GPU fragment shader.

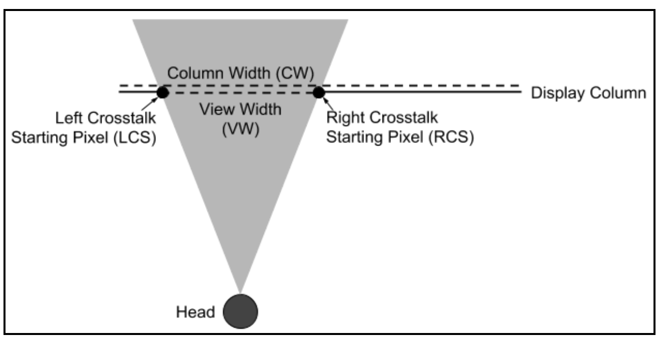

Figure 10. Left and right starting pixel determination.

Within the fragment shader the LCS and RCS are used to calculate where the crosstalk starts (CMin) and where the crosstalk stops (CMax). At CMax the user will view a clear picture with no crosstalk but the stereo has been reversed. Pixels closer to CMin will have a smaller percent change (PC) than those closer to CMax. Traversing from CMin to CMax, PC is linearly increased.

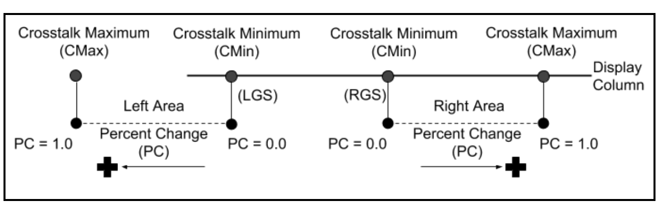

Figure 11. Crosstalk minimum and maximum within fragment shader.

The GPU shader samples intended pixel color (IPC) and the neighbor pixel, which is contributing to the crosstalk pixel color (CPC). The GPU shader then changes the pixel color using one of the following experimental formulas:

Formula 1 (F1)

$$PixelColor.RGB = ((1 + PC) \cdot IPC) - (PC \cdot CPC) \quad (3)$$

Formula 2 (F2)

$$PixelColor.R = ((1 + (PC \cdot 0.2126)) \cdot IPC) - ((PC \cdot 0.2126) \cdot CPC) \quad (4)$$

$$PixelColor.G = ((1 + (PC \cdot 0.7152)) \cdot IPC) - ((PC \cdot 0.7152) \cdot CPC) \quad (5)$$

$$PixelColor.B = ((1 + (PC \cdot 0.0722)) \cdot IPC) - ((PC \cdot 0.0722) \cdot CPC) \quad (6)$$

Formula 1 applies the percent change uniformly over the RGB channels while Formula 2 accounts for relative luminance; green light contributing the most to color perceived by humans and blue light the least [12]. The experiment described in the next section compares the efficacy of the two formulae.

Evaluation of Anti-Crosstalk Approaches

The experiment consisted of having seven subjects stand at a series of 13 locations within Destiny (Figure 12) and viewing the scene with and without the anti-crosstalk solutions enabled. At each location the subjects were allowed to freely switch between both options however they were not informed about which option was being viewed. They were then asked to choose which option they preferred or if they could not decide on a preference. Two different scenes were used, the first consisting of simple geometric shapes (Figure 13) and the second consisting of a more naturalistic jungle environment (Figure 14). The experiment was completed using each crosstalk formula independently.

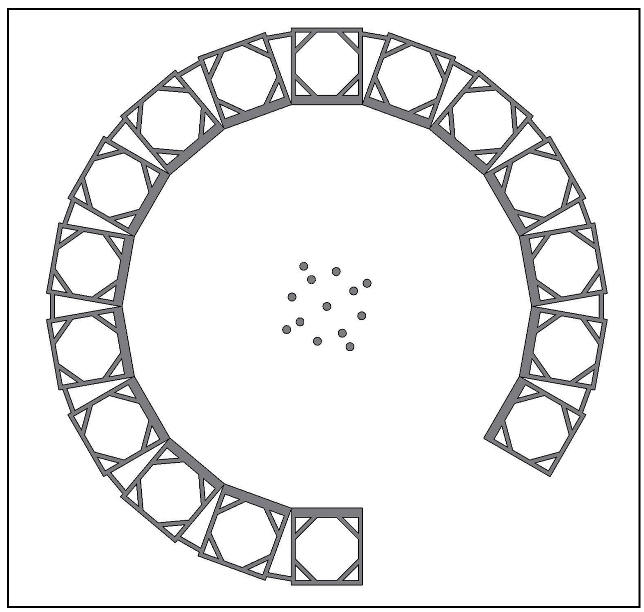

Figure 12. Top down view of Destiny and the positions used for each evaluation.

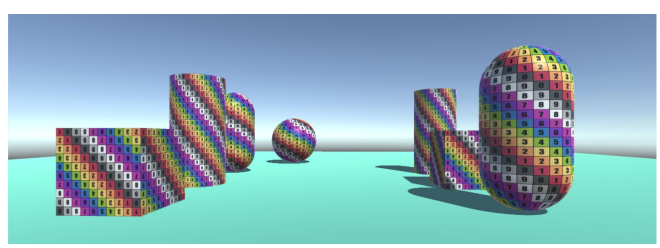

Figure 13. Geometric Shapes Scene.

Figure 14. Jungle Scene.

Results of Anti-Crosstalk Experiment

The graph below totals up the preferences (votes) for the different locations (Figure 15). The results suggest the anti-crosstalk options were preferred in all four tests. Formula 2 (F2) was favored over Formula 1 (F1) for both scenes.

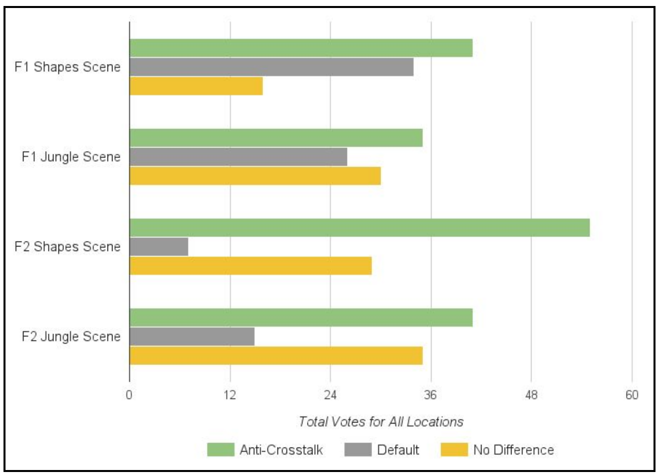

Figure 15. Comparison of the two anti-crosstalk approaches (F1 and F2) under two different scenes (Shapes Scene and Jungle Scene).

Conclusion

This paper outlined the design, construction, and evaluation of the Destiny-class CyberCANOE. The HoloLens was found to be useful in its initial design, although its use for construction training needs further study.

The GPU shader approach to crosstalk mitigation was found to be beneficial in a subjective experiment, with the approach that took relative luminance into account yielding better results. Additional experiments similar to the approach by Ponto [2] are underway to provide a more objective measure of improvement. Plans are also underway to extend the shader approach to resolve pseudo-stereo that occurs at severe off-axis viewing angles.

Acknowledgments

The project is funded in part by the National Science Foundation awards 1456638 and 1530873, and the Academy for Creative Media System (ACM) of the University of Hawaiʻi.

References

[1] A. Febretti, et al. "CAVE2: a hybrid reality environment for immersive simulation and information analysis." *IS&T/SPIE Electronic Imaging*. International Society for Optics and Photonics, 864903-864903 (2013).

[2] K. Ponto., et al. "DSCVR: designing a commodity hybrid virtual reality system." *Virtual Reality* 19.1, 57-70 (2015).

[3] K. Knabb., et al. "Scientific Visualization, 3D Immersive Virtual Reality Environments, and Archaeology in Jordan and the Near East." *Near Eastern Archaeology* (NEA) 77.3, 228-232 (2014).

[4] HoloLens. Microsoft. https://www.microsoft.com/microsoft-hololens/

[5] R, Soneira. (2015). *Flagship OLED and LCD TV Display Technology Shoot-Out*. [Online] Available: http://www.displaymate.com/TV_OLED_LCD_ShootOut_1.htm

[6] 8020 INC. https://www.8020.net/

[7] Autodesk AutoCAD. http://www.autodesk.es

[8] Unity3D. https://unity3d.com/

[9] C. Cruz-Neira., et al. "Surround-screen projection-based virtual reality: the design and implementation of the CAVE." *Proceedings of the 20th annual conference on Computer graphics and interactive techniques*. ACM, 135-142 (1993).

[10] A. Febretti, et al. "Omegalib: A multi-view application framework for hybrid reality display environments." *2014 IEEE Virtual Reality (VR)*. IEEE, 9-14 (2014).

[11] A. Woods. "Understanding crosstalk in stereoscopic displays." *Keynote Presentation at the Three-Dimensional Systems and Applications Conference, Tokyo, Japan*. 2010.

[12] R. Gonzalez, and R. Woods. *Digital image processing*. Upper Saddle River, NJ: Prentice Hall, 2008.

Author Biographies

Noel Kawano received his BS in electrical engineering from the University of Hawai i at Mānoa and is currently pursuing a MS in computer engineering.

Ryan Theriot is an undergraduate student at the University of Hawai i at Mānoa, pursuing a BS in Computer Science. Before attending UHM he served in the United States Coast Guard as Avionics Electrical Technician.

Jack Lam received his BS in Computer Science and is currently pursuing a MS in Computer Science from the University of Hawai'i at Mānoa.

Eric Wu received his BS in Computer Science from the University of Hawai i at Mānoa and is currently pursuing a MS in Computer Science.

Andrew Guagliardo is an undergraduate student at the University of Hawai i at Mānoa, pursuing a BA in Creative Media, with a focus on computer animation and programming.

Dylan Kobayashi received his MS in Computer Science from the University of Hawai i at Mānoa and is pursuing a PhD in Computer Science.

Alberto Gonzalez received his BS in Electrical Engineering and MS in Automation and Control Engineering from Universidad Politecnica de Madrid in Spain and a second MS in Computer Science from the University of Hawai i at Mānoa. He is currently pursuing a PhD in Computer Science.

Ken Uchida received his BS in Electrical Engineering from the University of Hawai i at Mānoa and is currently employed at Raytheon.

Jason Leigh received his PhD in Computer Science from the University of Illinois at Chicago (1998) where he is Director Emeritus of the Electronic Visualization Laboratory. Currently he is the Director of the Laboratory for Advanced Visualization and Applications (LAVA), and acting Director of the Academy for Creative Media System at the University of Hawai i at Mānoa.