

DSCVR: designing a commodity hybrid virtual reality system

Kevin Ponto · Joe Kohlmann · Ross Tredinnick

Received: date / Accepted: date

Abstract This paper presents the design considerations, specifications and lessons learned while building DSCVR, a commodity hybrid reality environment (HRE). Consumer technology has enabled a reduced cost for both 3D tracking and display, enabling a new means for the creation of immersive display environments. However, this technology also presents many challenges which need to be designed for and around. We compare the DSCVR system to other existing VR environments to analyze the tradeoffs being made.

Keywords Hybrid Reality · Virtual Reality · Display Wall · Immersive Systems · Commodity Hardware · 3D · High Resolution · Passive Stereo

1 Introduction

Recent advancements in consumer grade 3D display and gesture input technology have enabled new pathways for the creation of immersive Virtual Reality systems. Previous methods for constructing immersive display systems

Kevin Ponto
Wisconsin Institute for Discovery
Room 3176, 330 N. Orchard Street
Madison, WI 53715
Tel.: +(608) 316-4330
E-mail: kponto@discovery.wisc.edu

Joe Kohlmann
Wisconsin Institute for Discovery
330 N. Orchard Street
Madison, WI 53715
E-mail: jkohlmann@discovery.wisc.edu

Ross Tredinnick
Wisconsin Institute for Discovery
Room 1144B, 330 N. Orchard Street
Madison, WI 53715
E-mail: rtredinnick@discovery.wisc.edu

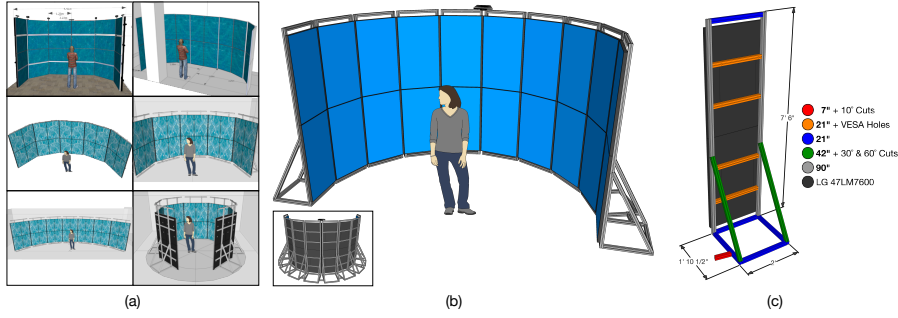


Fig. 1 Design schematics for the DSCVR System. (a) shows a subset of the designs considered, including horizontally-oriented displays, larger displays, a wider or tighter curvature, and positioning techniques to hide bezels. (b) shows the final, implemented structure, with the back of the system shown inset. (c) shows the 80/20 components of one of the final design’s ten columns.

required customized room configurations, tracking hardware, projections and screens. This in turn has meant most VR systems have become an exclusive enterprise, as the systems are thus designed and utilized by a limited number of privileged individuals.

Given this outlook, our mission is to develop a system entirely from commodity, off-the-shelf hardware that has comparable performance to commercially-built environments. However, as we discovered in the process of building this system, the current generation of commodity-grade technologies provides a significant number of challenges to creating effective immersive virtual environments.

In this paper we present the design considerations, specifications and lessons learned for building the “DSCVR System”, a hybrid reality environment (HRE) constructed from commodity-grade hardware. HREs, as defined by [19], enable the benefits of both tiled-display environments along with the immersive characteristics of Virtual Reality systems. Specific contributions include:

- Design guidelines for the construction of a virtual reality system utilizing commodity hardware, such as micro-polarization 3D displays.
- Quantification of attributes and performance of said system compared to professionally constructed virtual reality systems.
- Discussion of lessons learned and considerations for others attempting to create these types of systems.

1.1 Related Work

Many researchers have attempted to balance the tradeoffs between cost and fidelity in the creation of virtual reality systems. In 1991, Randy Pausch proposed building a VR system on the budget of five dollars a day [35]. The system was developed using a HMD with a Nintendo PowerGlove for interaction and a Polhemus Isotrak magnetic tracker for a cost around \$5000. Basua

et al. updated this concept in 2012, showcasing the ability to build a virtual reality system for a dollar a day [9]. Others such as Avery et al. have utilized custom HMDs to develop low-cost augmented reality [7].

Bowman and McMahan have posed the question of how much immersion is enough for the field of virtual reality [11]. This question has been studied from a variety of angles. Prabhat et al. have tried to study the difference between low-fidelity fishtank VR systems in relation to more immersive CAVE style systems [39]. Bacim et al. have attempted to study how the level of immersion in CAVE environments affect task performance [8]. Lahaet al. have studied the effects of immersion on the analysis of volumetric data in virtual environments [27]. Ragan et al. have studied how spatial judgment tasks were affected by stereo, head-tracking and field of regard [40]. Polys et al. studied how screen size and field of view affected performance using a tiled display environment [36]. Finally, McMahan et al. studied how immersion and fidelity affected performance in a first person shooter video game [33].

The video gaming industry has generated a large amount of motion-tracking hardware that has also spurred interest in low-cost virtual reality systems. For example, Schou et al. combined the Nintendo Wii Remote, multiple infrared sensor bars, and a two-wall immersive VR theatre [45]. Lange et al. also examined the use of a Microsoft Kinect motion-tracking sensor in a clinical VR rehabilitation task [28].

Immersive display environments have also seen several design iterations. Cruz proposed the original design and implementation of the CAVE in the early 1990s [13] [14]. This projection-based multi-wall design became the de facto standard for immersive, room-sized virtual reality environments. These types of systems range from a three-wall setup with a floor, to a fully immersive six-sided system. However, other designs which curve around the user have also been created for Virtual Reality, such as the Allosphere [5] and the i-Cone [47].

Tiled display walls rose in popularity in the mid-2000s for their ability to provide a large viewing area while maintaining a high image resolution. Systems such as HIPerWall at the University of California, Irvine [26], HIPerSpace at the University of California, San Diego [37], Stallion at the University of Texas [24] and the Reality Deck at New York's Stony Brook University [51] have shown the ability to create high-resolution data visualizations. However, these systems do not provide an efficient method to present stereoscopic imagery.

In 2011, DeFanti et al. proposed new methods for creating CAVE-style systems from the same components used in tiled display walls [16]. Since this time, new types of immersive display environments have been created, from the desk-sized, 3DTV-based HUVr device at University of California, San Diego [32] to the large scale NexCAVE at King Abdullah University of Science and Technology (KAUST) and CAVE2 at the University of Illinois at Chicago [19].

The development of CAVE2, apart from showcasing many advances in VR hardware and software, underscored the challenges of working with micro-polarization displays. The CAVE2 implementation used specialized filters and displays to address these issues. With these lessons in mind, one early goal for the DSCVR system was to recreate this type of system entirely with consumer-

grade hardware. As with all computer systems, many factors need to be taken into account in the design process [43]. We describe the design decisions made in the creation of the DSCVR system below.

2 Design

DSCVR’s design and implementation are ultimately a balance between the project’s financial, technological, and structural goals:

1. Implement a hybrid reality system in a cost-effective way, using unmodified, consumer-grade hardware, such that its performance rivals that of more expensive systems.
2. Reduce the appearance of bezels and stereo image crosstalk in users’ fields of view.
3. Build a frame that supports both display position adjustments and display upgrades.
4. Balance design tradeoffs between cost and performance.

To accomplish these goals, compromises were inevitably made, making DSCVR neither the best, worst, most expensive, or cheapest environment of its kind. However, as shown in Sections 4 and 5, the performance of the implemented system can equate or exceed that of much more expensive systems.

2.1 Display Technology

One of the early decisions while designing the DSCVR System was the choice of display technologies. Projectors were known to have substantial drawbacks, such as the need to replace bulbs, the need for specialized projection material, the need for extra space to accommodate throw distance, and the need for repeated color calibration. It was therefore prudent to utilize one of the increasingly capable consumer-grade 3D televisions available on the market.

Consumer 3DTVs currently use either active or passive stereo display technology. Active stereo, accomplished through synchronizing display frame swaps with shutter glasses, is the most common format in use today. Unfortunately, this technique becomes problematic when multiple TV screens are used, as each TV must show the same eye at the same time. Synchronizing this type of swapping would have required specialized and costly hardware, and thus was rejected in favor of passive stereo technologies.

While traditional passive stereo displays utilize linear polarizers, newer ones use micro-polarization technology. Febretti et al. provide a detailed description how this technology works [19]. Notably, these types of polarization filters can be easily produced for thin-bezel displays. The major problem with these types of displays is that, while the images are of good quality when viewed in the direction parallel to the micropolarizer lines, the images have substantial “crosstalk” when viewed from the direction perpendicular to the

micropolarizer lines. As these types of TVs provide the best image quality, contrast and 3D capability for the cost, we chose to tackle the hurdle of crosstalk through our arrangement. As of this writing, LG Electronics Inc. is the prominent manufacturer of consumer-grade, passive-stereo 3DTVs, so we focused on the capabilities of the LG LM7600 television. Additional details on model selection are given below.

2.2 Structure and Arrangement

The limitations of these LG TVs meant that the system’s shape and structure would have to accommodate the displays’ unmodifiable polarization filters, as well as their wider 1-inch bottom bezels. Early development on the system required accounting for uncertainties, such as the displays’ panel sizes—either 47 inches or 55 inches diagonal—and the unknown final location of the system. Therefore, regardless of the final shape, display quantity, or panel size, we decided to build the system as a series of modular, functionally independent columns, using 80/20 aluminum framing for the structure.

The first designs, modeled in Trimble SketchUp for its ease of use, adapted a tiled display wall layout into a cylindrical shape, making it a variation on systems such as the KAUST NexCAVE [15], in which users stand near the center of the display array’s radius. Landscape and portrait display orientations were both considered, as well as other techniques to hide the wider bezels behind the previous or next column of displays. These unimplemented designs, some of which used up to 28 displays, are shown in Figure 1(a).

Analyzing the tradeoffs posed by these design experiments informed our decisions for the final model. A wider, more gradual curving arrangement, for example, may have accommodated more viewers, but would have exacerbated the appearance of off-axis viewing artifacts, such as stereo image crosstalk. This curved layout has also been shown to have several benefits over a linear layout by Shupp et al. [46]. Framing constructions with four displays grouped together may have minimized the number of components, but would have reduced the modularity of the design. 55-inch displays may have made the system fill a larger field of view, but 47-inch displays would increase the pixel density, shorten the height of the system, and enable the mounting of tracking hardware on top of the framing while staying under the height of a standard office ceiling.

Portrait orientation became a particularly important consideration early on since the wide bezels of the landscape-oriented displays may have cut a tall and wide horizontal gap through the viewer’s entire field of view, regardless of their height. Though tucking away the wider bezels may have further reduced the appearance of gaps, positioning the displays edge-to-edge along the cylinder’s interior radius would instead minimize changes in display depth, enhancing the appearance of the array as a seamless, curved surface.

Understanding the limitations of the displays’ micropolarization filters further validated the use of a cylindrical, portrait-oriented layout. Purchasing one

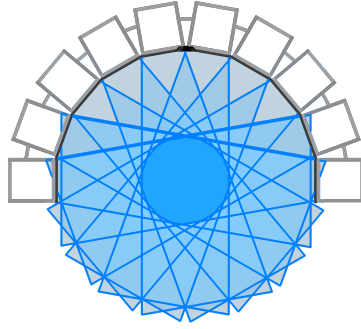


Fig. 2 An overhead view of the DSCVR System’s final design, showing the displays’ estimated viewing ranges and the center region in which all the viewing ranges overlap.

display allowed us to determine the range of horizontal and vertical viewing angles for which viewers could not observe any off-axis viewing artifacts. To accomplish this we used the method described in Section 4.3 to measure the level of crosstalk [52].

Measurements were taken from distances of 3, 5 and 10 feet from the TV. For each measurement, the instrument was positioned in the center of the monitor and was then slid parallel to the monitor until crosstalk occurred. The process was repeated 3 times at each distance, and was tested with the monitor in both landscape and portrait orientations. From these measurements, the angle from the edge of the TV was calculated for which the 3D effect would work correctly. From this initial test, it was determined that when the TV was positioned in a landscape orientation, the horizontal artifact-free field of view was approximately 170° s, while the the vertical artifact-free field of view was approximately 20° s.

Given the display’s very wide horizontal viewing range, mounting the displays in portrait would provide an artifact-free image to viewers of varying statures. The narrower vertical viewing range, however, indicated the need for a cylindrical arrangement small enough to have each column’s viewing ranges overlap, but large enough to support multiple viewers for 2D applications and spectator viewing.

This analysis allowed us to visualize a central “sweet spot” for our cylindrical models, shown in Figure 2. The understood constraints on viewing ranges, system size, and budget thus led to DSCVR’s final design, consisting of 10 columns and 20 displays arranged in a half-cylinder shape, with a “sweet spot” approximately 4 feet in diameter. The final design is shown in Figure 1(b).

One additional finding was that orienting the displays in portrait introduced a minor visual artifact when viewing stereo images through the included 3D glasses. Orienting the linear components of the glasses’ polarization filters *perpendicularly* to the filters on the displays produced an additional color-fringing artifact, most noticeable when viewing images with high-contrast edges, such as white text on a black background. We found that 3D glasses with circular-polarized filters at 90° left and 90° right completely eliminated

this artifact. After modifying a pair of included 3D glasses, we ordered two inexpensive pairs of built-to-order glasses with these orientation changes.

Finally, the system was created to only comprise half of a circle as opposed to being fully encompassing to enable benefits seen in tiled-display environments such as high resolution image and cinema viewing [38] [42] [37]. This arrangement also enabled an audience to easily observe a virtual experience from behind the participant in 3D. If one was inclined, extending the system to cover a full circle would simply be an extension of the described methods.

3 Implementation

Based on the design described above, the system was implemented as described below.

3.1 Framing

After finalizing the DSCVR System's shape, a single prototype column was constructed using 80/20 aluminum framing. 80/20, which is manufactured according to order, allowed us to develop a custom frame with greater utility and lower cost than any of the expensive, proprietary display stands we considered. Consumer-grade VESA mounts were initially used to attach two displays to the structure. Assessing the prototype allowed us to revise several components for the final construction, such as the frame's depth and its ease of assembly. We replaced the VESA mounts with custom horizontal pieces of 80/20 with precisely machined screw holes, enabling the displays to be mounted on the frame without sloping downwards.

The final single-column frame is shown in Figure 1(c). It stands 90 inches tall, 24 inches wide, and 22.5 inches deep, and uses inside-to-inside corner connectors to join the 80/20 pieces along their interior tracks. The 21-inch VESA mount pieces, with two precision holes drilled 200 mm apart, attach the displays to the frame. These pieces can be loosened, moved vertically along the inside of the frame, and reattached, thereby supporting precision height adjustments for both current and future displays. A separate 7-inch piece with 10° cuts attaches multiple columns to each other, simplifying inter-column alignment and increasing structural stability. An extra metal brace is attached between the tops of adjacent columns for even more stability.

3.2 Hardware

The DSCVR System utilizes several Alienware X51 mini gaming desktops, which were chosen for their high CPU and GPU performance, comparatively low power requirements, and comparatively low price. Each of 12 machines is equipped with an Intel Core i7-3770, 8 GB of 1600 MHz DDR3 SDRAM, a 1 TB SATA hard disk, gigabit Ethernet, an NVIDIA GTX 660 GPU with

1.5 GB of GDDR5 VRAM, and a 330W power supply. Ten “cluster nodes” drive the 20 displays, one “head node” hosts one or more VRPN tracking servers, and one “workstation” supports development and cluster control. A 13th “hot spare” machine is available to replace a malfunctioning one. Driving just two displays per cluster machine balances the capabilities of these single-GPU gaming PCs with the need to render a high-resolution, distributed, 3D viewport. The CentOS operating system provides a stable, UNIX-like software environment which also enables us to administer cluster commands via SSH and tentakel.

A USB to RS-232 serial interface connects each column’s two displays to its cluster machine, enabling programmatic control over display visibility and 3D modes. The serial cable is split once to send the same command to both displays simultaneously.

3.3 Tracking System

The Microsoft Kinect system was selected to track the user. Clark et al. found that Kinect, in combination with the Microsoft Kinect for Windows SDK, was able to provide data comparable to that of a commercial 3D motion analysis system [12]. The price of Kinect was substantially lower than specialized ultrasonic or optical tracking hardware. Furthermore, Kinect does not require the user to wear any specialized tracking equipment, such as a tracking bar, enabling users to easily move in and out of the tracking space.

As the “sweet spot” for the system is not overly large, the entire area can be easily covered by a single Kinect system mounted on top of the framing. On initialization, we use the Kinect SDK to tilt the sensor to its lowest possible level of declination. The accelerometer value is read to determine the actual tilt of the system. The Microsoft Kinect SDK’s skeleton tracking API¹ is used to determine the position and orientation of the user. All system-level transformations are “undone” before packing and sending the tracking data to client applications using VRPN [48]. In practice, simple temporal averaging was able to alleviate most jitter artifact. However, the steep downward tilting angle mixed with certain hair colors and styles has been shown to confuse hair regions and head regions. Future work will attempt to mitigate these effects.

To allow for multiple individuals to be in the space, we used the “sticky user” flag in the Kinect skeleton-tracking API. This allows a tracked user to utilize the space while other individuals are in the same area. A tracked user can “switch” with another user by simply walking out of the space, which allows Kinect to then detect and track the next available skeleton. This process occurs without the exchange of any glasses or equipment.

¹ Livingston et al. provide evaluations for the noise, accuracy, resolution, and latency of the skeleton tracking software provided by Kinect [30].

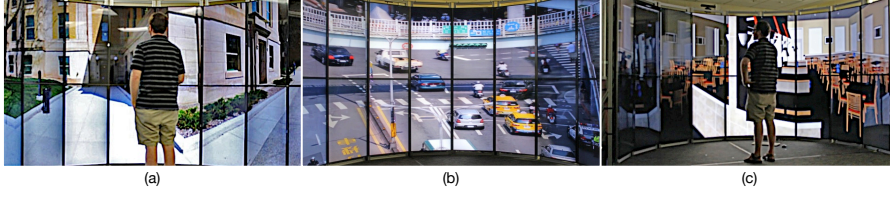


Fig. 3 The DSCVR System implementation, shown displaying (a) high-resolution panoramic imagery, (b) playing 4K stereo video at 30 FPS using custom software, and (c) rendering an interior environment in real time stereo 3D using the OGRE 3D engine.

3.4 Software

While the DSCVR System does not approach the resolution of environments such as Stallion, HIPerSpace, or Reality Desk [24][37][51], the resolution is still quite high compared to many still-image capture technologies. In this regard, the system is well suited for tiled display software, such as CGLX, Equalizer, Chromium, and others [17][18][29][31]. We demonstrate the ability to view this type of high-resolution content in Figure 3.

One feature of the DSCVR System compared to other high resolution displays [24][37][51] is its ability to display 3D media. Stereo panoramas, such as those outlined by Ainsworth et al. [4], are particularly well-suited to the display capabilities of the DSCVR System.

Another source of interesting data comes from 3D video, which has become readily available on the Internet. Unfortunately, tiled video players such as VideoBlaster and SAGE [38] [42] do not have native support for this 3D content. Having a desire to view this type of content, we developed a distributed video player application capable of playing back such 3D content.

The software is based on the VideoBlaster framework [38], which utilizes a message based protocol as opposed to a streaming-based technique. 3D media content can be provided in two ways: multiple video streams can be encapsulated in a single video file for each eye, or video frames can contain the left and right frames in a single video stream. The software takes the content for each eye and uploads the YUV frames to the graphics card. A single video rectangle can thus be moved around the display environment, with the appropriate content for each eye being shown in the appropriate location. This technique enables the playback of stereo 4K content at 30 FPS as shown in Figure 3(b).

The DSCVR System successfully enables several open source and free to use visualization and software applications such as the Unity 3D game engine [21] with stereoscopic rendering via side-by-side stereo as shown in Figure 4. In addition, DSCVR makes use of a custom-built software framework that runs a point cloud renderer, a volume renderer via the open source software Voreen [34], molecular visualization with the application VMD [23], and rendering of 3D models via the OGRE 3D engine [44], as shown in Figure 3(c). The software framework uses VRPN [48] for head tracking and generates asymmetric viewing frustums to create a seemingly seamless 3D viewport [13]. The virtual



Fig. 4 The DSCVR System running a scene built using the Unity 3D game engine.

binocular disparity (i.e. the distance between the virtual eyes) was set via a configuration file at startup. All input for these application is handled with a wireless PS3 Dual Analog controller. Future work will seek to add additional user input controls, such as the Leap Motion and continue to add new VR enabled applications.

4 Evaluation

The DSCVR System was designed with commodity-grade hardware in an effort to reduce costs. This effort introduced tradeoffs for a variety of factors such as resolution, field of regard, and latency.

As a reference, we compare our system against a professionally-built CAVE environment $2.93\text{m} \times 2.93\text{m} \times 2.93\text{m}$ in size. The CAVE system utilizes four workstations, each with $2 \times$ Quad-Core Intel Xeon processors and 2 NVIDIA Quadro 5000 GPUs. Two 3D projectors (Titan model 1080p 3D, Digital Projection), with a maximum brightness of 4500 lumens per projector, are used to generate projections with a resolution of 1920×1920 per display wall. The system utilizes an InterSense ultrasonic tracking system, VETracker Processor model IS-900 with MicroTrax model 100-91000-EWWD and MicroTrax model 100-91300-AWHT used for wand and head tracking, respectively.

We also compare the DSCVR System to the specifications to the CAVE2 system [19]. The CAVE2 system was selected as it implements a similar screen-based, cylindrical approach to immersive virtual reality. While direct compar-

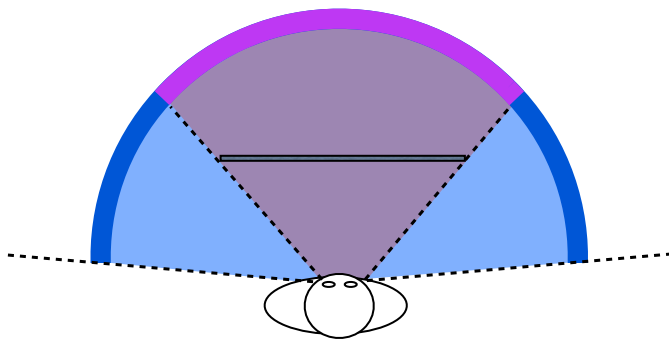


Fig. 5 The determination of coverage for a stationary position by determining the percent of view that the display covers (purple) compared to the human field of view (blue).

ison of these systems could not be achieved, as the authors did not have access to CAVE2, a comparison of DSCVR against published specifications.

4.1 Human Vision Factors

When comparing the immersiveness of different systems, several factors need to be accounted for simultaneously. For example, as someone moves their head closer to a screen, the amount the screen fills their field increases. However, as the user's eyes move closer to a screen, the size of the the pixels as projected onto their retinas increases, thus reducing the effective resolution. To this end, we analyze both field of view and resolution metrics simultaneously.

4.1.1 Stationary Viewpoint

For certain simulation tasks, users will generally stay in a fixed location viewing the virtual display environment. In these scenarios providing views behind the user is not considered important and performance can be estimated given a single view. Using previous literature, we can estimate the average human's field of view to be 175° horizontally and 135° vertically [6] [41] [50]. Using this knowledge we analyzed the following factors for each system:

1. **3D System Resolution:** the number of coordinated 3D megapixels which the system can display.
2. **System Viewable Area:** The percentage of the system which can be seen when the user is viewing a stereo image while standing in the center of the system.
3. **Viewable 3D Resolution:** the number of 3D megapixels that can be seen by the eye when the user is viewing a stereo image while standing in the center of the system.
4. **FOV Horizontal Coverage:** the percentage of the view which the display surface covers, using the average human's estimated *horizontal* field of view. See Figure 5

5. **FOV Vertical Coverage:** the percentage of the view which the display surface covers, using the average human’s estimated *vertical* field of view.
6. **Immersive Resolution:** 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.
7. **Refresh Per Eye:** a system specification describing the refresh rate per image seen by a single eye. As the CAVE uses frame interleaving to transmit left- and right-eye images, its value was nearly half the value of its counterparts.
8. **Immersive Bandwidth:** 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.

We show the comparison between the different systems in Table 1.

System	CAVE	CAVE2	DSCVR
3D System Resolution (MP)	22.1	36.2	20.7
System Viewable Area (%)	49%	54%	97%
Viewable 3D Resolution (MP)	10.8	19.7	20.1
FOV Horizontal Coverage (%)	100%	100%	100%
FOV Vertical Coverage (%)	100%	27%	62%
Immersive Resolution (MP)	10.8	5.4	12.5
Refresh Per Eye (Hz)	35.0	60.0	60.0
Immersive Bandwidth (MP/s)	378.0	322.2	750.0

Table 1 Eight human vision- and perception-based factors calculated for three virtual reality systems. The coloring refers to rank, with rank 1 colored green, rank 2 colored orange, rank 3 colored pink.

4.1.2 Moving Viewpoint

As shown, the DSCVR system performs admirably compared to the other three systems while the user is stationary. However, in other applications it may be important to for the user to look in different directions. This is often referred to as Field of Regard, being the range of the virtual environment that can be viewed with physical rotation [40].

Using this motivation we analyzed the following factors for each system:

1. **Horizontal Field of Regard:** the percentage of the horizontal view which the display surface covers for any viewing direction. See Figure 6
2. **Vertical Field of Regard:** the percentage of the vertical view which the display surface covers for any viewing direction.
3. **Motion Immersive Resolution:** the product of the Viewable 3D Resolution and the Vertical and Horizontal Field of Regard values. This attempts to balance how much the display surrounds the user, while also accounting for display resolution.

4. **Motion Immersive Bandwidth:** the product of the Motion Immersive Resolution and the Refresh per Eye values. This number accounts for frame interleaving by attempting to provide a moving-viewpoint measure of immersion.

We show the comparison between the systems in Table 2.

System	CAVE	CAVE2	DSCVR
Motion Horizontal Coverage (%)	100%	89%	50%
Motion Vertical Coverage (%)	100%	20%	46%
Motion Im. Resolution (MP)	10.8	3.5	4.6
Motion Im. Bandwidth (MP/s)	378.0	210.0	276.0

Table 2 Four human vision and perception-based factors calculated for three virtual reality systems. The coloring refers to rank, with rank 1 colored green, rank 2 colored orange, and rank 3 colored pink.

4.2 Latency

Latency is a common measurement of virtual reality systems. One common way to accomplish this is to use a pendulum model [49]. As our system utilizes Microsoft Kinect, we chose to use a variation on the method proposed by Livingston et al. [30].

The first step in the process was for the participant to orient arm parallel to the ground, setting the “zero point” in the virtual system and in the video.

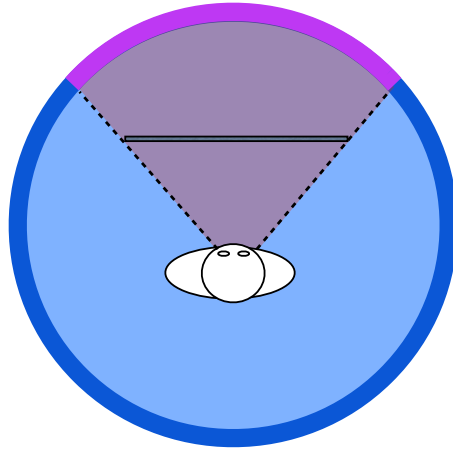


Fig. 6 The determination of field of regard or coverage for a moving position by determining the percent of view that the display covers (purple) compared to the field of view surrounding the user (blue).

The user then waved their arm up and down, mimicking the motion of a pendulum. On the screen in front of the user, the virtual height of the marker was displayed using colored rectangles. To make tracking easier, heights above the zero point were shown with a red rectangle, while heights below the zero point were shown with a blue rectangle. Images were captured with a GoPro Hero3 Black Edition camera, which was selected for its ability to capture images at a rate of 240 Hz at WGA resolution. Each video frame was extracted, and both marker height and virtual height were tagged in OpenCV, as demonstrated in Figure 7.

These tagged heights were then imported into statistical analysis software. While the height of physical and projector marker are not identical, the important component is the phase shift between the two signals. From this, the latency amount can be found, as shown in Figure 8.

Using 18 samples, we found an average latency of approximately 150 milliseconds with a standard deviation of 23 milliseconds. This result is similar to the latency of Kinect found by Livingston et al. of 146 ms [30]. As a comparison, we repeated this same procedure for the CAVE system. For the CAVE system, the position of the marker was tracked using the InterSense ultrasonic tracking system described previously. The CAVE system produced very similar results, with a latency of 150 milliseconds with a standard deviation of 24 milliseconds. This result is discussed further in Section 5.1.

Latency is not reported in Febretti et al.'s paper on the CAVE2 [19], so no direct comparisons can be made.

4.3 Stereo Crosstalk

As stated in Section 2.1, micro-polarization technology has a limited effective viewing range. When the viewer is not inside of the viewing range, crosstalk

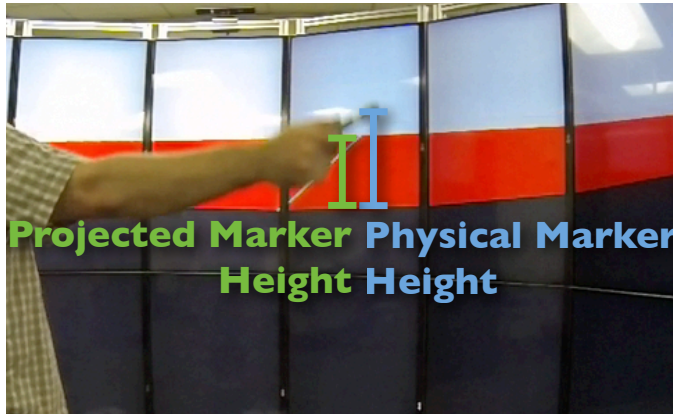


Fig. 7 The variation of the pendulum model used to evaluate DSCVR's latency. The user first sets the zero point with their arm straight out. The user then waves their arm up and down, like a pendulum.

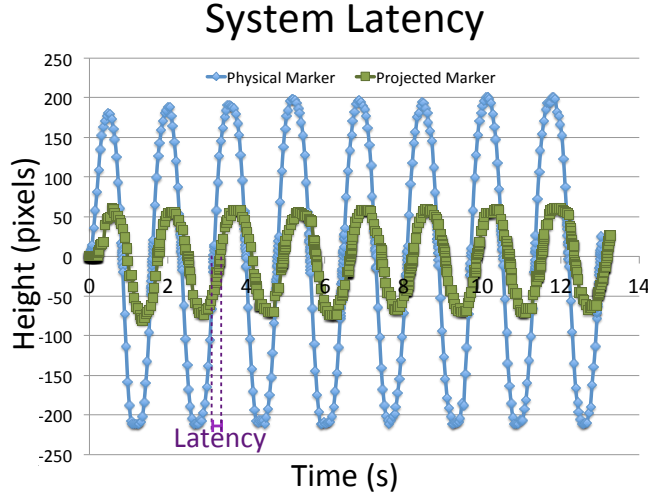


Fig. 8 The vertical position of the physical and projected markers for several iterations of the subject's movement. The latency is determined by the phase shift of the two waves.

between the stereo images occurs. DSCVR's arrangement, as shown in Figure 2, attempts to minimize crosstalk by creating a region in the center of the system for optimal viewing. We felt it was important to quantify the degradation of visual quantity outside of this zone. Early photographic analysis showed evidence of this phenomenon, as seen in Figure 9.

Previously, Febretti et al. attempted to measure crosstalk utilizing Weissman crosstalk patterns [19]. Unfortunately, this measurement approach requires a human's subjective assessment, meaning that precise measurements may require a large sample size. The process was also quite tedious as each monitor needed to be checked from each location, meaning each participant would need to make 500 evaluations. For these reasons we choose to use an optical approach [52].

To accomplish this we used a digital camera with an 8 megapixel sensor and 35 mm fixed focal length. As opposed to using patterns to assess stereo crosstalk, we used a luminance based approach similar to [22], however as opposed to separating the signals based on spatial locations, we instead separate the signals based on color [25]. We used red and blue, as these colors have an equal number of sensors which pass through a Bayer filter [10]. We chose 25 locations at which to sample the crosstalk amount for each of the ten columns. Three photographs were taken at each sample location for each column, with different configurations of left-eye / right-eye images: one with both images red (labeled R), one with both images blue (labeled B) and one in which one image was blue and the other was red (labeled T). To reduce indirect illumination from other displays, columns not being photographed were visually muted.

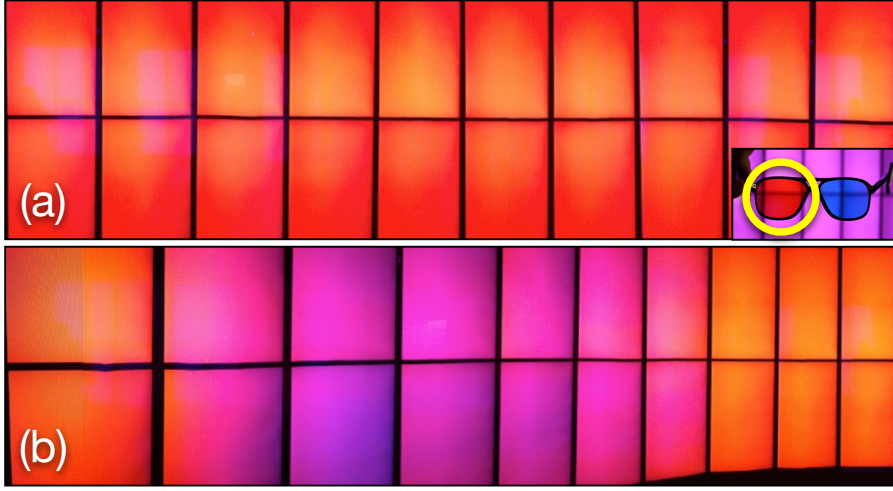


Fig. 9 Panoramic images, shot through a left-eye circular polarization filter (inset), showing stereo crosstalk artifacts. (a) shows a panorama captured at the center of the system, with all displays showing a red hue (apart from reflections). (b) shows a panorama taken to the left of center. Pink and purple hues indicate crosstalk due to off-axis viewing.

The three images were used to compute the amount of crosstalk for each column at each position. The first step was to compute the amount of crosstalk, done for the “red eye image” by measuring the the normalized signal loss of the red component and the normalized signal gain of the blue component across the display. (Equation 1). The second step was to determine the amount of the opposite eye’s image seen—which should not be seen under optimal conditions—by computing the gain in blue signal normalized to the difference between the blue component of the blue and red images (Equation 2). Finally, we computed the crosstalk amount as the sum of the red loss and blue gain (Equation 3).

Red Loss:

$$L = \frac{R_r - T_r}{R_r - B_r} \quad (1)$$

Blue Gain:

$$G = \frac{T_b - R_b}{B_b - R_b} \quad (2)$$

Crosstalk:

$$C = L + G \quad (3)$$

Figure 10 shows the average amount of stereo crosstalk for all columns in the system. As shown, the “sweet spot”, described in Section 2.2, is clearly visible at the center of the system. We found the average H value from Equation 3 to be 0.04. For positions extremely close to the system, however, values were close to 1.0, where the majority of columns were viewed off-axis. This result is further discussed in Section 5.2.

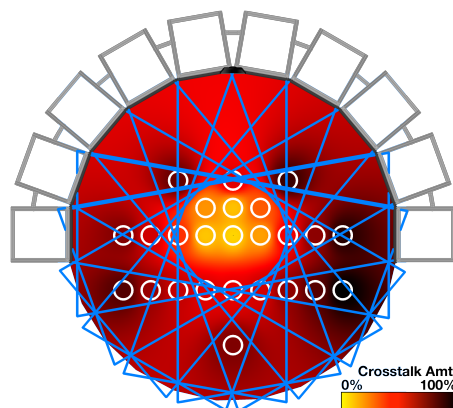


Fig. 10 A heat map showing the amount of crosstalk measured at 25 locations within the viewing ranges of the DSCVR System’s displays. Bright spots indicate low crosstalk, while darker spots indicate high crosstalk. Test locations and display viewing ranges are overlaid.

5 Discussion

We first provide a discussion of the results from the evaluation, then present a discussion of the challenges and future work.

5.1 Latency

Tracking using the Microsoft Kinect for Windows sensor was generally acceptable. While small jitters were sometimes evident, the flexibility of Kinect made it an excellent low-cost alternative to the InterSense system. As stated above, the small, centered optimal viewing area made the use of a single Kinect a viable option. However, multiple Kinects could provide a greater coverage area and a way to further improve the quality of the tracking data. Eventually replacing the single Kinect with the announced next-generation Kinect [20], featuring a higher-resolution sensor and lower latency, will likely have a substantial positive impact on the quality of the tracked user’s experience.

While the calculated latency for the DSCVR System met expectations, the determined latency for the CAVE was somewhat surprising. To verify the result, the test was performed using both TrackD and VRPN software [3] [48], and was performed on multiple software infrastructures. While the InterSense tracking system specifications reported very low latencies, our estimation is the smoothing parameters enabled by default on these trackers increased their latencies substantially.

5.2 Crosstalk

The results of the crosstalk test described in Section 4.3 adequately quantify and validate the convergence of viewing ranges we predicted during develop-

ment. As shown in Figure 10, the measured amounts of crosstalk were significantly less inside the area where all ten columns’ viewing ranges overlapped. While the measured average crosstalk amount inside of this “sweet spot” was never measured at 0, the results conform to previous studies, which have shown that a range of crosstalk in which 2-5% is considered to be very good, and a range of 5-8% is considered to be acceptable [19].

Several improvements to this test could improve its accuracy, however. The photos captured for the test may have been affected by the uneven lighting in DSCVR’s installation location, leading to variations in luminance between the left half and the right half. Furthermore, light from other sources, such as the windows on the right side of the room, produced reflections on the left half’s displays, contributing to luminance and hue variations. Though we tried to minimize these effects, improvements could be made to future versions of this test.

While the sweet spot in which all displays were without crosstalk was relatively small, the direct view facing the system was able to provide stereo imagery for spectators and audiences. Traditionally CAVE systems have enabled non-tracked viewers to share experiences through incorrect viewpoints, therefore we chose to focus on a single user virtual experience. However, synchronizable active stereo televisions would mitigate these crosstalk issues.

5.3 Comparison with Existing Systems

As shown throughout this paper, virtual reality systems present a plethora of trade-offs. In this regard, comparing and contrasting different virtual reality systems is extremely difficult. In [19], one metric used is cost per megapixel. We show this metric, along with overall cost and cost per immersive bandwidth (as described in Table 1) CAVE, CAVE2, and DSCVR in Table 3².

System	CAVE	CAVE2	DSCVR
Approximate Cost	\$1,000,000	\$926,000	\$41,000
Cost / 3D Megapixel	\$45,000	\$14,000 ⁴	\$2,040
Cost / Im. Bandwidth	\$2,600	\$2,875	\$55

Table 3 The cost of various virtual reality systems for different factors. The coloring refers to rank, with rank 1 colored green, rank 2 colored orange, and rank 3 colored pink.

However, even this simple comparison is somewhat problematic. For example, only sections of the CAVE and CAVE2 can ever be seen from a given viewpoint. On the other hand, the CAVE is the only system which provides total field of regard allowing the user to look in any direction. For the DSCVR

² Febretti et al. list the cost per megapixel as \$14,000 [19]. However, dividing the listed cost by the listed megapixels gives a number of approximately \$25,000. We have chosen to use the self-reported number in the paper for our analysis.

system we attempted to maximize the viewing characteristics from a single immersive viewpoint.

Our calculations of per-eye human vision characteristics (Table 1) show that DSCVR is competitive with similar systems, with a 3D system resolution only slightly below that of the CAVE, and a higher viewable 3D resolution than any other system evaluated. Furthermore, the immersive resolution and immersive bandwidth values show that DSCVR is, in fact, a superior high-bandwidth virtual reality environment to any of the other three systems. These performance characteristics are direct results of DSCVR’s use of higher-resolution 1080p stereo displays—perhaps an expected year-over-year improvement—and its smaller size—a deliberate choice, given the system’s design constraints. Finally, of the three systems surveyed, the DSCVR System has both the lowest cost per 3D megapixel and lowest cost per immersive bandwidth, demonstrating that a smaller, less expensive virtual reality environment can be a viable alternative to costly commercial-grade counterparts. As the quality of consumer grade technology continues to increase—and prices continue to decrease—we expect these systems to someday become commonplace.

5.4 Challenges and Future Work

Utilizing consumer-grade televisions provided many challenges in the design of the system. One of the unexpected challenges was that the consumer-grade LG displays shipped with many automatic image “optimization” features enabled by default. One particular setting, auto stereo adjustment, uses a “depth” value to shift 3D images in the horizontal direction. When the televisions were positioned in portrait orientation, this shift instead resulted in undesirable *vertical* image shifts. This problem was solved by setting the depth value to 10, apparently the “zero depth” point on a scale from 0 to 20. Additionally, the displays had a pattern detection feature enabled by default, which would shift the images in an attempt to find an ideal disparity. As the displays had been reoriented, this option needed to be disabled. Finally, like most modern televisions, latency-inducing image processing techniques had to be disabled by switching to the “Game” picture mode.

After evaluating both infrared and HDMI-CEC control methods, RS-232 communication was chosen for display communication because it offered the simplest, best-documented control scheme for these particular LG displays. Unfortunately, the displays, like the cluster machines, required relatively expensive USB serial adapters to access their embedded RS-232 hardware. HDMI-CEC appears to be a reluctant successor to decades-old serial control, but it is a divergent, undocumented, vendor-proprietary protocol as of this writing. Future developments in the field of consumer electronics may lead to better documentation and standardization of this protocol.

While the LG TVs’ bezels are significantly smaller than those shown in the NexCAVE [15], they are still noticeable. Three of the four bezels were

approximately 5mm across, but the forth bezel was five times larger, with a width of 25mm. While professional-grade televisions can be bought without this larger bezel, the cost of these displays is over eight times that of their consumer-grade counterparts at the time of this writing.

Furthermore, higher resolution 4K televisions have recently shown up in consumer markets. As the prices for these displays continue to fall, higher resolution HREs will be able to be built without substantial jumps in price. In the process of designing the DSCVR system, thought was put into how to make the system accessible for future upgrades. The DSCVR’s framing design, described in Sections 2.2 and 3.1, enable columns to be easily repositioned and adapted to new display hardware, offering the possibility to swap different sized monitors for system hardware upgrades.

Beyond TV, there is also a recent and earnest push towards consumer-grade virtual reality technology. The next-generation Kinect promises better resolution, a higher frame rate, and a more accurate sensor [20]. This technology will likely mitigate many of the issues currently raised by the utilization of a first-generation Kinect. Commodity input device technologies, such as the MYO wireless EMG armband developed by Thalmic Labs [1] or the STEM wireless, modular motion-tracking system [2], offer new means of virtual interaction at consumer-level pricing.

6 Conclusion

In this paper we have demonstrated the DSCVR System, a hybrid reality environment (HRE) built with commodity hardware. As part of the DSCVR System’s goals was to implement the system for under \$100,000, we present a breakdown of component costs and average energy consumption in Table 4. The final implementation of DSCVR cost just over \$ 40k and consumes slightly more than 3kW on average when in active use. The overall expenditures of the project, combined with the quality of the implementation, emphatically demonstrate that reasonably high-quality, large-scale HRE can be economically constructed from commodity off-the-shelf hardware.

While the price-point for DSCVR is much too high for most consumers, it is a very reasonable price for many small business and research labs. The ability to give clients a virtual walkthrough of an environment would be extremely useful for architects, real estate agents and interior designers, to name just a few beneficiaries. While virtual reality has been used in all of these fields, the cost has generally proved too high for smaller firms, limiting their interest in and utilization of VR and 3D user interfaces. We believe that, by significantly reducing the price-point of larger-scale immersive display environments, lower-cost systems like DSCVR will become commonplace in the future.

As with all virtual reality systems, many tradeoffs were considered during its development. While this approach has several shortcomings, such as crosstalk and display bezels, the DSCVR System has comparable and sometimes better performance characteristics than commercially-built systems, at

Hardware	Qty	Cost (USD)	Total Cost (USD)	Total Avg. Energy (kW)
LG 47LM7600 Displays	20	\$1078.60	\$21,571.99	1.172
Dell Alienware X51 Computers	12	\$1,057.51	\$12,690.12	2.080
Microsoft Kinect for Windows	1	\$234.25	\$234.25	0.012
3D Glasses ^a	2	\$21.95	\$43.90	—
80/20 Framing	—	—	\$3,938.35	—
Miscellaneous Equipment ^b	—	—	\$2,476.41	—
DSCVR System (Total)	—	—	\$40,955.02^c	3.264

^a Glasses use 90° Left & 90° Right Filters

^b Includes HDMI, RS-232, Ethernet cables; DisplayPort and USB serial adapters; and other equipment purchased during development.

^c A “hot spare” computer cost an additional \$1,057.51. This machine is not in active use, so we have omitted its cost and energy consumption from this table.

Table 4 Costs and energy consumption specifications for the DSCVR System and its individual components.

a fraction of their cost. As the quality of consumer-grade technology continues to increase while prices continue to decrease, it is likely that future consumer-grade HREs, using higher-resolution displays and higher-fidelity commodity tracking hardware, will have even better performance and lower costs than DSCVR. We see this as a democratizing trend that could enable new research and use cases in fields, industries and businesses that have previously been priced out of using such technology.

Acknowledgements We would like to acknowledge the support of the the Living Environments Laboratory, the School of Human Ecology, the UW-Madison Graduate School and the Wisconsin Institute for Discovery. We would specifically like to thank Vito Freese for his assistance with installation and Patricia Brennan, Kendra Kreutz, Andrew Wagner, John Hilgers and Roberto Rengel for their support and assistance in this project.

References

1. MYO - Gesture control armband by Thalmic Labs. URL <https://www.thalmic.com/en/myo/>
2. STEM System: The Best Way to Interact with Virtual Worlds by Sixense — Kickstarter. URL <http://www.kickstarter.com/projects/89577853/stem-system-the-best-way-to-interact-with-virtual>
3. trackd THE Device Driver Software for Immersive Displays. URL <http://www.mechdyne.com/trackd.aspx>
4. Ainsworth, R.A., Sandin, D.J., Schulze, J.P., Prudhomme, A., DeFanti, T.A., Srinivasan, M.: Acquisition of stereo panoramas for display in vr environments (2011)
5. Amatriain, X., Kuchera-Morin, J., Hollerer, T., Pope, S.T.: The allosphere: Immersive multimedia for scientific discovery and artistic exploration. *IEEE MultiMedia* **16**(2), 0064–75 (2009)
6. Arthur, K.: Effects of field of view on task performance with head-mounted displays. In: *Conference Companion on Human Factors in Computing Systems*, pp. 29–30. ACM (1996)
7. Avery, B., Thomas, B.H., Velikovsky, J., Piekarski, W.: Outdoor augmented reality gaming on five dollars a day. In: *Proceedings of the Sixth Australasian conference on*

- User interface - Volume 40, AUIC '05, pp. 79–88. Australian Computer Society, Inc., Darlinghurst, Australia, Australia (2005)
8. Bacim, F., Ragan, E., Scerbo, S., Polys, N.F., Setareh, M., Jones, B.D.: The effects of display fidelity, visual complexity, and task scope on spatial understanding of 3d graphs. In: *Proceedings of Graphics Interface 2013, GI '13*, pp. 25–32. Canadian Information Processing Society, Toronto, Ont., Canada, Canada (2013)
 9. Basu, A., Saupe, C., Refour, E., Raji, A., Johnsen, K.: Immersive 3d ui on one dollar a day. In: *3D User Interfaces (3DUI), 2012 IEEE Symposium on*, pp. 97–100 (2012)
 10. Bayer, B.E.: Color imaging array (1976). US Patent 3,971,065
 11. Bowman, D.A., McMahan, R.P.: Virtual reality: how much immersion is enough? *Computer* **40**(7), 36–43 (2007)
 12. Clark, R.A., Pua, Y.H., Fortin, K., Ritchie, C., Webster, K.E., Denehy, L., Bryant, A.L.: Validity of the microsoft kinect for assessment of postural control. *Gait & Posture* **36**(3), 372 – 377 (2012)
 13. Cruz-Neira, C., Sandin, D.J., DeFanti, T.A.: Surround-screen projection-based virtual reality: the design and implementation of the cave. In: *Proceedings of the 20th annual conference on Computer graphics and interactive techniques*, pp. 135–142. ACM (1993)
 14. CruzNeira, C., Sandin, D.J., DeFanti, T.A., Kenyon, R.V., Hart, J.C.: The cave: audio visual experience automatic virtual environment. *Communications of the ACM* **35**(6), 64–72 (1992)
 15. DeFanti, T., Acevedo, D., Ainsworth, R., Brown, M., Cutchin, S., Dawe, G., Doerr, K.U., Johnson, A., Knox, C., Kooima, R., Kuester, F., Leigh, J., Long, L., Otto, P., Petrovic, V., Ponto, K., Prudhomme, A., Rao, R., Renambot, L., Sandin, D., Schulze, J., Smarr, L., Srinivasan, M., Weber, P., Wickham, G.: The future of the cave. *Central European Journal of Engineering* **1**(1), 16–37 (2011)
 16. DeFanti, T.A., Acevedo, D., Ainsworth, R.A., Brown, M.D., Cutchin, S., Dawe, G., Doerr, K.U., Johnson, A., Knox, C., Kooima, R., et al.: The future of the cave. *Central European Journal of Engineering* **1**(1), 16–37 (2011)
 17. Doerr, K., Kuester, F.: Cglx: a scalable, high-performance visualization framework for networked display environments. *Visualization and Computer Graphics, IEEE Transactions on* **17**(3), 320–332 (2011)
 18. Eilemann, S., Makhinya, M., Pajarola, R.: Equalizer: A scalable parallel rendering framework. *Visualization and Computer Graphics, IEEE Transactions on* **15**(3), 436–452 (2009)
 19. Febretti, A., Nishimoto, A., Thigpen, T., Talandis, J., Long, L., Pirtle, J.D., Peterka, T., Verlo, A., Brown, M., Plepys, D., Sandin, D., Renambot, L., Johnson, A., Leigh, J.: CAVE2: a hybrid reality environment for immersive simulation and information analysis. In: M. Dolinsky, I.E. McDowall (eds.) *IS&T/SPIE Electronic Imaging*, pp. 864,903–864,903–12. SPIE (2013)
 20. Heddle, B.: The New Generation Kinect for Windows Sensor is Coming Next Year - Kinect for Windows Product Blog - Site Home - MSDN Blogs. URL <http://blogs.msdn.com/b/kinectforwindows/archive/2013/05/23/the-new-generation-kinect-for-windows-sensor-is-coming-next-year.aspx>
 21. Higgins, T.: Unity-3d game engine (2010)
 22. Hong, H., Jang, J., Lee, D., Lim, M., Shin, H.: Analysis of angular dependence of 3-d technology using polarized eyeglasses. *Journal of the Society for Information Display* **18**(1), 8–12 (2010)
 23. Humphrey, W., Dalke, A., Schulten, K.: Vmd: visual molecular dynamics. *Journal of molecular graphics* **14**(1), 33–38 (1996)
 24. Johnson, G.P., Abram, G.D., Westing, B., Navr'til, P., Gaither, K.: Displaycluster: An interactive visualization environment for tiled displays. In: *Cluster Computing (CLUSTER), 2012 IEEE International Conference on*, pp. 239–247. IEEE (2012)
 25. Kim, T., Ra, J.M., Lee, J.H., Moon, S.H., Choi, K.Y.: 3d crosstalk compensation to enhance 3d image quality of plasma display panel. *Consumer Electronics, IEEE Transactions on* **57**(4), 1471–1477 (2011). DOI 10.1109/TCE.2011.6131113
 26. Knox, C., Brown, M., Doerr, K., Jenks, S., Zender, C., Kuester, F.: Simultaneous visualization of the ipcc ar4 model ensemble on an extremely high resolution display wall (hiperwall). In: *AGU Fall Meeting Abstracts*, vol. 1, p. 1140 (2005)

27. Laha, B., Sensharma, K., Schiffbauer, J., Bowman, D.: Effects of immersion on visual analysis of volume data. *Visualization and Computer Graphics*, IEEE Transactions on **18**(4), 597–606 (2012). DOI 10.1109/TVCG.2012.42
28. Lange, B., Koenig, S., Chang, C.Y., McConnell, E., Suma, E., Bolas, M., Rizzo, A.: Designing informed game-based rehabilitation tasks leveraging advances in virtual reality. *Disabil Rehabil* **34**(22), 1863–1870 (2012)
29. Leigh, J., Johnson, A., Renambot, L., Peterka, T., Jeong, B., Sandin, D.J., Talandis, J., Jagodic, R., Nam, S., Hur, H., et al.: Scalable resolution display walls. *Proceedings of the IEEE* **101**(1), 115–129 (2013)
30. Livingston, M., Sebastian, J., Ai, Z., Decker, J.: Performance measurements for the microsoft kinect skeleton. In: *Virtual Reality Short Papers and Posters (VRW)*, 2012 IEEE, pp. 119–120 (2012)
31. Luo, J., Qin, K., Zhou, Y., Mao, M., Li, R.: Gpu rendering for tiled multi-projector autostereoscopic display based on chromium. *The Visual Computer* **26**(6-8), 457–465 (2010)
32. Margolis, T., DeFanti, T.A., Dawe, G.: Low cost heads-up virtual reality (HUVR) with optical tracking and haptic feedback. *SPIE-IS&T ...* (2011)
33. McMahan, R.P., Bowman, D.A., Zielinski, D.J., Brady, R.B.: Evaluating display fidelity and interaction fidelity in a virtual reality game. *Visualization and Computer Graphics*, IEEE Transactions on **18**(4), 626–633 (2012)
34. Meyer-Spradow, J., Ropinski, T., Mensmann, J., Hinrichs, K.: Voreen: A rapid-prototyping environment for ray-casting-based volume visualizations. *Computer Graphics and Applications*, IEEE **29**(6), 6–13 (2009)
35. Pausch, R.: Virtual reality on five dollars a day. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '91, pp. 265–270. ACM, New York, NY, USA (1991)
36. Polys, N.F., Kim, S., Bowman, D.A.: Effects of information layout, screen size, and field of view on user performance in information-rich virtual environments. *Computer Animation and Virtual Worlds* **18**(1), 19–38 (2007). DOI 10.1002/cav.159. URL <http://dx.doi.org/10.1002/cav.159>
37. Ponto, K., Doerr, K., Kuester, F.: Giga-stack: A method for visualizing giga-pixel layered imagery on massively tiled displays. *Future Generation Computer Systems* **26**(5), 693–700 (2010)
38. Ponto, K., Wypych, T., Doerr, K., Yamaoka, S., Kimball, J., Kuester, F.: Videoblaster: a distributed, low-network bandwidth method for multimedia playback on tiled display systems. In: *Multimedia, 2009. ISM'09. 11th IEEE International Symposium on*, pp. 201–206. IEEE (2009)
39. Prabhat, Forsberg, A., Katzourin, M., Wharton, K., Slater, M.: A comparative study of desktop, fishtank, and cave systems for the exploration of volume rendered confocal data sets. *Visualization and Computer Graphics*, IEEE Transactions on **14**(3), 551–563 (2008). DOI 10.1109/TVCG.2007.70433
40. Ragan, E., Kopper, R., Schuchardt, P., Bowman, D.: Studying the effects of stereo, head tracking, and field of regard on a small-scale spatial judgment task. *Visualization and Computer Graphics*, IEEE Transactions on **19**(5), 886–896 (2013). DOI 10.1109/TVCG.2012.163
41. Rash, C., McLean, W., Mozo, B., Licina, J., McEntire, B.: Human factors and performance concerns for the design of helmet-mounted displays. In: *RTO HFM Symposium on Current Aeromedical Issues in Rotary Wing Operation* (1999)
42. Renambot, L., Jeong, B., Hur, H., Johnson, A., Leigh, J.: Enabling high resolution collaborative visualization in display rich virtual organizations. *Future Generation Computer Systems* **25**(2), 161–168 (2009)
43. Rosson, M.B., Carroll, J.M.: *Usability engineering: scenario-based development of human-computer interaction*. Elsevier (2001)
44. Sampaio, P.N., de Freitas, R.I.C., Cardoso, G.N.P.: Ogre-multimedia: An api for the design of multimedia and virtual reality applications. In: *Knowledge-Based Intelligent Information and Engineering Systems*, pp. 465–472. Springer (2008)
45. Schou, T., Gardner, H.J.: A Wii remote, a game engine, five sensor bars and a virtual reality theatre. In: *OZCHI '07*, pp. 231–234. ACM Press, New York, New York, USA (2007)

46. Shupp, L., Andrews, C., Dickey-Kurdziolek, M., Yost, B., North, C.: Shaping the display of the future: The effects of display size and curvature on user performance and insights. *Human-Computer Interaction* **24**(1-2), 230–272 (2009)
47. Simon, A., Gobel, M.: The i-cone trade; - a panoramic display system for virtual environments. In: *Computer Graphics and Applications*, 2002. Proceedings. 10th Pacific Conference on, pp. 3–7 (2002). DOI 10.1109/PCCGA.2002.1167834
48. Taylor II, R.M., Hudson, T.C., Seeger, A., Weber, H., Juliano, J., Helser, A.T.: Vrpn: a device-independent, network-transparent vr peripheral system. In: *Proceedings of the ACM symposium on Virtual reality software and technology*, pp. 55–61. ACM (2001)
49. Teather, R.J., Pavlovych, A., Stuerzlinger, W., MacKenzie, I.S.: Effects of tracking technology, latency, and spatial jitter on object movement. In: *3D User Interfaces*, 2009. 3DUI 2009. IEEE Symposium on, pp. 43–50. IEEE (2009)
50. Wells, M.J., Venturino, M.: Performance and head movements using a helmet-mounted display with different sized fields-of-view. *Optical Engineering* **29**(8), 870–877 (1990)
51. Williams, S.C.: Immersive visualization. *Proceedings of the National Academy of Sciences* **110**(12), 4438–4438 (2013)
52. Woods, A.: Understanding crosstalk in stereoscopic displays. In: *Keynote Presentation at the Three-Dimensional Systems and Applications Conference*, Tokyo, Japan, pp. 19–21 (2010)