The Impact of Display Bezels on Stereoscopic Vision for Tiled Displays

Jürgen Grüninger*
Intel Visual Computing Institute

Jens Krüger†
High Performance Computing Group, University Duisburg-Essen
Scientific Computing and Imaging Institute, University of Utah

Abstract

In recent years high-resolution tiled display systems have gained significant attention in scientific and information visualization of large-scale data. Modern tiled display setups are based on either video projectors or LCD screens. While LCD screens are the preferred solution for monoscopic setups, stereoscopic displays almost exclusively consist of some kind of video projection. This is because projections can significantly reduce gaps between tiles, while LCD screens require a bezel around the panel. Projection setups, however, suffer from a number of maintenance issues that are avoided by LCD screens. For example, projector alignment is a very time-consuming task that needs to be repeated at intervals, and different aging states of lamps and filters cause color inconsistencies. The growing availability of inexpensive stereoscopic LCDs for television and gaming allows one to build high-resolution stereoscopic tiled display walls with the same dimensions and resolution as projection systems at a fraction of the cost, while avoiding the aforementioned issues. The only drawback is the increased gap size between tiles.

In this paper, we investigate the effects of bezels on the stereo perception with three surveys and show, that smaller LCD bezels and larger displays significantly increase stereo perception on display wall systems. We also show that the bezel color is not very important and that bezels can negatively affect the adaption times to the stereoscopic effect but improve task completion times. Finally, we present guidelines for the setup of tiled stereoscopic display wall systems.

CR Categories: B.4.2 [Input/Output and Data Communications]: Input/Output Devices—Image Display I.3.6 [Computer Graphics]: Methodology and Techniques—Ergonomics; I.3.7 [Computer Graphics]: Three-Dimensional Graphics and Realism—Virtual reality;

Keywords: stereoscopic vision, display bezels, graphical user interfaces, tiled displays, survey

1 Introduction

Large high-resolution display walls offer considerably more screen space for showcasing information, as compared to a standard desktop display. This is one reason why tiled display walls are already an ubiquitous element, e.g., at exhibitions, in shopping centers and in product presentations/advertisements. But even in scientific scenarios, tiled display walls are gaining more and more attention, for example, for control rooms and large-scale visualizations [Hiperwall 2010]. Recent research, however, shows that more screen space is not the only positive effect that can be achieved using large screens. They also improve collaboration [Birnholtz et al. 2007], allow faster physical navigation [Ball and North 2005; **?**] and support 3D navigation tasks [Tan et al. 2003; **?**; **?**]. Despite these positive arguments, tiled display systems suffer from the detrimental effect of monitor bezels disrupting the image when building a single large display by tiling multiple screens.

Although gaps between display tiles can be almost completely avoided using projection systems, these systems are very costly and require continuous maintenance. They either require the use of edge-blending, which causes black-level problems (particularly for scientific visualization where black backgrounds are often used), or a very complicated and fragile pixel-exact alignment is needed.

For stereoscopic displays, bezels cause two main problems. Since a stereoscopic image is displayed as two separate images with slightly different viewing positions, the complete image information for one point in space is separated to two points on the display. If a bezel covers one of these points, image information, and thus the stereoscopic effect, is lost for these parts (see Figure 1). For objects rendered behind the focal plane —i.e., objects that appear behind the screen— this effect is similar to a real-life phenomenon. For instance, standing close to a window, the window frames block the view to objects behind the glass. However, for objects rendered in front of the focal plane— i.e., objects that appear in front of the display—a very unnatural view is perceived. From the stereo disparity of the images, our visual system infers that the rendered object has to be in front of the display system. The fact that the bezels "occlude" the object, however, tells the visual system the opposite. For some observers, this visual paradox results in objects suddenly appearing "transparent" or "blurry" where bezels occlude parts. In this case, the occlusion is described as bezels "shining through" the object. For other observers, this paradox results in a complete breakdown of the stereo illusion, and the images for the left and the right eye are no longer interpreted as stereo images, but as separate, overlaid images of the scene.

These paradox effects are *not* limited to bezels around monitors of display walls. A similar effect can be experienced for mixed 2D and stereoscopic 3D content, such as a windowed stereoscopic environment that is embedded into a standard 2D graphical user interface. Also tiled projection-based systems can have seams or gaps when non-continuous projection surfaces are used or blending-zones between the tiles are large and/or misaligned. CAVE [Cruz-Neira et al. 1993] environments are a prominent example for such situations, where, due to their cube-like shape, seams can be perceived in each corner, especially when the projected image darkens too much towards the corners.

In this paper, we describe four surveys that are intended to better understand how feasible it is to use stereoscopic displays for large display walls, how such displays compare to stereoscopic projection systems in terms of image fidelity, and the impact of display bezels. As mentioned, a scene, in which the entire scene is rendered

*e-mail:grueninger@intel-vci.uni-saarland.de
†e-mail:krueger@intel-vci.uni-saarland.de

VRST 2013, October 6–9, 2013, Singapore.

ACM 978-1-4503-2379-6/13/10 $15.00

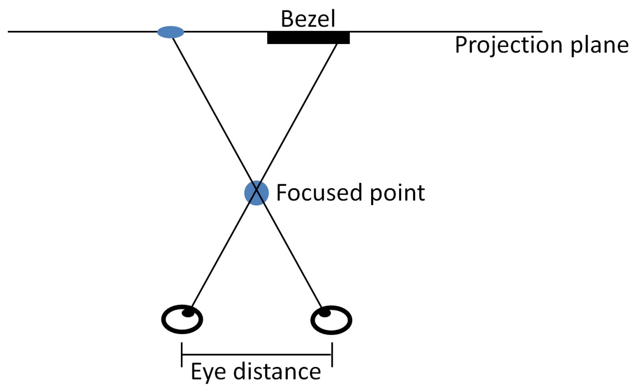

Figure 1: *Some information of the stereoscopic "double-image" is hidden by the bezel on the projection surface. The bezel is also visible behind the focused point/object, which does not reflect real world behaviour where the focused object would occlude the bezel.*

"behind" the physical screen ("fish tank stereo"), is much less critical than a scenario, in which the objects seemingly stick out of the display wall. To assess all aspects, we therefore focus our surveys on scenes that use the entire depth (behind, in, and in front of the image plane). Furthermore, we asked the participants of the second pre-survey to categorize their answers for near, middle, and far objects to get categorizable results for our assumption from experience. The main goal was to ascertain which controls (e.g., maximum acceptable bezel size, minimal screen size, blending-zones) are essential to the successful construction of stereoscopic display walls. Although we are aware that some of the general findings may conform to common sense, we find it still important to ascertain the boundaries and combinations of these controls, which allow for good results. These are not obvious at all at first glance. Therefore we carried out our studies with 71 participants and combined the outcome with previous findings about mixing stereoscopic images and 2D content.

2 Related Work

Previous research on display wall interaction and ergonomics has been explored mostly for monoscopic setups.

Ball and North [Ball and North 2005] conducted observational analyses, including direct observations and interviews over six months, to find how users change their behaviour when working with tiled displays over longer periods of time. In general, users in this analysis reported that the bezels are considered a negative hardware feature. In the process of adjusting to the environment, however, users employed the bezels to efficiently separate application windows. With training, some users were even able to ignore the bezels completely and use large windows that spanned multiple displays.

Robertson et al. [Robertson et al. 2005] examined how large displays can enhance workflows and support users. They reported that bezels introduce interaction and visual problems when a window or the mouse crosses them. This is due to missing information "behind" the bezels or discontinuous interactions (e.g. bent mouse paths). On the other hand, they state that bezels are helpful to organize activities by separating work on different displays, similar to Ball and North [Ball and North 2005]. To avoid these problems, Baudisch et al. developed "Mouse Ether" [Baudisch et al. 2004] for continuous mouse movement over multiple displays, as well as "OneSpace" for seamless image distribution.

Similar results were reported by Grudin [Grudin 2001] in his work on multiple monitor desktop systems, where users tended to use bezels as a logical delimiter for different workspaces and tasks.

Mackinlay and Heer [Mackinlay and Heer 2004] developed several user interface techniques for creating seam-aware applications that attempt to mitigate the detrimental effects of monitor bezels. They achieve this by regarding the entire display configuration as a large display and every individual monitor as a viewport into this larger space.

Tan et al. [Tan and Czerwinski 2003] conducted a study on the effects of information distribution and physical discontinuities across multiple displays. They placed two displays —either two monitors or one monitor and one projection screen— at different depths, with the same visual angle. With a divided-attention paradigm, they investigated how physical discontinuities or bezels affect user performance when responding to instant messenger pop-ups or comparing text. They showed that when information is separated with an offset in depth, small detrimental effects on the tested tasks occur. On the other hand, separation of information through monitor bezels alone did not affect user performance.

Recently, Bi et al. [Bi et al. 2010] analyzed user performance on display walls. They conducted several surveys and demonstrated that bezels do not affect visual search time or error rate, in contrast to splitting objects across bezels. They did, however, find that mouse-steering behaviour in these environments was impaired by bezels. Beyond that, no significant effects on the selection time of objects were found.

Ebert et al. [Ebert et al. 2010] developed a projection system called Tiled++, which allowed them to project lower-resolution images on display bezels to reveal information that would normally be hidden behind them. They showed that, in a non-stereoscopic setting, their system is capable of mitigating the drawbacks introduced by bezels on tiled displays.

The Varrier display system built by Sandin et al. [Sandin et al. 2005] consisting of 35 auto-stereoscopic LCD panels was one of the first large-scale display wall systems capable of displaying stereoscopic images. They reported that informal simulations of tiled borders in a CAVE environments showed no detrimental effect to user immersion. Tests and informal feedback from users of the Varrier system seemed to confirm this hypothesis, but they do acknowledge that, so far, no formal studies on the effect of bezels were conducted. This is the focus of the survey presented in this paper: to provide statistical evidence for and against assumptions that are often made about tiled stereoscopic display walls.

The remainder of this paper is structured as follows: In the next section we describe our survey setup, including the hardware used to conduct all tests. We first focus on a smaller survey that was conducted to make sure that the setup of our following surveys is valid. In section 5, we detail the parameters of another pre-survey, which tested different parameters of bezels, such as size and color. After that, the survey results are discussed. Section 6 shows the first timed, objective survey using very simple and hypothetical volumetric datasets to find more evidence of negative effects on user performance from bezels and to further asses scene complexity influences. We describe a last, more complex survey in section 7, which evaluates a more realistic scenario with a volumetric mandelbulb dataset. This is followed by our conclusions from the surveys and a note for future research.

3 Setup

As the target of this paper was to evaluate the general effects of a number of bezel parameters, including the size, color, and number,

Figure 2: *The SmartFactory scene used as a VR training scenario.*

we could not simply perform the evaluation using a single stereo tiled-screen system. In particular, we were also interested in the variation of these parameters outside the range of currently available LCD screens (e.g., make the bezels extremely small). Since real hardware setups would be impractical for this amount of variations, we chose to use rendered, virtual bezels that get displayed in the normal image. To confirm that this "virtualization" step does not invalidate our findings, we conducted a first pre-study, comparing real bezels to the virtual bezels and a second, exploring bezel variations.

During the pre-surveys, we utilized a $2.44m \times 1.83m$ stereo back-projection system consisting of two *Projectiondesign F20 sx+* projectors [Projectiondesign 2010] with a resolution of 1400×1050 pixels each. A passive stereo setup was used in which the projectors were outfitted with linear polarization filters and the participants wore polarization glasses. To satisfy virtual-reality (VR) definitions, an ART tracking system with four infrared tracking cameras was used to track the head position of the user with high precision in real-time. Here, we aimed for a "virtual bezel"-scenario similar to the one used by Bi et al. [Bi et al. 2010].

To render and control the VR environment, we used the "Lightning" [Landauer et al. 1997] software from the Fraunhofer Institute. *Lightning* uses a scene graph to represent the VR scene. To this scene graph, we added a new node that renders the screen bezels as a set of rectangles in front of the rest of the scene. To be able to modify the setup quickly, we wired this new node to a set of shortcut commands. As an example for a virtual scene, we selected a model of the inside of a factory building (see Figure 2) that is regularly used as a VR training and demo environment. For later tests it is important to note that the dominant color in this scene is gray due to the fact that aluminum profiles and unpainted walls are used in this factory. Also the users were advised to stand $1.5m$ away from the screen so that most of their field of view was covered by the screen.

For the two timed real task surveys, we used a bigger passive projection screen with the measurements of $4.48m \times 2.80m$ and two projectors with a resolution of 2560×1600 pixels each. Participants were asked to stand at a fixed point $3.5m$ away from the screen. To render and control the volumetric mandelbulb dataset, we used the visualization software "ImageVis3D" [Scientific Computing and Imaging Institute Utah 2012]. The dataset was rendered in grayscale colors, which also matched the dominant color in the pre-survey setups.

4 Virtual bezel pre-survey

To validate that the virtual bezels have a similar effect on user perception as real physical objects, we built a grid from corrugated cardboard pieces that served as a test dummy. To resemble real monitor bezels as closely as possible, we took photos of our monoscopic display wall, printed them on A3 paper sheets, and glued them on our cardboard grid. Putting this grid in front of a back projection screen allowed us to quickly switch between real bezels and virtual bezels by simply pushing the cardboard grid aside. Figure 3 shows both the cardboard grid as well as the virtual grid.

With this setup, we performed a smaller study with a total of 11 test subjects from the computer science department. All participants were male and their previous VR experience ranged from "none" to "using VR systems on a regular basis." Each user was given a short introduction into controlling the VR environment with a tracked computer mouse, used as a six-degrees-of-freedom interaction device. After that, every user was given the chance to fly around the factory in search for soap dispenser bottles which were placed near a production lane. Since this survey focused only on the visual appearance of the real and the virtual bezels, we kept the task simple and did not time it.

Each participant had to complete the task once with the real and once with the virtual bezels of the same geometry and color. At the end of the survey, each participant was asked if he similarly perceived the stereoscopic image in both scenarios or not.

Nine of the eleven participants stated that they perceived both scenarios with the same effects. Some pointed out that there were minor differences mainly because the cardboard grid was not completely flat and cast a shadow onto the screen. We used a Kolmogorov-Smirnov test [Eadie et al. 1983] to check if the answers significantly deviated from a standard normal distribution, which would have been the case if participants answered randomly. The test indicated that the empirical distribution significantly differed from the normal distribution ($Z = 2.714, p < 0.01$). The trend towards the positive answer therefore can be taken as significant and we can assume that the results based on our virtual grid are transferable to a real tiled screen.

5 Bezel feature pre-survey

After validating our settings with the pre-survey, we were able to define a set of hypotheses and scenarios for the bezel feature survey concerning the anticipated effects on depth perception.

Our *first hypothesis* is that a bezel-free image allows for significantly better depth perception with immersive images than an image with bezels. As mentioned earlier, we are mostly interested in the effects of bezels on a real VR scene. Consequently, our VR system, the scene, and the user's actions are designed to use the full depth range (i.e., objects are rendered to appear in front, in, and behind the image plane). Should this hypothesis prove to be wrong, no further test will be necessary.

If a significant difference can be demonstrated, we are looking for proof of our *second hypothesis* that thin bezels allow better depth perception than larger bezels with immersive images. Should this hypothesis be valid for bezel sizes larger than zero, then setting up a truly immersive environment with LCD screens will be only a matter of time, as display vendors have demonstrated prototype devices with extremely thin bezels [OLED-DISPLAY.net 2010].

Figure 3: *The cardboard grid (left) attached to the projection screen and the screen with virtual bezels (right). Our pre-survey shows that the appearance is very similar.*

Our *third hypothesis* is that bezels with the dominant color of the scene (in our case, gray) allow for better depth perception than black bezels. Here, we are trying to determine if blending bezels into the scene produces a similarly successful result as the approach of Ebert et al. [Ebert et al. 2010] for monoscopic images.

Our *fourth hypothesis* is that fewer bezels allow better depth perception. This would indicate that it makes sense to invest in fewer, larger, higher-resolution displays, which in practice means significantly higher cost.

Figure 4: *Close-up of the blending-zones used to smooth the boundaries of the screen bezels for testing hypothesis 5.*

Our *fifth hypothesis* is that blending zones (gradients) allow better depth perception with immersive images (see Figure 4). This hypothesis was based on previous informal tests that we conducted when combining 2D graphical user interfaces with embedded stereoscopic image areas. When using black gradients to blend the dark gray GUI with the stereoscopic area, the transition between these two elements seemed much smoother and caused fewer visual conflicts.

Since the survey was conducted as a within-subjects survey, we designed a relatively short questionnaire with nine scenarios to test all our hypotheses, and randomized the ordering of the scenarios during the actual survey, to counterbalance issues caused by fatigue or learning effects. The defined scenarios were as follows:

- **Scenario 1**: navigation of a bezel-free scene, serving as a training environment and as a basis of comparison for the following scenarios.
- **Scenario 2**: two horizontal and two vertical average/medium-sized bezels with black color.
- **Scenario 3**: two horizontal and two vertical thin bezels with black color.
- **Scenario 4**: two horizontal and two vertical thick bezels with black color.
- **Scenario 5-7**: same setup as scenario 2-4 but with a medium gray bezel color.
- **Scenario 8**: one horizontal and one vertical average/medium bezel with black color.
- **Scenario 9**: two horizontal and two vertical average/medium sized bezels with black color and an added blending zone on the insides of each bezel.

The measurement for average/medium-sized bezels was determined by considering the mean bezel thickness of 12 different LCD screens ranging from 20” to 30” display size. The mean size for the left and right bezel sides resulted in $2.485cm$, which was rounded up to $2.5cm$ for our survey use, resulting in a $5cm$ gap between two screens (see Figure 3). The thickest bezel size was based on the maximum bezel size found on the aforementioned 12 LCD screens and rounded up to $4cm$ ($8cm$ gap). The thinnest bezel size was a fictional bezel size of $0.6cm$. This size was chosen to be small enough that we suspect it to be sufficiently small to not be a distraction. Furthermore, we chose it to be larger than available high-end projection solutions (e.g., the “eyevis DLP® Cube” [eyevis 2010] with $0.6mm = 0.06cm$ total bezel size) and announced LCD screens ([OLED-DISPLAY.net 2010]). The black and gray colors—represented by the RGB values $[0, 0, 0]$ for black and $[0.5, 0.5, 0.5]$ for gray—were also based on the available bezel colors for the previously mentioned common LCD screens. As mentioned before, the grey color also allowed us to test our third hypothesis.

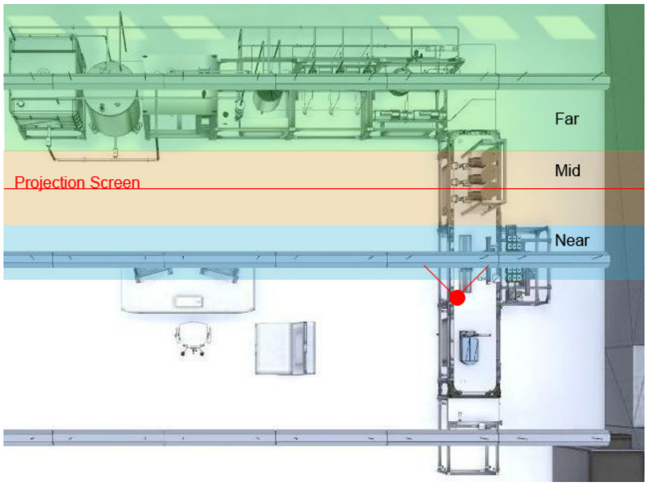

Figure 5: *The depth areas and the viewing position chosen for the survey with the SmartFactory scene. The red line shows the position of the zero parallax area/projection screen.*

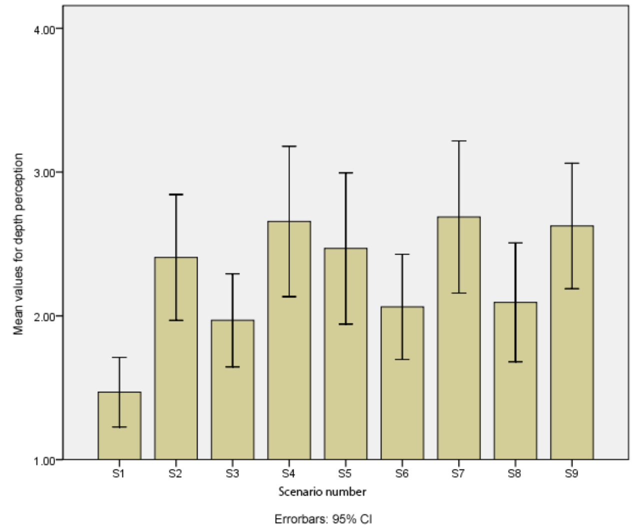

Figure 6: *Averaged score of depth perception relative to scenarios.*

Table 1: *Comparison between all estimated marginal means for general depth perception for scenes with two horizontal and two vertical bezel lines.*

Scenario	Mean	N
3 (thin/black/2×2)	1.9687	32
6 (thin/gray/2×2)	2.0625	32
2 (medium/black/2×2)	2.4063	32
5 (medium/gray/2×2)	2.4688	32
4 (thick/black/2×2)	2.6563	32
7 (thick/gray/2×2)	2.6875	32

The questionnaire was available in multiple languages and was designed to cover most of the possible effects that might occur in these scenarios. Two questions asked the participants to rate the quality of their depth perception on a seven point scale ranging from the values 1 to 7; two following questions were designed to determine whether the participants experienced any kind of "pressure" in their head. This pressure can be an indicator of misalignments in the image when the human visual system is processing confusing information. Three more questions asked participants to rate "blurry" image parts and identify the depth areas in which they occurred. In this pre-survey we chose these more subjective features over objective values since we were more interested in the general acceptance of a bezeled stereoscopic display than in the overall performance when working with the system.

Two questions about feelings of nausea and one free-text field for additional comments ended the questionnaire. For each scenario, the same questions were repeated.

As a reference position in our virtual-reality scene, we chose a camera position above a production line, which offered us the possibility to test multiple depth areas and motion at the same time (see Figure 2). In this production line, soap dispenser heads were used as a reference for near objects, a soap-filling hose was defined as the mid-depth area and a fluid tank was used for the far regions. A soap bottle moving on the production line and through all depth areas allowed us to test changing visual depths, despite the fixed viewing position (see Figure 5).

Thirty-five participants (13 female), aged 16 to 58, participated in this survey. One participant opted to abort the survey after the third scenario without a specific reason and the questionnaire was removed from the survey. Except for three experts, all other participants had little to no experience with stereoscopic content. All but one user with chronic uveitis (inflammation of the middle layer or interior of the eye) had normal or corrected-to-normal eyesight. The survey took about 30 minutes, and all participants were compensated for their participation.

5.1 Results

For checking the general effects of bezels on depth perception (hypothesis 1), we used a one-way repeated measures analysis of variance (RM-ANOVA) [Eadie et al. 1983]. Interactions between colors and bezel thickness were analyzed by using a two-way RM-ANOVA. Another one-way RM-ANOVA was used to compare the effects on depth perception between different bezel counts (scenarios 8, 2, 5) and the effects of a blending zone (scenarios 9, 2, 5).

For checking the general effects of bezels on depth perception (hypothesis 1), we used the aforementioned one-way RM-ANOVA to compare all nine scenarios. The test for the univariate contrast between scenario 1 and the remaining, bezeled scenarios showed a significant difference ($F(1,31) = 22.481, p < 0.01$). Looking at the means for general depth perception (shown in Figure 6), it can be stated that the depth perception in scenario 1 was rated significantly better than in other scenarios, which supports our first hypothesis.

We found a significant effect on depth perception for the bezel size as described in hypothesis 2 ($F(2,30) = 6.700, p = 0.004$). A univariate test to compare all scenarios with bezels and 9 image tiles showed a significant difference for scenarios with thin black, thin gray and thick gray bezels ($F(1,31) = 10.797, p = 0.003$ for scenario 3; $F(1,31) = 8.082, p = 0.008$ for scenario 6; $F(1,31) = 7.273, p = 0.011$ for scenario 7). Taking the means of scenario 6 ($M = 1.9687$, lower numbers represent better depth perception) and scenario 7 ($M = 2.6875$) into account, a desire for thinner bezels can be concluded (see Table 1). This also matches notes in the free-text question and statements from several participants.

Surprisingly, no significant effects on depth perception were found for the color of the bezels as assumed in hypothesis 3 ($F(1,31) =$

Table 2: *The first part of the table shows a comparison between the scores for general depth perception for scenes with similar bezel size but different amounts of bezels. The second part shows a comparison between the scenario with a blending-zone and scenarios with similar bezel size.*

SCENARIO	MEAN	N
8 (medium/black/1 $\times$ 1)	2.1515	33
2 (medium/black/2 $\times$ 2)	2.4545	33
5 (medium/gray/2 $\times$ 2)	2.5152	33
9 (medium/black/blending zone/2 $\times$ 2)	2.6250	32
2 (medium/black/2 $\times$ 2)	2.4063	32
5 (medium/gray/2 $\times$ 2)	2.4688	32

$0.326, p = 0.572$). It is, however, worth noting that some participants explicitly noted that they preferred the black bezels over the gray ones. Due to the mainly gray color of the used VR test scene, the gray bezels "vanished" in some parts of the image, while the black ones were continuous.

Also, no significant effects on depth perception were found for the interaction between bezel thickness and color ($F(2, 30) = 0.076, p = 0.927$), suggesting that the chosen color cannot compensate for a larger bezel size.

With the one-way RM-ANOVA, we found a significant effect in the univariate test for the bezel count (hypothesis 4) for scenario 8 compared to the other two tested scenarios ($F(1, 32) = 4.456, p = 0.043$). This suggests that fewer bezels improve depth perception. One observed effect of having larger image tiles was that users tended to focus on one tile at a time. Due to the increased size of the tiles in scenario 8, this covered more of their field of view and bezels were less present. Again taking mean values into account, scenario 8 attained the closest result for overall depth perception with medium-sized bezels compared to thin black or gray bezels ($M = 2.1515$, see Table 2).

Using the one-way RM-ANOVA for testing the blending zone (hypothesis 5), we found no significant better or worse values for scenario 9 compared to the two other scenarios ($F(1, 31) = 2.067, p = 0.161$). The estimated marginal means for this test are shown in the second part of Table 2. Although 4 users stated that the blending zone decreased the worsening effects of bezels on their depth perception, these results show no enhancement in depth perception using blending zones. This might be connected with the effect for bezel size, since the gradient for the blending zone virtually increased bezel size and slightly occluded more image information, which counterbalanced possible beneficial effects.

Also, 28 of 33 participants stated that, for most bezel scenarios, objects displayed above bezels and in front of the projection screen appeared "blurry". This matches the hypotheses for detrimental bezel effects mentioned in the introduction, and points towards the usage of "fishtank" stereoscopy—where all objects in front of the screen get clipped—on display walls or at least on systems with larger bezels. Although this would mean that some advantages of "real" VR are sacrificed, user performance and comfort could be increased. It is worth noting that this improvement in user performance and comfort has been validated by Demiralp [Demiralp et al. 2006] when comparing a CAVE system with a single fishtank VR display.

6 Timed task survey on a simple volumetric dataset

This survey consisted of 15 individual stereoscopic scenes showing multiple spheres rendered in red color on a light grey background. These colors were chosen due to the results of the second pre-survey, which indicated the preference for non-interfering features (objects/bezels/background). One scene was used for training; nine scenes were used for a spacial judgment task, where participants were asked to point out the sphere closest to them/sticking furthest out of the screen. The amount of spheres ranged from 3 to 25 where one half of the scenes consisted of regularly aligned spheres and the other half of randomly placed spheres. To rule out that participants merely used the size of the spheres as a distance indicator (e.g., big means close, small means far away), we normalized the screen size of all spheres to a constant value. This left stereo disparity as the only depth cue. The other five scenes were used for a counting task, showing a number of 11-19 spheres per scene. Each scene was presented once with and without bezels, resulting in 30 scenarios that were shown in a randomized order, resulting in a counterbalanced within-subjects survey design.

As bezel size we used $2.5cm$, which was the smallest size we knew of having an impact on stereo vision from the bezel feature pre-survey. The time used for adjusting to the stereo image, counting the features and also the result of the counting were recorded for each participant. The adjustment time was measured by asking the participants to turn away from the screen while scenarios were exchanged. When turning back to the screen, they gave a signal after having fully adjusted to the image and started counting. All times were measured in seconds.

Thirteen participants (2 female), aged 21 to 37, participated in this survey. Except for two experts, all other participants had little to no experience with stereoscopic content. All persons had normal or corrected-to-normal eyesight. All participants were compensated for their participation.

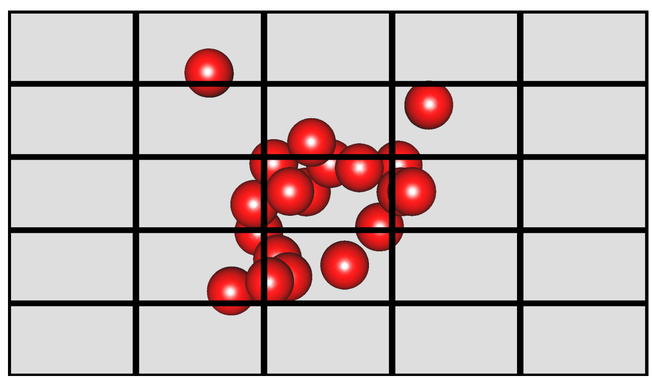

Figure 7: *A monoscopic view of scene 12 during a counting scenario with bezels.*

6.1 Results

A one-way RM-ANOVA analysis of the adjustment times for the scenario pairs in the spacial judgment task revealed significant results for three scenes: scene 4 ($F(1, 12) = 6.734$, $p = 0.023$), scene 5 ($F(1, 12) = 4.950$, $p = 0.046$) and scene 9 ($F(1, 12) = 8.015$, $p = 0.015$). All effects show a significantly better adjustment in scenarios where no bezels were present. Similar effects could be observed for the time the actual judging task took. A significant effect could be found for scene 1 ($F(1, 12) = 12.844$,

Table 3: *F and p values of the significant scenes in this survey.*

		Scene	F	p
Depth judging	Adjustment time	4	6.734	0.023
		5	4.950	0.046
		9	8.015	0.015
	Task time	1	12.844	0.004
		4	5.954	0.031
Counting	Adjustment time	12	7.417	0.018
		13	7.778	0.016
	Bezels * Scenes		10.626	0.007

$p = 0.004$) and scene 4 ($F(1, 12) = 5.954$, $p = 0.031$). A two-way ANOVA over all scenes did not reveal any significant interactions between bezels and scenes. Since scene 4 appeared with significant effects for adjustment times and the judging task, a closer look could help to understand why this happened: scene 4 consists of five spheres placed at random positions. Due to these random positions, most of the spheres ended up with only a slight overlap with the bezels and they never crossed vertically through the center region of the spheres. In previous informal tests, we also had the impression that horizontal bezels do not affect the stereoscopic image as much as vertical ones. This could be caused by the nature of stereoscopic projections, where the two images are shifted on the horizontal and not on the vertical axis. This results in the occlusion of the same parts of an object for horizontal bezels. Whereas a vertical bezel is hiding parts of the object further to the left in one stereo image and other parts further to the right in the other stereo image. It is also worth noting that, in our third scene, 10 of 13 participants gave at least one wrong answer which was the only significant effect for the correctness of the answers. Comparing the scene to all others, this was a scene consisting of five spheres aligned in a "X" shape, where three of the spheres were projected "in front" of the screen. Two of these spheres were placed at the same depth, which was half of the depth of the sphere placed closest to the participant. This could indicate, since distinguishing objects at the given distance from the participant was not entirely possible, that there were problems with the projection system or settings, resulting in a too large disparity between the two stereo images. When we asked the participants, however, all stated that they could clearly see all spheres. Also there were no problems with spheres this close to the participant in other scenes.

For the five remaining scenes that were used for the counting task, we found two significant effects for the time needed to adapt to the images: scene 12 ($F(1, 12) = 7.417$, $p = 0.018$) and scene 13 ($F(1, 12) = 7.778$, $p = 0.016$). Both again favored non-bezeled scenarios. There was also a significant effect for bezels ($F(1, 12) = 10.626$, $p = 0.007$) when we were looking for interactions between bezels and scenes. This effect was not present during the spatial judging task, which could indicate, that scenes with higher complexity (i.e., many depth cues as in "real world" scenes) see a higher detrimental effect from bezels than very simple, hypothetical scenes. On the other hand, there were no significant effects for the counting task completion time.

During the whole experiment, all participants described the bezels as "distracting" and "not helpful" and subjectively preferred the non-bezeled scenarios. One person felt "dizzy" after the experiment and had to take a break before leaving.

7 Timed task survey on a complex volumetric dataset

With the knowledge of the previous surveys and to further assess the effect of bezels on real user performance, we conducted a fourth, real task survey on the volumetric mandelbulb dataset. In this survey, participants were asked to count the "pin" shaped features on the "mushroom"-like structures found around the center of the dataset (see Figure 8). Flat and other non-pin-shaped features had to be excluded from the counting, which made the task intentionally much more complex but also more realistic than in the previous surveys. This dataset and task were chosen because of the mandelbulb being abstract enough to prevent people from using common sense or previous knowledge, yet the dataset shows simple enough structures to teach people without much effort. The mandelbulb dataset also filled most of the screenspace, so that in contrast to the previous timed survey there were many different depth cues visible at once. This is similar to the "real world" factory scene from the second pre-survey and allows us, when compared to the first timed task survey, to find indications, if scene complexity is an important factor. The survey consisted of four individual views on the mandelbulb dataset. One view was used for training and the other three were presented once with and without bezels, resulting in six timed scenarios that were shown in a randomized order during this within-subjects survey. As before, we used a bezel size of $2.5cm$. Again, the time used for adjusting to the stereo image and counting the features and also the result of the counting were recorded for each participant. Twelve participants (3 female), aged 22 to 37, participated in this survey. Except for two experts, all other participants had little or no experience with stereoscopic content. All persons had normal or corrected-to-normal eyesight. All participants were compensated for their participation.

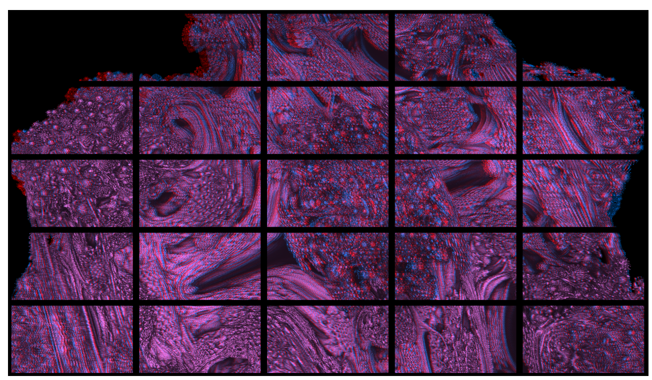

Figure 8: *An anaglyph stereo view on a "mushroom"-like structure on the mandelbulb with bezels.*

7.1 Results

To get a first general impression of the data, we calculated the tendency of each participant to produce counting errors for each of the three views. This was done by dividing the absolute difference of the participants' result to a ground truth number in the bezeled scenario with the sum of the absolute difference in the bezeled and non-bezeled scenario. This yields a number between 0 and 1, where 0 means all errors were made in the non-bezeled scenario, 1 means all errors occurred in the bezeled scenario and 0.5 stands for an equal amount of errors. The mean values for the different views (view 1: $M = 0.4794$, standard deviation $s = 0.3$; view 2: $M = 0.5254$, $s = 0.2812$; view 3: $M = 0.5404$, $s = 0.2525$) show no clear tendency towards bezeled or non-bezeled scenarios but a high standard deviation, indicating very individual results per participant.

Comparison of the time needed by the participants for adapting to the new scenario after turning back towards the screen showed a noticeable increase with bezeled scenarios. The time needed increased by a mean factor of $M = 2.31$ ($s = 1.2486$). One serious outlier, which saw an increase of a factor of 23.8, was left out of this calculation. A one-way RM-ANOVA showed a significant result for the first timed view ($F = (1, 6)\, 6.915$, $p = 0.039$), favoring the non-bezeled scenario over the bezeled. These findings support our results from the second pre-survey and the first timed task survey, where participants stated a subjective detrimental effect of bezels on stereo perception.

Interestingly, the times measured for the counting task seem to support the findings of Bi et al. [Bi et al. 2010] even for this stereoscopic 3D scene. For bezeled scenarios, the task completion time showed a tendency towards being faster (with mean factors of $M = 1.2839$ ($s = 0.295$) for view 1, $M = 1.0744$ ($s = 0.1826$) for view 2 and $M = 1.16$ ($s = 0.4322$) for view 3) as compared to the according non-bezeled scenarios. A two-way RM-ANOVA was used to show interactions of bezels and scenes, which demonstrated a significant effect for bezels over all scenes ($F\,(1, 11) = 5.845$, $p = 0.034$), favoring bezeled scenarios. A closer look at the individual scenario pairs revealed a single but very significant effect for bezels for view 2 ($F\,(1, 11) = 9.923$, $p = 0.009$), also in favour of the scenario with the bezels present. Overall, the increase in performance might be caused by the partitioning effect that the bezels introduce. People then can finish one "tile" at a time and have more prominent visual cues to orient themselves. This could also explain why we did not find the expected significant detrimental effects for the counting task in bezeled scenarios in the first timed survey. It is worth noting that all but two participants mentioned that they saw "double borders" and their eyes needed 1-2 seconds to adjust to a new "display tile" when bezels were present and their vision moved over such a border. This correlates to our findings in the previous study where subjects described parts of the image as "blurry". These findings also indicate increased stress on the visual system due to the constant refocussing of the eyes and therefore, although we did not look into this long-term issue, this effect might cause significantly more fatigue.

8 Conclusion

Our surveys revealed that all tested bezel sizes had a significant impact on overall depth perception—especially in the near field—compared to non-bezeled projections. However, bezels seem to help, or at least do not interfere, with counting tasks by partitioning the viewspace despite the less pleasing subjective appeal of the stereoscopic images. But there are options to at least minimize the detrimental perceptual effect for stereoscopic tiled display systems and therefore possible fatiguing and disturbing effects.

- **In general smaller bezels are better:** With the data gathered from our survey, we can verify that—as common sense suggests—smaller bezels perform better than bigger ones. Less obvious is the fact that, with gaps between displays of $1.2cm$ or smaller ($0.6cm$ bezel on each side), one can still achieve satisfying results. This is corroborated by the aforementioned findings of Sandin et al. [Sandin et al. 2005] where bezels seemed to have no effect on user perception.

- **Large tiles are better than smaller tiles:** Having large tiles allows users to get closer to the screen, covering a large part of their field of view without seeing bezels. This might be especially true for the high-resolution (foveal vision) part of our visual system. The surrounding displays then act as a contextual view, supporting the focused information, such as with the foveal display built by Baudisch et al. [Baudisch et al. 2002]. It is worth noting that within the boundaries of our survey, variations in bezel size and tile size have equal influence on depth perception.

- **Bezel colors matter to a lesser degree:** Statistically, our survey did not show a significant effect for bezel colors. But we found it surprising that, for depth perception, at least some users preferred the black bezels over gray bezels in a mostly gray VR scene. We believe that this finding requires further study as it contradicts other findings for mono-scenarios on 2D tiled displays, such as those of Ebert et al. [Ebert et al. 2010].

- **"Fishtank" VR as an option to avoid conflicting depth cues:** Since most perception problems are experienced for objects displayed as if they were in front of the screen, clipping objects with a negative parallax seems worth considering if bezel sizes with less than $1.2cm$ are not available. The whole scene then appears as if viewed through a window, and mismatching depth cues are not possible.

9 Future Work

One area for an in-depth analysis may be finding the largest acceptable bezel size. In our tests, we have shown that our minimum bezel size appears to suffice. For decisions on which specific hardware to install, it may be useful to analyze the probability-distribution function for the largest acceptable bezels. However, in order to achieve stable results, this would require a much larger sample size of participants.

Another interesting area for further research is the different behaviour of vertical and horizontal bezels as additionally indicated during our first timed survey in section 6. While horizontal bezels appear as a solid line, vertical bezels appear —due to the stereopsis— as a "double border" when focusing on an object displayed in front of the screen. As mentioned in the introduction, objects then seem to appear as transparent or borders seemingly disappear due to our brain piecing together and filling in information to reconstruct a "correct" image. Important parameters would be how much information loss can be compensated for and how the affected region is shaped in 3D so one could determine, for example, which parts of a volumetric dataset are not visible.

We intend to do further tests with varying gradients as blending zones. In particular, we are interested in why no significant improvement can be seen for display bezels, while we did see significant improvements for other related tasks (e.g., a blending zone around a 2D monoscopic text inlay in a 3D stereo scene).

Acknowledgements

The work presented in this paper has been co-financed by the Intel Visual Computing Institute and by the NIH/NCRR Center for Integrative Biomedical Computing, P41-RR12553-10 and by Award Number R01EB007688 from the National Institute Of Biomedical Imaging And Bioengineering. The content is under sole responsibility of the authors.

References

BALL, R., AND NORTH, C. 2005. An analysis of user behavior on high-resolution tiled displays. In *In Interact 2005 Tenth IFIP TC13 International Conference on Human-Computer Interaction*, Springer, 350–363.

BAUDISCH, P., GOOD, N., BELLOTTI, V., AND SCHRAEDLEY, P. 2002. Keeping things in context: a comparative evaluation of focus plus context screens, overviews, and zooming. In *Proceedings of the SIGCHI conference on Human factors in computing systems: Changing our world, changing ourselves*, ACM, New York, NY, USA, CHI '02, 259–266.

BAUDISCH, P., CUTRELL, E., HINCKLEY, K., AND GRUEN, R. 2004. Mouse ether: accelerating the acquisition of targets across multi-monitor displays. In *CHI '04 extended abstracts on Human factors in computing systems*, ACM, New York, NY, USA, CHI '04, 1379–1382.

BI, X., BAE, S.-H., AND BALAKRISHNAN, R. 2010. Effects of interior bezels of tiled-monitor large displays on visual search, tunnel steering, and target selection. In *Proceedings of the 28th international conference on Human factors in computing systems*, ACM, New York, NY, USA, CHI '10, 65–74.

BIRNHOLTZ, J. P., GROSSMAN, T., MAK, C., AND BALAKRISHNAN, R. 2007. An exploratory study of input configuration and group process in a negotiation task using a large display. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, ACM, New York, NY, USA, CHI '07, 91–100.

CRUZ-NEIRA, C., SANDIN, D. J., AND DEFANTI, T. A. 1993. 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*, ACM, New York, NY, USA, SIGGRAPH '93, 135–142.

DEMIRALP, C., JACKSON, C. D., KARELITZ, D. B., ZHANG, S., AND LAIDLAW, D. H. 2006. Cave and fishtank virtual-reality displays: A qualitative and quantitative comparison. *IEEE Transactions on Visualization and Computer Graphics 12* (May), 323–330.

EADIE, W., DRIJARD, D., JAMES, F., ROOS, M., AND SADOULET, B. 1983. *Statistical Methods in Experimental Physics*. Elsevier Science Ltd, June.

EBERT, A., THELEN, S., OLECH, P.-S., MEYER, J., AND HAGEN, H. 2010. Tiled++: An enhanced tiled hi-res display wall. *IEEE Transactions on Visualization and Computer Graphics 16*, 120–132.

EYEVIS, 2010. Dlp®cube. Website, http://www.eyevis.de/index.php?clang=1.

GRUDIN, J. 2001. Partitioning digital worlds: focal and peripheral awareness in multiple monitor use. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, ACM, New York, NY, USA, CHI '01, 458–465.

HIPERWALL, 2010. Hiperwall - high resolution, large format, display systems. Website, http://hiperwall.com/.

LANDAUER, J., BLACH, R., BUES, M., RÖSCH, A., AND SIMON, A. 1997. Toward next generation virtual reality systems. *Multimedia Computing and Systems, International Conference on 0*, 581.

MACKINLAY, J. D., AND HEER, J. 2004. Wideband displays: mitigating multiple monitor seams. In *CHI '04 extended abstracts on Human factors in computing systems*, ACM, New York, NY, USA, CHI '04, 1521–1524.

SCIENTIFIC COMPUTING AND IMAGING INSTITUTE UTAH, 2012. Imagevis3d. Website, http://www.sci.utah.edu/cibc/software/41-imagevis3d.html.

OLED-DISPLAY.NET, 2010. LG display unveils LCD panel featuring the world's narrowest bezel. Website, http://www.oled-display.net/lg-display-unveils-lcd-panel-featuring-the-world

PROJECTIONDESIGN, 2010. F20 SX+ product webpage. Website, http://www.projectiondesign.com/Default.asp?CatID=1522.

ROBERTSON, G., CZERWINSKI, M., BAUDISCH, P., MEYERS, B., ROBBINS, D., SMITH, G., AND TAN, D. 2005. The large-display user experience. *IEEE Comput. Graph. Appl. 25* (July), 44–51.

SANDIN, D. J., MARGOLIS, T., GE, J., GIRADO, J., PETERKA, T., AND DEFANTI, T. A. 2005. The varrier™autostereoscopic virtual reality display. In *ACM SIGGRAPH 2005 Papers*, ACM, New York, NY, USA, SIGGRAPH '05, 894–903.

TAN, D. S., AND CZERWINSKI, M. 2003. Effects of visual separation and physical discontinuities when distributing information across multiple displays. In *In Proceedings of Interact 2003*, 252–255.

TAN, D. S., CZERWINSKI, M., AND ROBERTSON, G. 2003. Women go with the (optical) flow. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, ACM, New York, NY, USA, CHI '03, 209–215.