Gaze-Contingent Depth of Field in Realistic Scenes: The User Experience

Margarita Vinnikov
York University
mvinni@cse.yorku.ca

Robert S. Allison
York University
allison@cse.yorku.ca

Abstract

Computer-generated objects presented on a display typically have the same focal distance regardless of the monocular and binocular depth cues used to portray a 3D scene. This is because they are presented on a flat screen display that has a fixed physical location. In a stereoscopic 3D display, accommodation (focus) of the eyes should always be at the distance of the screen for clear vision regardless of the depth portrayed; this fixed accommodation conflicts with vergence eye movements that the user must make to fuse stimuli located off the screen. This is known as accommodation-vergence conflict and is detrimental for user experience of stereoscopic virtual environments (VE), as it can cause visual discomfort and diplopia during use of a stereoscopic display. It is believed that, by artificially simulating focal blur and natural accommodation, it is possible to compensate for the vergence-accommodation conflict and alleviate these symptoms. We hypothesized that it is possible to compensate for conflict with a fixed accommodation cue by adding simulated focal blur according to instantaneous fixation.

We examined gaze-contingent depth of field (DOF) when used in stereoscopic and non-stereoscopic 3D displays. We asked our participants to compare different conditions in terms of depth perception, image quality and viewing comfort. As expected, we found that monocular DOF gave a stronger impression of depth than no depth of field, stereoscopic cues were stronger than any kind of monocular cues, but adding depth of field to stereo displays did not enhance depth impressions. The opposite was true for image comfort. People thought that DOF impaired image quality in monocular viewing. We also observed that comfort was affected by DOF and display mode in similar fashion as image quality. However, the magnitude of the effects of DOF simulation on image quality depended on whether people associated image quality with depth or not. These results suggest that studies evaluating DOF effectiveness need to consider the type of task, type of image and questions asked.

CR Categories: **H.5.1** [Information Interfaces and Presentation]: Multimedia Information Systems - Artificial, augmented, and virtual realities; **H.5.2** [Information Interfaces and Presentation]: User Interfaces - Graphical user interfaces (GUI); **I.3.7** [Computer Graphics] Three-Dimensional Graphics and Realism - Virtual reality.

Keywords: Gaze-contingent display, visual simulations, depth-of-field.

1 Introduction

Despite many attempts to simulate realistic depth in computer displays, there remains a sense that the simulated space does not match the real or modeled scene. In many cases, the simulated space appears flattened relative to the original space. One reason for this is that while the display supports selected depth cues, it omits others. The omitted cues typically indicate that the display is flat and can significantly alter depth perception [Watt et al. 2005; Thompson et al. 2004]. A common deficiency is the inability to simulate the eye's focus. In the real world, the clarity of the retinal image of an object depends on its relation to fixation in the scene. The retinal image is the sharpest for objects at the focal distance, and is increasingly blurred as the depth of the object from the focal distance increases. In contrast, in typical computer-generated displays everything is physically presented on the screen and thus at a fixed focal distance. There is no focal blur associated with objects located away from fixation. Image blur can be introduced to simulate the user's depth of field (DOF). DOF is the depth range, centred at the focus point, where the eyes perceive the image as sharp. In other words, DOF is a measure of the range of depth over which focal blur is imperceptible; outside the DOF objects have noticeable blur [Gullstrand 1910] and this perceptible blur can be simulated graphically. However, simulated DOF will not change naturally with fixation changes and will only be correct if the user looks at the modeled fixation point.

The mechanism that controls the focus state of the eye, via the ciliary muscles adjusting the shape of the lens, is known as accommodation. Accommodation provides a physiological cue to distance because the accommodation required for clear vision is directly related to the distance of the fixation point. As discussed above, objects presented on a normal display have the same focal distance regardless of portrayed depth and thus accommodation should not change when fixating different 3D objects. In a stereoscopic 3D display, this conflict is increased because the eyes need to make vergence eye movements to fuse disparate stimuli located off the screen. As vergence and accommodation are normally physiologically linked and tightly coupled [Muller 1846], this discrepancy creates a conflict known as the accommodation-vergence conflict. Accommodation-vergence conflict is usually associated with negative side effects such an eyestrain [Luebke et al. 2003; Mon-Williams and Wann 1998; Wann et al. 1995]. It is believed that, by artificially simulating focal blur and natural accommodation, it is possible to alleviate symptoms of simulator sickness and diplopia resulting from accommodation-convergence conflict [Drascic and Milgram 1996; Villarruel 2006]. While stimulating the accommodation cue requires special displays [Akeley et al. 2004] current technology can easily supports simulation of DOF.

In computer graphics, there have been attempts to use simulated DOF for realistic rendering. One of the biggest challenges in such an approach is determining which object in the scene should be in focus since the user is free to look about the image. One possible solution is to estimate the user's gaze and centre the DOF on the object that the user is fixating (the point of regard). For example, Brooker and Sharkey [2001] proposed a system that generated a synthetic DOF linked to the user's point of regard. Their experimental results did not show a significant performance improvement. However, it is possible that the system did not measurably enhance depth perception due to technology or methodology limitations including system latency, low precision of the eye tracking device and inappropriate task (the task was depth estimation of textured surfaces that lacked any additional depth cues). Although a combination of slanted textured surfaces and DOF can provide strong depth, the constructed maze in this experiment did not have any slanted planes; instead all planes were parallel to the screen. Another source of information would have been occlusion of textures but it appeared that the textures were identical, causing difficulty in occlusion detection [Wanger et al. 1992; Gibson 1950]. Hence, it is possible that the hypothesized benefits could be demonstrated by improving the system and the experimental task. Van der Linde [2004] also developed a gaze-contingent system that simulated stereoscopic DOF combined with a foveation technique. The foveation technique restricted the size of the region that needed to be rendered with high detail such that it was bound by the size of the fovea and the depth of focus. While the system did produce a DOF effect, the primary purpose of the system was to reduce rendering complexity. However, the technique used only post processing algorithms and did not reduce the rendering volumes. In addition, their work was not validated either theoretically or experimentally and perceptual impact on depth perception was not considered.

Despite the conceptual and practical issues involved in rendering DOF without knowing the fixation point, there have been a few studies that looked at (non-gaze-contingent) DOF and its effects [Barsky and Kosloff 2008; Hillaire et al. 2008b; Hillaire et al. 2008a; Krivanek et al. 2003; Mulder and Van Liere 2000]. In their paper, Hillaire et al. [2008] looked at user performance, when the user played a first-person-shooter game with simulated static DOF and without (infinite DOF). They reported that DOF blur simulation degraded participants' overall performance in the game (the number of dead enemies, the number of times the player died, and shooting precision). Interestingly, they also reported that DOF improved the subjective impression of fun, presence, and virtual environment (VE) realism for some users. They explained their results by the fact that blur hides many important high-frequency features and hence hides imperfections of the game while reducing precision at the same time. However, their experiment was not gaze-contingent and people could not compensate for the lack of precision by shifting their gaze to objects of interest. In a different study, Hillaire et al. [2008] looked at gaze-contingent DOF in context of navigation and user experience. Once again, they reported that the users preferred gaze-contingent DOF on the basis of the positive user experience metrics that also included impression of depth. Unfortunately, they did not report any psychophysical results. Leroy et al. [2012] looked at DOF and user comfort in context of a pointing task; in particular, they looked at eye strain reduction over time. They reported that their system was successful in reducing eye-strain and that errors in pointing performance were unaffected by the simulated blur. However, the implemented system was not gaze-contingent. Furthermore, they used a stereoscopic display, which might explain why they found no effect

of DOF on pointing performance, while Hillaire et al. [2008], who did not use stereo displays, reported significant performance degradation. As stereo is a very strong depth cue, it could have provided enough information to achieve a high degree of pointing precision for both trials with and without DOF.

In summary, while some studies have attempted to simulate instantaneous focal blur via DOF simulation, there has been very little research done to systematically evaluate the user experience of real time DOF simulation. In our study, we tried to evaluate the benefits of DOF as a depth cue and determine the crucial perceptual factors that might impact the strength of DOF as a cue. Furthermore, for focal blur simulation to mimic human experience it must reflect current fixation and thus we implemented and evaluated gaze-contingent DOF that simulated DOF at different focal distances as provided from tracking the user's point of regard. Similar to Leroy et al. [2012], we are interested in determining whether simulation of DOF can improve a user's perceptual performance. However, our primary interest is in exploring whether and how gaze-contingent DOF improves depth perception. As a first step, we wanted to evaluate whether DOF can help to improve the general subjective sense of depth perception. Hence, we assessed people's impressions of depth perception, general comfort and aesthetic preferences. In addition, most of the previous systems operated in monocular or Cyclopean-eye modes; but, binocular vision can provide important additional cues for depth. Therefore, we considered it important to explore DOF with and without stereoscopic depth cues. We discuss the results obtained in terms of the depth cue support required in virtual-reality simulations and games as well as in terms of issues that need to be solved when introducing DOF simulation into monocular and binocular displays.

2 Gaze-Contingent System

For the purpose of exploring DOF in the context of gaze-contingent display, we built a gaze-contingent experimental platform that allowed us to setup different psychophysical experiments with complex virtual scenes. For our experimental environment platform we chose Unity3D [2013]. This is a multi-platform game development engine that allows rapid development of virtual-reality environments. This platform allows for easy customization and integration of components such as physics engines and collision systems. It also supports scripting languages such as JavaScript and C#, while ensuring fast compilation and performance with convenient interfaces to .NET. Most importantly, by integrating DirectX and OpenGL, Unity3D supports advanced lighting and shadows activation and configuration. As a result, our system is very modular and script driven but still supports high-performance graphics. We can easily adjust different aspects of the experiment without affecting other parts of the system. Our system architecture is presented in Figure 1.

2.1 System Architecture

Unity3D uses an objects paradigm to tie different components together. Each object is associated with a scene and can have different properties. For example, there will always be at least one object that will describe the camera components. Other objects can be models of simulated physical objects that will be viewed by the user; these objects will have different physical properties. In our system, most of the scripts are assigned to a MainCamera object Figure 1. MainCamera is responsible for reading experiment setup scripts and building an appropriate

scene specified by the experimenter. It is also responsible for connecting to the eye-tracking system and calibrating it. In addition, MainCamera includes an experimental script that sequences experimental trial execution, monitors time intervals allocated for each experimental stage and records explicit (direct responses from concrete questions) and implicit data (behavior data such as gaze patterns) collected from an observer. Finally, the MainCamera object is also responsible for presenting the properly simulated image of the virtual scene.
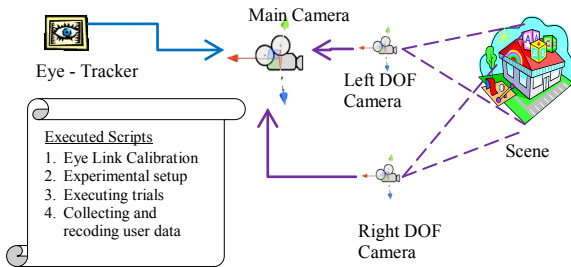

Figure 1 *Platform setup*

2.1 Binocular DOF Simulation

In order to simulate DOF, we incorporated gaze-contingent image processing. Unity3D supports a DOF post-processing effect that simulates the properties of a camera lens. The effect is achieved by applying Cg shaders [Fernando and Kilgard 2003] onto the image that is acquired by a virtual camera, in other words *the RenderedTexture.* The general idea relies on determining the circle of confusion (CoC) that determines the extent of blur due to defocus. To achieve this we alter the acuity within the CoC, in other words, we blur the image pixels based on the depth buffer associated with the observed scene. In order to generate different levels of blur, the DOF script in Unity generates several images of the same scene at different levels of detail, it then interpolates between these levels to achieve smooth blur based on the depth coordinates. Because it is an additive image-processing operation there might be a problem of over exposure and the scene becoming brighter or darker than expected. To avoid this problem, contrast scaling based on the aperture size was used. To make DOF gaze-contingent, we change the focal distance of our camera in real-time. We determine the focal distance, by determining the point of regard that is the point in 3D, where the user's gaze intersects with a virtual object (Figure 2).
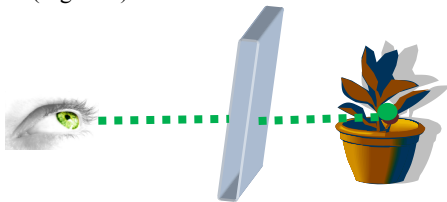

Figure 2 *Focal distance*

We extended our system to support stereoscopic rendering. To achieve this, we had to ensure that there was a proper DOF simulation applied for each eye. Hence, we actually had two cameras in our scene that simulated the DOF for each eye as well as the binocular parallax arising from the difference in vantage point for each eye. In addition to DOF the system can simulate different visual defects or characteristics that include visual field blur simulation and visual distortions. More details about defect rendering techniques are described in Vinnikov et al. [2008].

3 Experimental Methods (DOF)

3.1 General Methodology

We hypothesized that the scene with stereo and DOF cues would provide more compelling depth perception than scenes rendered with stereo or DOF alone, which in turn should be more compelling than simulations without stereoscopic or DOF rendering. However, image quality and comfort could be impaired by simulating these depth cues. To test our hypotheses, we performed a mixed between-subjects/within-subjects experiment. The between-subjects variable was the judgments performed by the observer and the within-subjects variable was the type of depth cues provided. During each trial, the observer was presented with two images of the same real life scene rendered with different depth cues. Depending on the group, a participant was randomly assigned, she had to either answer questions about depth and image quality or about comfort and image quality.

3.2 Stimuli

The virtual scene consisted of several 3D models of everyday objects, such as books. In all cases, the objects were placed on a horizontal plane so that the participant could interpret the objects as lying on a table-top. The scenes were designed to resemble those encountered in everyday interaction with objects placed on common surfaces within our living space. Figure 4 shows the three scenes presented in the experiment. The scene with the flowers was chosen because it also had many vertical elements; conversely the scene with stationery items had many details distributed across the horizontal plane, such as the keyboard and cigars. We also wanted to see the effect of blur on textures and hence for the last scene, we chose to show pizzas that had a lot of textured fine detail. Fine detail was also present on the pages of the book with fine print for the second scene. Each of these scenes was presented with or without stereo and with or without DOF.

3.2.1 Procedure

The overall experimental flow is outlined in Figure 3. Each trial consisted of two images of a given scene, presented sequentially with different depth cues –monocular cues, monocular and combined with DOF, stereo without DOF, and stereo and DOF cues.
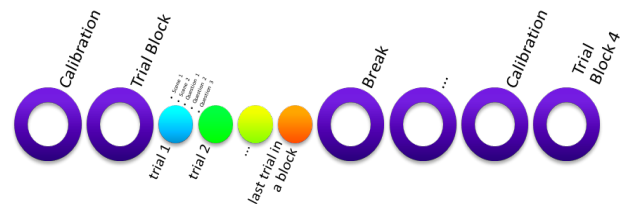

Figure 3 *Experimental flow*

For the first experiment, by permutation there were 12 (4*3) pairings of the different cue conditions, 3 different scenes and 4 repeats for each, providing 144 trials in total per participant. Each scene was presented to user for 8 seconds (4 per condition). We divided our experiment into 4 blocks to ensure adequate calibration at all times. Therefore, a calibration procedure was performed prior to each block.
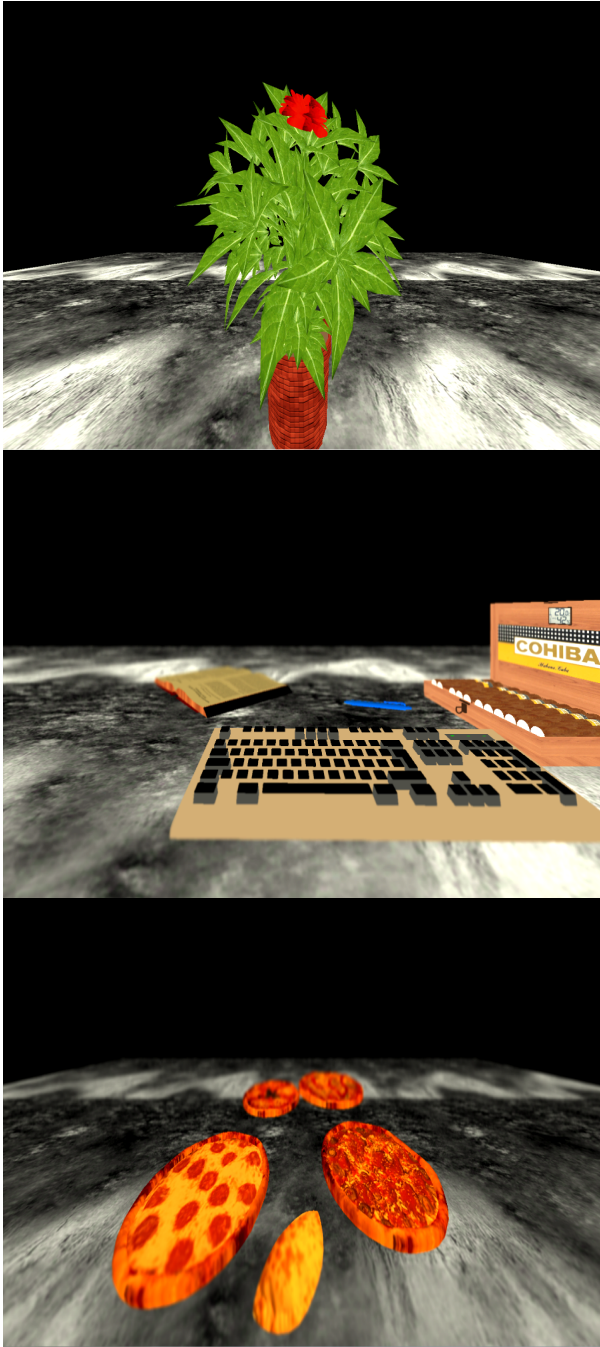
Figure 4 *Scenes: (a) Flower (b) Stationery (c) Pizza*

The calibration involved sequentially fixating 9 points displayed on the screen in pseudo-random order. Each calibration step was followed by a validation step where another set of 9 points randomly appeared on the screen in sequence. The distances between expected gaze location and the estimated location acquired during this step, indicated whether calibration was successful or not.

In the case that the tracking error was above the required threshold, the system was recalibrated after necessary adjustments to the tracker were made.

Participants answered a set of 2 questions following each trial. The first group of participants was asked the following questions:

1.   Which scene had the better image quality?
2.   Which scene had the most compelling depth?

The second group was asked the following questions:

1.   Which scene was most comfortable to view?
2.   Which scene had the better image quality?

Participants were required to choose one or the other interval even if there was little apparent difference (a forced-choice procedure).

Participants were instructed to inspect the scenes thoroughly so they could compare the image pairs and make relative judgments. To ensure that the subjects were paying attention to the task and actively visually exploring the images, the experimenter monitored participants' eye movements online during the trial.
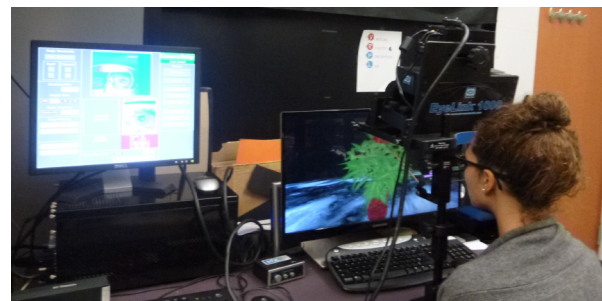

Figure 5 *Experimental setup*

3.1   Apparatus:

Visual displays were generated on a desktop computer with AMD FirePro W9000 FireGL, Windows 7 Enterprise, Intel ® Core™ CPU, 3.50 GHz, 3.50 GB Ram. The stimulus was presented on a Samsung 950 Series 3D monitor (27", with a pixel resolution of 1920 H x 1080 V and a refresh rate of 120Hz). In both experiments, the screen was viewed binocularly at a distance of 0.35 m, and the stimulus subtended a horizontal visual angle of 80.6° (Figure 5). A real-time gaze-contingent system was built by incorporating an EyeLink 1000 [SR Research] eye-tracker. Because of the limitation of the near distance tracking, we had to use the tower mount setup. A chinrest was used to stabilize the head and to maintain viewing distance with the midpoint between the right or left eyes, centered on and at the level of the center of the screen. The experiment was conducted in a darkened room.

3.2   Participants

The first group consisted of 9 university students (6 females and 3 males, ranging in age from 22 to 29, average age 24) and the second group consisted of 9 university students (6 females and 3 male, ranging in age from 19 to 24, average age 21). All participants had uncorrected distance visual acuity of 20/30 or better and could see clearly at the viewing distance without glasses.

Prior to the experiment, we screened participants for stereopsis, which is their ability to binocularly distinguish distance between objects. For this purpose we used RANDOT stereo-test [Chicago, USA]. All participants were able to perform at 50 seconds of arc or better at distance of 40.6 cm. Written informed consent was obtained from all participants in accordance with a protocol approved by the York University Ethics Board.

4   Results

We looked at the data in terms of the questions asked during each session and considered the effects of pairings of questions.

4.1   Depth Perception

In the first group, we asked people to compare different depth cue conditions in terms of depth perception and image quality. A Cochran Q test revealed that there was a significant difference in depth rendering preference between conditions ($\chi^2(5) = 35.31$, $p < .001$). The mean percentage of trials that one condition in a pair was preferred over the other is shown in Figure 6. The figure shows these biases, averaged across observers (the error bars show standard error of the mean), as a function of the conditions paired. As expected, stereoscopic displays gave typically stronger depth impressions than non-stereoscopic conditions. We wanted the scenes to be as realistic as possible and hence used rich depth cue environments. Thus it is likely that some participants did not notice singnificant differences in depth rendering. Also participants sat very close to the screen and hence with stereo they could experience some discomfort and fusion issues that they did not experience when viewing non-stereoscopic display. DOF provided a stronger sense of depth when added to non-stereoscopic conditions, but this was not the case when it was combined with stereo.
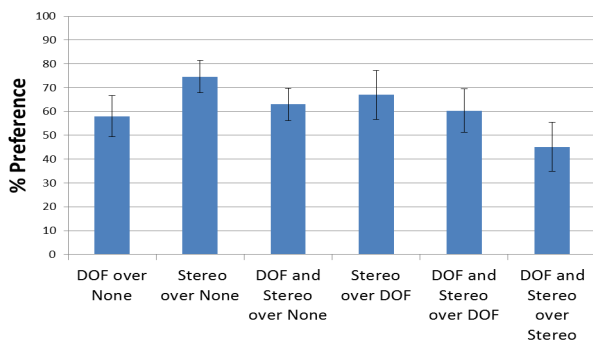

Figure 6 *Depth perception*

Figure 7 shows the depth perception preferences for different scenes. One can see that people had a preference for DOF in non-stereoscopic cases for the scene with flowers and it correlated with personal accounts from participants. They reported they liked the DOF effect for variability in features at different distances in depth. Overall, on 25% of the trials people prefered monocular presentation over stereo, which can be attributed particularly to the response to the stationary scene that had, very strong monocular cues such as perspective, size difference, texture gradient and common sizes.

4.2   Viewing Comfort

The second group of participants was asked to compare image pairs in terms of viewing comfort and image quality. A Cochran
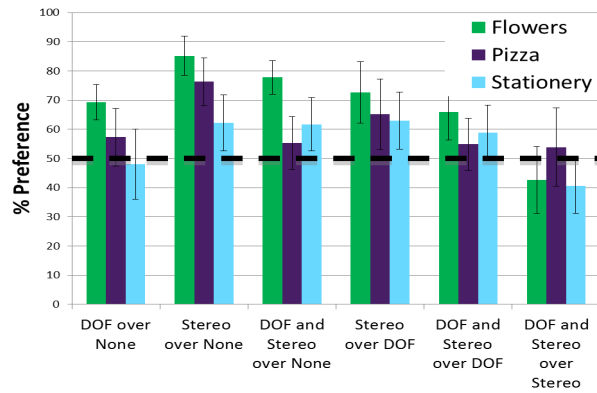

Figure 7 *Depth perception for different scenes*

Q test revealed that the there was a significant difference in comfort between conditions ($\chi^2(5) = 555.162$, $p < .001$). The means and standard error for the comfort comparisons are shown in Figure 8. The pattern of biases for viewing comfort preference was opposite to depth preferences. In other words, DOF was typically considered the less comfortable condition to view. This was true even when combined with stereoscopic rendering and stereoscopic rendering was judged more comfortable than DOF and stereo. It is likely that the DOF simulation determined these ratings since comparisons of DOF and stereo versus DOF gave no indication of difference in comfort (preferences were at 50%).

Viewing comfort, for different scenes is presented in Figure 9. Interestingly, there was no difference between scenes in terms of viewing comfort. This implies that viewing comfort does not depend on features presented in the scene but rather on depth cues that were presented.
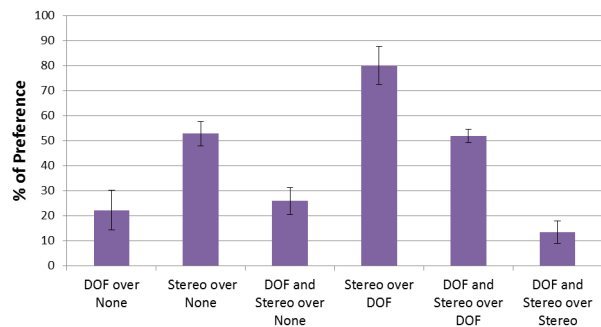

Figure 8 *Viewing comfort*

4.3   Image Quality

Image quality preference data are presented in Figure 10. Overall, we observed that the pattern of image quality ratings closely resembled the pattern of comfort ratings. This suggests that image quality and image blur were likely important drivers of comfort judgments in our displays. As hypothesized, conditions with image blur added to simulate DOF were judged as having lower image quality. Interestingly, stereoscopic rendering seemed to have little influence on image quality since the comparisons between stereo and none versus stereo and DOF conditions and between stereo and DOF versus DOF conditions showed no measurable bias in preferences: preferences were near 50%, indicating subjects were equally likely to prefer either of the two conditions.
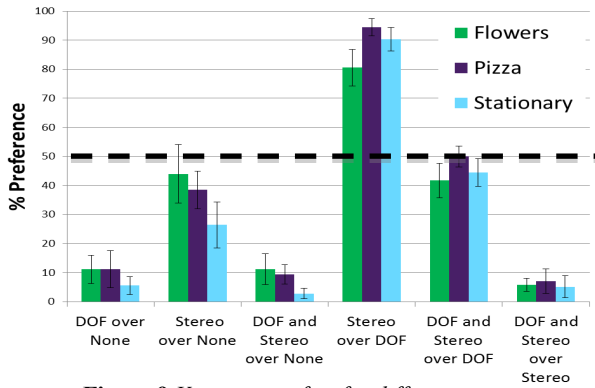
Figure 9 *Viewing comfort for different scenes*

One can observe that preference biases for each condition pair were very similar in the two experimental groups (i.e., when image quality was queried together with depth preference compared to when queried with comfort preference). However, when observers judged quality and comfort, they showed stronger quality biases for stereo over DOF and for stereo over stereo and DOF than did observers who judged quality along with depth. This implies that people associate image quality with many factors and hence can be biased in their responses. Drawing attention to the enhanced depth impressions seemed to increase the perceived quality of the DOF displays. Nevertheless, DOF rendering still resulted in reduced image quality even for this group, albeit to a lesser extent.
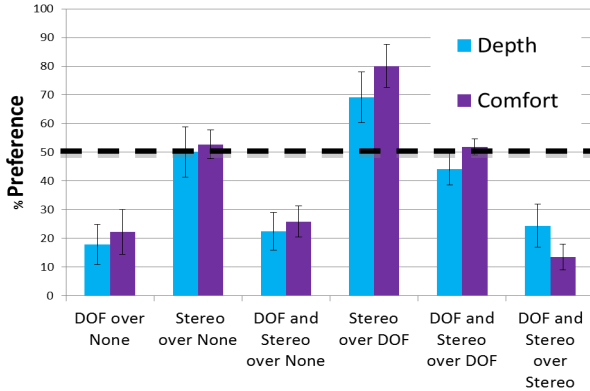

Figure 10 *Image quality preferences for each experimental group (Depth Question or Comfort Question)*

5   Discussion

Our results indicate that DOF contributes to subjective impressions of depth perception when added to other depth cues. However, it is known that focal blur is not a very strong depth cue as it does not provide direction and hence can be overpowered by stereopsis [Grossmann 1987; Pentland 1987; Marshall et al. 1996; Nguyen et al. 2005]. Nevertheless, our data suggest that it does add qualitatively to the depth provided by non-stereoscopic cues, which should enhance the overall 3D experience.

There is no well-defined definition for visual discomfort. An unpleasant feeling, often occuring instantaneously, is typically associated with viewing a stimulus that produces *visual stress*. As with other physiological stressors, visual stressors are a challenge to the visual system that can produce *visual strain* or *eyestrain*. The effects can build over time and produce *visual fatigue* often involving visual discomfort both when the stimulus

is present and even when it is removed. There are many papers that address and research the impact of visual discomfort using a variety of different techniques and measures. The short-term and long-term effects of visual stress on visual discomfort and visual fatigue are not always distinguished. Many studies present images to the subjects for several seconds (0.5 to 15 s is typical) and have them judge or compare the images in terms of viewing comfort [Blum et al., 2010; Kooi & Toet, 2004; O'Hare & Hibbard, 2013]. To study visual fatigue extended exposures are needed to see how the visual system responds to the potential visual stress [Hoffman et al., 2008; Mon-Williams & Wann, 1998; Ukai & Howarth, 2008].

In this study, we used the former approach to assess the effects of gaze-contingent depth of field on visual discomfort in images and our data do not speak directly to the effects on visual fatigue. It is possible that the short-exposure and long-exposure effects differ because image blur has been associated with immediate visual discomfort [O'Hare & Hibbard, 2013] while appropriate optical blur in a display reduces visual fatigue by reducing acommodative-vergence conflict [Hoffman et al., 2008]. We are conducting experiments with extended viewing to assess the effects of dynamic DOF cues on visual fatigue.

In terms of short-term viewing comfort, we found that both viewing comfort and image quality were reduced when DOF was added as a depth cue. This can be explained by the fact that people notice the blurriness that results from DOF simulation. Although we did our best to match the blur to a realistic DOF we could not account for individual optic differences. Furthermore, the DOF of simulation modeled focal blur that results from viewing through a finite sized pupil but not the effects of aberrations and micro-fluctuations of accommodation that are believed to be important to disambiguate depth from accommodative blur [Nguyen et al. 2005]. Theoretically, if simulated depth of field matched a natural depth of field, the simulation should be indistinguishable from a real scene (in terms of image sharpness). Dispite our best effort to minimize latency of the system, in some cases, the change in focus was still noticed by participants and it unfortunately reduced the comfort [Kooi and Toet 2004; O'Hare and Hibbard 2013]. Kooi and Toet [2004] showed noticeable imperfections in stereoscopic image can be described as very uncomfortable for an observer. It is possible that these technical limitations on spatial and temporal precision reduced image quality from expectations of natural viewing.

However, it is also possible that users actually preferred the image quality of the non-realistic scene with infinite DOF over their natural viewing. This is a reasonable supposition, since the image quality of the infinite DOF is better than expected from natural viewing in terms of image sharpness and image information. Consequently, high-frequency content in the scene is preserved everywhere in the image. It is possible that users judged image quality as better in these 'hyper realistic' images than in more realistic but blurry DOF simulations.

We also observed differences in the utility of the depth provided by the DOF simulation between the scenes presented to the users. We expected this pattern of results because the scenes were designed to contain differences in variation in depth across different features and across space. The scene with a flower had large differences in depth between neighboring features in the image; hence, by looking from one leaf to another, users constantly shifted their simulated focal distance. This resulted in the largest effect of DOF simulation on the image. On the other

hand, the scene included many small features varying dramatically in depth so that accurate DOF simulation required very precise calibration. Thus, people were more sensitive to any inconsistencies with simulation, particularly when combined with stereo. The other two scenes consisted of objects located on a horizontal plane and thus depth varied smoothly and monotonically in the image. As a result, the effect of DOF was weaker than in the scene with the flower. Yet in case of the pizza scene, when stereo combined with DOF was compared with either non-stereoscopic DOF or stereo alone there was no significant bias in the judgments, indicating that people did not have a preference. This equivalence of depth cue effects can be attributed to the fact that such richly textured scenes will have a more noticeable blur gradient and hence a better balance between stereo and DOF. Lastly, we expected that the stationary scene would also benefit from adding DOF as a depth cue since it included many fine details. However, this was not the case. This can be explained by the several contributing factors, such as these high detail regions were very sensitive to any kind of calibration error and they were not the most viewed regions. In conclusion, we think that the content of the rendered scene plays a very important role in determining if DOF should be added to the virtual reality system or a game, based on what features it would include, where these features are located, and on the accuracy of the calibration.

6   Future Work and Conclusions

In summary, we built a system that uses a gaze-contingent DOF to simulate the natural change in image focus resulting from accommodation, that we normal experience when looking about a real scene. While such a system does not stimulate accommodation it does produce a natural relationship between retinal image blur and binocular disparity. Therefore, we hypothesized that it might enhance depth perception and alleviate the accommodation-convergence conflict. We demonstrated that simulated gaze-contingent DOF could indeed enhance depth perception when combined with other depth cues. However, it does not seem to enhance viewing comfort and has strong impact on perception of image quality. This implies that DOF is not a straight-forward solution for accommodation-convergence conflict. The effects on viewing comfort may be short-term discomfort effects related to the degradation in perceived image quality due to perceptible blur. Ongoing experiments are assessing whether gaze-contingent DOF can help to alleviate the visual fatigue associated with extended viewing of stereoscopic displays.

In other words, our results suggest that studies evaluating DOF effectiveness need to consider the type of task and questions asked. Specifically our study supports the notion that DOF should be added to virtual-reality simulations and video games both for monocular and binocular displays if precise gaze-contingent display can be achieved and strong depth impressions are desired. However, special care is required from the developers to make sure that the image quality and comfort preferences also improve for displays with gaze-contingent DOF. To achieve this, we think that individual tuning of the display is required to systematically evaluate the tradeoff between the amount of simulated blur and image quality degradation for DOF simulations at different distances. The accuracy and latency of the gaze-contingent DOF display should always be minimized as much as possible. Other additional factors that help to bridge the gap between comfort of viewing and DOF preferences should also be considered.

Our next step in the gaze-contingent DOF system development will be to work on quantitative DOF evaluation methodologies to complement these experiments on subjective user experiences. We also plan to examine gaze patterns and behaviors associated with DOF simulations. In our current study, we used a model that describes an average pupil size and utilized a generalized DOF model. We verified the acceptability of these generic settings with pilot studies. In our future study, we would like to customize DOF for each participant based on their individual pupil size and consequently tailored for personal spherical and chromatic aberration. In addition, we would like to investigate the potential of more extreme DOF effects to enhance and perhaps exaggerate depth. These could potentially produce stronger depth effects but, if too extreme, could degrade image quality and produce artifacts. Such artifacts include the puppet-theater effect [Yamanoue et al. 2006] (where extensive blur makes objects appear unusually small) or less commonly the gigantism effect, when too little blur cause objects to look abnormally large. We plan to explore the potential and limits of exaggerated DOF in the future. Finally, we would like to examine other age groups. Specifically, because accommodative capacity degrades over the lifespan, typically reducing an accommodative range of 10-12 diopters at 25 years of age to only 0.5-1 diopters at 60 years old, we would plan to extend our work for the aging population. Particularly, we would like to explore whether gaze-contingent DOF can be beneficial for people who have actually lost their capacity for accommodation later in life.

7   References

AKELEY, K., WATT, S.J., GIRSHICK, A.R., AND BANKS, M.S. 2004. A Stereo Display Prototype with Multiple Focal Distances. *Proceedings of ACM SIGGRAPH 2004 23*, 3, 804–813.

BARSKY, B.A. AND KOSLOFF, T.J. 2008. Algorithms for Rendering Depth of Field Effects in Computer Graphics. *World Scientific and Engineering Academy and Society (WSEAS)*, 999–1010.

BLUM, T., BIAS, WIECZOREK, M., AICHERT, A., TIBREWAL, R., AND NAVAB, N. 2010. The effect of out-of-focus blur on visual discomfort when using stereo displays. *2010 IEEE International Symposium on Mixed and Augmented Reality*, 13–17.

BROOKER, J.P. AND SHARKEY, P.M. 2001. Operator performance evaluation of controlled depth of field in a stereographically displayed virtual environment. *Stereoscopic Displays and Virtual Reality Systems VIII 4297*, 1, 408–417.

DRASCIC, D. AND MILGRAM, P. 1996. Perceptual issues in augmented reality. *Stereoscopic Displays and Virtual Reality Systems III 2653*, 123–134.

FERNANDO, R. AND KILGARD, M. 2003. *The Cg Tutorial: The Definitive Guide to Programmable Real-Time Graphics*. Addison-Wesley.

GIBSON, J.J. 1950. *The perception of the visual world*. Houghton Mifflin, Boston, MA, USA.

GROSSMANN, P. 1987. Depth from focus. *Pattern Recognition Letters 5*, i1, 63–69.

GULLSTRAND, A. 1910. The optical system of the eye. *Helmholtz's Treatise on Physiological Optics 1*, 350–358. http://poseidon.sunyopt.edu/helmholtz/OCRVolume3.pdf.

HILLAIRE, S., LECUYER, A., COZOT, R., AND CASIEZ, G. 2008a. Using an Eye-Tracking System to Improve Camera Motions and Depth-of-Field Blur Effects in Virtual Environments. *Virtual Reality Conference, VR '08. IEEE*, 47–50.

HILLAIRE, S., LÉCUYER, A., COZOT, R., AND CASIEZ, G. 2008b. Depth-of-field blur effects for first-person navigation in virtual environments. *IEEE computer graphics and applications 28*, 6, 47–55.

HOFFMAN, D.M., GIRSHICK, A.R., BANKS, M.S., AND AKELEY, K. 2008. Vergence–accommodation conflicts hinder visual performance and cause visual fatigue. *Journal of Vision 8*, 3, 1–30.

KOOI, F. AND TOET, A. 2004. Visual comfort of binocular and 3D displays. *Displays 25*, 2-3, 99–108.

KRIVANEK, J., ZARA, J., AND BOUATOUCH, K. 2003. Fast depth of field rendering with surface splatting. *Computer Graphics International*, 196–201.

LEROY, L., FUCHS, P., AND MOREAU, G. 2012. Real-time adaptive blur for reducing eye strain in stereoscopic displays. *ACM Transactions on Applied Perception 9*, 2, 1–18.

VAN DER LINDE, I. 2004. Multiresolution image compression using image foveation and simulated depth of field for stereoscopic displays. *Proc. SPIE 5291, Stereoscopic Displays and Virtual Reality Systems XI 5291*, 71–80.

LUEBKE, D., REDDY, M., COHEN, J., VARSHNEY, A., AND WATSON,, B. 2003. *Level of Detail for 3D Graphics*. Morgan Kaufmann.

MARSHALL, J.A., BURBECK, C.A., ARIELY, D., ROLLAND, J.P., AND MARTIN, K.E. 1996. Occlusion edge blur- A cue to relative visual depth.pdf. *Jurnal of the Optical Society of America A*, 13, 681–688.

MON-WILLIAMS, M. AND WANN, J.J.P. 1998. Binocular Virtual Reality Displays: When Problems Do and Don't Occur. *Human Factors: The Journal of the Human Factors and Ergonomics Society 40*, 1, 42–49.

MULDER, J.D. AND VAN LIERE, R. 2000. Fast perception-based depth of field rendering. *VRST '00 Proceedings of the ACM symposium on Virtual reality software and technology*, 129–133.

MULLER, J. 1846. *Elements of Physiology*. Lea and Blanchard, Philadelphia.

NGUYEN, V.A., HOWARD, I.P., AND ALLISON, R.S. 2005. Detection of the depth order of defocused images. *Vision Research 45*, 8, 1003–1011.

O'HARE, L. AND HIBBARD, P. 2013. Visual discomfort and blur. *Journal of Vision 13*, 5, 1–12.

PENTLAND, A.P. 1987. A new sense for depth of field. *Pattern Analysis and Machine Intelligence, IEEE Transactions on* 4, 523–531.

SR RESEARCH. http://www.sr-research.com/solutions_gaz_con.html.

STEREO OPTICAL CO., INC. http://www.stereooptical.com/products/stereotests.

THOMPSON, W.B., WILLEMSEN, P., GOOCH, A. A., CREEM-REGEHR, S.H., LOOMIS, J.M., AND BEALL, A.C. 2004. Does the Quality of the Computer Graphics Matter when Judging Distances in Visually Immersive Environments? *Presence: Teleoperators and Virtual Environments 13*, 5, 560–571.

UKAI, K. AND HOWARTH, P. A. 2008. Visual fatigue caused by viewing stereoscopic motion images: Background, theories, and observations. *Displays 29*, 2, 106–116.

UNITY - GAME ENGINE. http://unity3d.com/.

VILLARRUEL, C. 2006. Computer Graphics and Human Depth Perception with Gaze-Contingent Depth of Field. https://ida.mtholyoke.edu/jspui/handle/10166/688.

VINNIKOV, M., ALLISON, R.S., AND SWIERAD, D. 2008. Real-time simulation of visual defects with gaze-contingent display. ACM, 127.

WANGER, L.R., FERWERDA, J.A., AND GREENBERG, D.P. 1992. Perceiving spatial relationships in computer-generated images. *Computer Graphics and Applications, IEEE 12*, 3, 44–58.

WANN, J.P., RUSHTON, S., AND MON-WILLIAMS, M. 1995. Natural problems for stereoscopic depth perception in virtual environments. *Vision Research 35*, 19, 2731–2736.

WATT, S.S.J., AKELEY, K., ERNST, M.M.O., AND BANKS, M.M.S. 2005. Focus cues affect perceived depth. *Journal of Vision 5*, 10.

YAMANOUE, H., OKUI, M., AND OKANO, F. 2006. Geometrical analysis of puppet-theater and cardboard effects in stereoscopic HDTV images. *Circuits and Systems for Video Technology, IEEE Transactions on 16*, 6, 744–752.

126