

Depth Perception with Gaze-contingent Depth of Field

Michael Mauderer*, Simone Conte*, Miguel A. Nacenta* and Dhanraj Vishwanath†

*SACHI, School of Computer Science †School of Psychology

University of St Andrews, St Andrews, United Kingdom

{mm285, sic2, mans, dv10}@st-andrews.ac.uk

ABSTRACT

Blur in images can create the sensation of depth because it emulates an optical property of the eye; namely, the limited depth of field created by the eye's lens. When the human eye looks at an object, this object appears sharp on the retina, but objects at different distances appear blurred. Advances in gaze-tracking technologies enable us to reproduce dynamic depth of field in regular displays, providing an alternative way of conveying depth. In this paper we investigate *gaze-contingent depth of field* as a method to produce realistic 3D images, and analyze how effectively people can use it to perceive depth. We found that GC DOF increases subjective perceived realism and depth and can contribute to the perception of ordinal depth and distance between objects, but it is limited in its accuracy.

Author Keywords

Blur; depth cues; depth perception; depth-of-field; eye tracking; gaze-contingent display; three-dimensional graphics and realism

ACM Classification Keywords

H.1.2 Models and Principles: User/machine Systems; I.3.7 Computer Graphics: Three-dimensional Graphics and Realism

INTRODUCTION

Representing 3D information in flat displays can be valuable to improve the realism of scenes and to increase the amount of visual information that is conveyed to viewers (depth can be considered an additional visual variable). One common technique for providing depth in flat displays is through binocular disparity: different images are presented to each eye, and the visual systems derives depth information from the differences [4]. Although binocular disparity is commonly used in current 3D technologies, it has significant problems: a substantial percentage of people (up to 14% in some investigations [26]) have difficulty using it and it can lead to increased visual fatigue, especially when conflicting with other depth cues (e.g., vergence or accommodation), which is commonly the case with flat displays [13].

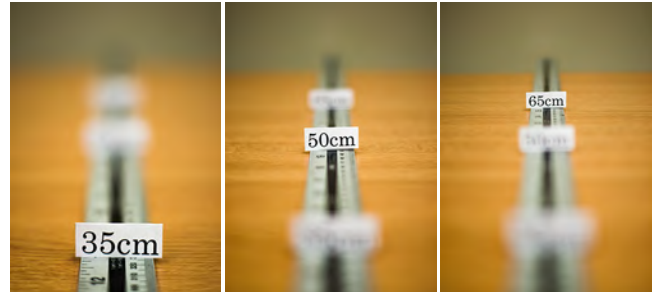


Figure 1: Scene captured using an optical systems with limited DOF focused on different distances showing the resulting defocus blur. Images © Michael Mauderer.

Fortunately, other depth cues exist that can also be used to convey depth. One such depth cue that is also thought to influence the perception of depth is defocus blur from depth of field (DOF). In this work we explore the value of simulating DOF to induce the perception of depth.

When we look at an object, other objects located at different depths can appear blurred (see Fig. 1). This is due to the constrained depth of field of the human eye; objects appear more or less blurred depending on their position, and the position of the focal plane [11,25]. Blur patterns can affect the perceived distance between objects [19] and convey information about the distance of the focused object from the observer [32].

If a display system has up-to-date information about the gaze location of the viewer, it can recreate the blur pattern generated by an optical system: objects in the same plane as the object being looked at will be rendered sharply, while objects at different distances will be blurred to varying degrees. This approach, that we call gaze-contingent depth of field (GC DOF), has been suggested for virtual reality to increase realism [9, 16]; however, it is not clear whether GC DOF can affect the perception of depth. Existing evidence is not in complete agreement ([24] vs. [29]), and to our knowledge nobody has conducted an experiment quantifying the degree to which GC DOF can convey depth information.

In this paper we contribute to the understanding of GC DOF, its influence on perceived depth and realism and to what extent it conveys quantitative depth information. Our findings show that GC DOF can enhance subjective perceived realism and depth and can contribute to the perception of ordering and distance between objects. However, although GC DOF blur could be a useful addition to systems trying to convey realistic scenes, e.g., games, it is probably not sufficiently reliable to serve as an additional visual variable in information visualization applications that require accurate perception.

RELATED WORK

In the following sections we present the current state of knowledge about the role of blur in depth perception and how it is applied in research display systems. We focus on approaches that emphasize depth perception or perceived realism rather than on approaches that use it, e.g., for directing the viewer's attention [12].

Blur as a Depth Cue

Blur has been studied in the perceptual literature as an isolated depth cue by means of static textures. Blur is thought to only provide information about ordinal depth. For example it has been shown that relative blur differences affect the judgment of ordinal depth but not quantitative depth [17, 18, 21, 23]. However, the presence of blur has also been shown to alter quantitative depth perceived on the basis of other cues such as perspective or binocular disparity [18, 22, 34, 35].

The overall pattern of blur created by DOF is a quantitative cue for perceived egocentric distance [32] and perceived size of objects [7]. The egocentric distance information provided by defocus blur may also serve to scale other depth cues [32]. Finally, blur can contribute to speed and accuracy in depth perception [20] and in some situations, blur differences between objects separated in depth are more discriminable than binocular disparity between them [6].

Dynamic Blur

Sun and Holliman [29] investigated the differences in depth perception quality of three non-gaze-contingent blur conditions: a movie scene with no blur, the same scene with a fixed focal plane, and scene where the focal plane changed in a pre-determined fashion. A Likert questionnaire was used for evaluation and they found that the scenes with the focal plane changes and the all-on-focus scene were not statistically different, but the fixed focus scene was worse.

Toyoura et al. created a head-mounted display in which the DOF was controlled by the user to convey depth [31]. The effect of DOF was evaluated using a spatial dexterity task. The results showed an improvement in task completion time for some participants, but no overall improvement was reported.

Hillaire et al. [8] compared two ways of using DOF in a monitor-based VR system. The first was a dynamic adaptation of focus to the object in the screen center. The other used the gaze position of users, an idea first suggested by Rokita [27]. While not the main goal of their investigations, they found participants anecdotally reporting an increased sense of depth due to the DOF. Others have also found subjective increases of realism or immersion using GC DOF [9, 16].

Otani et al. [24] conducted a study on the influence of GC DOF on on participant's qualitative perception of depth. They evaluated real-time rendered gaze-contingent DOF with a self defined questionnaire. They found that using the GC DOF led to an increase in perceived depth but also a decreased sense of realism. None of the studies above have reported objective measures of depth perception accuracy.

Finally, stereoscopic displays relying on binocular disparity have been combined with blur to reduce eye strain and discomfort. The simplest approach applies blur to static scenes producing a fixed and static DOF [10, 14]. Other approaches allow dynamic modification of the DOF position via manual input [1, 14] or utilize GC DOF [2, 15].

GAZE-CONTINGENT DEPTH OF FIELD

To create gaze-contingent depth of field we use two main elements: a sequence of images with a range of focal points, and an algorithm to present the images depending on gaze location and timing. In this section we describe how the different elements come together to simulate DOF. We assume the use of a flat display and an eye tracker that can provide the location of gaze within the coordinate space of the display with appropriate accuracy and timeliness. Details of the hardware that we used can be found in the apparatus sections of the experimental descriptions.

In order to achieve GC DOF the display will present an image that is sharp at the viewer's gaze location and appropriately blurred in other regions. A different image is presented for each of the distances of the objects within the scene. These images can be rendered in advance, generated on the fly from a 3D model (as in [16]), or obtained photographically from a real scene, e.g., through a sequence of exposures with different lens adjustments (focus bracketing) or through a light-field camera (e.g., Lytro [5]).

In our system we can use both photographic and rendered images, but we use pre-rendered images for our experimental setup since they provide us with more control over the scene. We do not render the images in real-time to minimize delays in the presentation of the image; generating highly realistic 3D scenes with correct DOF (e.g., using physical based ray-tracing approaches [3] as used by software like POVRay and Blender Cycles) in real time requires significant processing power and/or specialized hardware.

The presentation algorithm determines which of the images has to be presented for a given gaze location. For this, it uses a mapping that provides the index of the correct image for each possible gaze location. This mapping can be derived from the depth map of the scene, which is easy to generate from a synthetic 3D scene as depth information is usually readily available (e.g., through the z-buffer). However, mapping the values from the depth map directly to an image suffers from problems caused by two factors: the inherent delay of focus changes (i.e., accommodation) in the human eye, and the lack of a precise gaze point delivered by the eye tracker.

In the human eye, a process called accommodation changes the lens properties in order to focus. This process is not instantaneous but takes different amounts of time, depending on various internal and external factors (e.g., age or distance) [28, 30]. If this process is ignored or the required time is underestimated this can negatively affect the viewing experience [16, 24]. While it is beyond this work to create a comprehensive simulation of the human visual accommodation process, our algorithm alleviates this problem by gradually adapting the focus using a transition function.

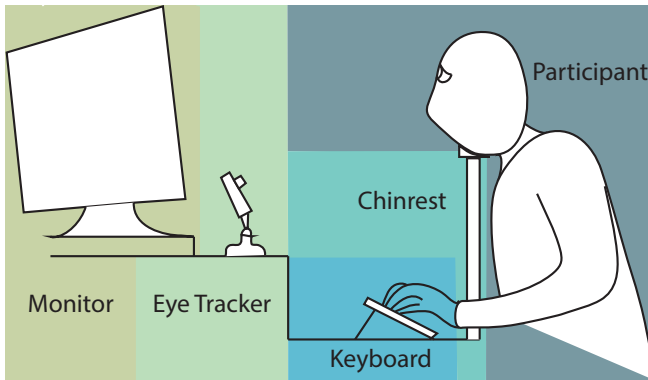


Figure 2: Overview of the hardware setup used in Experiment 1.

The precision and usefulness of the gaze point are limited by two factors: a) imprecision due to technical limitations (current state of the art eye trackers, using a fixed head position and calibration procedure, can deliver only a precision of up to 0.5° , which, depending on the distance, can lead to considerable noise in the measured data); and b) the gaze center, i.e., the fovea, is not a single point but an area. These can lead to ambiguity about which part of this area is currently focused on, especially at edges between areas with different depth.

If the assumed focus is rapidly changing between areas, e.g., due to noise in the gaze point, or if the wrong area is chosen, the viewing experience is disrupted. To solve these problems we use a form of hysteresis: we can assume that small changes in gaze location are probably local and should not introduce a change in focus depth. Only larger changes should be taken into account. Therefore our algorithm uses mappings that are state dependent and initiate a change of focus only if the current area is left by a certain threshold. This is achieved by dilating the active depth region in the map using a circular kernel with a 20 px radius.

EXPERIMENT 1 — QUALITATIVE DEPTH PERCEPTION

We designed our first experiment with two main objectives. First, we needed to validate our implementation of gaze-contingent DOF to show that our setup is able to produce *at least* similar results as seen in related systems. Second, we wanted to provide a more reliable replication of previous work by using validated methodology and solid statistical analysis. For this, we designed an experiment that compares the presentation of realistic scenes presented through GC DOF and a baseline all-on-focus image. The subjective measurements and procedure were modified from previous studies in the psychology literature that compare subjective perceptions of realism and depth produced by different (not gaze-based) stereoscopic presentation techniques [33].

Apparatus

The gaze-contingent display is implemented through an Eye-Link 1000 eye tracker that provides data with a rate of up to 2000 Hz with a nominal 1.4 ms delay¹. The display is a Iiyama HM204DT 22 inch CRT display with a resolution of 1280 px \times 1024 px running at 100 Hz. A CRT was used instead of the now more common LCD displays because of their

¹http://www.sr-research.com/EL_1000.html

better timing properties (CRTs are still preferred to LCDs and other technologies in perceptual research). To stabilize the participants' head we provided a chin rest to keep their face at constant 60 cm and perpendicular to the screen. The participants provided responses through a keyboard placed in a comfortable position close to the resting position of their hands. See Fig. 2 for an overview of the setup.

We implemented the DOF algorithm as described earlier with Matlab and OpenGL. To assure timely delivery of the images, they were pre-loaded as textures before each trial.

Stimuli

The main goal of this experiment was to measure subjective impressions of realism and depth, therefore we chose to present realistic scenes: A kitchen and a patio (see Fig. 3)². The models were pre-rendered using POV-Ray at a resolution of 1152 px \times 864 px. The virtual camera positions and scene scales were adjusted to produce a rendering compatible to what a camera would record if placed central in front of a window of the screen's size and looking at a real scene with normal-sized objects. The aperture was set at a value of 0.8 for the kitchen and 1.0 for the patio scene. The objects in the kitchen scene are located in an area between 70 and 146 units from the camera. For the patio scene we used larger distances between 286 and 900 (distances in POV-Ray units).

We rendered 30 different focal images for each scene by changing the focal point in the POV-Ray camera model. The focal points are spaced throughout the scene using an exponentially distributed density function, with higher density in the foreground, where there is more detail in the scene, and lower density at the back. This corresponds to the geometry and perceptual characteristics of blur—blur does not change linearly with distance and produces more noticeable differences for closer focal planes [25].

Procedure and Experimental Design

15 participants (seven female, age 19 to 31, all with normal or corrected-to-normal vision) took part in the experiment for a £5 compensation. After providing written consent, participants sat in front of the display and the experimenter performed the eye-tracker calibration with them. The experimenter provided initial instructions about the experiment and answered questions.

In the main section of the experiment, participants went through two trials, the first with the kitchen scene, the second with the patio scene (see Fig. 3). During each trial the participants had access to the two conditions by pressing the left and right control keys, which they could reach without having to look at the keyboard. One of the keys (randomly assigned) would activate GC DOF, the other would show a static all-in-focus image of the scene. To reinforce the transition between conditions, the system emitted a beep for every condition change. Participants were instructed to view each condition for at least five seconds, to switch at least five times, and to look for differences in depth. If no key or two keys were pressed the screen would remain blank.

²Both models are kindly provided by Jaime Vives Piqueres (<http://www.ignorancia.org/>) under a CC BY-SA 3.0 license

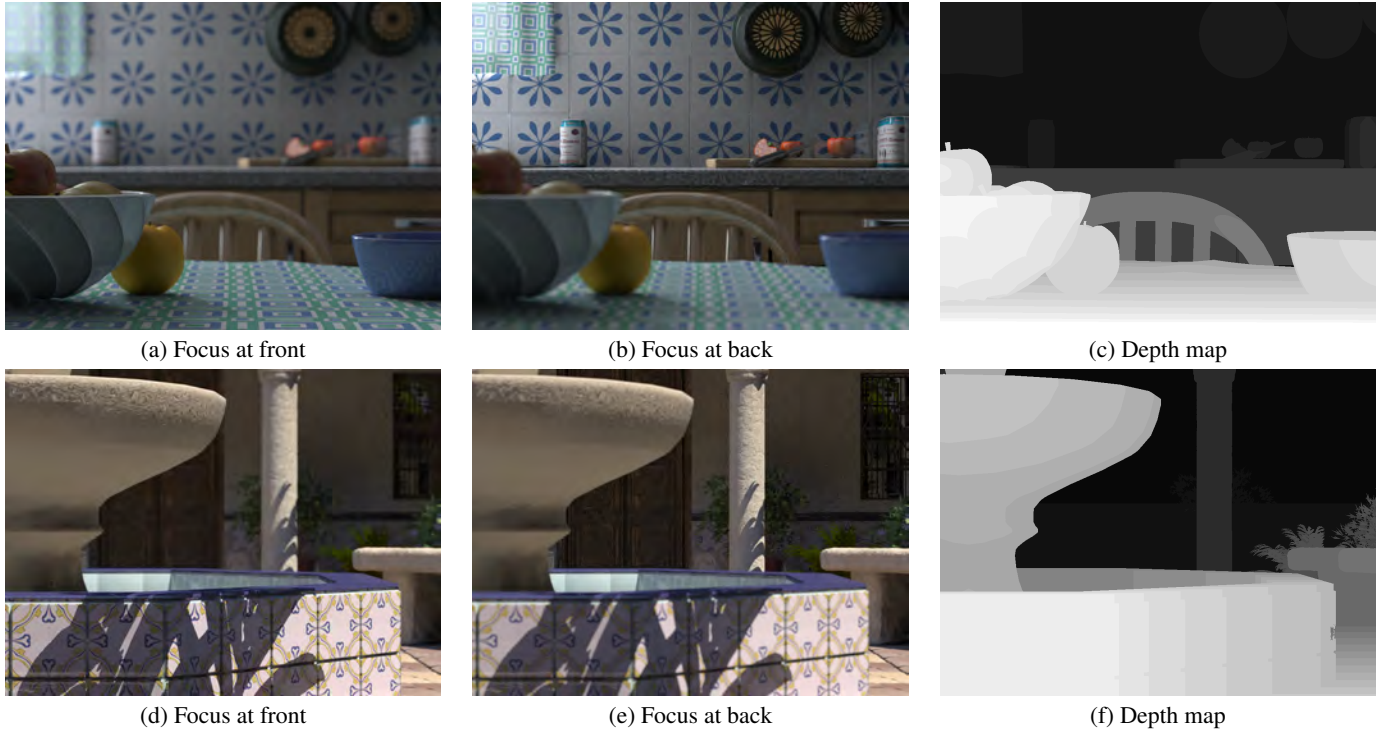


Figure 3: Stimuli used in Experiment 1 and their depth maps. The top row shows the kitchen scene, the bottom row the patio scene.

After viewing both conditions participants reported whether they perceived any difference and rated a battery of statements displayed on the computer screen (see next section) by using the same control keys. The statements did not refer directly to “gaze-contingent” and “static” conditions; instead, they referred to the “Left” or “Right” conditions, which corresponded to the key used for activation.

The first trial was followed by a second trial depicting the other scene. After they finished the experiment, participants performed an acuity and stereo vision test.

Measures

The participants assessed the perceived 3D-effect and realism of the conditions by rating the statements listed in Tbl. 1 on a seven point Likert-scale from “Strongly Agree” to “Strongly Disagree”. Each statement (except for S10 and S11) was included twice, once referring to the “Left” and once referring to the “Right” condition. The statements are based on a battery of statements from Vishwanath and Hibbard [33], and include four catch-statements (S8-S11) referring to properties that are not related to depth perception (e.g., whether objects appear distorted or translucent). This allowed us to assess whether participants are biased towards ascribing properties to one of the conditions independently of whether these are related to depth perception. The questions were presented in four randomized orders balanced over all participants.

Results

We analyzed the answers to statements S1 to S11 with Wilcoxon signed-rank tests to determine if there were any differences between the conditions. All the tests for regular statements (S1-S7) were statistically significant, and all

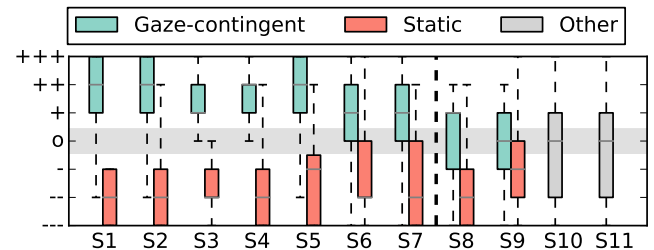


Figure 4: Results for each question of the questionnaire. Rating ranged from “Strongly Agree” (+++) to “Strongly Disagree” (---).

tests for the catch questions were non-significant except for S8. Results of the statistical tests are summarized in Tbl. 1; p-values were corrected for multiple comparisons using the Bonferroni-Holm method. The tests for S10 and S11 are slightly different (single-sample Wilcoxon signed-rank tested against the median of the rating scale) as they compare both conditions within the same question and are designed only to detect departures from the neutral response.

As can be seen also in Fig. 4, the results of the tests provide strong evidence that, in the gaze-contingent condition, participants judged objects to be more “like looking at real objects”, appear to “stick out of the screen”, have a more defined depth, and appear more separate from each other. The results for the catch statements also show that, with the exception of S8 (“objects appear distorted”), participants did not generally think that objects appeared larger, more transparent and translucent, or that objects had different shapes in the two conditions.

Discussion

The results suggest that gaze-contingent DOF allows viewers to perceive images more realistically and with an increased sense of depth. There is also a strong distinction between the answers of the assessing statements and the catch statements. While S8, one of the designated catch trials, shows a difference between conditions, we are not concerned about the overall validity of the results. Examining the distribution of the ratings it becomes clear that there was no strong trend to identify the gaze-contingent condition with the presented concept (“distortion”), merely indecision and a strong opinion about the absence of distortion in the static condition.

The results of this experiment are useful in two main ways. First, they provide replication of previous research indicating that GC DOF increases impression of depth and realism compared to regular images. Importantly, the experiment uses a set of questions validated in previous research and statistical analyses that build upon existing evidence.

Second, the replication of these effects with our own setup provides necessary validation of our implementation and algorithms. This allows us to confidently go beyond previous research and test new aspects of the gaze-contingent DOF technique to produce depth perception without having to doubt whether the results are simply due to differences in the used hardware or implementation.

EXPERIMENT 2 — QUANTITATIVE DEPTH PERCEPTION

Experiment 1 provides assurance that there is a subjective effect of gaze-contingent DOF for the perception of realistic scenes. However, this does not answer the question of whether GC DOF is an effective and reliable method to convey depth information. In other words, can people extract depth information accurately from GC DOF? This is important for applications where precise information is important (e.g., in information visualization).

To answer this question we designed a quantitative, controlled experiment that investigates depth perception accuracy through a depth comparison task (i.e., asking participants to estimate the depth of objects). The experiment is closer in nature to experiments in the perception literature, and uses abstract objects in a synthetic space to control for possible confounds (e.g., other depth cues).

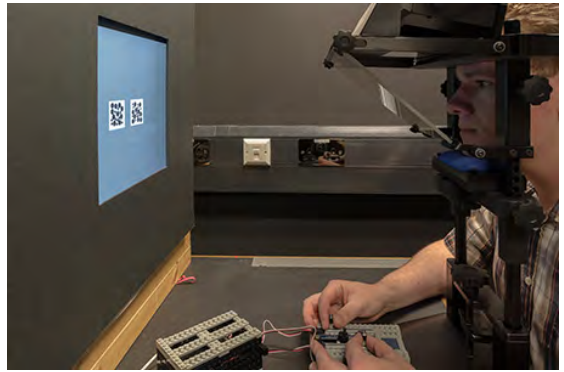


Figure 5: Overview of the setup used in Experiment 2

Apparatus and Software

The hardware used for this experiment was the same as in Exp. 1, except that the eye-tracker was mounted vertically with a tower mount (see Fig. 5), which provides easier calibration and better quality gaze-tracking. The distance between headrest center and screen was 40 cm. Additionally, we darkened the room for the trials and covered the area surrounding the display with a flat black surface to avoid visual distraction and possible interference due to the surrounding area of the monitor.

Input responses required recording estimated relative positions of two objects in the visual field; for this we replaced the keyboard with a custom-built Phidget device with two sliders that record positions in a fixed range, and a capacitive plate that allows participants to indicate the end of a trial.

We reimplemented the same DOF algorithm but with two differences: (a) we disconnected the hysteric adjustment of the depth map because with the simpler, widely spaced objects it is not required, (b) the gradual adjustment of blur in accommodation was exponential instead of sinusoidal (the distance between current focused point and target focus point was halved each frame and rounded down). The new implementation was programmed in Python 2.7 using PsychoPy 1.75.01³ (integrated with OpenGL, allowing for efficient use of graphics card memory); gaze information was retrieved using the Eyelink Python API (Pylink 1.0.0.37).

³<http://www.psychopy.org/>

#	Statement text	Z	p	r
S1	Depth is more defined and clearer in [L/R] condition	-3.39	.007*	-.65
S2	Objects appear more 3-dimensional in [L/R] condition	-3.17	.012*	-.61
S3	It feels more like looking at real objects than a picture of objects in [L/R] condition	-2.94	.013*	-.57
S4	Objects appear to stick out/come out of the screen in [L/R] condition	-3.24	.011*	-.63
S5	There is a more definitive sense of separation in [L/R] condition	-3.73	.002*	-.72
S6	In [L/R] condition it feels though you could reach out and touch things	-3.02	.015*	-.58
S7	There is a greater amount of separation between objects in [L/R] condition	-3.16	.011*	-.61
S8	Objects appear distorted in [L/R] condition	-2.98	.014*	-.57
S9	Objects appear larger in [L/R] condition	-1.43	.308	-.27
S10	Objects appear more transparent and translucent in [L/R] condition compared to [R/L] condition	-0.68	.495	-.13
S11	The shape of objects is different between [L/R] and [R/L] condition	-1.52	.384	-.29

Table 1: List of statements used to assess participants’ perception in Experiment 1 and the results of their statistical evaluation between conditions. Significant results ($\alpha = .05$) are marked by an asterisk.

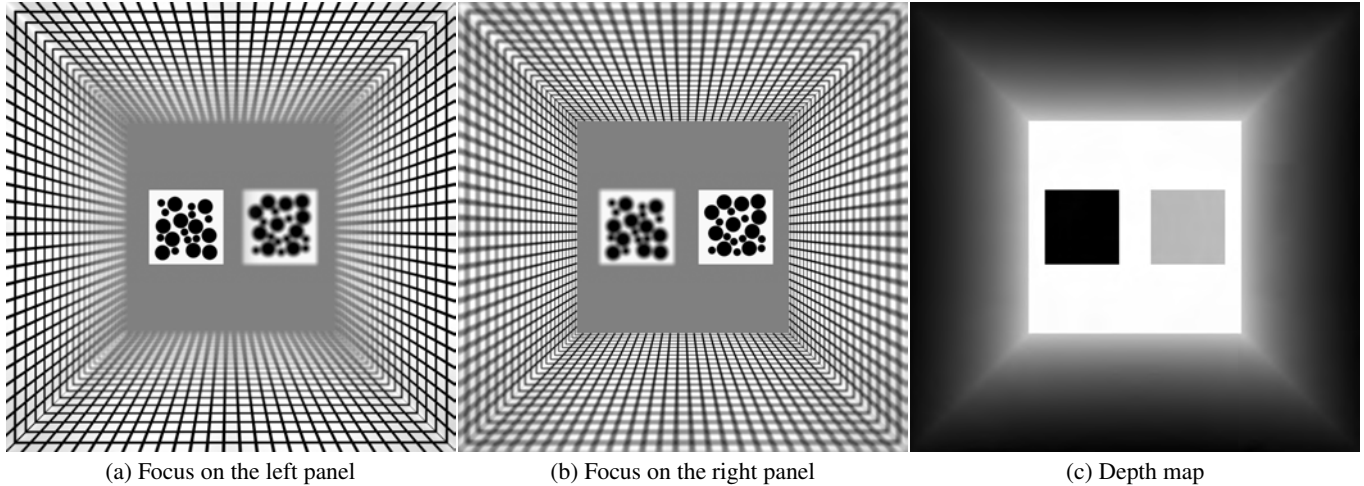


Figure 6: Example of stimuli used in the practice of Experiment 2 and its depth map. The left tile is positioned at 10 %, the right tile at 90 %.

Task and Stimuli

Participants were presented with two abstract square objects that contain a black-and-white circle pattern that was different between tiles, but contained the same number of circles of equal size to make sure that both tiles have equivalent perceived brightness, contrast and information from blur. The objects are perpendicular to the screen plane, aligned with the screen edge, and are floating in mid air. In half the conditions there also was a patterned square tunnel with a depth of 60 cm (see Fig. 6).

The main task was to estimate the positions of the two objects. Participants used the left and right sliders of the purpose-build Phidget input device to indicate the relative positions of the left and right objects. The end ranges of the sliders represented the front and back ends of the tunnel. We chose relative indirect input for this task because of the possible effect of overlapped motor activity and because seeing their own hand could affect perception of distance (see e.g., [36]).

As in the previous experiment, the virtual scene was created to match the physical setup, i.e., objects were displayed at the same distance from the virtual camera as they were supposed to be located from the user. Each scene had 20 different focal points spaced throughout the area of interest, rendered through Blender v2.64 using the built-in Cycles ray-tracing engine. We chose to switch to Blender due to its more efficient GPU rendering capabilities. The aperture radius of the camera model was set to 0.005 and the field of view was 50° . Since depth values in Blender are calculated as distance from the camera position, a correction was applied to produce distances from the image plane instead.

While the tiles were displayed at varying depths, their size was always adjusted to occupy the same visual angle to avoid introducing size as a depth cue confound. The elements in the screen (tunnel and tiles) did not overlap and were rendered separately. To avoid the interference of binocular cues participants wore a one-eye patch on their non-dominant eye.

Procedure and Experimental Design

16 participants (nine female, aged 18 to 33, all had normal or corrected-to-normal vision, nine were right-eye dominant) took part in the experiment. After filling a demographic questionnaire, testing for eye dominance (Miles test) and acuity (Snellen chart) participants put on an eye patch to eliminate binocular depth cues. They then learned the task and performed one block of practice trials and one block for all four experimental conditions.

There were ten practice trials, in which two tiles were shown in the tunnel, close to the opposite extremes (back and front) and blur was adjusted according to gaze. The other four blocks consisted of variations of two factors with two levels each: gaze-contingency (gaze-contingent *GC* vs. non-gaze-contingent *NGC*) and background visibility (visible background *VB* vs. no background *NB*).

We obviously introduced the gaze-contingency factor to compare it to the baseline. However, the introduction of the background factor corresponds to a less obvious goal. There is a running discussion in the blur perception literature regarding the nature of the type of information provided by defocus blur in natural perception (ordinal vs. quantitative) and whether the context blur of the field/scene provides valuable information [6, 32]. Notice that an appropriately blurred background does contain information about the depth of an object (objects on focus are at the same depth than the parts of the background that is on focus). Including this factor allows us a) a wider generalization of our results to applications with different backgrounds, b) to draw useful conclusions about the importance of background in *GC* *DOF*, and c) contribute to the ongoing discussion on the perceptual mechanisms of blur.

Notice also that, unlike in Exp. 1, the static conditions did have blur (one of the objects appeared permanently in focus). Making the non-gaze-contingent conditions all-in-focus would make no sense, since the objects' location would be indistinguishable regardless of their depth.

Trials differed in the position of the two stimulus objects (Pos_S). The front tile could appear at 10, 30 and 50% of the box depth (6 cm, 18 cm and 30 cm behind the screen), and the other tile 0, 20 and 40% box depth (0 cm, 12 cm and 24 cm) behind the first one. Each configuration was shown twice, which resulted in 36 trials per condition. In static (non-gaze-contingent) conditions, each of the two trial repetitions for a given position and distance had a different object on focus. The condition presentation order was balanced but the first two conditions would be either both gaze contingent or non-gaze contingent to avoid unnecessary switches and calibration problems (a total of eight possible orders).

Participants had no time limit to complete a trial, and indicated the end of each trial by tapping on a large capacitive sensor attached to the near end of the slider input device. Participants were allowed to rest between blocks.

Measures

The main raw measurement taken during the experiment is the position of the sliders when participants indicated the end of a trial (at the moment of tap). For practical purposes we report slider positions on a range from 0 to 100 (Pos_I). However, since participants used a variable range of positions in the sliders, and because our measures are relative anyway, we chose to normalize each participant's measurements so that their closest and furthest reported locations over the whole experiment correspond to 0 and 100 respectively.

From the raw position data we derived several additional measures: The difference between the input sliders ($|Diff_I|$), the absolute difference between the input sliders ($I_{|D|}$), and the perceived ordering in depth of the tiles (i.e., whether the left tile was perceived to be in front or behind the right tile). These additional measures are useful to determine whether the perceived depth information in a condition is ordinal (just what is behind what), quantitative (how far an object is from the other), or both (order and estimated distance).

Results

The experiment has a general repeated measures design with two factors that have two levels each: gaze contingency (G) (gaze contingent GC vs. non-gaze contingent NGC) and the presence of background (B) (background VB vs. no-background NB). Importantly, the analyses also use a *position* factor to control for stimulus object position; this factor is what differentiates the different analyses below and can be one of: a) the absolute position of objects (10, 30, 50, 70 or 90), b) the distance between objects ($-40, -20, 0, 20$ or 40), c) the absolute distance between objects (0, 20 or 40), or d) the ordering of objects (left in front, right in front). Tests that do not show an effect related to the position factor (e.g., timing) are of little interest to us, since they do not relate to the participant's ability to use the stimuli information to discriminate different depths and thus we omit discussing them.

Repeated trials in the same cell were averaged before analysis, and when the assumption of sphericity was broken (i.e., a Mauchly's sphericity test was significant) we applied the Greenhouse-Geisser correction. Error bars in plots represent 95% confidence intervals unless otherwise stated.

#	Factors	df	\tilde{df}	F	p	η_p^2
Absolute Position						
A1	G	1	15	1.26	.280	.08
A2	B	1	15	16.36	.001**	.52
A3	Pos_S	2.05	30.75	8.85	< .001**	.37
A4	$B \times G$	1	15	3.64	.076	.20
A5	$G \times Pos_S$	2.00	29.81	22.83	.702	.02
A6	$B \times Pos_S$	2.36	35.33	5.08	.008**	.25
A7	$B \times G \times Pos_S$	2.68	40.19	0.42	.716	.03
Depth Difference						
B1	G	1	15	3.93	.066	.21
B2	B	1	15	0.94	.349	.06
B3	$Diff_S$	2.51	37.64	1.80	.172	.11
B4	$B \times G$	1	15	5.21	.038*	.26
B5	$G \times Diff_S$	2.25	33.81	2.37	.103	.14
B6	$B \times Diff_S$	4	60	1.91	.121	.11
B7	$B \times G \times Diff_S$	2.30	34.57	0.68	.534	.04
Unordered Depth Difference						
C1	G	1	15	4.73	.046*	.24
C2	B	1	15	5.05	.040*	.25
C3	$ Diff_S $	1.12	16.77	34.83	< .001**	.70
C4	$B \times G$	1	15	0.16	.698	.01
C5	$G \times Diff_S $	1.11	16.67	5.78	.025*	.28
C6	$B \times Diff_S $	2	30	0.21	.810	.01
C7	$B \times G \times Diff_S $	2	30	0.02	.982	.00
Depth Order						
D1	G	1	15	8.57	.010*	.36
D2	B	1	15	0.40	.539	.03
D3	$B \times G$	1	15	0.05	.833	.00
Meta-Analysis of Correlations						
E1	G	1	15	26.26	< .001**	.64
E2	B	1	15	0.71	.412	.05
E2	$G \times B$	1	15	0.56	.465	.04

Table 2: Results of the statistical analysis of the experimental measurements. Significant results ($\alpha = .05$) are marked by an asterisk, highly significant results ($\alpha = .01$) by two asterisks.

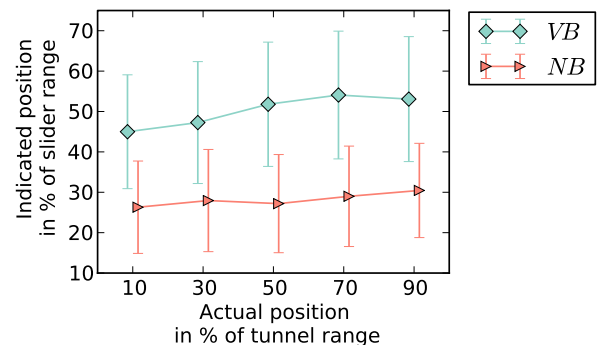


Figure 7: Indicated positions Pos_I against actual positions Pos_S in the two background conditions.

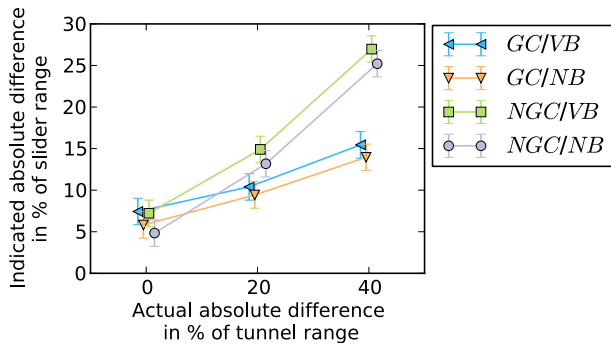


Figure 8: Absolute stimulus difference $|Diff_S|$ against the absolute indicated input difference $I_{|D|}$.

Absolute Position

Our first test is designed to estimate whether participants estimated positions differently depending on the virtual positions of the stimuli; in other words, whether participants were able to extract information from the stimuli to judge the absolute positions of the tiles. For this, we performed a $2 \times 2 \times 5$ repeated-measures ANOVA of the individual position responses for the tiles. We expected to see strong main effects of the stimuli positions on the perceived positions of objects, if not in all cases, at least for certain conditions.

We did find a significant main effect for Pos_S (Tbl. 2: A3) and the background condition (Tbl. 2: A2) as well as an interaction between them (Tbl. 2: A6). Contrasts reveal a linear relationship for Pos_S ($F(1, 15) = 16.77, p < 0.001, \eta_p^2 = .53$). However, we did not find an effect involving G (Tbl. 2: A1) or G and Pos_S (Tbl. 2: A5, A7).

These results suggest that some absolute information is extracted from blur overall: Stimuli representing objects that are further result in further judgments, with a slight improvement when there is background. However, there are no significant effects of gaze contingency, and a quick look at Fig. 7 shows that the linear relationship is not very pronounced and is indeed small compared to the amount of perceptual noise (evident from the large 95% confidence intervals) for both background and non background conditions.

Depth Difference

The absolute measurements could have been affected by the indirectness of the input device, and since many perceptual effects are relative, we tested also the difference in the judgments of the left and right object as compared to the difference in position of the tiles. This analysis is analogous to that of the previous section, but position is instead encoded as relative differences in object position in both the dependent variable, and the $Diff_S$ factor.

Surprisingly, none of the tests involving $Diff_S$ (Tbl. 2: B3, B5, B6, B7) were significant. This indicates that participants were even worse at comparing the distance between objects.

Unordered Depth Difference

A possible explanation for the previous result is that blur might be good for judging the magnitude of depth differences between objects, but not to estimate the direction of this difference. To test this we used the unordered depth distances in

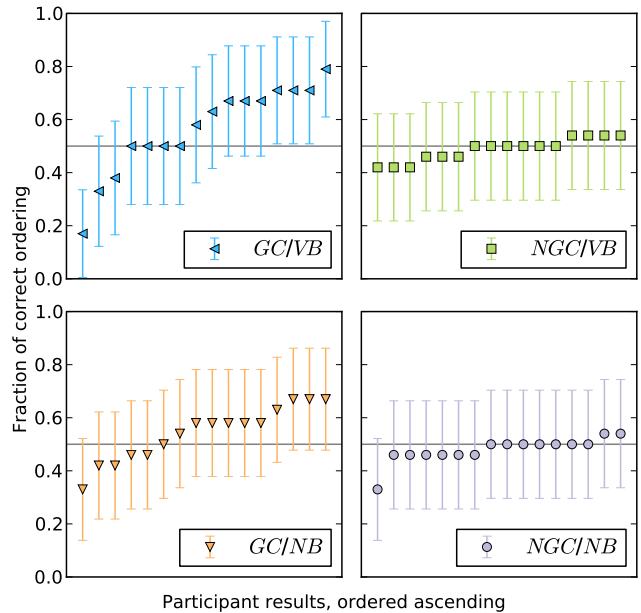


Figure 9: Individual performances for depth order discrimination.

the stimuli ($|Diff_S|$ - 0, 20 and 40) with a similar transformation in the dependent variable: $I_{|D|}$.

The results (see Fig. 8) indicate that this is indeed the case: all main effects were significant (Tbl. 2: C1, C2, C3), as well as the interaction between $|Diff_S|$ and G (Tbl. 2: C5). A look at Fig. 8 and the interaction reveals that gaze-contingent conditions resulted in a smaller range of responses: static representations of blur result thus in a stronger impression of depth.

Depth Order

Since the previous test eliminates the ordinal information, it makes sense to run its counterpart to test whether participants could extract pure order information (i.e., determine which object is in front). We first converted the positions of objects to a binary location factor (left tile in front vs. right tile in front) as well as the participant responses (left slider in front vs. right slider in front), and determined the agreement between them. For this analysis we excluded trials for which $Diff_S = 0$, as these do not allow a correct ordering decision.

Results indicate a significant main effect for G (Tbl. 2: D1) with GC ($N = 32, M = 0.55, SD = .02$) showing a higher rate of correct ordering than NGC ($N = 32, M = 0.48, SD = .01$). It is important to notice, however, that the correctness rate for GC is not much above chance.

Individual Differences

It is also possible that beyond the overall effect of the different conditions, individual differences between observers might explain partly the fragmented results; in other words, could certain participants extract information very effectively while others could not?

To test this we evaluated the accuracy of individual participants at depth ordering. The results are displayed in Fig. 9,

and the results paint an interesting picture: individual differences are highly visible in the gaze-contingent with background condition, with four participants scoring significantly above chance, and one significantly below chance (average correctness is above chance when the 95% interval of confidence does not touch the 50% line). These results show that the ability to perceive depth ordering with GC DOF is not universal, and at least one person interpreted the information in a consistently erroneous way.

Meta-Analysis of Correlations

As a complementary post-hoc analysis to the initial investigation of estimated depth difference, we analyzed the correlations between estimated depth difference and displayed depth difference in the different conditions. We conducted a 2×2 ($G \times B$) RM-ANOVA over each participants' coefficient of determination R^2 between $|Diff_I|$ and $Diff_S$ in each of the conditions. We chose R^2 over r as we are more concerned about consistency and explanatory power than sign of the relationship. The results show a significant main effect for G (Tbl. 2: E1) indicating that there was a stronger correlation in the gaze-contingent condition ($N = 32$, $M = .082$, $SD = .01$) than in the non gaze-contingent condition ($N = 32$, $M = .012$, $SD = .01$).

Discussion

Experiment 2 contributes results that extend our knowledge about the gaze-contingent DOF approach to represent 3D information and about blur in general.

Blur and GC DOF Blur's Contribution

Our findings indicate that gaze-contingent blur contributes to depth perception; this is the first quantitative evidence of depth perception from blur with a GC DOF setup. When considering both static and gaze-contingent blur, we found a general contribution of blur to the perception of absolute depth, and to the perception of depth differences between objects if we only consider magnitude. GC DOF also allows better than chance discrimination in the spatial ordering of objects, although static blur seems to have a more pronounced effect in the detection of depth differences between two objects.

The effects, however, are small. When considered across all participants and conditions, the ability of people to judge ordinal depth is significant above chance, but just by 5 percentage points. Similarly, quantitative judgments of depth are also statistically significant, but small if compared with the degree of variability seen in the accuracy of the depth judgments.

How Depth from Blur Works

Our results provide valuable insights in how blur provides information about depth. If we consider first only static stimuli, it is not surprising to find that blur does not provide ordering information, since it is sign-ambiguous: The more blurred object could be in front of or behind the focused object [6, 17]. However the improvement in ordinal depth perception in the gaze-contingent conditions suggests that ordinal depth perception could be informed by the dynamic blur transitions between focal planes.

We also found that the presence of GC DOF yields a better correlation between observed object distances and actual object distances in the presence of contextual information. With background, gaze-contingent depth of field yielded greater absolute magnitudes of perceived depth. These results are consistent with previous findings reported in the perception literature reporting that blur difference between isolated objects is by itself insufficient to convey quantitative depth [17], but now we know that blur can influence perceived depth in the presence of context without the need for a shared border.

However, this result needs to be followed up by further experiments as the backgrounds we used all had defocus blur consistent with the viewing parameters. To know whether the advantage is caused by the background blur pattern, or just merely the presence of an additional context, we need to compare an additional condition with unblurred background.

Individual Differences

Further analysis revealed that depth from blur is not universal; although some participants were particularly good at deriving depth from the GC DOF, others seemed to be unable to extract information from it. It is unclear whether this can be taught or, as in binocular displays, it is inherent to the perceptual processing capability of some people [26].

GENERAL DISCUSSION

The results from our two experiments are complementary. The first experiment confirms previous results that gaze-contingent DOF provides a subjectively more complete perception of 3D than images with static monocular cues. This results suggests that GC DOF could be a viable technology to enhance 3D perception, or to complement other existing ways of providing 3D such as binocular disparity.

Our results have important implications for possible applications of GC DOF: although the results from previous literature and our first study might seem strong support for the application of this 3D cue in more displays, the second experiment also shows that, while GC DOF provides some information of depth (absolute position, unordered depth distance, ordinal) the low levels of accuracy make it undesirable for applications that require accurate perception of depth information. This means that, while depth of field has been used for benefit in information visualization for focus+context tasks [12], it can probably not be reliably used to convey a quantitative visual information variable on its own. However, games, virtual reality applications, or realistic image displays might be able to benefit from the increase in perceived depth and realism alone.

Whether static and GC DOF might be more useful in combination with other depth cues requires further research; however, it is likely that, in this case, GC DOF with a rich context will be preferable. GC DOF provides better depth order information and, although possibly less pronounced in its magnitude than static blur, has the advantage that whatever is being looked at is automatically in focus. Additionally, further studies are necessary to validate the results of Experiment 2 in more naturalistic scenarios including other common cues like perspective or size.

CONCLUSION

In this paper we presented an experimental investigation of gaze-contingent depth of field, studying its influence on qualitative and quantitative depth information. For this purpose we conducted two experiments using a GC DOF algorithm based on pre-rendered images.

We contribute evidence that GC DOF increases perceived realism and depth and adds validity to prior research by confirming these findings using methodology validated in perception research. We also demonstrate that GC DOF does contain information about depth order and relative spatial position of objects. This information is, however, limited. Our findings show that GC DOF can be a valuable addition to systems that aim to convey realistic scenes (e.g., games) but has to be used with caution in applications that try to communicate accurate information such as information visualizations.

ACKNOWLEDGEMENTS

This research is supported by the European Union's Marie Curie Program: CIG - 303780.

REFERENCES

1. Blum, T., Wieczorek, M., Aichert, A., Tibrewal, R., and Navab, N. The effect of out-of-focus blur on visual discomfort when using stereo displays. In *Proc. ISMAR*, IEEE (2010), 13–17.
2. Brooker, J. P., and Sharkey, P. M. Operator performance evaluation of controlled depth of field in a stereographically displayed virtual environment. In *Proc of Electronic Imaging*, SPIE (2001), 408–417.
3. Cook, R. L., Porter, T., and Carpenter, L. Distributed ray tracing. In *Proc. SIGGRAPH*, vol. 18, ACM (1984), 137–145.
4. Cutting, J. E., and Vishton, P. M. Perceiving layout and knowing distances: The integration, relative potency, and contextual use of different information about depth. In *Perception of Space and Motion*. AP, 1995, 69–117.
5. Georgiev, Todor and Yu, Zhan and Lumsdaine, Andrew and Goma, S. Lytro camera technology: Theory, algorithms, performance analysis. In *Proc. of Electronic Imaging*, SPIE (2013), 86671J–86671J.
6. Held, R. T., Cooper, E. a., and Banks, M. S. Blur and disparity are complementary cues to depth. *Current Biology* 22, 5 (2012), 426–31.
7. Held, R. T., Cooper, E. A., Banks, M. S., and O'Brien, J. F. Using blur to affect perceived distance and size. In *Transactions on Graphics*, ACM (2010), 19.
8. Hillaire, S., Lécuyer, A., Cozot, R., and Casiez, G. Depth-of-field blur effects for first-person navigation in virtual environments. *CG&A* (2008), 47–55.
9. Hillaire, S., Lecuyer, A., Cozot, R., and Casiez, G. Using an eye-tracking system to improve camera motions and depth-of-field blur effects in virtual environments. In *Proc. VR*, IEEE (2008), 47–50.
10. Jung, Y., Sohn, H., Lee, S.-i., Speranza, F., and Ro, Y. Visual importance- and discomfort region-selective low-pass filtering for reducing visual discomfort in stereoscopic displays. In *TCSVT*, IEEE (2013).
11. Kolb, C., Mitchell, D., and Hanrahan, P. A realistic camera model for computer graphics. In *Proc. SIGGRAPH*, ACM (1995), 317–324.
12. Kosara, R., Miksch, S., and Hauser, H. Semantic depth of field. In *Proc. InfoVis*, IEEE (2001), 97.
13. Lambooj, M., IJsselstein, W., Fortuin, M., and Heynderickx, I. Visual discomfort and visual fatigue of stereoscopic displays: A review. *Journal of Imaging Science and Technology* 53, 3 (2009), 30201–1.
14. Leroy, L., Fuchs, P., and Moreau, G. Real-time adaptive blur for reducing eye strain in stereoscopic displays. In *TAP*, ACM (2012), 9.
15. Leroy, L., Fuchs, P., and Moreau, G. Visual fatigue reduction for immersive stereoscopic displays by disparity, content, and focus-point adapted blur. In *TIE*, IEEE (2012), 3998–4004.
16. Mantiuk, R., Bazyluk, B., and Tomaszewska, A. Gaze-dependent depth-of-field effect rendering in virtual environments. In *Proc. SGDA*, vol. 6944, Springer (2011), 1–12.
17. Marshall, J. a., Burbeck, C. a., Ariely, D., Rolland, J. P., and Martin, K. E. Occlusion edge blur: A cue to relative visual depth. *JOSA A* 13, 4 (1996), 81–8.
18. Mather, G. Image blur as a pictorial depth cue. *Royal Society of London Proceedings Series B* 263, 1367 (1996), 169–72.
19. Mather, G. The use of image blur as a depth cue. *Perception* 26, 9 (1997), 1147–58.
20. Mather, G., and Smith, D. R. Combining depth cues: Effects upon accuracy and speed of performance in a depth-ordering task. *Vision Research* 44, 6 (2004), 557–562.
21. Mather, G., and Smith, D. R. R. Blur discrimination and its relation to blur-mediated depth perception. *Perception* 31, 10 (2002), 1211–1219.
22. Nefs, H. T. Depth of field affects perceived depth-width ratios in photographs of natural scenes. *Seeing and Perceiving* 25, 6 (2012), 577–595.
23. O'Shea, R. P., Govan, D. G., and Sekuler, R. Blur and contrast as pictorial depth cues. *Perception* 26, 5 (1997), 599–612.
24. Otani, K., Yoshida, T., Nii, H., and Kawakami, N. Active focus control with eye gaze detection for depth cognition. In *Proc. ASIAGRAPH* (2008).
25. Potmesil, M., and Chakravarty, I. A lens and aperture camera model for synthetic image generation. In *Proc. SIGGRAPH*, ACM (1981), 297–305.
26. Richards, W. Stereopsis and stereoblindness. *Experimental Brain Research* 10, 4 (1970), 380–8.
27. Rokita, P. Generating depth of-field effects in virtual reality applications. *CG&A* (1996), 18–21.
28. Schaeffel, F., Wilhelm, H., and Zrenner, E. Inter-individual variability in the dynamics of natural accommodation in humans: relation to age and refractive errors. *The Journal of Physiology* 32 (1993), 301–320.
29. Sun, G., and Holliman, N. Evaluating methods for controlling depth perception in stereoscopic cinematography. In *Proc. SD&A*, SPIE (2009).
30. Temme, L. A., and Morris, A. Speed of accommodation and age. *Optometry & Vision Science* (1989).
31. Toyoura, M., Kashiwagi, K., Sugiura, A., and Mao, X. Mono-glass for providing distance information for people losing sight in one eye. In *Proc. VRCAI*, ACM (2012), 39.
32. Vishwanath, D., and Blaser, E. Retinal blur and the perception of egocentric distance. *Journal of Vision* 10 (2010), 1–16.
33. Vishwanath, D., and Hibbard, P. B. Seeing in 3-D with just one eye: Stereopsis without binocular vision. *Psychological Science* (2013).
34. Wang, J., Barkowsky, M., Ricordel, V., and Le Callet, P. Quantifying how the combination of blur and disparity affects the perceived depth. In *Proc. Electronic Imaging*, SPIE (2011).
35. Watt, S. J., Akeley, K., Ernst, M. O., and Banks, M. S. Focus cues affect perceived depth. *Journal of Vision* 5, 10 (2005), 834–62.
36. Witt, J. K., Proffitt, D. R., and Epstein, W. Tool use affects perceived distance, but only when you intend to use it. *Journal of Experimental Psychology: Human Perception and Performance* 31, 5 (2005), 880.