## Blur and the perception of depth at occlusions

Vision Science Program, University of California, Berkeley,
Berkeley, CA, USA
Present address: Oculus Research, Redmond, WA, USA
Department of Physics, Durham University, Durham, UK
Department of Electrical Engineering and Computer Science, University of California Berkeley, Berkeley, CA, USA Present address: Department of Computer Science and Engineering, University of Minnesota, Minneapolis, MN, USA
Vision Science Program, University of California, Berkeley, Berkeley, CA, USA School of Optometry, University of California, Berkeley, Berkeley, CA, USA

The depth ordering of two surfaces, one occluding the other, can in principle be determined from the correlation between the occlusion border's blur and the blur of the two surfaces. If the border is blurred, the blurrier surface is nearer; if the border is sharp, the sharper surface is nearer. Previous research has found that observers do not use this informative cue. We reexamined this finding. Using a multiplane display, we confirmed the previous finding: Our observers did not accurately judge depth order when the blur was rendered and the stimulus presented on one plane. We then presented the same simulated scenes on multiple planes, each at a different focal distance, so the blur was created by the optics of the eye. Performance was now much better, which shows that depth order can be reliably determined from blur information but only when the optical effects are similar to those in natural viewing. We asked what the critical differences were in the single- and multiplane cases. We found that chromatic aberration provides useful information but accommodative microfluctuations do not. In addition, we examined how image formation is affected by occlusions and observed some interesting phenomena that allow the eye to see around and through occluding objects and may allow observers to estimate depth in da Vinci stereopsis, where one eye's view is blocked. Finally, we evaluated how accurately different rendering and displaying techniques reproduce the

retinal images that occur in real occlusions. We discuss implications for computer graphics.

## Introduction

The problem of how we see in three dimensions is interesting because one dimension-depth-is lost in the projection of the environment onto the retina. Vision scientists conceive of the experience of the depth dimension as a construction based on a variety of depth cues—i.e., properties of the retinal image that signify variations in depth. It is useful to categorize depth cues according to their cause: (a) cues based on triangulation (i.e., seeing the world from different vantage points), (b) cues based on perspective projection (e.g., linear perspective, texture gradient, relative size), and (c) cues based on light transport and interaction with materials (e.g., shading, atmospheric effects, occlusion).

Occlusion occurs when one object partially blocks the view of another object. The conventional wisdom is that occlusion indicates the order of distances to the occluding and occluded objects, but nothing more (but see Burge, Fowlkes, & Banks, 2010). Even if it provides only order information, occlusion is nonetheless a powerful depth cue. An example is provided by the pseudoscope, an optical instrument that presents the

Citation: Zannoli, M., Love, G. D., Narain, R., & Banks, M. S. (2016). Blur and the perception of depth at occlusions. Journal of Vision, 16(6):17, 1–25, doi:10.1167/16.6.17.

doi: 10.1167/16.6.17

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License. Downloaded From: http://jov.arvojournals.org/pdfaccess.ashx?url=/data/Journals/JOV/935165/ on 05/19/2016

1

 $\sim$ 

2

left eye's image to the right eye and the right eye's image to the left-i.e., it reverses the sign of binocular disparity. Many people who view the natural world through a pseudoscope do not notice that something is amiss. Figure 1 provides an example. With cross fusing, the upper panels provide correct disparities and the lower one reversed disparities. The disparity reversal is evident in some parts of the image but not others. For example, the large statue, which is actually farther than the woman, appears nearer than the woman as dictated by the reversed disparities. But for many viewers the bookcase, which is farther than the woman, continues to look farther because the woman occludes the background. In this case, occlusion appears to have more influence on the depth interpretation than disparity.

But when an occlusion border is detected (due to a difference in texture, color, etc.), the depth order still has to be determined. How is this done? If the surfaces are viewed binocularly, disparity indicates the order and therefore which surface is the occluder. However, the disparity gradient (the rate of change of disparity relative to change in position) is very large near the border. When the gradient exceeds a value of  $\sim 1$ , thereby exceeding the disparity-gradient limit (Burt & Julesz, 1980), disparity cannot be estimated, and as a consequence, depth cannot be estimated either (Filippini & Banks, 2009). The estimation failure will occur near the occlusion border, so the depths have to be inferred from background points displaced from the border. If the viewer is moving or one of the surfaces is moving, the accretion and deletion of texture near the border between the surfaces also indicates which one is the occluder (Gibson, 1966). But in some cases, neither disparity nor motion parallax is available, so a viewer has to rely on other information to determine the depth order. T-junctions can be informative (Cavanagh, 1987), but such junctions are often not well delineated in the retinal image. Blur is potentially useful because it is in principle nearly always informative about which surface is nearer and therefore which surface is occluding the other. In computer-generated imagery in particular, the blur in the image can be directly controlled, so it would therefore be useful to know whether reproducing the properties of natural retinal blur improves the realism of synthetic images. In this article we examine blur around occlusion borders and ask whether and how human observers use this signal to determine depth order. We also investigate some other phenomena in image formation near an occlusion border.

For now we consider geometric optics only to describe how occlusions affect the formation of the retinal image. Figure 2 illustrates the geometry. The eye is represented by a single lens, an aperture, and an image plane. It is focused at distance  $z_0$ . An object at

that distance creates a sharp image in the image plane. Objects nearer or farther than the focal plane create blurred retinal images. Image blur is quantified by the diameter of the retinal image of a point object. The diameter of the blur circle is

$$b = As_0 |\frac{1}{z_0} - \frac{1}{z_1}| = As_0 |\Delta D|, \quad (1)$$

where A is pupil diameter,  $s_0$  is the distance from the lens plane to the image plane,  $z_1$  is the distance to the object creating the blurred image (all in meters), and  $\Delta D$  is the difference (in diopters) between the distances  $z_0$  and  $z_1$ . The absolute value of the scene term  $\Delta D$  is used because, from a geometric standpoint, blur is unsigned. We can simplify Equation 1 by using the small-angle approximation:

$$\beta \approx A |\Delta D|, \quad (2)$$

where  $\beta$  is the blur-circle diameter in radians.

The blurs created by  $z_1$  (red) and  $z_2$  (blue) in Figure 2 are the same even though the image of  $z_1$  is formed behind the retina and that of  $z_2$  is formed in front. Because defocus blur is unsigned, it cannot by itself indicate whether the object creating the blurred image is nearer or farther than the object on which the eye is focused. But depth order can in principle be determined when an occlusion is present (Marshall, Burbeck, Ariely, Rolland, & Martin, 1996; Mather, 1996). Figure 3 illustrates this by showing image formation when the eve is focused on the background plane or on the occluding plane. When the eye is focused on the background (upper panel), the image of the texture of that surface is sharp while the image of the texture on the occluding surface is blurred. The edge of the occluding surface is also out of focus, so its image is blurred. If the eye is focused at the distance of the occluding surface (lower panel), the images of the texture on that surface and the occlusion border are sharp. Thus, there is a generally reliable relationship between the relative blur of the images of the two surfaces, the blur of the occlusion border, and depth order. (In the unlikely case that the eye is focused at the dioptric midpoint between the background and occluding planes, the blur of the occlusion border is the same no matter which surface is the occluder, so border blur becomes uninformative.)

Several researchers have investigated whether human viewers use the blur information at occlusions to determine depth order. For example, Marshall et al. (1996) presented stimuli that depicted two planes, one occluding the other, with a vertical border in between as in Figure 3. Blur was rendered using a Gaussian blur kernel. Participants viewed the stimuli monocularly and reported which of the two sides appeared closer (equivalent to asking which is the occluding surface). Despite the fact that the blur of the occlusion border



Figure 1. Stereograms with normal and reversed disparities. If one cross fuses the images, the upper panel has the correct relationship between disparity and other depth cues, and the lower panel has reversed disparities that put them in conflict with other depth cues. Some parts of the stereogram with reversed disparities have a notably different depth interpretation than the corresponding parts of the stereogram with nonreversed disparities. For example, the large statue between the chest of drawers and the woman appears to be nearer than the woman when the disparities are reversed and farther than the woman when they are not reversed. Many parts of the reversed-disparity stereogram, however, have a similar depth interpretation to the corresponding parts in the nonreversed stereogram. For example, the woman's position relative to the bookcase behind her seems similar in both stereograms. The animal carpet appears nearer than the textured carpet in both stereograms. When occlusion is present (the woman occluding the bookcase, the animal carpet occluding the larger carpet), the depth interpretation tends to be consistent with occlusion and not disparity. (Produced by Underwood & Underwood. Available at: http://loc.gov/pictures/resource/ppmsca.08781.)



Figure 2. Image formation in a simple eye around an occlusion border. The diagram is a top view, which will be adopted for all such diagrams in this article. The value  $z_0$  is the focal distance of the eye given focal length f and distance  $s_0$  from the lens to the image plane. The black lines represent the light rays entering the eye to form a sharp image. The value  $z_1$  is the distance to the occluding border, and  $s_1$  is the distance to where the image of the border is formed. Those distances are represented by red arrows, and the light rays by dashed red lines. The value  $z_2$  is the distance to the background, and  $s_2$  is the distance to where the image of the background is formed. The pupil diameter is A; is the diameter of the blur circle of the image of a point at distance  $z_1$  is b. Those distances are represented by blue arrows, and the light rays by dashed blue lines.

was a completely reliable indicator, the majority of participants could not perform the task reliably. Two of the five participants were strongly biased to report that the surface with the blurred texture was closer, whether the border was blurred or not. One of the five had the opposite bias, reporting that the surface with the sharp texture was closer. Mather and Smith (2002) and Palmer and Brooks (2008) conducted similar experiments with similar stimuli. Their participants also had difficulty performing the task. Many reported that the surface with the sharp texture was closer regardless of whether the occlusion border was blurred or not. In summary, these three studies reported inconsistent depth-order judgments, with different participants exhibiting different biases. The findings are surprising because the relative blur of the occlusion border and the surface textures were easily discriminable, so depth order was in principle easy to determine.

Some properties of the images in these studies are not representative of the retinal images created by occlusions in natural viewing.

First, the stimuli were displayed on a single plane, so blur was rendered into the stimulus rather than created by the eye. In rendering the stimuli, the researchers set the focal distance to either the distance of the near occluding plane or the far background plane. The retinal image could thus only be correct if the viewer accommodated to the distance of the display screen. If the viewer accommodated nearer or farther than that, the images of the occluding and background planes would both become blurrier. Of course, this is not what happens in natural vision when a viewer accommodates on one plane and then the other.

Second, the blur kernels employed to render the stimuli were Gaussian, which is not appropriate for simulating defocus blur. The aim in these studies and the current one is to produce stimuli that will yield a retinal image similar to that produced by viewing a real scene. The image of a real scene is affected by defocus blur as well as by diffraction and higher order aberrations. As the eye becomes more and more defocused, the defocus component becomes the dominant effect, while the other effects remain roughly constant (Wilson, Decker, & Roorda, 2002). To a first approximation, the latter effects are independent of defocus, so the total blur is a combination of the two. Then, assuming that participants accurately focus on the stimulus, one should add only blurring due to simulated defocus, because the other effects will be inserted by the viewer's eye. In particular, if the simulated scene is in focus, one should not insert any blur into the stimulus. In the general case, to determine the blur due to defocus alone, consider geometric optics and imaging of a point object in the world. The light rays from the object form a cone in the eye, the crosssection of which is a circle. This effect is captured by the cylinder function, not the Gaussian.

Third, the effects of other optical aberrations were not accurately simulated. Consider, for example, longitudinal chromatic aberration (Bedford & Wyszecki, 1957). In natural vision, the color fringes that are produced by the eye's chromatic aberration differ according to the distance of the object relative to the



5



Figure 3. Defocus blur in the presence of occlusion. The upper and lower panels indicate retinal-image formation when the eye is focused, respectively, on the background plane and the occluding plane. In the upper panel, the retinal image of the texture of the background is sharp and the occlusion border is blurred. The rays associated with the sharply focused background are represented by the black lines. The rays associated with the blurred occluder are represented by the dashed red lines. In the lower panel, the retinal image of the texture of the occluder is sharp and the border is sharp. The rays associated with the sharply focused occluder are represented by the black lines. In the lower panel, the retinal image of the texture of the occluder is sharp and the border is sharp. The rays associated with the sharply focused occluder are represented by the black lines. Those associated with the blurred background are represented by the dashed blue lines (thinner for the ray that would in reality be blocked by the occluder).

focal plane. We know that longitudinal chromatic aberration can be used to drive accommodation (Fincham, 1951; Kruger, Mathews, Aggarwala, & Sanchez, 1993), so it is quite plausible that appropriate chromatic effects affect the perception of depth in general and at occlusion borders in particular. With the single-plane displays of the previous studies, the chromatic effects are consistent with a scene consisting of one plane only.

Because of these potentially important departures from the properties of retinal images in the natural environment, we reexamined the perception of depth at occlusions by comparing the ability to determine depth order with single-plane displays and with a multiplane display (Love et al., 2009; Narain et al., 2015) in which the occluding and background surfaces are presented at different focal distances and the viewer's eye creates the blur.

## **Experiment 1**

#### Methods

#### Observers

Thirteen young adults, 24–32 years of age, participated. They gave informed consent under a protocol approved by the Institutional Review Board of the University of California, Berkeley, consistent with the Declaration of Helsinki. The data from four of them were discarded because they could not perform the task consistently. Thus, the results presented here are from the remaining nine participants. Of them, five had myopia, one had emmetropia, and three had hyperopia. The people with myopia wore their optical corrections while doing the experiment. Six of the nine participants were male, and the other three were female.



Figure 4. Schematic of the multiplane display system. The switchable-lens systems (indicated by rectangles) consist of two birefringent (calcite) lenses (blue), two ferroelectric liquid-crystal polarization modulators, a linear polarizer, and a glass ophthalmic lens. Each eye views a CRT display via the switchable-lens system and a prism with a front-surface mirror. The lens control unit detects light pulses in the corner of each CRT to synchronize the changes in the focal power of the lens system to the displays.

#### Apparatus

To investigate the perception of occlusions, we needed to be able to present accurate focus cues. To this end, we used the multiplane display described by Love et al. (2009). Figure 4 is a schematic of the apparatus. The display is stereoscopic, but in the experiments reported here, all images were viewed monocularly. Images were presented on a cathode-ray tube (CRT) display (56-cm Iiyama HM204DT) viewed with a front-surface mirror. A switchable lens system was positioned between the eye and mirror to enable manipulation of focal distance. The key element is a birefringent lens. Birefringent materials have two indices of refraction, one for light polarized along one crystalline axis and the other for light polarized along the orthogonal axis. When such material is cut into a lens shape, it can take on one of two focal powers depending on the light's linear polarization angle. To implement the change in focal power, we manipulate the polarization angle using liquid-crystal polarization modulators. By stacking two modulator-lens pairs, we obtain four discrete focal powers separated by 0.6 D. We synchronize the switchable lens system to the CRT so that the system adjusts focal distance to an assigned value as an image is displayed on the CRT at the same time. The displayed image at a given time contains the range of distances in the simulated scene that is appropriate for the current focal state of the lens system. By cycling the lens and imagery at 180 Hz, the full volume is displayed at 45 Hz. The display's workspace covers 1.8 D, but that space can be translated forward and backward by adding a fixed lens.

Because the displayed images are a discrete approximation to a volume of light, the display creates nearly correct focus cues no matter where in the workspace the viewer's eye is accommodated. Thus, defocus blur is created within the eye and varies appropriately with changes in accommodation. Stimuli presented in such displays drive accommodation effectively even when they are presented between presentation planes (Mac-Kenzie, Hoffman, & Watt, 2010).

In the central  $10^{\circ}$  of the visual field, image quality (assessed by measurements of the modulation transfer function) is comparable to that of a high-quality singlelens reflex camera, although the quality varies somewhat across presentation planes (Love et al., 2009). The optics of the system has minimal longitudinal chromatic aberration: The difference in focal distance from 450 to 650 nm is less than 0.05 D.

Participants were positioned relative to the apparatus with a bite bar in order to place the viewing eye correctly on the optical path. This was accomplished by a combination of a sighting technique to locate the eye relative to the bite bar (Hillis & Banks, 2001), positioning of the bite bar relative to the display system, and software alignment once the participant was in place (Akeley, Watt, Girshick, & Banks, 2004).

For all but the quite unlikely case that the distance of a point in the simulated scene coincides exactly with the distance of one of the presentation planes, a rule is required to assign image intensities to presentation planes. In previous work we used a depth-weighted blending rule in which the image intensity at each presentation plane is weighted according to the dioptric distance of the simulated point from that plane (Akeley et al., 2004; Love et al., 2009). This per-pixel blending rule works well for diffuse surfaces in scenes in which depth varies slowly across the image, but it does not produce accurate results at occlusion boundaries.

To generate more accurate results for occlusions, we developed a new blending algorithm that is described in detail by Narain et al. (2015). The goal of the algorithm is to best reproduce the retinal images that would occur when a person views a real three-dimensional scene and accommodates through it. Using a model of image formation in the eye, we obtain the focal stack of retinal images that would be seen by the viewer when accommodating to different distances. We then optimize the assignment of light intensities to presentation planes so that the retinal images seen in the display, predicted with the image-formation model, are as close as possible to the retinal images of the original scene across a range of accommodative distances. This approach greatly minimizes visible artifacts at occlu-



Figure 5. Stimuli in Experiment 1. The left side of the figure illustrates the presentation of the single-plane stimuli. The upper part of the left side illustrates the presentation when the sharp texture was on the near surface and the blurred texture on the far surface. The stimulus in this case was presented on the near presentation plane at 3.2 D. The upper panel in the middle provides an example of what that stimulus would look like when the viewer accommodates to 3.2 D. The lower part of the left side of the figure illustrates the presentation when the blurred texture was on the near surface and the sharp one on the far surface. The stimulus in this case was presented on the far presentation plane at 2.0 D. The lower panel in the middle provides an example of how that stimulus would appear when the viewer accommodates to 2.0 D. The right part of the figure illustrates the presentation of the multiplane stimuli. The two surfaces are presented on different presentation planes at 3.2 and 2.0 D. The upper part of the right side illustrates the situation when the viewer accommodates to the near surface at 3.2 D. The upper panel in the middle provides an example of how that stimulus would appear. The lower part of the right side illustrates the situation when the viewer panel of the right side illustrates the situation when the viewer panel of the right side illustrates the situation when the viewer accommodates to the near surface at 3.2 D. The upper panel in the middle provides an example of how that stimulus would appear. The lower part of the right side illustrates the situation when the viewer accommodates to the far surface at 2.0 D. The lower panel in the middle provides an example of how that stimulus would appear. The lower part of the right side illustrates the situation when the viewer accommodates to the far surface at 2.0 D. The lower panel in the middle provides an example of how that stimulus would appear. The lower panel in the middle is an example of how that stimulus would appear. The green sh

sion boundaries, so we used it in the experiments reported here.

The multiplane display can of course be used as a conventional single-plane display. We do this by displaying the stimulus on just one of the four presentation planes.

#### Stimuli

The experimental stimuli (Figure 5) depicted two opaque frontoparallel surfaces with distinct textures. The textures on the surfaces were chosen randomly from four precomputed ones: Three were generated from photographs of food and one from a Voronoi diagram. They had the same space-average luminance and contrast energy, and similar amplitude spectra. One surface occluded the other, and the occlusion border was sinusoidal. The stimuli were viewed through a 10° circular aperture. All three primaries in the CRT were illuminated so that the stimulus appeared gray.

The stimuli were presented in two ways. The first presentation method employed conventional single-

plane rendering and display, as shown on the left in Figure 5. In this case, the stimuli were presented on one presentation plane at either 2.0 or 3.2 D. The stimuli were generated using Mitsuba, a conventional ray tracer (Jakob, 2010). The virtual camera was given an aperture of 4, 5, 6, or 7 mm to encompass the pupil diameters we measured in situ in our participants. We used different diameters to be sure that one of the simulated values matched the participant's actual diameter. The aperture is simulated by sampling a disk with 100-200 points (Cook, Porter, & Carpenter, 1984). The samples are randomly jittered to avoid alignmentrelated artifacts such as aliasing. Each sample is directed to the image plane such that scene points nearer or farther than the focused distance generate blur. The resulting blur kernel is a cylinder, which is a better approximation than a Gaussian function for modeling defocus in the human eye (as discussed earlier). The left and right halves of the stimulus were displayed on the same presentation plane. To simulate accommodation to the near surface at 3.2 D, the scene was rendered by focusing the virtual camera to that

distance and then displaying the stimulus on the presentation plane at 3.2 D. To simulate accommodation to the far surface at 2.0 D, the camera was focused to that distance and the stimulus displayed on the presentation plane at 2.0 D. This rendering and display technique produces retinal images that are nearly correct, provided that the viewer is accommodated to the presentation plane. For example, when the virtual camera is focused on the near surface, its texture is sharp, the occlusion border is sharp, and the far texture is noticeably blurred. This creates a reasonable approximation to what a viewer would see with a real occlusion and with accommodation to the near surface. (We point out later that chromatic effects in singleplane rendering and display are incorrect even when accommodation is appropriate.) If the viewer does not accommodate to the distance of the presentation plane, defocus blur is introduced to all parts of the stimulus. and the retinal images become quite incorrect for the simulated scene.

The second presentation method employed multiplane rendering and display, as shown in the right part of Figure 5. In generating the multiplane stimuli, we used pupil diameters of 4, 5, 6, and 7 mm in the optimized blending algorithm. Again, we did this to be sure that one simulated diameter matched the participant's actual diameter. The simulated distances of the near and far surfaces were 3.2 and 2.0 D, respectively, corresponding to the distances of the first and third presentation planes. Although the distances of the simulated surfaces correspond to the distances of the first and third planes in the apparatus, the optimized blending algorithm required significant illumination of pixels in the second plane (and even some in the fourth) to create a realistic impression of an occlusion. The 1.2-D separation between the near and far surfaces is much larger than the minimum separation required to produce discriminable differences in image sharpness (Campbell & Westheimer, 1958; Sebastian, Burge, & Geisler, 2015) and much larger than is required to drive accommodation (Campbell & Westheimer, 1959; Kotulak & Schor, 1986; MacKenzie et al., 2010). When participants accommodated to the near surface (3.2 D), the far surface (2.0 D) was noticeably blurred and the occlusion border appeared sharp. When they accommodated to the far surface, the near surface and border were noticeably blurred.

The retinal images produced in the single- and multiplane conditions were similar when the eye was accommodated to the appropriate distance and the actual pupil diameter corresponded to the one used to create the stimuli. The space-average luminance of the single- and multiplane stimuli was  $0.95 \text{ cd/m}^2$ .

On each trial, the textures for the two surfaces were randomly selected from the precomputed textures with the constraint that the textures on the two sides were never the same. On each trial, the side containing the occluder (left or right) was chosen randomly.

#### Procedure

Observers viewed the stimuli monocularly with their preferred eye. Before each stimulus presentation, an accommodation and fixation stimulus was presented. It was a small black "E" surrounded by black lines of random size and orientation on a gray background. The space-average luminance of the accommodation and fixation stimulus was the same as that of the experimental stimuli. The accommodation and fixation stimulus was presented for 1 s at either 2.0 or 3.2 D. Participants were told to look at the "E" and make it sharp. The accommodation and fixation stimulus was then replaced by the experimental stimulus, which was presented for 300 ms or 3 s. Observers then indicated with a key press whether the left or right half of the stimulus appeared nearer. No feedback was provided. One session was run for each participant. Each session lasted  $\sim 60$  min.

In a separate session, each participant's pupil diameter was measured during viewing of the experimental stimuli. Participants first adapted to the illumination of the experimental room for a few minutes. Then they viewed the experimental stimulus with the eye they used in the main experiment while we photographed the nonviewing eye. The average pupil diameter across participants was 6.2 mm (SD = 0.6 mm).

#### **Experimental** conditions

As already mentioned, there were two types of presentation (single- and multiplane), two accommodative distances (2.0 and 3.2 D), four simulated pupil diameters (4, 5, 6, and 7 mm), and two stimulus durations (300 ms and 3 s). The latency for voluntary accommodation is 300-500 ms (Kasthurirangan, Vilupuru, & Glasser, 2003; Schor, Lott, Pope, & Graham, 1999). Thus, the longer duration of 3 s was intended to allow changes in voluntary accommodation that could enable participants to determine depth ordering in the multiplane presentations by perceiving the relationship between accommodation and blur. The shorter duration of 300 ms was designed to not allow voluntary accommodation, so participants could not use the accommodation-blur relationship to determine depth ordering. Of course, they may have been able to use image changes due to accommodative microfluctuations, which have a period of 500-1000 ms (Charman & Heron, 2015). In all, there were 32 types of trials. Each type was presented 10 times, for a total of 320 trials per participant. Those trials were divided into four blocks, two with the short duration and two with the long, and



Figure 6. Depth-order judgments in the single-plane condition. Left: The proportion of correct judgments of depth order is plotted for each participant as a function of whether the occlusion border was blurred or sharp. The black circles represent the individual participant data and the red ones the averages across participants. Chance performance would be 0.5. Right: The proportion correct in the single-plane condition when the occlusion border is blurred versus sharp. The abscissa is the proportion correct when the occlusion border was blurred, and the ordinate is the proportion correct when the border was sharp. The black circles represent the data for each participant and the red circle the average across participants. If participants always responded correctly, the data would lie in the upper right corner at (1, 1). If they responded randomly with no bias, the data would lie in the middle at (0.5, 0.5).

those 40-trial blocks were presented in random order. Participants completed the experiment in one session of approximately 40 min.

#### Results

We first conducted a repeated-measures ANOVA on the proportion of correct responses with presentation type (single- vs. multiplane), fixation distance (3.2 vs. 2.0 D), simulated pupil size (4, 5, 6, or 7 mm), and stimulus duration (300 ms vs. 3 s) as within-participant variables. There was no significant effect of duration, F(1, 8) = 3.62, p = 0.09, which means that participants did not perform significantly better at the longer duration, although they had a tendency to do so. There was also no significant effect of simulated pupil size, F(3, 24) = 0.73, p = 0.54, which means that the amount of blur had no systematic effect on the depth-order judgment. There was also no significant effect of fixation distance, F(1, 8) = 0.16, p = 0.70, which means that there was no overall tendency to see surfaces with blurred or sharp textures as near. There was, however, a very significant effect of presentation type, F(1, 8) =29.6, p < 0.001: The proportion of correct responses was consistently greater when the stimuli were presented on multiple planes than when they were presented on one plane.

It is illuminating to examine the single-plane results by themselves. Recall that there is a perfectly reliable cue for determining which surface is in front in this condition: If the occlusion border is blurred, the surface with the blurred texture is very likely to be in front; if the border is sharp, the surface with the sharp texture must be in front. The left panel of Figure 6 plots for each participant the proportion of correct responses as a function of whether the occlusion border was blurred or sharp. The data in the figure have been combined across duration and pupil size. Chance performance is 0.5. The mean proportions correct were 0.62 and 0.57 when the border was, respectively, blurred and sharp, so participants performed only slightly better than chance. These results are in general agreement with those of Marshall et al. (1996), Mather and Smith (2002), and Palmer and Brooks (2008). We conclude, as they did, that participants have difficulty making reliable judgments of depth order from rendered focus information alone.

There were distinct differences in judgments across participants. Some participants were very likely to report that the surface with the blurred texture was near, so they were correct on nearly every trial when the occlusion border was blurred and incorrect on nearly every trial when it was sharp. Other participants had the opposite tendency, so they performed much better when the border was sharp. The differing biases are



Figure 7. Proportion of correct depth-order judgments in the multiplane condition. Left: The proportion of correct judgments of depth order is plotted for each participant as a function of whether the occlusion border was blurred or sharp. The black circles represent the individual participant data and the red ones the averages across participants. Chance performance would be 0.5. Right: Proportions correct in the multi- and single-plane conditions when the occlusion border is blurred versus sharp. The abscissa is the proportion correct when the occlusion border was blurred, and the ordinate is the proportion correct when the border was sharp. The unfilled and filled black circles represent the single- and multiplane data, respectively, for each participant, and the unfilled and filled red circles the averages across participants in each condition (also plotted in Figure 6). The filled black circles represent the multiplane data for each participant, and the filled red circle the average across participants in that condition. The individual participant data in the two presentation conditions are connected by lines. The thin gray lines represent expected data for different biases and amounts of edge usage. The diagonals from upper left to lower right represent different values of  $P_{useEdge}$ , and the lines that converge in the upper right corner represent different values of  $P_{blurNear}$ .

evident when the data are replotted with the abscissa representing the proportion of correct responses when the border was blurred and the ordinate the proportion of correct responses when the border was sharp (Figure 6, right). The data for participants with a bias to see blurred texture as near fall in the lower right quadrant, and data for participants with the opposite bias fall in the upper left quadrant.

One can model the participants' behavior by assuming that they use the information in the border on some proportion of trials and that on the other trials they base judgments on their bias (e.g., to see blurred textures as near). Under those assumptions, the proportion correct on trials in which the occlusion border is blurred is given by

$$PC_{\text{blurredEdge}} = P_{\text{useEdge}} + (1 - P_{\text{useEdge}})P_{\text{blurNear}},$$
(3)

where  $P_{\text{useEdge}}$  is the proportion of trials in which they use the edge information (and therefore respond correctly; we describe an alternative at the end of the Results) and  $P_{\text{blurNear}}$  is the bias to see blurred textures as near. The value of  $P_{\text{blurNear}}$  varies from 0 to 1. Likewise, the proportion correct on trials in which the occlusion border is sharp is given by

$$PC_{\text{sharpEdge}} = P_{\text{useEdge}} + (1 - P_{\text{useEdge}})(1 - P_{\text{blurNear}}).$$
(4)

If participants always used the edge information, the data in the right panel of Figure 6 would lie in the upper right corner at (1, 1). If they never used the information, the data would lie along a diagonal from (0, 1) to (1, 0).

We next examine the results from the multiplane condition. The left panel of Figure 7 plots the proportion correct when the occlusion border was blurred or sharp. It was blurred when the accommodative stimulus was on the far plane (2.0 D) and sharp when the accommodative stimulus was on the near plane (3.2 D). The proportion correct was consistently higher in multiplane presentations (0.83 and 0.82 for blurred and sharp, respectively) than in single-plane (0.57 and 0.62). As the aforementioned ANOVA results show, the improvement in performance was statistically significant.

The right panel of Figure 7 plots the multiplane data from each participant along with their single-plane data. Clearly, each participant's performance was



Figure 8. Blur bias and edge usage in the single- and multiplane conditions. Left: The abscissa is the estimated blur bias for each participant in the single-plane condition, and the ordinate is the estimated bias in the multiplane condition. Right: The abscissa is the estimated edge usage in the single-plane condition, and the ordinate is the estimated edge usage in the multiplane condition.

better in the multiplane condition than in the singleplane. The improvement even occurred at the shorter duration (300 ms), which is too brief for voluntary accommodation to have occurred. These data show that when the blur is created by the participant's eye rather than in the rendering of the stimulus, people are much better at judging depth order.

Participants exhibited the same bias in both presentation conditions. This is evident in the left panel of Figure 8, which plots estimated  $P_{\text{blurNear}}$  for each participant in the two conditions. The correlation between the two values was significant, r(7) = 0.863, p < 0.01.

The overall improvement in performance was presumably due to participants using the additional focus information created in the eye rather than rendered into the stimulus. The right panel of Figure 8 illustrates this by plotting estimated  $P_{\text{useEdge}}$  for each participant in the two presentation conditions. The value of  $P_{\text{useEdge}}$  is systematically higher in the multiplane condition than in the single-plane condition.

The fact that there was no effect of simulated pupil size may seem surprising, but we do not think it is. The simulated diameter has a large effect on the single-plane stimuli (larger diameters leading to larger rendered blurs). However, participant performance with singleplane stimuli hovered around chance, suggesting that participants did not make good use of the blur whether it was small or large. The simulated pupil size affected the rendering of the multiplane stimuli, but the differences were small and all near the occlusion border. They were small because the two simulated surfaces were presented at different focal distances, so to first approximation, the blur in the retinal image was determined by the viewer's own pupil size and not by the simulated size.

We have described the multiplane results as if they show that viewers use the blur information near the occlusion border to determine depth order. But the results could also be interpreted as showing that viewers use the blur signals from the textured surfaces rather than from near the border. This second interpretation cannot be ruled out. One could test it by blacking out the region between the two surfaces so that no occlusion boundaries were visible. We point out, however, that depth changes in the natural environment are very commonly accompanied by occlusion and that depth changes with no information between two surfaces are very uncommon. In addition, we have evidence from a pilot experiment that the occlusion border matters to the judgment. In preliminary testing, we presented multiplane stimuli with linear depth-weighted blending (Akeley et al., 2004) and optimized blending (Narain et al., 2015). The retinal images near the occlusion border differed significantly between linear and optimized blending, but the images were essentially identical away from the border. Although we did not collect data formally. subjects reported having difficulty judging depth order with linear blending but not with optimized blending. This suggests that the occlusion itself is important to making depth estimates in the data reported here. So we choose to describe the multiplane results as due to use of edge information even though we cannot rule out the other logical interpretation. In either case, the results show convincingly that blur created in the viewer's eye yields much better depth ordering than blur rendered into the stimulus.

We next turn to the question of what additional focus information was available and used in the multiplane condition. Specifically, we examine two cues that plausibly produce more useful information in the multiplane condition than in the single-plane: longitudinal chromatic aberration and microfluctuations of accommodation.

## **Experiment 2**

Longitudinal chromatic aberration occurs in humans because the eye's refracting elements have different refractive indices for different wavelengths. The change in focal power from 400 to 700 nm is  $\sim 2$  D (Bedford & Wyszecki, 1957). When the eye views a depth-varying scene, chromatic aberration produces different color fringes for different object distances relative to the eye's focal distance. A bright object at a distance greater than the current focal distance yields a sharper image for long wavelengths (e.g., red), so the color fringes produced by a bright object are blue. Likewise, an object at a distance shorter than the focal distance yields red fringes. These color fringes are generally not perceived, but there is clear evidence that they affect visual function. For example, most viewers cannot accommodate to monochromatic stimuli even though they are perfectly able to accommodate to polychromatic stimuli (Fincham, 1951; Kruger et al., 1993). In addition, most viewers are better able to determine which of two real surfaces is nearer when the edges are illuminated with polychromatic light rather than monochromatic light (Nguyen, Howard, & Allison, 2005). Generally, single-plane stimuli contain no chromatic aberration effects, but those effects are introduced by the viewer's eye. The effects will be similar throughout the image, so this is information that the object is a plane that does not vary in depth. Chromatic aberration is appropriately reproduced in our multiplane display system (which has almost no chromatic aberrations of its own) because objects at different simulated distances are presented at different focal distances. If the eye changes accommodation, the changes in chromatic effects are approximately correct. Thus it is plausible that the nearly correct handling of chromatic aberration in the multiplane display provided the additional information required for depthorder judgments.

Microfluctuations of accommodation are involuntary variations in focal power with an amplitude of ~0.25–0.5 D (Campbell, Robson, & Westheimer, 1959; Winn & Gilmartin, 1992). There are high-frequency variations at 1–2 Hz and low-frequency ones (~0.6 Hz) with somewhat greater amplitudes. The high-frequency fluctuations are probably driven by cardiopulmonary responses, while the lower frequency ones seem to be driven by the ciliary muscle's action on the crystalline lens (Charman & Heron, 2015). With single-plane displays, changes in the retinal image due to accommodative microfluctuations provide no information about depth variation in the scene, because all image regions would come in and out of focus together no matter what the variation in scene depth is. In multiplane displays, however, image changes due to the fluctuations are potentially informative about variation in scene depth. They might therefore aid depth-order judgments when their effects on the retinal image are appropriate.

In Experiment 2 we investigated whether longitudinal chromatic aberration and accommodative microfluctuations affect depth-order judgments by manipulating the spectral bandwidth of the stimulus and the viewer's ability to accommodate.

## Methods

#### Observers

Seven young adults, one of whom participated in Experiment 1, participated in Experiment 2. They were 23–26 years of age. Six had myopia and one had emmetropia. Two were male and the other five were female. They gave informed consent under a protocol approved by the Institutional Review Board of the University of California, Berkeley.

#### Stimuli and procedure

The stimuli were very similar to those in Experiment 1 with a few important exceptions. To investigate whether chromatic aberration provides useful information for the perception of depth order, we manipulated the spectral bandwidth of the stimulus. Half of the stimuli were broadband with contributions from the R, G, and B primaries of the CRTs as in Experiment 1. The other half were less broadband, consisting of only the G primary. The space-average luminance of the broad- and narrower band stimuli was  $0.95 \text{ cd/m}^2$ , the same as in Experiment 1. The wavelength spectra of the two types of stimuli are shown in Figure 9. The green stimulus had a somewhat narrower spectrum than the gray stimulus, which reduces chromatic aberration fairly significantly because the largest dioptric difference is between short and long wavelengths, both of which were greatly attenuated by turning off the B and R primaries. Unfortunately, we could not present a narrower spectrum than the G primary, because doing so would have reduced the luminance to the mesopic range. Thus we were only able to reduce, not eliminate, contributions of longitudinal chromatic aberration.



Figure 9. Spectral distributions of the broadband (gray) and narrower band (green) stimuli. Normalized photopic luminance is plotted as a function of wavelength. The gray curve represents the distribution of the broadband stimulus, and the green curve the distribution of the narrower band stimulus.

To investigate whether accommodative microfluctuations provide useful information for the perception of depth order, we conducted the experiment with and without inducing cycloplegia (paralysis of accommodation). The tests with and without cycloplegia were done in separate sessions. In the cycloplegic sessions, we induced nearly complete cycloplegia in both eyes by topical application of two drops of tropicamide (1% ophthalmic solution). This eliminates voluntary accommodation and greatly reduces or eliminates the low-frequency microfluctuations. It is unclear, however, whether cycloplegia eliminates the high-frequency fluctuations of 1–2 Hz (Charman & Heron, 2015). Because accommodation was nearly paralyzed, participants could not voluntarily accommodate to different fixation distances as they did in Experiment 1. To allow them to focus on the near occluding plane or the far background plane as appropriate on a given trial, we set one eye's focal distance to the distance of the near plane (3.2 D) and the other eye's focal distance to that of the far plane (2.0 D). We accomplished this by inserting ophthalmic lenses in the optical path for each eye. Stimuli were presented randomly to one eye or the other. Halfway through the session (after  $\sim 25$  min), another drop of tropicamide was applied to both eyes to maintain cycloplegia, and the focal distances were swapped so that the left eye now viewed stimuli at the distance that the right eye had and vice versa. Participants were unaware of the focal-distance swapping. In the noncycloplegic sessions, we of course did not apply tropicamide to either eye and did not insert ophthalmic lenses into the optical path of the eye. We

attempted to make the viewing conditions as similar as possible to those in the cycloplegic sessions by having two fixation distances (2.0 and 3.2 D), one for each eye. As in the cycloplegic sessions, focal distances were swapped halfway through the experiment. Participants were of course free to make accommodative responses in the noncycloplegic sessions. All participants ran the tropicamide and no-tropicamide sessions on separate days. To avoid any confounding effects of training, half of them went through the no-tropicamide condition first. Because we did not find a significant effect of pupil size in Experiment 1, the same simulated pupil diameter (6 mm) was used in the tropicamide and no-tropicamide conditions. On every trial, a fixation target was first presented for 2 s to the eye that would be stimulated on the upcoming trial. The stimulus was then displayed for 300 ms or 3 s.

#### **Experimental conditions**

On each trial, there were random assignments of the occluding plane on the left or right, the left or right eye being stimulated (one at 2.0 D and the other at 3.2 D), single- or multiplane presentation, and broad- or narrower band spectrum. The textures on the occluding and background planes were also chosen randomly. The two stimulus durations were presented in blocks of 96 trials. Half of the observers performed the cycloplegic condition first; the other half performed it second. All possible combinations of these conditions were randomly interleaved. Each observer saw 24 repetitions of each possible combination of cycloplegic state, presentation type, fixation distance, wavelength spectrum, and stimulus duration.

#### Predictions

Experiment 2 was designed to determine if chromatic aberration or accommodative microfluctuations provide useful information for judging depth order. If chromatic aberration provides useful information, we expect performance with multiplane presentation to become poorer when the spectral bandwidth is decreased (i.e., going from gray to green). If microfluctuations provide useful information, we expect multiplane performance to be poorer with cycloplegia than without. If both cues are used, and either is individually sufficient, removal of both cues would cause a drop in performance but removal of either cue alone would not necessarily do so.

#### Results

We conducted a repeated-measures ANOVA on the proportion of correct responses with presentation type



Figure 10. Results from Experiment 2. Left: The proportion correct is plotted for single-plane and multiplane presentations. The data for each participant (black filled and unfilled circles) have been averaged across stimulus duration, fixation distance, spectral bandwidth, and cycloplegic state. The red symbols represent the across-participants averages. Right: The difference in proportion correct between the broad- and narrower band conditions in the single- and multiplane conditions. The black unfilled symbols represent the difference for individual participants in the single-plane condition, and the black filled symbols the difference for the multiplane condition. The red symbols represent the means. Larger values indicate that performance was worse with the narrower band stimulus (green) than with the broadband stimulus (gray). The difference between presentation conditions was statistically significant (Wilcoxon signed-rank test: p < 0.01, one tailed).

(single- vs. multiplane), stimulus duration (300 ms vs. 3) s), fixation distance (2.0 vs. 3.2 D), bandwidth (gray vs. green), and cycloplegic state (tropicamide vs. no tropicamide) as within-participant variables. There was a significant effect of presentation type, F(1, 7) = 18.39, p < 0.01: The proportion of correct responses was consistently greater when the stimuli were presented on multiple planes than when they were presented on one plane. This is illustrated in the left panel of Figure 10, which plots proportion correct for single- and multiplane presentations combined across the other variables. There were no significant effects of duration, F(1,7) = 0.00, p = 0.96, fixation distance, F(1, 7) = 0.233, p = 0.2330.64, or cycloplegic state, F(1, 7) = 0.15, p = 0.70. There was a significant effect of stimulus bandwidth, F(1, 7) =51.54, p < 0.001: The proportion correct was greater when the stimulus was broadband (gray) than when it was narrower band (green). Importantly, there was a significant interaction between presentation type and stimulus bandwidth, F(1, 7) = 13.1, p < 0.05, because participants performed the same with the gray and green stimuli when they were presented on a single plane but performed better with the gray stimulus when they were presented on multiple planes. The right panel of Figure 10 plots the difference in proportion correct between the broadband (gray) and narrower band (green) stimuli when presented on single and multiple planes. Higher values indicate that performance was poorer with the narrower band stimulus. Although the effect is small, it is very consistent: The difference is larger for every participant in the multiplane condition than in the single-plane condition. The difference is

statistically significant (Wilcoxon signed-rank test: p < 0.01, one tailed). This means that reducing bandwidth adversely affected the ability to make depth-order judgments in the multiplane but not the single-plane condition. Presumably, the drop in performance in the multiplane condition would have been larger if we could have presented a very narrowband stimulus that would have eliminated chromatic aberration. The results thus indicate that chromatic aberration provides useful information for making depth-order judgments.

We did not observe a statistically reliable effect of manipulating the ability to accommodate. Specifically, performance in all conditions was similar whether cycloplegia had been induced or not. This suggests that accommodative microfluctuations do not play a role in the perception of depth order. But this conclusion may be unwarranted. Tropicamide affects the ciliary muscle, which drives the low-frequency component of microfluctuations. But tropicamide may not affect the highfrequency fluctuations, because they are driven by the cardiopulmonary system and therefore may not be affected by cycloplegic agents (Collins, Davis, & Wood, 1995; Davies, Wolffsohn, & Gilmartin, 2009).

Performance in the multiplane condition was generally lower in Experiment 2 than in Experiment 1. This was likely due to the dichoptic presentation involved in Experiment 2, which participants found somewhat challenging.

We used the model described by Equations 3 and 4 to estimate the blur bias  $P_{\text{blurNear}}$  of each participant. As in Experiment 1, there was a significant correlation between the estimates from the single- and multiplane



Figure 11. Video demonstrating depth-dependent chromatic fringes. The scene is two white squares on a dark-gray background. The left and right squares are at distances of 3.2 and 2.0 D, respectively-i.e., the one on the left is nearer. The rendered field of view is 3°. The model eye was given the longitudinal chromatic aberration of a typical human eye (Marimont & Wandell, 1994). Pupil diameter was 5 mm. Diffraction was not incorporated. The eye is initially focused on the far surface at 2.0 D. As the video plays, accommodation becomes nearer and then farther. Depending on where the eye is focused relative to the two surfaces, the colors of the fringes change. For example, when the eye is focused in between the surfaces, the fringes are red on the left and blue on the right because short wavelengths are in better focus on the left and long wavelengths are in better focus on the right. This visualization was created using the RGB primaries of the CRTs used in Experiments 1 and 2. Its appearance will of course vary depending on the primaries of the reader's display.

conditions, r = 0.92, p = 0.003, which shows that people retain their biases to report blurred or sharp textures as near even while their performance improves in the multiplane condition.

## Discussion

# Chromatic aberration and depth-order information

The key observation from Experiment 1 is that people are more able to determine depth order when the stimulus is presented on multiple planes rather than on a single plane. In Experiment 2 we investigated what additional signals available in multiplane presentation are used in making depth-order judgments. We found evidence that longitudinal chromatic aberration is used and that accommodative microfluctuations may not be. Here we further analyze the effects of chromatic aberration on the retinal images of depth-varying scenes, particularly scenes with occlusions.

Chromatic aberration occurs because the eye's refracting elements have different refractive indices for different wavelengths. When presented with spectrally broadband stimuli, the eye is usually best focused at  $\sim$ 580 nm. In that case, focal distance as a function of wavelength for the typical eye is, to a close approximation,

$$D(\lambda) = 1.7312 - \frac{633.46}{\lambda - 214.10}, \quad (5)$$

where  $\lambda$  is wavelength in nanometers and  $D(\lambda)$  is diopters as a function of wavelength (Marimont & Wandell, 1994).

When the eye views a depth-varying scene, longitudinal chromatic aberration produces different fringes for different object distances relative to the eye's focal distance. These color fringes are generally not perceived, but there is clear evidence that they affect visual function. For example, Fincham (1951) found that narrowing the wavelength spectrum adversely affected accommodation: Most viewers could not accommodate to monochromatic stimuli even though they were perfectly able to accommodate to polychromatic stimuli. Kruger and colleagues confirmed this and also showed that reversing the eye's usual longitudinal chromatic aberration causes some viewers to accommodate in the wrong direction (Aggarwala, Nowbotsing, & Kruger, 1995; Kruger et al., 1993; Lee, Stark, Cohen, & Kruger, 1999).

In natural viewing, chromatic aberration provides depth-order information at an occlusion border. For example, if the viewer's eye is focused on the occluding surface, bright points on the background create blue fringes, which signals that that surface is farther than the occluder. If the eye is focused on the background, bright points on the near occluding surface produce red fringes. These effects are not appropriately reproduced in conventional rendering for single-plane presentation. Generally, single-plane stimuli contain no chromatic aberration effects, so those effects are introduced by the viewer's eye when looking at the stimulus. The effects will be similar throughout the image, so this is information that the stimulus is a plane that does not vary in depth. Chromatic aberration is appropriately reproduced in our multiplane display because objects at different simulated distances are presented at different focal distances (and because the device itself introduces essentially no longitudinal chromatic aberration). When the image is formed by the eye's optics, the appropriate distance-dependent color fringes are produced. If the eye changes accommodation, the changes in the fringes are again approximately correct.

Figure 11 illustrates longitudinal chromatic aberration created by the typical eye. To generate the video, we rendered a simple scene consisting of two white



Figure 12. Partial occlusion. The values  $z_1$  and  $z_2$  are the distances to the occluding plane and background plane, respectively, and A is the pupil diameter. The shaded region on the background plane is partially occluded: Light rays from this region enter only part of the eye aperture, thereby altering the effective aperture. The value w is the width of this zone of partial occlusion and is measured in the direction orthogonal to the orientation of the occlusion border.

squares (left one at 3.2 D and right one at 2 D) on a dark-gray background. Wavelength-dependent blur was introduced by adjusting the distance between the model eye and the rendered scene based on the eye's chromatic aberration (Equation 5). The scene was rendered for wavelengths ranging from 400 to 700 nm in steps of 20 nm and for focal distances ranging from 1.8 to 3.4 D in steps of 0.2 D. For each focal distance, a composite RGB image was created by filtering the stack of wavelength-dependent rendered images with the RGB spectral responses of the CRTs used in Experiments 1 and 2. Therefore, appearance in this demonstration depends on the primaries of the reader's display or printer. Color fringes are predominantly red and blue for objects located respectively nearer and farther than the focus distance of the eye. As a result, the predominant colors of the fringes on the two squares are opposite when the eye is focused in between.

## **Partial occlusion**

In the presence of an occlusion, some of the background produces light rays that enter the whole pupil, so image formation from that region is unaffected by the presence of the occluding surface. Other parts of the background are, of course, completely blocked from view, so no rays enter the pupil. Background parts in between are partially occluded: Light rays enter only part of the pupil. As a consequence, the effective aperture is changed. Here we quantify those changes and their effect on the pointspread function (PSF) for defocus blur. We then describe some optical phenomena that occur near the occlusion border.

The geometry for partial occlusion is schematized in Figure 12. The width w of the partial occlusion zone is measured orthogonal to the orientation of the occlusion border:

$$v = A\left(\frac{z_2}{z_1} - 1\right).$$
 (6)

ı

Thus, the partially occluded region becomes larger as the aperture expands, the distance to the background increases, and the distance to the occluder decreases. It is unaffected by the eye's focal distance.

The PSF for defocus blur changes accordingly from a circle of diameter *b* in Equation 1 for nonoccluded points to a segmented circle with successively larger portions cut off for background points that are more and more occluded. This effect is schematized in Figure 13: A background point is partially occluded and its associated PSF is cut off to half the diameter of the circle. The eye is focused on the occluder at distance  $z_0$ . The width of the PSF (in radians; Equation 2) as a function of position *x* on the background is

$$\beta \approx A \left| \frac{1}{z_0} - \frac{1}{z_2} \right| \qquad \text{for } x \ge \frac{w}{2}$$
  
$$\beta \approx A \left| \frac{1}{z_0} - \frac{1}{z_2} \right| \left( \frac{x}{w} + \frac{1}{2} \right), \quad \text{for } \frac{-w}{2} < x < \frac{w}{2}. \quad (7)$$
  
$$\beta = 0 \qquad \text{for } x \le \frac{-w}{2}$$

The PSFs become anisotropic and therefore create astigmatic depth of field much like what happens with



Figure 13. Point-spread function (PSF) for defocus blur and partial occlusion. The eye is focused on the occluder at distance  $z_0$ . The partially occluded region on the background is indicated by yellow shading. Light rays from the indicated point on the background enter half of the pupillary aperture, so the PSF is cut off on one side creating a segmented circle. The width of the resulting PSF is  $b_{\text{partial}}$ .

the elongated pupils of many animal eyes (Banks, Sprague, Schmoll, Parnell, & Love, 2015). One can observe this effect in photographs with the camera not focused at the background distance: Contours on the background that are parallel to the occlusion border are less blurred than contours that are perpendicular to the border.

Images generated with a finite aperture contain spatial information that is lost in generation with a pinhole aperture. For instance, a camera with a finite aperture allows one to see behind occlusions, as in Figure 14. In the left panel the camera is focused on the near occluding rosebuds, and much of the sunflower in the background is blocked from view. In the right panel the camera is focused on the background sunflower and the whole flower becomes visible. As evident from Figure 13, this "see-through" effect depends on the size of the aperture, where the viewing device is focused, the size of the occluding object, and its distance relative to the background. With a pinhole aperture, the sunflower would be occluded no matter where the camera was focused.

To further examine image formation in the presence of an occlusion, we produced a physical optics model of



Figure 14. Demonstration of partial occlusion with a finite aperture. The photograph is of rosebuds in the foreground (the occluder) and a sunflower in the background. Camera focal length was 200 mm, with an *f* ratio of 5.6. The occluder is 1 D closer than the background. When the camera is focused on the occluder (left), the sunflower is occluded. When it is focused on the background (right), the whole sunflower becomes visible due to the partial occlusion effect in Figure 13.



Figure 15. Videos comparing two rendering and displaying techniques with natural viewing using the HDR-VDP-2 visibility metric. The output of HDR-VDP-2 was computed for the scene in Experiment 1. Simulated pupil diameter was 6 mm. The retinal images were calculated for white light. Longitudinal chromatic aberration was modeled; diffraction was not. The focal distance of the eye varied from 1.8 to 3.4 D in steps of 0.2 D. The reference images in all comparisons were the focal stack of retinal images of the real-world scene used in the optimization procedure. Those images are represented by luminance. The map of probability of discrimination is overlaid: Blue, green, and red denote increasing probabilities. Each video contains a schematic on the left showing the eye, the occluding and background planes, and where the eye is focused moment to moment. (A) Single-plane compared to natural viewing. The single-plane images were computed for an eye focused on the near plane at 3.2 D. As the eye focuses through the image stack, the discrimination probabilities are large except when the eye's focus is  $\sim$ 2.0 D. (B) Multiplane presentation with optimized blending compared to natural viewing. The discrimination probabilities are low for all focus distances and all regions of the stimulus.

an imaging system consisting of a simple eye (lens and detector at image plane), a distant background plane, and a closer occluding plane. The model includes the following features:

- It simulates the effects of diffraction, occlusion of the pupil, longitudinal chromatic aberration, and defocus due to objects being at different distances.
- It does not simulate monochromatic aberrations, nor off-axis aberrations because the imaging angles are small. The effects of these aberrations would be small compared to the optical effects of defocus and the occlusion border.
- It calculates the appropriate PSF for each point in the object when the eye is focused on the background and when it is focused on the occluder. The PSF is calculated by taking the square modulus of the Fourier transform of the aperture function, taking into account the effective pupil size, where the eye is focused, and longitudinal chromatic aberration.

Supplementary Movie S1 is a video that shows the PSFs generated by points on a background plane when they are not occluded or partially occluded, and how those PSFs change when the eye focuses at different distances. The video shows that when a point on the background is well away from the occluder, the eye's effective aperture is a disk and the PSF is not affected by the occlusion. Without occlusion, the PSF is small and greenish when the eye is focused on the background (because the eye was focused at 580 nm); the PSF when the eye is focused on the occluder is much larger and reddish (because long wavelengths are relatively better focused than medium and short wavelengths when the eye is focused nearer than the object). As the point moves closer to the occlusion boundary, the effective aperture becomes vignetted and the PSF is affected. The chromatic effects one sees are a consequence of rays passing through an eccentric part of the aperture such that longitudinal chromatic aberration produces a lateral effect. As the point moves yet farther behind the occluder, the effective aperture becomes highly vignetted and diffraction dominates both PSFs.

## Evaluating rendering and presenting techniques

We had to develop a rendering and display technique that produces the retinal images created by occlusions in the natural environment. Linear blending and a multiplane display work well for diffuse surfaces in a scene where depth varies slowly across the image. But it produces unacceptable haloing artifacts around occlusions (Akeley et al., 2004; Ravikumar, Akeley, & Banks, 2011). The optimized blending technique we used yields a much better approximation to real occlusions (Narain et al., 2015).

To quantify the accuracy of different rendering and display techniques, we assessed differences in the retinal images created by conventional single-plane rendering and displaying in comparison to the images created in natural viewing, and we also compared images created by optimized blending and multiplane displaying to natural viewing. To make the comparisons, we used the HDR-VDP-2 visibility metric (Mantiuk, Kim, Rempel, & Heidrich, 2011). HDR-VDP-2 estimates the discriminability of a pair of images (a reference and a test image). In particular, it estimates the probability that differences between two images will be visible to a typical human observer. Its output is a probability-ofdiscrimination map, where points in the map correspond to different regions in the stimuli. HDR-VDP-2 takes into account some important properties of human vision, including the contrast sensitivity function (and therefore the optical transfer function and neural transfer function), nonlinearities in the response to luminance, local light adaptation, and the change in effective contrast sensitivity at suprathreshold contrasts (Georgeson & Sullivan, 1975).

The results of the analyses are shown in the videos of Figure 15. Figure 15A shows discrimination probabilities when comparing single-plane presentation to natural viewing. Figure 15B shows the probabilities when comparing multiplane presentation with optimized blending to natural viewing. The real-world scene is represented by luminance, and hue represents discriminability. Red regions indicate parts of the stimulus that produce retinal images that are very discriminable from the images created in natural viewing.

The comparison of single-plane and natural yields highly discriminable results from many regions in the stimulus except when the eye's focus is  $\sim 2$  D, the distance assumed in generating the single-plane stimulus. Discriminability never goes to zero because chromatic aberration produces depth-dependent effect in natural viewing that do not occur in the single-plane stimulus. This video illustrates that the images produced in single-plane rendering and display are not good approximations of those generated by the real world. Multiplane with optimized blending versus natural viewing yields low discrimination probabilities for all focus distances and regions of the stimulus. These results show that the rendering and presentation technique we used in our experiments generates nearly correct focus cues even when depth variation is large. That is, it reproduces the optical effects at occlusion borders nearly correctly.

#### Implications for computer graphics

In computer graphics, rendering techniques that model defocus blur have traditionally sought to reproduce photographic or cinematic appearance by modeling the depth of field of a camera (Cook et al., 1984; Kolb, Mitchell, & Hanrahan, 1995) rather than blur in the human eye. But with the growing prevalence of stereoscopic and virtual-reality displays, the need to consider optical phenomena in the viewer's eyes is increasing. As the focus of computer-graphics applications shifts from reproducing photographic imagery to depicting an immersive three-dimensional scene to be directly observed by the viewer, modeling the optical phenomena in the human eye will be necessary to achieve effective, comfortable, and realistic viewing.

For example, stereoscopic displays that present different images to the two eyes are now commonplace in movie theaters, commercially available in consumer televisions, and employed in virtual-reality headsets. While these displays use binocular disparity to indicate varying scene distances, the accommodation distance to produce a sharp retinal image remains fixed at the distance to the display surface. This vergence-accommodation conflict causes perceptual distortions (Watt, Akeley, Ernst, & Banks, 2005), difficulty in simultaneously fusing and focusing the image (Akeley et al., 2004; Hoffman, Girshick, Akeley, & Banks, 2008), and viewer discomfort and fatigue (Emoto, Niida, & Okano, 2005; Hoffman et al., 2008; Lambooij, Fortuin, Heynderickx, & IJsselsteijn, 2009; Shibata, Kim, Hoffman, & Banks, 2011). In practice, content creators often try to minimize these effects by composing scenes so that the main subject of the scene is presented with a disparity of zero or close to zero (Mendiburu, 2009), or they modify the disparities after scene composition by warping the disparity map to reduce large disparities (Lang et al., 2010; Didyk, Ritschel, Eisemann, Myszkowski, & Seidel, 2011). But the minimal-disparity heuristics limit scene composition and still produce conflict when other objects in the scene are fixated.

Additionally, defocus effects due to finite aperture must be statically included in the presented images. When such defocus effects are computed using a thinlens model that approximates a camera, natural optical phenomena such as chromatic aberration are not reproduced, and as we have shown here, this can yield erroneous depth percepts. Inaccurate reproduction of retinal blur has been shown to create artifacts and incorrect scale cues (Held, Cooper, O'Brien, & Banks, 2010). Thus, presenting imagery that is more faithful to what the human viewer would see when observing a real three-dimensional scene should improve depth perception and enhance the sense of immersion.

A challenging issue arises with augmented- and mixed-reality technologies. In these displays, which are mostly head-mounted, the user's real environment is registered using cameras and virtual objects are incorporated into the visual scene using a see-through



Figure 16. The occlusion conditions for which background blur should be discriminable. The eye is focused at distance  $z_0$ , which is either on the occluding surface (right side) or on the background plane (left side). The surface creating the blurred image is at distance  $z_1$  (the occluder on the left side, the background on the right side). The combinations of occluder distance and relative distance  $(z_1/z_0)$  that should produce discriminable blur are represented by the shaded regions. The blue and red shaded regions represent those combinations when the just-discriminable change in distance is 0.2 D. The red shaded region alone represents the combinations for a distance change of 1.2 D, as in our experiment.

display. The virtual objects are often meant to look like part of the real environment, so occlusions must be handled appropriately. If, for example, a virtual object is partially occluded by a real object, the current approach is to cut out the occluded portion and align the border between the virtual and physical objects as accurately as possible. Even with excellent alignment, the result is quite similar to our observations with linear, per-pixel blending in multiplane displays (Narain et al., 2015); as we said, this produces objectionable haloing and gaps. The research reported here shows that better handing of occlusions yields better perceptual outcomes, so we hope that this work will inform the way in which virtual and physical objects are incorporated into the visual experience.

#### Viewing conditions that create useful blur

We wondered what viewing conditions produce sufficient blur for it to be a useful depth-order cue at occlusions. We make the conservative assumption that the blur must at least exceed the detection threshold to be useful in depth ordering. For photopic viewing conditions and pupil diameters of 3–4.5 mm, the justnoticeable change in distance is 0.15-0.25 D at the fovea when blur is the only available information (Kotulak & Schor, 1986; Sebastian et al., 2015; Walsh & Charman, 1988). In our analysis, we assumed that textured surfaces have natural statistics so that optical defocus can be determined (Burge & Geisler, 2011). We also assumed that the viewer is focused on a near occluding surface at distance  $z_0$  and that there is a farther background at distance  $z_1$ . The distance change is of course

$$\Delta D = \frac{1}{z_0} - \frac{1}{z_1}.$$
 (8)

Absolute value is not required because  $z_1$  is greater than  $z_0$ . Solving for  $z_0$ , we obtain

$$z_0 = \frac{z_1}{\Delta D z_1 + 1}.$$
 (9)

We substitute 0.2 D for  $\Delta D$  and find the combinations of occluder distance  $z_0$  and relative distance  $z_1/z_0$  that would yield just-discriminable blur. Those combinations are represented in the right half of Figure 16 by the blue and red shaded regions.

We next assumed that the viewer is focused on the far background surface at a distance of  $z_0$ , and that there is a nearer occluding surface at distance  $z_1$ . Solving for  $z_1$ , we obtain

$$z_1 = \frac{z_0}{1 - \Delta D z_0}.$$
 (10)

We again substitute 0.2 D for  $\Delta D$  and find the combinations of occluder distance  $z_1$  and relative distance  $z_1/z_0$  that would yield just-discriminable blur. Those are shown in the left half of Figure 16.

Two general observations can be made. First, when the distance to the occluder is greater than 5 m, there are no conditions that produce useful blur. Second, when the occluder's distance is much less than 4 m, there are numerous conditions that should produce noticeable blur. Thus, the blur information examined in this article may be useful for near- to medium-range viewing, where by "medium-range" we mean up to two or three times arm's length. It is not useful for long-range viewing. But that analysis is based on the assumption that blurs that just exceed detection threshold can be used in depth ordering. It could be that larger blurs, like the 1.2-D values used in our experiments, are required for ordering. The combinations of distances that would yield such a change are indicated by the red contours and shaded regions in the figure.

#### Realizability of blur in previous studies

As mentioned earlier, three studies—Marshall et al. (1996), Mather and Smith (2002), and Palmer and

Zannoli, Love, Narain, & Banks

Study	Viewing distance (cm)	Standard deviation of blur kernel (arcmin)	Diameter of blur kernel (arcmin)	Diameter at infinity (arcmin)
Marshall et al. (1996)	50	14.4	36.1	30.9
Palmer & Brooks (2008)	60	10.8	27.1	25.8
Mather & Smith (2002)	57	8.0	20.7	27.1
Mather & Smith (2002)	57	16.0	41.3	27.1
Mather & Smith (2002)	114	4.0	10.3	13.6
Mather & Smith (2002)	114	8.0	20.7	13.6

Table 1. Blur magnitudes in previous studies and plausibility in natural viewing.

Brooks (2008)—examined the ability to determine depth order at occlusions from blur. The stimuli were presented on single planes. All three studies inserted large amounts of blur into their stimuli. Here we ask what natural viewing situations would create such blur magnitudes.

To do this, we first converted the Gaussian blur kernels they used into the most similar cylinders (in the least-squares sense) using the formula  $\beta = \sigma/0.399$ , where  $\beta$  is the cylinder diameter and  $\sigma$  is the Gaussian standard deviation. Making this conversion allows us to relate blur magnitudes to scene and eye parameters using Equation 2. We evaluated the situation in which the eye is focused on the near occluding plane, so the background plane is blurred. We set the distance of the occluding plane to the experimental viewing distances because participants were presumably accommodated to that distance. The space-average luminance for Mather and Smith (2002) was 37.5  $cd/m^2$ ; the other studies did not state the luminance. We assumed a diameter of 4.5 mm, which is appropriate for 37.5 cd/ m<sup>2</sup> (Spring & Stiles, 1948), and then calculated the value of  $\beta$  when the background plane is at infinite distance. From Equation 2,  $\beta$  (in radians) for that situation (i.e.,  $z_1 = \infty$ ) is

$$\beta \approx \frac{A}{z_0}, \quad (11)$$

where A is pupil diameter and  $z_0$  is the distance of the near occluding plane.

Table 1 shows the results. From left to right, the rows provide the name of the study, viewing distance, standard deviation of the Gaussian blur kernel, diameter of the most similar cylinder, and diameter of the blur kernel for a background at infinite distance. When the value in the fourth column is greater than the one in the fifth (red), the blur in the stimulus could not be created by a physically realizable stimulus. This occurs in four of the six conditions in these studies. In two conditions, the difference between the possible and experimental blur magnitudes is small; in the other two, it is large. Thus, in some conditions in these studies the retinal images could not be observed in natural viewing.

We also evaluated the situation in which the eye is focused on the background plane and the near occluding surface is blurred. In that case, there is always an occluding surface that would be close enough to the eye to create the observed blur. Thus, these stimuli are all consistent with physically possible viewing conditions.

The observation that impossible blurs occurred when the simulated focus was on the near occluding surface might help explain why participants performed unreliably in these experiments. For example, in the study by Marshall et al. (1996), participants were more likely to report that the blurred surface was nearer than the sharp surface than to report that the blurred surface was farther than the sharp one (see their figure 6). One might expect this because the blur in the former situation cannot occur in natural viewing, while blur in the latter situation can.

#### Blur and estimating distance in half occlusions

Half occlusion is the situation in which one eye can see an object point and the other eye cannot (Gillam & Borsting, 1988). In this case, there is no computable disparity, so the distance to the half-occluded object is not specified. But several authors have noted that there are geometric constraints on the possible three-dimensional positions of such an object (Nakayama & Shimojo, 1990; Tsirlin, Wilcox, & Allison, 2014; Zannoli & Mamassian, 2011). Figure 17 illustrates the geometry. The blue point is visible to the right eye but not the left. The possible positions of the object are indicated by the thick blue arrow, which runs off to infinity. Interestingly, participants tend to perceive the monocular object near the leading edge of the depthconstraint zone, provided that the object is not more than 20-30 arcmin from the occlusion border. When the object's distance from the border is greater than that, perceived distance regresses toward the distance of the occluder (Nakayama & Shimojo, 1990; Zannoli & Mamassian, 2011). To our knowledge, there is no persuasive argument about why the perceived distance is near the leading edge for small displacements and regresses toward the occluder for large displacements.

The geometrical analysis in these studies is based on a pinhole-camera model (i.e., pupil diameter is zero).



Figure 17. The geometry for half occlusions with pinhole apertures: A plan view of the viewer with the right eye above and the left eye below. The eyes are separated by the interocular distance *I*. An occluder is blocking views of object space for the eyes. The green lines indicate where the occluded region begins for the two eyes. Because of the occluder, the object point (blue dot) is visible to the right eye but not the left eye. (The occluder is infinitely wide, so all left-eye views are blocked once the object is behind the occluder from that eye's perspective.) The shaded area represents the depth-constraint zone (Nakayama & Shimojo, 1990). The leading edge of that zone is the labeled green line. The possible positions of the object are indicated by the thick blue line.

This is of course not realistic, because pupil diameters in human eyes are never zero. Here we consider the additional information provided in the half-occlusion situation by the eyes' nonzero apertures.

When the pinhole is replaced by a finite aperture, the images formed in the unoccluded eye are changed. In Figure 18, the eye is focused at the distance of the occluder. The half-occluded object point creates a blurred retinal image with a blur-circle diameter of b. This diameter can of course provide information about how distant the object point is. If b is small, the object must be at roughly the same distance as the



Figure 18. The geometry for half occlusions with nonzero apertures. The eyes are separated by the same distance *I*, but the eyes' apertures have diameter *A*. The eyes are focused on the occluder at distance  $z_0$ . The retinal image of the partially occluded object point seen by the right eye creates a blur circle of diameter *b*. The left eye may receive rays from the object point in a portion of that eye's aperture. The thin red line indicates the directions of object points that will just receive rays in the margin of the aperture. It is the leading edge of the depth-constraint zone, once the eye's finite aperture is taken into account. Notice that it is rotated relative to the leading edge in Figure 17. Objects farther to the left will not illuminate the pupil; objects farther to the right will illuminate more of the pupil. The darkly shaded area indicates the region in which some rays will enter the left eye. The thick red line indicates the possible object positions if some rays enter the left eye. The thick green line (which continues to infinite distance) indicates possible positions if no rays enter the left eye. In fact, the partial occlusion of the left eye's view provides more information than shown in the figure, because those rays will also give an indication of the direction of the object from the left eye's vantage point.

occluder; if it is large, the object is probably at a significantly greater distance than the occluder. Hoffman and Banks (2010) demonstrated just such an effect. They showed that the perceived distance of a half-occluded object increased as the object became progressively blurred.

There is additional information in the eye with the occluded view. In the pinhole model, the blockage of the object point in that eye is precipitous: The point is either seen or not seen. In reality, the transition from seen to unseen is gradual, as shown in Figure 18. Here we consider rays from the object point that enter part of the aperture. If some rays enter the occluded eye, the possible distances are restricted to positions along the red line. If no rays enter that eye, the possibilities are on the green line. In addition, when some rays enter, the size of the resulting PSF provides information about the half-occluded point's distance relative to where the eye is focused.

In the previous studies the half-occluded object was always sharp (Nakayama & Shimojo, 1990; Tsirlin et al., 2014; Zannoli & Mamassian, 2011). We offer two hypotheses concerning the data they reported. The first hypothesis explains why participants perceived the halfoccluded object at the leading edge of the depthconstraint zone when the displacement was less than 20-30 arcmin. With small displacements, the combination of observing no blur in the unblocked eye and nothing in the blocked eye is physically possible, but the object must be no farther than the nearest part of the constraint zone for it not to be blurred. The second hypothesis explains why the perceived distance of the half-occluded object regresses toward the occluder when the displacement is large. With large displacements, the blur of the object should increase in the unblocked eye and the probability of some light leaking into the blocked eye should increase. Depending on the viewing distance, the observation of no blur in the unblocked eye and nothing in the blocked eye becomes impossible. Then the depth percept ends up being a compromise between the observed blur (which indicates it is at the same distance as the occluder) and the observed half occlusion (which indicates it is in the depth-constraint zone).

Keywords: occlusions, depth perception, blur, accommodation, chromatic aberrations, micro-fluctuations of accommodation

## Acknowledgments

We thank Rachel Albert, Abdullah Bulbul, and Luiza Rocha Araújo for assistance in conducting Experiment 1 and Sylvain Reissier for assistance in conducting Experiment 2. We thank Steven Cholewiak, Robin Held, and George Koulieris for comments on an earlier draft. Research was supported by the National Science Foundation.

Commercial relationships: none.

Corresponding author: Martin S. Banks.

Email: martybanks@berkeley.edu.

Address: Vision Science Program, University of California, Berkeley, Berkeley, CA, USA.

## References

- Aggarwala, K. R., Nowbotsing, S., & Kruger, P. B. (1995). Accommodation to monochromatic and white-light targets. *Investigative Ophthalmology & Visual Science*, 36(13), 2695–2705. [PubMed] [Article]
- Akeley, K., Watt, S. J., Girshick, A. R., & Banks, M. S. (2004). A stereo display prototype with multiple focal distances. ACM Transactions on Graphics, 23, 804–813.
- Banks, M. S., Sprague, W. W., Schmoll, J., Parnell, J. A. Q., & Love, G. D. (2015). Why do animal eyes have pupils of different shapes? *Science Advances*, 1(9), 1–9.
- Bedford, R. E., & Wyszecki, G. (1957). Axial chromatic aberration of the human eye. *Journal of the Optical Society of America*, 47, 564–565.
- Burge, J. L., Fowlkes, C. C., & Banks, M. S. (2010). Natural-scene statistics predict the influence of the figure-ground cue of convexity on depth human depth perception. *Journal of Neuroscience*, 30, 7269–7280.
- Burge, J., & Geisler, W. S. (2011). Optimal defocus estimation in individual natural images. *Proceedings of the National Academy of Sciences, USA, 108*, 16849–16854.
- Burt, P., & Julesz, B. (1980). A disparity gradient limit for binocular fusion. *Science*, 208, 615–617.
- Campbell, F. W., Robson, J. G., & Westheimer, G. (1959). Fluctuations of accommodation under steady viewing conditions. *Journal of Physiology*, 145, 579–594.
- Campbell, F. W., & Westheimer, G. (1958). Sensitivity of the eye to differences in focus. *Journal of Physiology*, 143, 18.
- Campbell, F. W., & Westheimer, G. (1959). Factors influencing accommodation responses of the human eye. *Journal of the Optical Society of America*, 49, 568–571.
- Cavanagh, P. (1987). Reconstructing the third dimension: Interactions between color, texture, motion,

binocular disparity and shape. Computer Vision, Graphics, and Image Processing, 37, 171–195.

- Charman, W. N., & Heron, G. (2015). Microfluctuations in accommodation: An update on their characteristics and possible role. *Ophthalmic & Physiological Optics*, 35, 476–499.
- Collins, M., Davis, B., & Wood, J. (1995). Microfluctuations of steady-state accommodation and the cardiopulmonary system. *Vision Research*, *35*, 2491–2502.
- Cook, R. L., Porter, T., & Carpenter, L. (1984). Distributed ray tracing. *ACM Transactions on Graphics*, 18(3), 137–145.
- Davies, L. N., Wolffsohn, J. S., & Gilmartin, B. (2009). Autonomic correlates of ocular accommodation and cardiovascular function. *Ophthalmic & Physi*ological Optics, 29, 427–435.
- Didyk, P., Ritschel, T., Eisemann, E., Myszkowski, K., & Seidel, H.-P. (2011). A perceptual model for disparity. ACM Transactions on Graphics, 30(4), 96.
- Emoto, M., Niida, T., & Okano, F. (2005). Repeated vergence adaptation causes the decline of visual functions in watching stereoscopic television. *Journal of Display Technology*, 1(2), 328–340.
- Filippini, H. R., & Banks, M. S. (2009). Limits of stereopsis explained by local cross-correlation. *Journal of Vision*, 9(1):8, 1–18, doi:10.1167/9.1.8. [PubMed] [Article]
- Fincham, E. F. (1951). The accommodation reflex and its stimulus. *British Journal of Ophthalmology*, 35, 381–393.
- Georgeson, M. A., & Sullivan, G. D. (1975). Contrast constancy: Deblurring in human vision by spatial frequency channels. *Journal of Physiology*, 252, 627–656.
- Gibson, J. J. (1966). *The senses considered as perceptual systems*. Boston: Houghton-Mifflin.
- Gillam, B., & Borsting, E. (1988). The role of monocular regions in stereoscopic displays. *Perception*, 17, 603–608.
- Held, R. T., Cooper, E. A., O'Brien, J. F., & Banks, M. S. (2010). Using blur to affect perceived distance and size. ACM Transactions on Graphics, 29(2).
- Hillis, J. M., & Banks, M. S. (2001). Are corresponding points fixed? *Vision Research*, 41, 2457–2473.
- Hoffman, D. M., & Banks, M. S. (2010). Focus information is used to interpret binocular images. *Journal of Vision*, 10(5):13, 1–17, doi:10.1167/10.5.
  13. [PubMed] [Article]
- Hoffman, D. M., Girshick, A. R., Akeley, K., & Banks, M. S. (2008). Vergence–accommodation conflicts hinder visual performance and cause visual fatigue.

*Journal of Vision*, 8(3):33, 1–30, doi:10.1167/8.3.33. [PubMed] [Article]

- Jakob, W. (2010). Mitsuba renderer. http://www. mitsubarenderer.org
- Kasthurirangan, S., Vilupuru, A. S., & Glasser, A. (2003). Amplitude dependent accommodative dynamics in humans. *Vision Research*, 43, 2945–2956.
- Kolb, C., Mitchell, D., & Hanrahan, P. (1995, Sept). A realistic camera model for computer graphics. In *Proceedings of the 22nd annual conference on computer graphics and interactive techniques* (pp. 317–324). Los Angeles, CA: ACM.
- Kotulak, J. C., & Schor, C. M. (1986). The accommodative response to subthreshold blur and to perceptual fading during the Troxler phenomenon. *Perception*, 15, 7–15.
- Kruger, P. B., Mathews, S. Aggarwala, K., & Sanchez, N. (1993). Chromatic aberration and ocular focus: Fincham revisited. *Vision Research*, 33, 1397–1411.
- Lambooij, M., Fortuin, M., Heynderickx, I., & IJsselsteijn, W. (2009). Visual discomfort and visual fatigue of stereoscopic displays: A review. *Journal* of Imaging Science and Technology, 53(3), 30201-1.
- Lang, M., Hornung, A., Wang, O., Poulakos, S. Smolic, A., & Gross, M. (2010, July). Nonlinear disparity mapping for stereoscopic 3D. ACM Transactions on Graphics, 29(4), 75.
- Lee, J. H., Stark, L. R., Cohen, S., & Kruger, P. B. (1999). Accommodation to static chromatic simulations of blurred retinal images. *Ophthalmic & Physiological Optics*, 19, 223–235.
- Love, G. D., Hoffman, D. M., Hands, P. J., Gao, J., Kirby, A. K., & Banks, M. S. (2009). High-speed switchable lens enables the development of a volumetric stereoscopic display. *Optics Express*, 17, 15716–15725.
- MacKenzie, K. J., Hoffman, D. M., & Watt, S. J. (2010). Accommodation to multiple-focal-plane displays: Implications for improving stereoscopic displays and for accommodation control. *Journal of Vision, 10*(8):22, 1–20, doi:10.1167/10.8.22. [PubMed] [Article]
- Mantiuk, R., Kim, K. J., Rempel, A. G., & Heidrich, W. (2011). HDR-VDP-2: A calibrated visual metric for visibility and quality predictions in all luminance conditions. ACM Transactions on Graphics, 30, 1–14.
- Marimont, D. H., & Wandell, B. A. (1994). Matching color images: The effects of axial chromatic aberration. *Journal of the Optical Society of America A*, 11, 3113–3122.
- Marshall, J. A., Burbeck, C. A., Ariely, D., Rolland, J. P., & Martin, K. E. (1996). Occlusion edge blur: A

cue to relative visual depth. *Journal of the Optical Society of America A, 13,* 681–688.

- Mather, G. (1996). Image blur as a pictorial depth cue. Proceedings of the Royal Society B, 263, 169–172.
- Mather, G., & Smith, D. R. R. (2002). Blur discrimination and its relation to blur-mediated depth perception. *Perception*, *31*, 1211–1219.
- Mendiburu, B. (2009). *3D movie making: Stereoscopic digital cinema from script to screen*. Waltham, MA: Focal Press.
- Nakayama, K., & Shimojo, S. (1990). Da Vinci stereopsis: Depth and subjective occluding contours from unpaired image points. *Vision Research*, 30, 1811–1825.
- Narain, R., Albert, R. A., Bulbul, A., Ward, G. J., Banks, M. S., & O'Brien, J. F. (2015). Optimal presentation of imagery with focus cues on multiplane displays. ACM Transactions on Graphics, 34(4), 59.
- Nguyen, V. A., Howard, I. P., & Allison, R. S. (2005). Detection of the depth order of defocused images. *Vision Research*, 45, 1003–1011.
- Palmer, S. E., & Brooks, J. L. (2008). Edge-region grouping in figure-ground organization and depth perception. *Journal of Experimental Psychology: Human Perception & Performance*, 34, 1353–1371.
- Ravikumar, S., Akeley, K., & Banks, M. S. (2011). Creating effective focus cues in multi-plane 3D displays. *Optics Express*, 19, 20940–20952.
- Schor, C. M., Lott, L. A., Pope, D., & Graham, A. D. (1999). Saccades reduce latency and increase velocity of ocular accommodation. *Vision Research*, 39, 3769–3795.
- Sebastian, S., Burge, J., & Geisler, W. S. (2015).

Defocus blur discrimination in natural images with natural optics. *Journal of Vision*, *15*(5):16, 1–17, doi:10.1167/15.5.16. [PubMed] [Article]

- Shibata, T., Kim, J., Hoffman, D. M., & Banks, M. S. (2011). The zone of comfort: Predicting visual discomfort with stereo displays. *Journal of Vision*, *11*(8):11, 1–29, doi:10.1167/11.8.11. [PubMed] [Article]
- Spring, K. H., & Stiles, W. S. (1948). Variation of pupil size with change in the angle at which the light stimulus strikes the retina. *British Journal of Ophthalmology*, 32, 340–346.
- Tsirlin, I., Wilcox, L. M., & Allison, R. S. (2014). A computational theory of da Vinci stereopsis. *Journal of Vision*, 14(7):5, 1–26, doi:10.1167/14.7.5. [PubMed] [Article]
- Walsh, G., & Charman, W. N. (1988). Visual sensitivity to temporal change in focus and its relevance to the accommodation response. *Vision Research*, 28, 1207–1221.
- Watt, S. J., Akeley, K., Ernst, M. O., & Banks, M. S. (2005). Focus cues affect perceived depth. *Journal* of Vision, 5(10):7, 832–865, doi:10.1167/5.10.7. [PubMed] [Article]
- Wilson, B. J., Decker, K. E., & Roorda, A. (2002). Monochromatic aberrations provide an odd-error cue to focus direction. *Journal of the Optical Society* of America A, 19, 833–839.
- Winn, B., & Gilmartin, B. (1992). Current perspective on microfluctuations of accommodation. *Ophthalmic & Physiological Optics*, 12, 252–256.
- Zannoli, M., & Mamassian, P. (2011). The role of transparency in da Vinci stereopsis. *Vision Research*, *51*, 2186–2197.