Vision Research 107 (2015) 15–21

Contents lists available at ScienceDirect

Vision Research

journal homepage: www.elsevier.com/locate/visres

Are blur and disparity complementary cues to depth?

Michael S. Langer *, Ryan A. Siciliano

School of Computer Science, McGill University, Montreal H3A039, Canada

ARTICLE INFO

Article history:
Received 22 August 2014
Received in revised form 20 October 2014
Available online 4 December 2014

Keywords:
Binocular disparity
Blur
Defocus
Depth perception
Cue combinations

ABSTRACT

The image blur and binocular disparity of a 3D scene point both increase with distance in depth away from fixation. Perceived depth from disparity has been studied extensively and is known to be most precise near fixation. Perceived depth from blur is much less well understood. A recent experiment (Held, R. T, Cooper, E. A., & Banks, M. S. (2012). *Current Biology, 22*, 426–431) which used a volumetric stereo display found evidence that blur and disparity are complementary cues to depth, namely the disparity cue dominates over the blur cue near the fixation depth and blur dominates over disparity at depths that are far from fixation. Here we present a similar experiment but which used a traditional 3D display so that blur was produced by image processing rather than by the subjects' optics. Contrary to Held et al., we found that subjects did not rely more on blur to discriminate depth at distances far from fixation, even though a sufficient level of blur was available to do so. The discrepancy between the findings of the two studies can be explained in at least two ways. First, Held et al.'s subjects received trial-to-trial feedback in a training phase and may have learned how to perform the task using blur discrimination. Second, Held et al.'s volumetric stereo display may have provided other optical cues that indicated that the blur was real rather than rendered. The latter possibility would have significant implications about how depth is perceived from blur under different viewing conditions.

1. Introduction

When our eyes fixate on a point in 3D space, they both accommodate and converge to that point. Accommodation brings the point into sharp focus on each retina. Vergence brings the point to the center of each fovea where spatial resolution is highest. When accommodation and vergence are correct, the fixated point is in sharp focus and has zero binocular disparity. For scene points that are depths other than the fixation depth, their blur and disparity are proportional to the inverse distance (diopters) from the fixated point, with the disparity being roughly an order of magnitude larger than the blur width (Schechner & Kiryati, 2000).

Although blur and disparity both vary with inverse distance from fixation, there are differences in the visual system's sensitivity to these cues and how the visual system uses these cues in depth perception. Depth discrimination from disparity is very accurate near the fixation distance but it worsens rapidly with increasing distance from fixation, especially once diplopia occurs (Howard & Rogers, 2012). Depth discrimination from blur is much less well understood as we will discuss later. Blur discrimination itself is most accurate, not at the fixation depth, but rather at depths that are in front of and behind the fixation depth. JND's for blur obey a dipper function which achieves a minimum when the blur radius[1] is about 1 arcmin (Watson & Ahumada, 2011). In particular, there is a considerable range of depths around fixation over which all surfaces appear in focus, the so-called depth of field region.

This paper concerns the range of depths beyond the depth of field for which both disparity and blur cues are present. One might expect that over this range, the visual system combines the disparity and blur cues, for example, in a linear cue combination scheme (Landy et al., 1995). Mather and Smith examined disparities up to the limits of the fusion range but found little evidence for cue combination (Mather & Smith, 2000). This led them to an alternative hypothesis. Rather than estimating depth by combining blur and disparity cues, the visual system relies on disparities over the small depth range in which that cue is reliable, and it relies on blur to infer depth beyond that depth range. In this sense, disparity and blur would be complementary cues to depth.

---

* Corresponding author.
E-mail addresses: michael.langer@mcgill.ca (M.S. Langer), ryan.siciliano@mail.mcgill.ca (R.A. Siciliano).

[1] It is common to define the level of blur by the standard deviation of a Gaussian blur kernel, though we will also refer to a nominal blur "width". See Table 1 and surrounding text.

http://dx.doi.org/10.1016/j.visres.2014.10.036

Held, Cooper and Banks carried out a depth discrimination experiment that explored this hypothesis further (Held, Cooper, & Banks, 2012). Subjects discriminated the depths of two textured surfaces that both lay beyond fixation depth. Three cue combinations were tested: disparity-only, blur-only (monocular), and disparity-and-blur. At small disparity pedestals, the JND's were lower for the disparity-only condition than for the blur-only condition. This order reversed when the disparity pedestal became large. Most interesting is that, when both disparity and blur cues were present, the JND's followed the lower of the JND's for the disparity-only and blur-only condition. This suggests that subjects were relying more on the disparity cue for distances close to fixation and more on the blur cue for distances well beyond fixation[2] which is consistent with the hypothesis that blur and disparity are complementary cues to depth.

One concern that has been raised about Held et al. study is that JND's measure the precision of depth perception, but not the accuracy (Vishwanath, 2012). This is a well known distinction and indeed most studies of depth from disparity also addressed precision rather than accuracy (Ogle, 1953; Blakemore, 1970; McKee, Levi, & Bowne, 1990; Wilcox & Allison, 2009). The question of accuracy should not be neglected, however. For example, one of Held et al.'s stated motivations for studying depth perception for surfaces that are from the fixation depth is that these depth percepts would be needed for making eye movements and reaching movements.[3] But such movements surely require a high level of accuracy, not just precision.

Indeed there is evidence that, when disparities are large, depth perception becomes not merely imprecise but it also becomes inaccurate. For example, Richards and Kaye showed that perceived depth from disparity is not a monotonic function of physical disparity (Richards & Kaye, 1974). Rather it is a unimodal function: as the disparity increases, perceived depth at first increases but then it decreases to zero. A similar idea was discussed by Ogle (1952) who distinguished "patent stereopsis", where perceived depth increases as disparity increases, from "qualitative stereopsis" where only the sign of depth relative to fixation is perceived.[4] Ogle also noted that for sufficiently large disparities, no depth is perceived i.e. not even the sign.

In this paper, we present an experiment in which we attempted to confirm the findings of Held et al. Our experiment different from Held et al.'s in a few key ways, however. First, we used a conventional stereo display whereas they used a volumetric stereo display (Love et al., 2009). Second, our subjects had only a few minutes of training and were not given any feedback, whereas their subjects had 30 min of training with trial-to-trial feedback in all three conditions. We were concerned that the training given to Held et al.'s subjects may have led them to perform the task based on perceived blur when it was present, rather than on perceived depth (Vishwanath, 2012). Indeed Held et al. reported that one of the two naïve subjects was aware of the correlation between blur and depth and in the blur-only condition sometimes judged the blurrier stimulus as farther.

Our experiment consisted of two parts. The first part was a depth discrimination task which corresponded to the experiment of Held et al., with some differences mentioned above and others that will be described later. The second part was a blur discrimination task. The purpose was to verify that there was a sufficient amount of blur present in the stimuli for subjects to use the blur cue in the first part, where the task was to discriminate depth.

2. Methods

2.1. Stimuli

The experiment was run on a Dell Precision M6700 laptop. The stimuli were generated and controlled using PsychoPy (Peirce, 2007). Stereo images were presented using 3D Vision shutter glasses by NVidia. The display screen was 1920×1080 pixels. Viewing distance was 63 cm. At this distance, each pixel subtended about $1'$ (one arcmin) of visual angle. From now on, we use units of pixels and arcmin interchangeably. The gamma of the monitor was measured to be 2.0. The images were gamma corrected so that luminance was proportional to digital gray level.

The stimuli were similar to those used by Held et al. Each image consisted of a foreground occluder and two background surfaces. The occluder was a texture composed of a grid of square tiles. Each tile was of size $64' \times 64'$. The occluder contained a fixation cross. See Fig. 1. The occluder also defined two window panels, each $512' \times 128'$ through which a background reference and test surface were shown. These background images each consisted of white squares randomly placed on a black background. Each square was $16' \times 16'$ and the density was 4 squares/deg^2. The size and density of squares was similar to the stimuli used in Held et al.

The background textures were defined offline prior to the experiment, as follows. First, a background texture of size $512' \times 512'$ was generated by placing small white squares uniformly randomly on a black background. This background texture was then blurred by a set of 2D Gaussians with varying standard deviations σ and these blurred textures were stored. On each trial of the experiment, a random cropped window from a blurred background texture was selected for the reference and for the test. Disparities were produced by selecting a cropped region for the left eye and a shifted cropped window for the right eye.

The set of reference disparities used in the experiment are listed in the first column of Table 1. These reference disparities ranged from 0 to $96'$ in steps of $24'$. For each disparity value, we define a nominal blur width ω such that the disparity to blur width ratio is 12:1, which corresponds to the ratio of the interocular distance to the pupil diameter, assuming a pillbox blur kernel (Held, Cooper, & Banks, 2012). Rather than using a pillbox kernel for blur

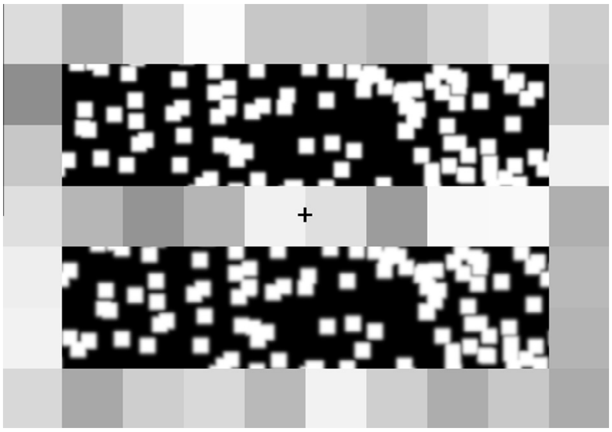

Fig. 1. Example stimulus. The top and bottom windows were rendered with a blur width ω of $6'$ and $7'$ respectively. (See Table 1.) The image should be viewed such that each foreground tile spans $64' \times 64'$, so width of just over 1 deg. See text for details.

---

[2] Held et al. noted that they did not have sufficient statistical power to distinguish a cue switching strategy from an optimal cue combination strategy.

[3] Strictly speaking, eye movements and accommodation do not require a depth estimate. Rather they just require a disparity estimate or blur estimate, respectively. Reaching movements do require a depth estimate though.

[4] Richards and Kaye's plots are not entirely consistent with Ogle's characterization. We assume that patent stereopsis corresponds roughly to the increasing segment of the Richards and Kaye plots and qualitative stereopsis corresponds to the downward sloping segments of Richards and Kaye's plot.

Table 1
Reference values of blur and disparity and their corresponding depths.

Disparity (arcmin)	Blur width ω (arcmin)	Blur radius σ (arcmin)	Depth (cm)	Depth (diopters)	Δ Depth (diopters)
0	0	0	63	1.59	0
24	2	0.58	68	1.48	.11
48	4	1.15	73	1.37	.21
72	6	1.73	79	1.26	.33
96	8	2.33	87	1.16	.43

rendering, however, we used a Gaussian kernel whose standard deviation was slightly less than half of the blur width.[5] Although the Gaussian kernel is only a crude approximation to the actual point spread function of human eyes, it is the most commonly model for rendering blur in blur perception studies (Watson & Ahumada, 2011).

The second column of Table 1 lists the set of nominal reference blur widths which are from 0′ to 8′ in steps of 2′. The third column lists the corresponding standard deviations of the Gaussian blurs that were used to generate the stimuli. The relation between disparity and blur were mentioned in a footnote above. The last three columns of Table 1 list the depths that correspond to these reference disparity (and blur) values. These depths Z that are given in the table were determined by isolating Z in the formula

$$\text{disparity} = 60 * 180/\pi * \text{IOD} * (1/Z - 1/Z0)$$

where the disparity values are given in Table 1, the interocular distance IOD was 6.5 cm, and the fixation distance $Z0$ was 63 cm. The values of Z that are given the table have been rounded to the nearest cm. These depth values and their inverted values (diopters) play no direct role in the plots or discussions that follow. We list them mainly who wishes to compare our viewing distances to those of Held et al.

2.2. Subjects

Fifteen subjects were recruited. Each was paid $10. Each had normal or corrected-to-normal vision. Subjects had little if any experience with psychophysics experiments. All were unaware of the purpose of the experiments beyond what was explained to them (see Section 2.4). Informed consent was obtained. The experiment was carried out in accordance with the Code of Ethics of the World Medical Association (Declaration of Helsinki).

2.3. Design

Disparity and blur cues were combined in one of three ways, which corresponded roughly to Held et al.'s disparity-only, disparity-and-blur, and blur-only conditions:

(1) only the disparities differed,
(2) both the disparities and blurs differed,
(3) only the blurs differed.

The depth discrimination experiment used all three cue combinations. The blur discrimination experiment used only cue condition (3).

In each trial, the reference disparity was one of 24′, 48′, 72′, or 96′ and the corresponding reference blur width was 2′, 4′, 6′ or 8′, respectively, as was shown in Table 1. Note that these values satisfy the 12:1 disparity to blur width ω ratio mentioned in Section 2.1. For condition (1), the disparity increments were 6′, 12′, or 18′, and there was no blur increment. For condition (3), the blur increments were 0.5′, 1.0′, or 1.5′ and there was no disparity increment. For condition (2), there were both disparity and blur increments always paired with a 12:1 ratio.

The depth discrimination task consisted of 432 trials, namely 12 trials in each of 36 conditions (3 cue combinations × 4 reference values of blur and disparity × 3 levels of blur and/or disparity increment). The blur discrimination task consisted of 144 trials, namely 12 trials in each of 12 conditions (1 cue combination × 4 reference values of blur × 3 levels of blur increment).

There are a few differences worth mentioning between our Stimuli and our Design and that of Held et al. First, our reference stimuli always had a non-zero value of blur and disparity which was not the case for Held et al. Held et al.'s disparity-only conditions did not have any blur, and their blur-only conditions did not have any disparity. Indeed their blur-only condition was monocular whereas ours was binocular. We justify these differences as follows.

For our disparity-only condition, i.e. cue combination (1), we could have rendered the squares without any blur. We chose not to do so because we wanted this condition to be as consistent as possible with the disparity-and-blur conditions, i.e. (2), in which blur manipulated. We were concerned that if our disparity-only trials had used sharply focussed images, then subjects might have learned that disparity does not imply blur and they might have been more likely to ignore the blur cue in condition (2). Of course, we wanted to avoid this. Thus, we included the appropriate level of blur to the reference stimuli in condition (1).

For the blur-only condition i.e. (3), we used binocular viewing with a non-zero disparity. This is different from Held et al. who used monocular viewing. There are advantages and disadvantages to each. The disadvantage of binocular viewing is that the disparities create an unnecessary cue conflict: they specify that the reference and test are at the same depth. This cue conflict is a more of an issue for the smaller reference disparities (24′ and 36′) since small disparities provide reliable cues to depth. According to linear cue combination theory (Landy et al., 1995), subjects should give a large weight to the disparity cue which specifies that the depths are equal, and thus would drive performance towards chance. The cue conflict should be less of an issue at larger reference disparities (48′ and 60′) since large disparities are less reliable and so subjects should give them a low weight, as desired for the blur-only condition.

Held et al. avoided this cue conflict issue by using monocular viewing. However, monocular viewing has its own disadvantages. First, the monocular condition needs to be run in separate blocks since it is too distracting for subjects if monocular and binocular conditions are randomly interleaved (as we discovered in a pilot experiment). The problem is that blocking the conditions makes subjects more aware of which condition they are running in each block which increases the likelihood that subjects will adopt different strategies in the different conditions. Our design uses randomly interleaved conditions, and so subjects are more likely to adopt a common strategy for all conditions.

---

[5] We used a Gaussian kernel having standard deviation $\sigma = \omega/\sqrt{12}$, namely the standard deviation of a 1D uniform kernel of width ω. We used a Gaussian rather than a pillbox because the former is less sensitive to pixel sampling. Also, for the smallest values of σ, we used a 3×3 kernel with standard deviation equal to that of a continuous Gaussian. Note that the ratio of disparity to σ was $12\sqrt{12} : 1$ and that this ratio is somewhat arbitrary since it is based on an approximation of the pupil diameter and the eye's point spread function.

A second disadvantage of monocular viewing is that when one removes the disparity cue entirely, the weight of the blur cue should increase but the weight of any other depth cue that is present should also increase. In particular, the stimuli contain a size cue: the same size squares are used in both reference and test, which is a cue that the reference and test are at the same depth. Thus, while monocular viewing removes the cue conflict from disparity, it potentially increases the weight of the size cue conflict.

A third disadvantage of monocular viewing is that it is more susceptible to the twofold ambiguity in depth from blur. With monocular viewing, the only cue for disambiguating the depth sign is the occlusion cue which is known to be weak (Marshall et al., 1996). Indeed in one of our pilot studies which used monocular viewing, we found that many subjects responded as if the blurrier stimulus was closer. This was the reason why we originally switched to binocular viewing design for all conditions.

Another difference between Held et al.'s experiment and ours is that our stimuli covered a smaller range of disparities. Held et al. used disparities up to about 4 deg whereas our disparities all were below 2 deg. There are two reasons why we limited our disparities in this way. First, according to our calculations based on Fig. 2 of Held, Cooper, and Banks (2012), the crossover of their disparity-only and blur-only curves occurred at about 90′ of disparity. By 120′ of disparity the two curves were well separated and remained so at greater disparity values.[6] Second, a study that used random dot stimuli found that, for dot densities similar to ours,[7] the threshold for determining sign of depth (called Dmax) was reached when the disparities were about 90′ (Glennerster, 1998). When both reference and test stimuli are beyond Dmax, it becomes impossible to discriminate depth since even the sign of depth is not perceived. For these two reasons, we felt that our smaller range of disparities was sufficient.

A final difference is worth noting here. We used depth increments to define our test stimuli, whereas Held et al. used decrements.[8] The reason they did so is that in pilot studies they had been unable to compute 75% correct thresholds using depth increments when reference depths were beyond some limit. This difficulty was due in part to the limited range of their volumetric display, but more fundamentally it was due to the limitations on stereopsis that were described by Ogle, and that we mentioned in the Introduction. The idea is as follows. When a reference surface is beyond the limit of patent stereopsis and a depth increment is used, the test surface will be also beyond the limit of patent stereopsis. In this case, according to Ogle, an observer would at best perceive the (identical) signs of these two stimuli and would be unable to discriminate their depths – certainly not at a 75% correct threshold level. Held et al. avoided this problem by using depth decrements rather than increments: for a sufficiently large decrement, the test surface would fall into the range of patent stereopsis and thus would become discriminable from the test. Using decrements allowed Held et al. to compute thresholds for all reference depths. In our experiments, the distinction between increments and decrements is less important since we compute percent correct scores only. As we will see later, these are sufficient for us to make our arguments.

2.4. Procedure

Prior to the experiment, subjects were given a brief introduction to the purpose of the study. It was explained how binocular vision allows one to perceive depth differences and how there are two different ranges of depth differences that are interesting to study, namely when binocular vision is fused versus when it is diplopic. Subjects were shown examples of fingers held at different distances and how one finger becomes diplopic when it is sufficiently far from the fixated finger. It was also mentioned that some of the stimuli would appear blurred and that blur arises because it is only possible to focus on one depth at a time. We did not explain the relationship between blur and disparity.

Subjects were then tested for binocular stereopsis using the Randot test. All subjects passed this test with a stereo acuity of less than a minute of arc. They were then shown examples of the stimuli used in the experiment and had a chance to practice the task for several minutes. The task was as follows. In each trial, a reference and test background was assigned randomly to the upper and lower window. Subjects had to judge which of the two appeared farther in depth. They responded by selecting either the *up* or *down* arrow on the keyboard. No trial-to-trial feedback was given either in the practice or the test phases.

The practice phase consisted of two parts. In the first part, the presentation time was several seconds which gave subjects a chance to make eye movements so that they could fuse the reference and test surfaces. When choosing the stimuli for these long duration practice trials, we decided to not include blur. The reason was that when subjects made eye movements to binocularly fuse the images, such blur would have remained and we were concerned that the blur was not due to defocus. Of course, not having any blur at all in these long duration practice trials was also incorrect but we felt it was the lesser of two evils.

Once it was determined that subjects could perform at near 100% for the long duration examples in which they could make eye moves, they were given 36 practice trials with the stimuli used in the experiment (which contained blur) and with the presentation time (200 ms) used in the experiment. This short presentation time did not allow subjects to make eye movements.

During the long duration practice trials, subjects were told they could look anywhere they wished. In the short duration practice trials and in the experiment itself, subjects were instructed to look at the fixation cross. The fixation cross and tiled window frame were shown continuously throughout the experiment. The two windows were gray except for the period in which the reference and test surfaces were shown.

Subjects first ran the depth discrimination experiment and then they ran the blur discrimination experiment.

3. Results

3.1. Depth discrimination

Fig. 2a–c shows the means of the percent correct scores for the three cue combinations of the depth discrimination experiment. These data include only 11 of 15 subjects, namely those who achieved an arbitrary performance criterion of at least 70% correct when the reference disparity was smallest (24′) and the reference and test disparities differed, i.e. cue combinations (1) and (2).

In general for cue combinations (1) and (2), performance fell as the reference disparity increased. In particular, for a fixed value of Δ disparity, performance decreased as the reference disparity increases. This was expected for cue combination (1) since depth from disparity is known to be less precise for larger disparities

---

[6] The crossover occurs in their plots when their reference depth is about 31 cm and the curves are well separated beyond a depth of 32 cm. To compute their disparities, we used a fixation distance of 27.5 cm and 6.5 cm interocular distance.

[7] By "dot density" here, we mean 4 dots/deg^2 whereas Glennerster used the term to refer to the percentage of pixels. For the unit conversion, we note that Glennerster's dots were 6′ wide and ours were 16′. Also note that Glennerster's "dots" were in fact squares.

[8] Although Held et al. reported in their paper that they used depth increments to define the test depths, in fact they used depth decrements (R. Held, personal communication).

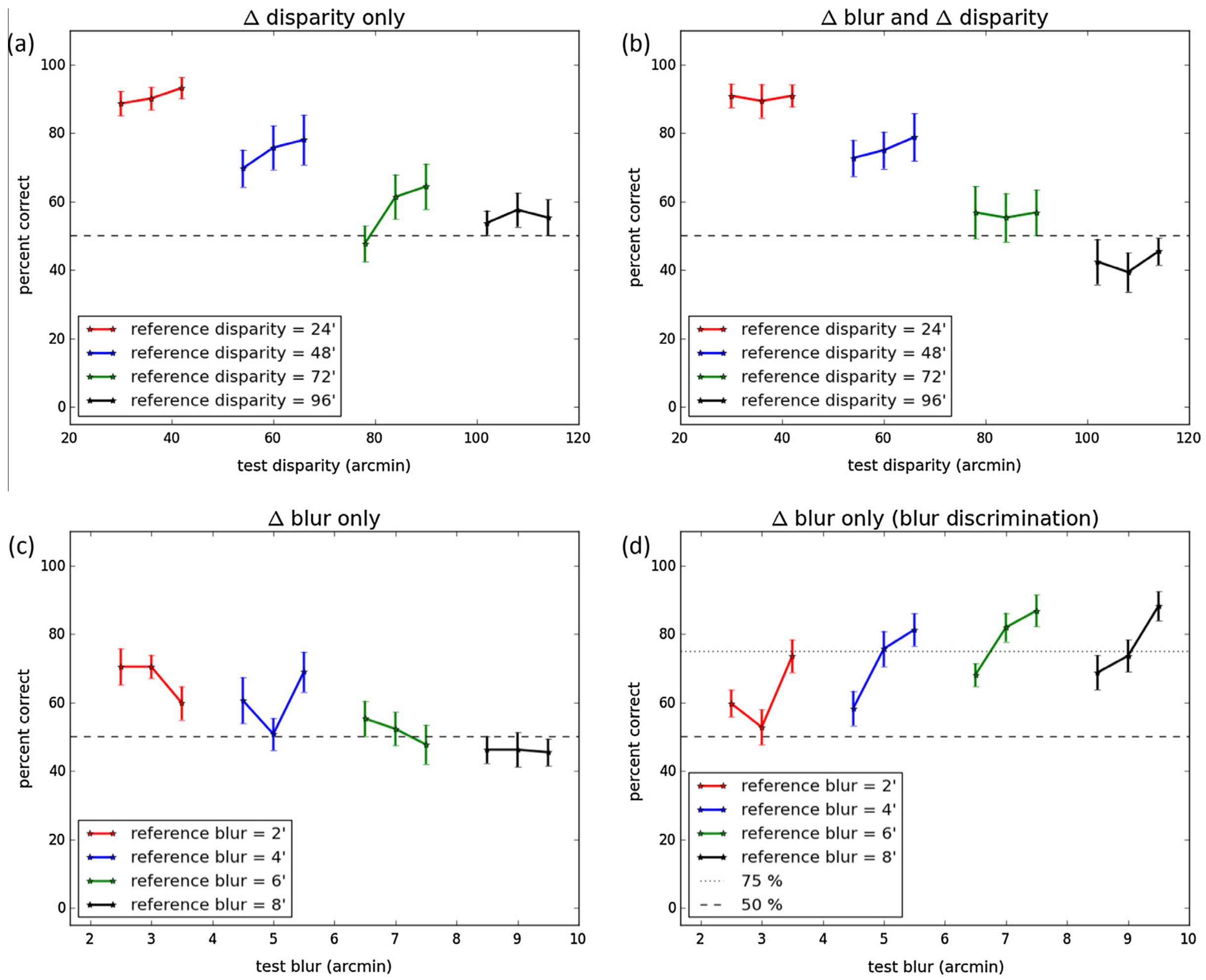

Fig. 2. (a–c) Percent correct scores for cue combinations (1)–(3) in the depth discrimination task. (d) Percent correct scores for the blur discrimination task which used cue combination stimuli (3). Error bars show the standard error of the mean. Individual subject data (not shown) is considerably more variable as each subject had only 12 trials per condition.

(recall the Introduction). Our data are also consistent with Held et al.'s data for cue combination (1). In their plots for the disparity-only condition, JND's increased with reference disparity.

Performance for cue combination (1) was similar to cue combination (2). See Fig. 2a and b. This suggests that our subjects were not using the blur cue that was available in the latter condition to discriminate depths when the reference disparity was large. These results are fundamentally different from those of Held et al. They found that, at large reference depths, subjects performed better in the blur-and-disparity condition than in the disparity-only condition. Held et al.'s subjects were able to use the blur cue to judge which surface was farther.

Results for cue combination (3) are shown in Fig. 2c. Here the "correct response" was defined such that that blurrier means farther. For the lower two values of reference blur, subjects were slightly above chance. This suggests that subjects may have used the blur cue when judging depth. Moreover, performance for these low values of reference blur was lower than in cue combinations (1) and (2) which is consistent with the results of Held et al.; when their reference blur and disparity values were small, the JNDs for blur were larger than for disparity. However, for the two larger values of reference blur, our subjects' performance fell to chance. This is inconsistent with Held et al.'s findings in the blur-only condition, namely their JND's were nearly constant across reference blur level in that condition.

3.2. Blur discrimination

One concern about the differences between our results and those of Held et al. might be whether or not there was a sufficient amount of blur present for our subjects to have used it in the depth discrimination task. Fig. 2d shows that there was. This figure shows the results of the blur discrimination task, which used the stimuli from cue combination (3). When the reference blur widths were $4'$, $6'$ and $8'$, subjects achieved 75% correct when the Δblur increment was roughly $1'$. For reference blur widths of $6'$ and $8'$, these performance levels were far greater than in cue combination conditions (2) and (3). Thus there was indeed ample blur information available to subjects in the depth discrimination task, especially in the cases that the reference disparities were large.

As an aside, we note that the 75% thresholds in Fig. 2d decreased slightly as the reference blur increased. This may seem surprising but in fact it is consistent with the blur discrimination literature. For example, blur discrimination JND's near the fovea follow a dipper function where the smallest JND's are achieved at a Gaussian standard deviation (σ) of roughly $1'$ to $2'$ (Watson & Ahumada, 2011). For more peripheral stimuli, a dipper function also exists and the JND minima occur at larger pedestal blurs than those near the fovea (Wang & Ciuffreda, 2005). For our results, the 75% thresholds decrease as the reference blur increases, which suggests our data lie on the downward slope of such a dipper function.

4. Discussion

The visual system is adept at discriminating blur, both in the fovea (Watson & Ahumada, 2011) and in the periphery (Wang & Ciuffreda, 2005). Since defocus blur of a 3D scene point varies with the distance from the focal plane, such blur carries information about depth. Indeed several studies have shown that blur information can be used in depth perception. For example, blur can be used to disambiguate the ordinal depth relationship at an image boundary (Marshall et al., 1996; Mather, 1996, 1997; Mather & Smith, 2002). Blur can also aid in segmentation of surfaces that lie at different depths (Hoffman & Banks, 2010). Finally, blur gradients that arise on slanted planes can be combined with binocular disparity and perspective to yield more accurate perception of surface slant (Watt et al., 2005) and absolute distance (Held et al., 2010; Vishwanath & Blaser, 2010).

Despite the above examples of how blur can be used in depth perception, blur is often regarded as a weak and unreliable depth cue only. Why is this so? One reason is that discriminating defocus blur is not the same as perceiving depth from blur. In particular, depth from defocus blur is limited by several ambiguities. First, there is a sign ambiguity: although scene points are blurred in proportion to the dioptric distance from the plane of focus, the sign information still needs to be provided. The sign can be provided by geometric cues such as disparity, occlusions and perspective when these cues are available and reliable.[9] Temporal fluctuations in accommodation (Charmon & Heron, 1988; Kotulak & Schor, 1986) also can be used to resolve the depth sign ambiguity. A second ambiguity for depth from blur, which is important for depth discrimination, is that the relationship between blur and distance depends on pupil size and on the distance to the focal plane. These variables are known only approximately in typical situations. A third ambiguity is that blur-like image patterns often arise for reasons other than de-focus. For example, shadow boundaries produce blurred image edges which are due to a penumbra, and so does smooth shading that is due to surface orientation variations (Elder & Zucker, 1995). Material changes across a surface can be gradual as well and hence can give rise to blur-like patterns.

While the above ambiguities imply there are fundamental limits on how much information about depth is available from blur, these ambiguities only concern the magnitude of blur, i.e. defocus. When other blur cues such as chromatic aberration are present, these cues provide additional information about depth. It is well known that chromatic aberration can be used to resolve the two-fold depth ambiguity mentioned earlier (Flitcroft, 1990; Kruger et al., 1993). There is also evidence that chromatic aberration can be used for depth discrimination (Nguyen, Howard, & Allison, 2005). Indeed a recent computational model has shown how chromatic aberration, as well as detailed spatial properties of the eye's point spread function such as astigmatism could allow the visual system to obtain quite accurate estimates of depth (Burge & Geisler, 2011).

Can these "higher order" blur cues be the reason why our experimental findings differed from those of Held et al? The blur in our stimuli were rendered by convolving a sharp image with a Gaussian whereas the blur in Held's et al.'s stimuli was produced by their subjects' optics. Although our stimuli contained sufficient defocus blur to perform the depth discrimination task, which we know from the results of the blur discrimination experiment, perhaps the rendered blur in our stimuli was perceived by our subjects as artificial and thereby not associated with depth, whereas the blur experienced by Held et al.'s subjects was due to their own optics and thereby was perceived as due to depth.

On the one hand, it would be not be surprising if subjects could distinguish the rendered blur used in our stimuli from real blur that is due to their own optics, as there is evidence that the visual system can discriminate quite subtle differences in blur. For example, Artal et al. found that if subjects are presented with monochromatic images that have been blurred using their own point spread function, then these images appear sharper than images that have been blurred using a rotated version of the their point spread function (Artel et al., 2004). This suggests that each person's visual system is adapted to the spatial properties of its own blur, similar to how it is adapted to chromatic aberration i.e. people do not perceive color fringes that occur at black and white edges.

On the other hand, it would be surprising if the visual system did not associate blur with depth, just because the blur was rendered rather than due to the subject's optics. Many studies of depth perception use images that are rendered using models that only crudely approximate the physics of image formation. For example, studies of shape from texture, contour, shading, and motion often use rendered stimuli that are obviously artificial, yet still give rise to 3D percepts. Why would the visual system be so finicky when it comes to associating blur with depth?

An alternative reason for why we obtained different results than Held et al. is that their subjects might not have been using depth perception to perform the task, but rather they might just have been discriminating blur. Held et al.'s naïve subjects were trained with trial-to-trial feedback prior to running the experiment, and through that training they may have learned (possibly subconsciously) to respond that the blurrier stimulus was farther. A simple way to rule out this explanation would be to repeat Held et al.'s experiment, again using a volumetric stereo display, but now to use naïve subjects who were not trained with trial-to-trial feedback.

In general, further experiments are needed to show which aspects of blur are used in depth perception, and to what extent the visual system treats rendered blur differently from real blur. Volumetric displays are a very good tool for such studies since they allow one to control blur cues and other cues. A recent experiment addressed ordinal depth judgments at edges and found that performance was better for stimuli presented with real optical blur on a volumetric display than on a single plane display (Zannoli et al., 2014). More of these studies are needed.

A second method for addressing the effects of real versus rendered blur could be to use a traditional stereo display but to use a more accurate rendering model (Barsky, 2011). To produce an accurate rendering would be very challenging, since there is considerable variation in the point spread functions between subjects, not merely in the lower order blur terms that are corrected with glasses but in higher order terms as well (Porter & et al., 2001). Thus, one would need to measure each subject's PSF and render different stimuli for different subjects. Measuring each subject's PSF at a single wavelength as in Wilson, Decker, and Roorda (2002) would not be enough since chromatic aberrations vary between subjects as well. Moreover, the rendering model would need to take account of the spectral emission of the display since the blur would be rendered for each of the display's RGB channels, not for each wavelength. Another complication is that if the stimuli span several degrees of visual angle as in our experiments, then one would need to measure the chromatic PSF in many different visual directions since the PSF varies considerably with visual direction. Finally, the subject's optics will produce retinal blur even for a focussed image and so when rendering the blur of a defo-

---

[9] Note that for our experiment, disparity and occlusion cues were available and we can assume that the sign of depth was correctly perceived, based on previous studies. Specifically, in Held et al.'s disparity-only condition, JNDs (75% correct) for depth discrimination were obtained for all disparity levels, which is only possible if the sign ambiguity has been resolved. Also, the Dmax data from Glennerster (1998) which was mentioned earlier suggests that sign of depth can be unambiguously perceived from our stimuli.

cussed image, one would need only to render the additional component that is due to the defocus. Of course, taking all of the above factors above is probably infeasible and one would have to use approximations. But it is unclear in advance which approximations to make: that is part of the research question being addressed.

If it should turn out that having correct spatial and chromatic properties of retinal blur is crucial for depth perception from blur, then this would explain why blur has thus far been regarded as a weak cue to depth and it would open up new lines of investigation into this relatively neglected depth cue.

Acknowledgments

The authors would like to thank Robert Allison, Laurie Wilcox, Robin Held, Emily Cooper, and Martin Banks for helpful discussions. This research was supported by an NSERC Discovery Grant to ML.

References

Artal, P., Chen, L., Fernández, E. J., Singer, B., Manzanera, S., & Williams, D. R. (2004). Neural compensation for the eye's optical aberrations. *Journal of Vision, 4*, 281–287.

Barsky, B. A. (2011). Vision-realistic rendering: Simulation of the scanned foveal image with elimination of artifacts due to occlusion and discretization. In P. Richard & J. Braz (Eds.). *Computer vision, imaging and computer graphics: Theory and applications. Communications in computer and information science* (Vol. 229, pp. 3–27). Springer-Verlag.

Blakemore, C. (1970). The range and scope of binocular depth discrimination in man. *Journal of Physiology, 211*, 599–622.

Burge, J., & Geisler, W. S. (2011). Optimal defocus estimation in individual natural images. *Proceedings of the National Academy of Sciences, 108*(40), 16849–16854.

Charmon, W. N., & Heron, G. (1988). Fluctuations in accommodation: A review. *Ophthalmic and Physiological Optics, 8*(2), 153–164.

Elder, J.H., & Zucker, S.W. (1998). Local scale control for edge detection and blur estimation.

Flitcroft, D. I. (1990). A neural and computational model for the chromatic control of accommodation. *Visual Neuroscience, 5*(6), 547–555.

Glennerster, A. (1998). Dmax for stereopsis and motion in random dot displays. *Vision Research, 38*(6), 925–935.

Held, R. T., Cooper, E. A., & Banks, M. S. (2012). Blur and disparity are complementary cues to depth. *Current Biology, 22*, 426–431.

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*, 1–16.

Hoffman, D. M., & Banks, M. S. (2010). Focus information is used to interpret binocular images. *Journal of Vision, 10*(5), 1–17 (article no. 13).

Howard, I. P., & Rogers, B. (2012). *Perceiving in depth: Stereoscopic vision* (Vol. 2). Oxford University Press.

Kotulak, J. C., & Schor, C. M. (1986). A computational model of the error detector of the human visual system. *Biological Cybernetics, 54*, 189–194.

Kruger, P. B., Mathews, S., Aggarwala, K. R., & Sanchez, N. (1993). Chromatic aberration and ocular focus: Fincham revisited. *Vision Research, 33*(10), 1397–1411.

Landy, M. S., Maloney, L. T., Johnston, E. B., & Young, M. (1995). Measurement and modelling of depth cue combination: In defense of weak fusion. *Vision Research, 35*(3), 389–412.

Love, G. D., Hoffman, D. M., Hands, P. J. W., 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.

Marshall, J. A., Burbeck, C. A., Ariely, D., Rolland, J. P., & Martin, K. E. (1996). Occlusion edge blur: A cue relative visual depth. *Journal of the Optical Society of America, 13*, 681–688.

Mather, G. (1996). Image blur as a pictorial depth cue. *Proceedings of the Royal Society of London, 263*, 169–172.

Mather, G. (1997). The use of image blur as a depth cue. *Perception, 26*, 1147–1158.

Mather, G., & Smith, D.R. (2000). Depth cue integration: Stereopsis and image blur.

Mather, G., & Smith, D. R. R. (2002). Blur discrimination and its relation to blur-mediated depth perception. *Perception, 31*, 1211–1219.

McKee, S. P., Levi, D. M., & Bowne, S. F. (1990). The imprecision of stereopsis. *Vision Research, 30*(11), 1763–1779.

Nguyen, V. A., Howard, I. P., & Allison, R. S. (2005). Detection of depth order of defocused images. *Vision Research, 45*, 1003–1011.

Ogle, K. N. (1952). Disparity limits of stereopsis. *AMA Archives of Ophthalmology, 48*(1), 50–60.

Ogle, K. N. (1953). Precision and validity in stereoscopic depth perception from double images. *Journal of the Optical Society of America, 43*, 906–913.

Peirce, J. W. (2007). Psychopy-Psychophysics software in Python. *Journal of Neuroscience Methods, 162*, 8–13.

Porter, J., Guirao, A., Cox, I. G., & Williams, D. R. (2001). Monochromatic aberrations of the human eye in a large population. *Journal of the Optical Society of America, 18*(8), 179–1803.

Richards, W., & Kaye, M. G. (1974). Local versus global stereopsis: Two mechanisms? *Vision Research, 14*, 1345–1347.

Schechner, Y. Y., & Kiryati, N. (2000). Depth from defocus vs. stereo: How different are they really? *International Journal of Computer Vision, 39*, 141–162.

Vishwanath, D. (2012). The utility of defocus blur in binocular depth perception. *i-Perception, 3*, 541–546.

Vishwanath, D., & Blaser, E. (2010). Retinal blur and the perception of egocentric distance. *Journal of Vision, 10*(10), 1–16 (article no. 26).

Wang, B., & Ciuffreda, K. J. (2005). Blur discrimination of the human eye in the near retinal periphery. *Optometry and Vision Science, 82*, 52–58.

Watson, A. B., & Ahumada, A. J. (2011). Blur clarified: A review and synthesis of blur discrimination. *Journal of Vision, 11*(5).

Watt, S. J., Akeley, K., Ernst, M. O., & Banks, M. S. (2005). Focus cues affect perceived depth. *Journal of Vision, 5*, 834–862.

Wilcox, L. M., & Allison, R. S. (2009). Coarse-fine dichotomies in human stereopsis. *Vision Research, 49*, 2653–2665.

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*(5), 833–839.

Zannoli, M., Albert, R. A., Bulbul, A., Narain, R., O'Brien, J. P., & Banks, M. S. (2014). Correct blur and accommodation information is a reliable cue to depth ordering. In *Vision science society annual meeting (Abstract)*, May 2014.