Blur and Disparity
Are Complementary Cues to Depth

Robert T. Held,1,* Emily A. Cooper,2 and Martin S. Banks3,2,4
1Computer Science Division
2Helen Wills Neuroscience Institute
3Vision Science Program
4Department of Psychology
University of California, Berkeley, Berkeley, CA 94720, USA

Summary

Estimating depth from binocular disparity is extremely precise, and the cue does not depend on statistical regularities in the environment. Thus, disparity is commonly regarded as the best visual cue for determining 3D layout. But depth from disparity is only precise near where one is looking; it is quite imprecise elsewhere [1–4]. Away from fixation, vision resorts to using other depth cues—e.g., linear perspective, familiar size, aerial perspective. But those cues depend on statistical regularities in the environment and are therefore not always reliable [5]. Depth from defocus blur relies on fewer assumptions and has the same geometric constraints as disparity [6] but different physiological constraints [7–14]. Blur could in principle fill in the parts of visual space where disparity is imprecise [15]. We tested this possibility with a depth-discrimination experiment. Disparity was more precise near fixation and blur was indeed more precise away from fixation. When both cues were available, observers relied on the more informative one. Blur appears to play an important, previously unrecognized [16, 17] role in depth perception. Our findings lead to a new hypothesis about the evolution of slit-shaped pupils and have implications for the design and implementation of stereo 3D displays.

Results

Assume an observer fixates and focuses on a point at distance z0 (Figures 1A–1C). Another point at z1 is imaged onto the two retinae. Horizontal disparity is the horizontal difference in the projected positions of that point, and is determined by distances z0 and z1 and some eye parameters:

\[ d = \frac{l}{z_0} \left( \frac{1}{z_0} - \frac{1}{z_1} \right) \]  

(1)

where \( d \) is in units of distance, \( l \) is interocular distance, and \( s \) is the distance between the eye’s optical center and the retina. Using the small-angle approximation to convert into radians and rearranging,

\[ \delta = l \left( \frac{1}{z_0} - \frac{1}{z_1} \right). \]  

(2)

We define blur as the diameter of the circle over which the point at \( z_1 \) is imaged at the retina. The blur-circle size in radians is determined by distances \( z_0 \) and \( z_1 \) and some eye parameters [6]:

\[ \beta = A \left| \frac{1}{z_0} - \frac{1}{z_1} \right| \]  

(3)

where \( A \) is pupil diameter. The analysis summarized by Equation 3 incorporates geometric blur due to defocus and not blur due to diffraction and higher-order aberrations [18]. Incorporating diffraction and aberrations would yield more blur but only for object distances at or very close to the focal distance. We are most interested in blur caused by significant defocus where geometric blur is the dominant source [19].

\( \delta \) and \( \beta \) are proportional to the difference between the reciprocals of \( z_0 \) and \( z_1 \) (i.e., the difference in diopters). Disparity and blur have very similar dependencies on scene layout because both are based on triangulation: disparity derives from the different positions of the two eyes and blur from light rays entering different parts of the pupil. Combining Equations 2 and 3 yields the relationship between disparity and blur for \( z_1 \):

\[ \frac{\beta}{\delta} = \frac{A}{T} \]  

(4)

The ratio \( A/T \) is \( \sim 1/12 \) for typical steady-state viewing situations [6, 20], so the magnitude of blur is generally much smaller than that of disparity. But this does not mean that depth estimation from blur is necessarily less precise than depth from disparity, because relative precision is also dependent on how the cues are processed physiologically.

The just-noticeable change in disparity is very small (~10 arcseconds) at fixation but increases dramatically in front of and behind fixation [1]. To reduce computational load, the visual cortex has many neurons with small receptive fields devoted to encoding small disparities (near fixation) and fewer neurons with larger receptive fields for encoding large disparities (far from fixation) [7, 8]. This strategy is manifest in the size-disparity correlation [10, 11]. The just-noticeable change in blur does not increase rapidly with base blur [12]. Not much is known about how the visual system encodes blur, but models have been developed that rely on pooling the responses of spatial-frequency-selective filters or neurons [13, 14]. One such model can, with few filters, achieve near-constant precision across a wide range of defocus levels [14]. Thus, the computational load for encoding changes in blur for different amounts of base blur may be relatively low, allowing the visual system to maintain roughy equal precision across a wide range of blurs. From these considerations, we hypothesize that depth from blur is more precise than depth from disparity for the parts of visual space in front of and behind where one is looking [15]. Such complementarity could be involved in conscious perception of depth and in programming of motor behavior such as eye movements and reaching.

We tested the complementarity hypothesis in a psychophysical experiment. Subjects indicated which of two stimuli appeared farther in three conditions: (1) blur alone (monocular viewing of stimuli whose focal distance varied), (2) disparity
alone (binocular viewing of stimuli whose focal distance was the same), and (3) disparity and blur (binocular viewing of stimuli whose focal distance varied). To do this, we used a unique stereoscopic, volumetric display developed in our laboratory [21]. This display allows the presentation of correct focus cues over a range of distances; without it, the current study would not be possible. In the apparatus, blur in the retinal image is created solely by the differences between the subject’s focus distance and the stimulus distance (i.e., not by rendering blurred images on a display screen).

The stimulus and results are shown in Figure 2. Figure 2B plots the just noticeable change in distance as a function of the distance of the nearer stimulus. The disparity-alone results confirm previous work showing that depth discrimination from disparity worsens very rapidly away from fixation [1]. The blur-alone results reveal that depth discrimination from blur is much poorer at fixation than depth from disparity but does not worsen significantly with increasing distance from fixation. At greater distance, depth from blur was actually more precise than depth from disparity. When both cues were present, subjects generally based discrimination on the more precise of the two, thereby yielding much better depth discrimination than if they had relied exclusively on disparity. The experiment was not designed to determine whether subjects integrated the two cues optimally [22], but we nonetheless calculated what the two-cue discrimination thresholds would be if optimal integration occurred. A sign test yielded no significant difference between optimal and observed two-cue performance (p = 0.34), but we cannot definitively determine whether the results reflect optimal cue combination or cue switching. Although we did not formally measure discrimination for points nearer than fixation, pilot testing showed that blur plays a similar role in that region of visual space.

Discussion

Here we consider the usefulness of blur in natural viewing and in the design of stereo 3D media, how blur can help guide motor behavior such as an upcoming eye movement, and how our theoretical and experimental results lead to a new hypothesis concerning the evolution of slit pupils.

Usefulness of Blur

We do not normally experience changes in the precision of depth estimates behind and in front of where we are looking. This state is often achieved by using other depth cues to fill in the gaps left by disparity. But the usefulness of some cues is quite dependent on viewing situation. For example, the utility of perspective depends on geometric regularities in the world. In contrast, blur is nearly always informative [14, 23]. Our results show that this generally available cue is indeed used to make depth estimation significantly more precise throughout visual space. This is surprising given that previous researchers argued that blur is not a very useful cue to depth [16, 24]. Blur has traditionally been regarded as a weak cue for two reasons.

(1) Defocus blur does not in any obvious way indicate the sign of a change in distance: i.e., whether an out-of-focus object is nearer or farther than an in-focus object. However, the visual system does clearly solve the sign-ambiguity problem. For depth estimation, the system solves the problem by using other depth cues that do not provide metric depth information. For example, the blur of an occluding contour determines whether an adjoining blurred region is perceived as near or far [16, 17]. Furthermore, perspective cues, which specify relative distance, provide disambiguating sign information, so blur plus perspective can be used to estimate...
absolute distance [6]. For driving accommodation in the correct direction, the system solves the sign-ambiguity problem using sign information contained in chromatic aberration [25], higher-order aberrations [19], and accommodative microfluctuations [26, 27].

(2) The relationship between distance and blur depends on pupil size (Equation 3). There is no evidence that people can measure their own pupil diameter, so the relationship between measured blur and specified distance is uncertain. But steady-state pupil size does not vary much under typical daylight conditions. Specifically, intrasubject pupil diameters vary over a range of 2.8 mm for luminance levels between 0.40 and 1,600 cd/m² [28], yielding an uncertainty in the estimate of z₀ of only 66% over a luminance range of 200,000%.

Stereoscopic 3D media is becoming increasingly commonplace. Our work shows that disparity and depth-of-field blur have the same underlying geometry and therefore that blur is roughly a fixed proportion of disparity (Equation 4). Given that the two cues complement each other, stereo 3D media should be constructed with this natural relationship in mind. When the natural relationship is violated, the puppet-theater effect (characters perceived as too small because of too much blur) or the gigantism effect (characters perceived as too large because of too little blur) may ensue [29].

Motor Behavior

We showed that depth from disparity is very precise near fixation but quite imprecise in front of and behind fixation [1]. It is also known that the precision of depth from disparity falls dramatically with increasing retinal eccentricity: above and below fixation [30, 31] and left and right of fixation [1, 30]. Importantly, blur-discrimination thresholds do not increase significantly with retinal eccentricity [32], so it is quite likely that depth from blur is more precise than depth from disparity above and below and left and right of fixation as well. Thus, there is a small 3D volume surrounding the current fixation in which depth from disparity can be estimated quite precisely. Outside this volume, the visual system must rely on other cues to estimate 3D structure. Having shown that blur fills in the void behind and in front of fixation, we hypothesize that it also does so left and right and above and below fixation. Alternatively, the viewer can make an eye movement to move the volume of high precision to a region of interest. However, when determining the movement required to fixate a new point in space, distance must be estimated to determine whether the eyes need to converge or diverge and by how much. Because disparity is imprecise away from fixation, blur may provide very useful information for programming the upcoming vergence eye movement.

To guide other motor behavior such as reaching and grasping, we must estimate metric distance. Can distances z₀ and z₁ be estimated from disparity and blur? Absolute distances can in fact be estimated from disparity. An extraretinal, eye-position signal is used to estimate the eyes’ vergence [33] and thereby estimate z₀. Because interocular distance f and eye length s are known, z₁ can also be estimated. This problem is also solved by using vertical disparity [34]. But can z₀ and z₁ be estimated from blur? If the pupil diameter A is known approximately, one would only have to estimate the eye’s current focal distance z₀. This distance could be estimated crudely from proprioceptive signals arising from structures controlling the focal power of the crystalline lens [35–38]. It could also be estimated from the eyes’ vergence if the eyes are fixated and focused on the same point at z₀. Finally, blur can act as an absolute cue to distance when combined with depth cues that provide relative depth information [6].

Slit Pupils

The pupils of many species are circular when dilated but slit-like when constricted. There have been three hypotheses about the utility of slit pupils: (1) they provide larger adjustments in area with simple musculature, which enables visual function in day and night [39], (2) they produce better image

Figure 2. Stimulus and Data

(A) Cross-fusable stereo pair of an example stimulus. Observers fixated the central cross and reported whether the lower or upper patch was farther away. Blur has been added to simulate the appearance of a stimulus with disparity and focus cues available.

(B) Depth-discrimination thresholds plotted as just-noticeable differences for disparity, blur, and both cues conditions. Subjects fixated and focused at a distance of 27.5 cm (dotted line). Pedestal stimuli are indicated on the abscissae. The ordinates represent the just-noticeable difference in depth for each pedestal distance. Blue, red, and black lines represent performance using disparity only, blur only, and both blur and disparity. The gray line represents the predicted behavior if the visual system combined the cues optimally [22]. Disparity outperformed blur when the pedestal stimulus was within 3 mm of fixation, whereas blur outperformed disparity at greater distances. Four of the panels show individual subject data and the fifth shows the across-subject average data. Error bars represent standard error.
quality for contours perpendicular to the pupil’s long axis [40], and (3) they preserve chromatic-aberration correction in some lenses when the pupil is constricted [41, 42]. As far as we know, slit or elliptical pupils are always either vertical or horizontal relative to the upright head. Species with vertical slits (listed in Figure 3 caption) are all nocturnal predators and nearly all of them hunt on the ground. Species with horizontal slits or ellipses (listed in caption) are all terrestrial grazers with laterally placed eyes. Hypotheses 1 and 3 above do not explain why slits are always vertical or horizontal nor why they are vertical in terrestrial predators and horizontal in terrestrial grazers. Hypothesis 2 predicts that pupils should be perpendicular to the horizon but has the effect of pupil diameter on visual resolution backward [40]. Our results showing the importance of blur for depth discrimination lead to a new hypothesis.

Consider a slit pupil with height $A_v$ and width $A_h$. With the eye focused at distance $z_0$, the retinal images of the limbs of a cross at $z_1$ would be blurred differently: the blur of the horizontal and vertical limbs would be determined by $A_v$ and $A_h$, respectively:

$$
\beta_h = A_h \left( \frac{1}{z_0} - \frac{1}{z_1} \right); \quad \beta_v = A_v \left( \frac{1}{z_0} - \frac{1}{z_1} \right). \quad (5 \text{ and } 6)
$$

Combining the two equations:

$$
\frac{\beta_v}{\beta_h} = \frac{A_h}{A_v}. \quad (7)
$$

For vertical slits, $A_v > A_h$, so $\beta_v > \beta_h$. For horizontal slits, $A_v < A_h$, so $\beta_v < \beta_h$. Thus, such eyes have astigmatic depth of field. With vertical slits, depth of field is smaller (i.e., blur due to misaccommodation is greater) for horizontal than for vertical contours; with horizontal slits, the opposite obtains. Figure 3A illustrates this point by showing the retinal images associated with crosses at different distances for a vertical-slit pupil. We hypothesize that slit pupils provide an effective means for controlling the amount of light striking the retina by enabling large changes in pupil area while also providing short depth of field for contours of one orientation (horizontal contours for vertical slits), which is useful for estimating distances of those contours. We thus predict that animals with vertical-slit pupils are better able to utilize the blur of horizontal contours to estimate depth than the blur of vertical contours.

The ground is a common and important part of the visual environment for terrestrial predators and grazers. With the head upright, the ground is foreshortened vertically in the retinal images, which increases the prevalence of horizontal or nearly horizontal contours in those images [43]. The vertical slit of many terrestrial predators aligns the orientation of shorter depth of field (Equation 7) with horizontal contours, which should allow finer depth discrimination of contours on the ground. This seems advantageous for their ecological niche (Figure 3B). Another observation is consistent with this hypothesis: The eyes of some reptiles with vertical-slit pupils rotate about the line of sight when the head is pitched downward or upward, which keeps the pupil’s long axis roughly perpendicular to the ground [40]. What about terrestrial grazers with horizontal slits? The eyes of these species are laterally positioned in the head, so when they pitch the head downward to graze, their pupils are roughly vertical relative to the ground. Again this arrangement aligns the orientation of the shorter depth of field with horizontal contours along the ground, which seems advantageous for their niche at least while grazing.

There is another potential advantage of vertical-slit pupils for terrestrial predators. Most of these animals have forward-facing eyes and stereopsis (unlike terrestrial grazers, who have lateral-facing eyes and minimal stereopsis). Vertical contours are critical for the computation of horizontal disparity, which underlies stereopsis. A large depth of field for vertical contours aids the estimation of depth from disparity, whereas a small depth of field for horizontals aids depth from blur for horizontal contours that are commonplace when viewing across the ground. This may be another sense in which disparity and blur are used in complementary fashion to perceive 3D layout.
Conclusions
We demonstrated through theoretical analysis and experimentation that blur provides greater depth precision than disparity away from where one is looking. These results are inconsistent with the previous view that blur is a weak, ordinal depth cue. They will aid the design of more effect stereoscopic 3D media, and also lead to a new hypothesis concerning the evolution of slit pupils.

Experimental Procedures

Subjects

Four subjects participated. R.T.H. (28 years old) and E.A.C. (23) were authors. P.V.J. (24) and J.O.L. (70) were unaware of the experimental goals. Before formal data collection began, each subject was given 30 min of training in the experimental task with trial-by-trial feedback. Subject protocol was approved by the University of California, Berkeley, Institutional Review Board.

Apparatus

The experiments were conducted on a stereoscopic, multplane display that provides nearly correct focus cues [21]. The display contains four image planes per eye. The planes are separated by 0.6 diopters. Distances inbetween planes are simulated by an interpolation algorithm that produces retinal images that are in most cases indistinguishable from real images [44–46]. The display allowed us to independently manipulate disparity and blur. The subjects’ eyes were fixed in focus at 27.5 cm (the distance of the nearest image plane) by use of cyclopia or temporary paralyzation of accommodation and ophthalmic lenses. Cyclopia causes pupil dilation, so to mimic natural pupils, we had subjects wear contact lenses with 4.5 mm diameter apertures. Subject J.O.L. was presbyopic (and therefore unable to accommodate), so his eyes were nontocycloped, and he viewed stimuli with natural pupils of 3.5 mm diameter.

Task and Stimulus

Stimuli were two rectangular patches of random-dot patterns (dot density = 4.2 dots/deg²). The patches were presented above and below a fixation cross that was centered in vision and was partially occluded by a solid frame (400 × 200 arcminutes) that bounded the stimulus region (see Figure 2A). The cross and frame were always presented at a distance of 27.5 cm. We included the frame to maintain a clear from occlusion that the stimuli were always farther than fixation. On each trial, the random-dot stimuli were presented simultaneously for 250 ms: one stimulus—the standard—had a distance of 27.5, 29.5, 31.5, 33.5, 35.5, 37.5, or 39.5 cm. An increment in distance was added to the other stimulus—the test—according to the method of constant stimuli. Regardless of distance, the stimuli had the same luminance and subtended the same visual angle at the eye. Subjects indicated with a key press which of the two stimuli appeared farther. The distance of the standard stimulus, the increment of the test stimulus, and which stimulus was above or below the fixation cross were randomized across trials. After a response was recorded, the next stimulus was presented after a delay of 500 ms.

Conditions

Three experimental conditions were presented: disparity only, blur only, and disparity and blur together. Subjects viewed the stimuli binocularly in the disparity-only condition. In that case, the multplane feature of the display was disabled, so the disparity of the standard and test stimuli differed, but the focal distances were the same. Subjects viewed the stimuli unocularly in the blur-only condition. The multplane feature of the display was enabled, so the focal distances of the standard and test stimuli differed. Subjects viewed stimuli binocularly in the disparity-and-blur condition with the multplane feature enabled, so the standard and test stimuli differed in disparity and focal distance. Subject J.O.L. frequently perceived the farther stimulus in the blur-only condition as blurrier than farther, and in those cases he responded by picking the blurrier stimulus.

Analysis

The psychometric data rose from ~50% to ~100% as the distance between the standard and test stimuli increased. We fit cumulative Gaussians to these data for each subject in each condition using a maximum-likelihood criterion [47]. The mean of the fitted function—the 75% point—was the estimate of the discrimination threshold for that subject and condition. Average data were calculated by fitting psychometric functions to the data from all of the subjects in a given condition.

Acknowledgments

We thank David Hoffman for his generous help with programming the experiment and Peter Vangorp for rendering the slit-pupil image in Figure 3B. The research was supported by National Institutes of Health grant EY012851, a National Science Foundation graduate fellowship to R.T.H., and a National Defense Science and Engineering Graduate Fellowship to E.A.C.

Received: November 5, 2011
Revised: December 3, 2011
Accepted: January 17, 2012
Published online: February 9, 2012

References