The natural statistics of blur

William W. Sprague Vision Science Program, University of California, Berkeley, Berkeley, CA, USA ✉

Emily A. Cooper Department of Psychological & Brain Sciences, Dartmouth College, Hanover, NH, USA ✉

Sylvain Reissier Engineering, Université Pierre et Marie Curie, Paris, France ✉

Baladitya Yellapragada Vision Science Program, University of California, Berkeley, Berkeley, CA, USA ✉

Martin S. Banks Vision Science Program, University of California, Berkeley, Berkeley, CA, USA
School of Optometry, University of California, Berkeley, Berkeley, CA, USA
Helen Wills Neuroscience Institute, University of California, Berkeley, Berkeley, CA, USA ⌂ ✉

Blur from defocus can be both useful and detrimental for visual perception: It can be useful as a source of depth information and detrimental because it degrades image quality. We examined these aspects of blur by measuring the natural statistics of defocus blur across the visual field. Participants wore an eye-and-scene tracker that measured gaze direction, pupil diameter, and scene distances as they performed everyday tasks. We found that blur magnitude increases with increasing eccentricity. There is a vertical gradient in the distances that generate defocus blur: Blur below the fovea is generally due to scene points nearer than fixation; blur above the fovea is mostly due to points farther than fixation. There is no systematic horizontal gradient. Large blurs are generally caused by points farther rather than nearer than fixation. Consistent with the statistics, participants in a perceptual experiment perceived vertical blur gradients as slanted top-back whereas horizontal gradients were perceived equally as left-back and right-back. The tendency for people to see sharp as near and blurred as far is also consistent with the observed statistics. We calculated how many observations will be perceived as unsharp and found that perceptible blur is rare. Finally, we found that eye shape in ground-dwelling animals conforms to that required to put likely distances in best focus.

Introduction

When a scene is imaged with a pinhole camera, all points in the image are sharp. But the light throughput of pinholes is very limited, so pinhole systems are generally impractical. Instead, consumer cameras and biological visual systems employ larger, adjustable apertures to increase light throughput and adjustable optics to bend the diverging light rays traveling through the aperture into a focused image. As a result, cameras and eyes with finite apertures and adjustable optics have a finite depth of field: Objects closer and farther than the focal distance create blurred images.

This *defocus blur* reduces the overall sharpness of retinal images, but it is also a potential source of information about 3D scene layout because blur magnitude is monotonically related to the difference in distance between an object and the point of fixation (i.e., where the eye is focused). The effects of defocus blur on image quality and the usefulness of blur as a depth cue are not well understood. The lack of
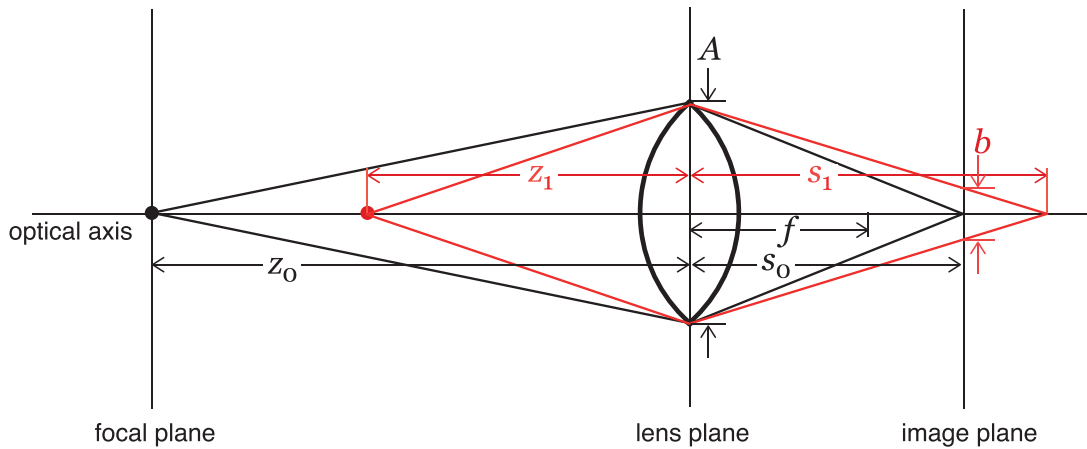
Figure 1. Defocus blur in a simple eye. z_0 is the focal distance of the eye given the focal length f and the distance from the lens to the image plane s_0. The relationship between these parameters is given by $1/s_0 + 1/z_0 = 1/f$. Thus, $f = s_0$ when $z_0 =$ infinity. An object at distance z_1 creates a blur circle of diameter b given the pupil diameter A.

understanding derives in large part from uncertainty about the patterns of defocus blur experienced in day-to-day life. Here we investigate both topics (informativeness about depth and effects on image quality) from the perspective of natural-scene statistics.

Geometry of defocus blur

Figure 1 illustrates the geometry that underlies defocus blur. 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 (black lines). Objects nearer or farther than the focal plane create blurred retinal images (z_1; red lines). Defocus blur is quantified by the diameter of the retinal image of a point of light. The diameter of the blur circle is

$$b = As_0 \left| \frac{1}{z_0} - \frac{1}{z_1} \right| \quad (1)$$

where A is pupil diameter, s_0 is the distance from the lens plane to the image plane, and z_1 is the distance to the object creating the blurred image. We can simplify Equation 1 by using the small-angle approximation:

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

where β is the blur-circle diameter in radians and ΔD is the difference in the distances z_0 and z_1 in diopters (inverse meters). The absolute value is used because defocus blur is unsigned: An object of a given relative distance from fixation will have the same defocus blur whether it is farther or nearer. Thus, defocus blur cannot by itself indicate whether the object creating the blurred image is farther or nearer than the current fixation distance.

Defocus as a depth cue

The systematic relationship between relative depth and defocus in Equations 1 and 2 suggests that the amount of blur could be a useful cue to depth. In support of this, Burge and Geisler (2011) showed that it is possible to estimate defocus from natural images without knowing the contents of the object (apart from statistically). Furthermore, the magnitude of blur can have a powerful effect on the perceived scale of a scene, an effect that has been called tilt-shift miniaturization (Held, Cooper, O'Brien, & Banks, 2010; Nefs, 2012; Okatani & Deguchi, 2007; Vishwanath & Blaser, 2010). In tilt-shift images, the sign of defocus blur does not need to be determined from blur signals because it is provided by other cues. In contrast to these findings, others have found that manipulations of defocus blur have inconsistent effects on judgments of depth order (Marshall, Burbeck, Ariely, Rolland, & Martin, 1996; Mather, 1996; Mather & Smith, 2002; Palmer & Brooks, 2008; Zannoli, Love, Narain, & Banks, 2016), on judgments of which of two images is closer to the viewer (Grosset, Schott, Bonneau, & Hansen, 2013; Maiello, Chessa, Solari, & Bex, 2015), and on judgments of amount of depth in a scene (Zhang, O'Hare, Hibbard, Nefs, & Heynderickx, 2014). In most of those cases, the sign of defocus blur had to be determined from blur signals in order to perform the task. Interestingly, when the blur is gaze contingent (i.e., the focal plane changes depending on where the viewer is looking), depth-order judgments improve presumably because the sign ambiguity is solved by the contingency (Mauderer, Conte, Nacenta, & Vishwanath, 2014).

Here we examine whether the sign ambiguity of defocus blur can be resolved through knowledge of the natural distribution of distances relative to fixation. We

do so by measuring the distributions of fixation distances, object distances, and defocus blur as humans perform natural tasks. We find that the distribution of object distances most likely to be associated with a given amount of defocus is not uniform across the visual field nor across fixation distances. We also report the results of a perceptual experiment that shows that observers take advantage of these statistical regularities in interpreting images with ambiguous blur.

Effect of defocus blur on perceived image quality

Pentland (1987) made an intriguing observation in his classic article on blur and depth estimation:

Our subjective impression is that we view our surroundings in sharp, clear focus. . . . Our feeling of a sharply focused world seems to have led vision researchers . . . to largely ignore the fact that in biological systems the images that fall on the retina are typically quite *badly* focused everywhere except within the central fovea. . . . It is puzzling that biological visual systems first employ an optical system that produces a degraded image, and then go to great lengths to undo this blurring and present us with a subjective impression of sharp focus. (p. 523)

There are a number of possible explanations for why the world appears sharp. First, apparent sharpness could be due to variations in visual attention. Attending to an object is usually accompanied by an eye movement to fixate the object and accommodation to focus its image. According to this view, detectable blur may be common but not noticed because it is unattended (Saarinen, 1993; Shulman, Sheehy, & Wilson, 1986). Second, apparent sharpness may be the consequence of neural mechanisms that sharpen the appearance of blurred objects, particularly in the peripheral visual field (Galvin, O'Shea, Squire, & Govan, 1997). That is, retinal images may often be quite blurred away from the fovea, but the blur is not perceived because of deblurring in subsequent neural processing. A third possibility is that as humans interact with natural scenes, the combination of fixation and accommodation produces retinal images that are reasonably sharp relative to the visual system's sensitivity to blur. In this view, blur is generally not detectable because larger blur magnitudes tend to occur outside the fovea where blur-detection thresholds are high (Wang & Ciuffreda, 2005).

Here we examine these possibilities by comparing the natural distribution of defocus blur in different parts of the visual field to the eccentricity-dependent thresholds for detecting the presence of blur.

Natural statistics experiment

We measured the distributions of blur due to defocus for each point in the central visual field as people engaged in natural tasks. To do this, we needed to measure the 3D geometry of visual scenes, the places in those scenes that people fixate, and pupil diameter. We made these measurements with a mobile eye- and scene-tracking system. This tracking system, along with a previous analysis of other aspects of the data set, is described in detail in Sprague, Cooper, Tošić, and Banks (2015). Here we provide an overview of the device and means of data collection and analysis.

Methods

Participants and tasks

Three young adults (all men, 21–27 years old) participated. The human subjects protocol was approved by the Institutional Review Board of the University of California, Berkeley. All participants gave informed consent. They all had normal best-corrected visual acuity and normal binocular vision. If they typically used an optical correction, they wore it during experimental sessions. The participants performed four tasks as they wore the eye- and scene-tracking system. Tasks were selected to represent a broad range of everyday visual experiences: walking outside in a natural area (*outside walk*), walking inside a building (*inside walk*), ordering coffee at a cafe (*order coffee*), and making a peanut butter and jelly sandwich (*make sandwich*). The sandwich-making task emphasizes near-work and has been used previously (Land & Hayhoe, 2001). The tasks were performed in separate sessions.

In many of the analyses reported here, we combined data across the four tasks. We wanted the combination to be as representative of typical visual experiences as possible. To achieve this, we used the American Time Use Survey (ATUS) from the U.S. Bureau of Labor Statistics. ATUS provides data from a large sample of the U.S. population, indicating how people spend their time on an average day. We assigned a set of weights to each secondary activity (in ATUS Table A1) corresponding to our four tasks (time asleep was given zero weight). Those weight assignments were done before we analyzed the eye-tracking and scene data, so there were no free parameters in weight assignment. The resulting weights represent estimates of the percentage of awake time that an average person spends doing activities similar to our tasks. They were 0.16 for outside walk,

0.10 for inside walk, 0.53 for order coffee, and 0.21 for make sandwich. We used these values to compute weighted-combination distributions of the relative distances between objects and fixation, and defocus blur. We did this by randomly sampling data from each task with probability proportional to the weights.

Apparatus

The eye- and scene-tracking system (Figure 2) has two outward-facing cameras (green box) that capture stereoscopic images of the scene in front of the participant and a binocular eye tracker that measures gaze direction for each eye (orange box). Two laptop computers were used, one for the eye tracker and one for the cameras and calibration (Sprague et al., 2015). The two computers were placed in a backpack worn by the participant and ran on battery power. With no cables tethered to external devices, participants had nearly full mobility while performing the experimental tasks.

Eye tracker

The eye tracker was a modified Eyelink II (SR Research), a head-mounted, video-based binocular eye-tracking system. The manufacturer reports an RMS angular error of 0.022° and field of view of 40° × 36°. The tracker was calibrated at the beginning of each session using the standard SR Research Application software. During calibration, the participant wore the device while positioned on a custom bite bar such that the eyes were at known positions relative to the display panel. The participants sequentially fixated nine small targets on the display screen in front of them. These fixations were used to map the eye-tracker data into a gaze direction for each eye in a head-centered coordinate system. Gaze direction at each time point for each eye was the vector from the eye's rotation center to the gaze coordinate on a plane 100 cm in front of the cyclopean eye (the midpoint between the eyes' rotation centers). Not surprisingly, the angular error between the fixation targets with known positions and the estimates of gaze direction for our whole pipeline was greater than the manufacturer's stated RMS error. In our whole system, the cumulative distribution function of the error magnitude across all subjects, gaze angles, and distances showed that 90% of the errors were less than 0.83°. For details, see Sprague et al. (2015).

Stereo cameras

The stereo cameras were Sony XCD-MV6 digital video cameras. Each has a 640 × 480 monochrome CMOS sensor, fixed focal-length (3.5 mm) lens, and 69°

× 54° field of view. Focus was set at infinity. Frame captures were triggered in a master/slave relationship with the master camera set to free run at 30 Hz. Each capture in the master camera sent a hardware signal triggering a capture in the slave camera. The two cameras were attached to the eye-tracker head mount on a rigid bar just above the tracker's infrared lights. The optical axes of the camera were parallel, separated by 6.5 cm, and pitched 10° downward to maximize the field of view that was common to the subject and cameras. The tracker and camera data were synchronized in time with a signal sent from the master camera to the tracker host computer. This signal was recorded on the host computer in the Eyelink Data File.

The intrinsic parameters of the individual cameras as well as their relative positions in the stereo rig were estimated prior to each experimental session using camera-calibration routines from OpenCV (v2.3.2; opencv.org) and a procedure described in Sprague et al. (2015). We discarded the data from an experimental session whenever the parameter estimates yielded an RMS error greater than 0.3 pixels between predicted and actual pixel positions of the calibration pattern.

Depth maps

The stereo images were used to generate depth maps of the visual scene. Disparities between each pair of images were estimated using the semi-global block matching algorithm in OpenCV (Hirschmüller, 2008). This method finds the disparity that minimizes a cost function on intensity differences over a region surrounding each pixel in the stereo pair. For details, see Sprague et al. (2015). Final disparity maps were converted to 3D coordinates, and the resulting matrices contained the 3D coordinates of each point in the scene in a coordinate system with the left camera of the stereo rig at the origin.

Pupil diameter

The Eyelink II reports pupil area in arbitrary units for each frame. To convert those units into diameters in millimeters, we measured the pupil diameter of each subject by placing a ruler in the pupil plane of the eye and taking a photograph. We then measured the width of the pupil in the image in pixels and used the image of the ruler to determine the corresponding width in millimeters. Photos were taken without a flash while the subject looked at target 4 m away in the calibration room with the overhead lights on and the calibration display off. At the beginning of the task, the room was returned to this condition, and the eye-tracking data were marked for reference. The pupil area reported at this reference point was used to map the eye tracker's
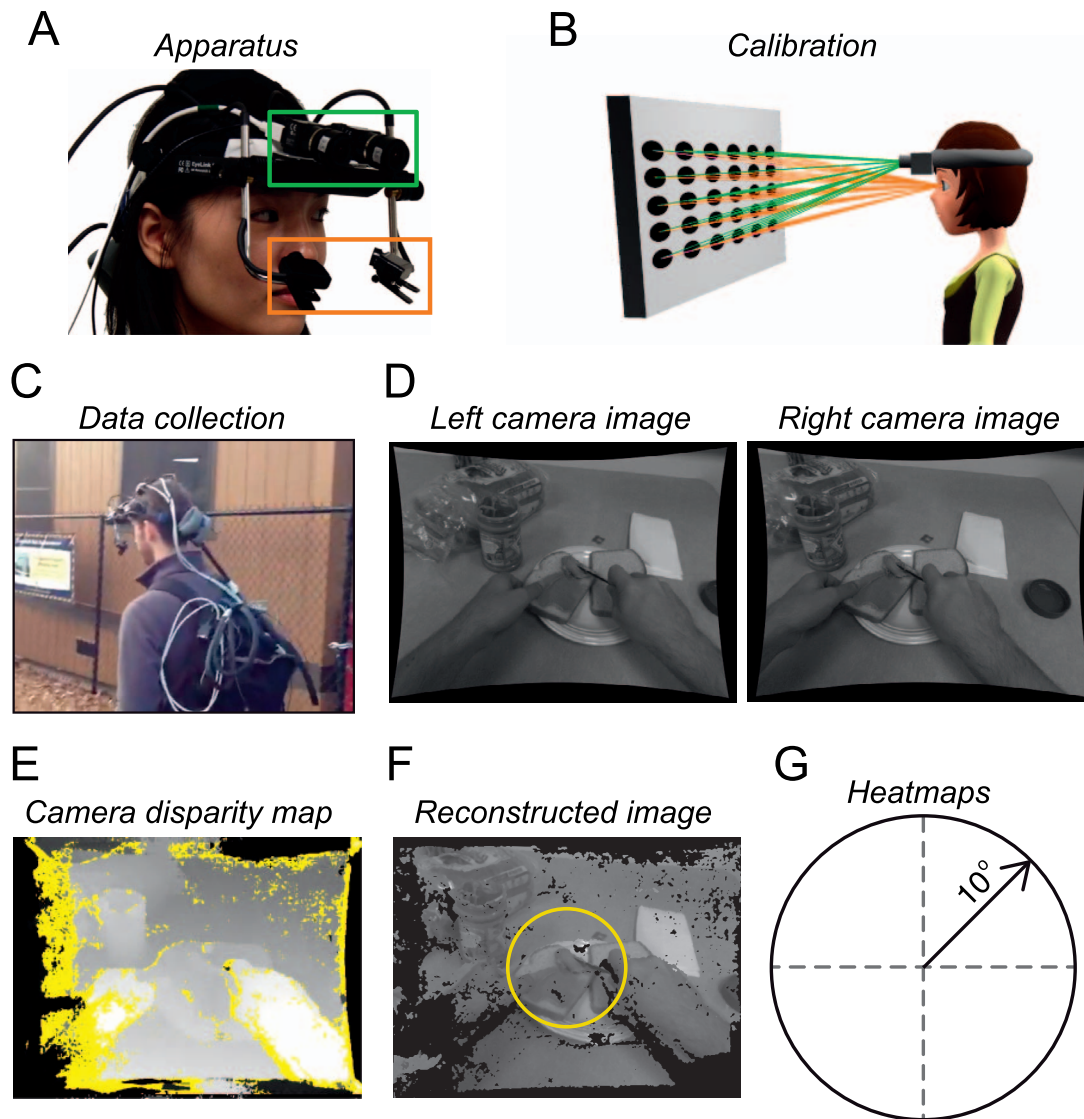
Figure 2. Apparatus and method for determining defocus blur when participants are engaged in everyday tasks. (A) The apparatus. A head-mounted binocular eye tracker (orange box) was modified to include outward-facing stereo cameras (green box). Eye-tracking and stereo-camera data collection were synchronized, and disparity maps were computed offline from the stereo images. (B) Calibration. To determine the distances of scene points, we had to transform the image data in camera coordinates to eye-centered coordinates. The participant was positioned with a bite bar to place the eyes at known positions relative to a large display. A calibration pattern was then displayed for the cameras and used to determine the transformation between camera viewpoints and the eyes' known positions. (C) Apparatus and participant during data collection. Data from the tracker and cameras were stored on mobile computers in a backpack worn by the participant. (D, E) Two example images from the stereo cameras and a disparity map. Images were warped to remove lens distortion, resulting in bowed edges. The disparities between these two images were used to reconstruct the 3D geometry of the scene. Those disparities are shown as a grayscale image, with bright points representing near pixels and dark points representing far pixels. Yellow indicates regions in which disparity could not be computed due to occlusions and lack of texture. (F) The 3D points from the cameras were reprojected to the two eyes. The example is a reprojected image for the left eye. The yellow circle (20° diameter) indicates the area over which statistical analyses were performed. (G) Data are plotted in retinal coordinates with the fovea at the center. Upward in the visual field is upward and leftward is leftward. Azimuth and elevation are calculated in Helmholtz coordinates.

arbitrary area units into diameters in millimeters using the previously measured pupil width.

We were concerned that the tracker's reported pupil areas were affected by gaze direction (Brisson et al., 2013; Gagl, Hawelka, & Hutzler, 2011). To examine this, we extracted the tracker's reported area when the pupil was fixed in size for a wide variety of gaze directions. We cyclopleged the right eye of one subject

with tropicamide. This agent dilates the pupil, making it unresponsive to changes in light level and accommodative state. We then had the subject sequentially fixate the nine calibration targets on the display screen at a distance of 100 cm. With each fixation, we recorded the tracker's estimate of pupil area. We found that the diameter estimates derived from the area estimates were reasonably constant (within ~8%) for all but one gaze direction. When the subject looked 12° up from straight ahead, the diameter estimate increased by ~25% relative to the mean estimate for the other gaze directions. But subjects very rarely fixated that far upward in the experimental session, so we decided not to correct for this in the estimates of pupil diameter.

Prior work has shown that eye-tracker estimates of gaze direction are affected by fluctuations in pupil size (Choe, Blake, & Lee, 2016; Ivanov & Blanche, 2011). These effects are contingent on gaze direction and thus are difficult to correct systematically. Because reported effects are generally much less than 1° of error for a change in pupil area of up to 30% (Choe et al., 2016), we did not attempt to measure and correct for potential pupil-size effects in our estimates of gaze direction.

Full system calibration

The three-dimensional (3D) coordinates estimated from the stereo cameras were in camera coordinates, so we needed to translate and rotate these points into eye coordinates. To determine the required transformation from cameras to eyes, we performed a calibration procedure. The participant wore the tracker and camera system while positioned precisely with the bite bar. The left camera captured images of a calibration pattern presented at 50, 100, and 450 cm from the participant's cyclopean eye. OpenCV routines were then used to estimate the translation and rotation between the centers of projection of the cameras and a cyclopean-eye–referenced coordinate system. Because gaze direction was also measured in cyclopean-eye coordinates, we could then transform the 3D locations in the scene into cyclopean-eye coordinates for each point in time captured by the cameras.

We measured system error immediately before and after each participant performed a task. While positioned on the bite bar at three screen distances (50, 100, and 450 cm), participants fixated a series of targets in a manner similar to the standard Eyelink calibration. Each target pattern contained a small letter "E" surrounded by a small circle. We instructed participants to fixate the center of the "E" and to focus accurately on it. The participant indicated with a button press when he or she was fixating and focusing the pattern, and eye-tracking data were gathered for the next 500 ms. At the same time, the stereo cameras captured images of each target pattern.

We measured the total error of the reconstruction by comparing the expected location of each "E" in each eye during fixation (i.e., the foveas) to the system's estimate, including all stages of 3D reconstruction, gaze estimation, and coordinate transformation. Performing the pre- and posttests enabled us to identify sessions in which the tracker moved substantially during task performance, resulting in large errors at the end of the task. Sessions with large reconstruction errors (~25%) were rerun. The median error magnitude (in version) was 0.29° across all participants, distances, and eccentricities. Ninety-five percent of the errors were less than 0.82°. Once all sessions were completed for each participant and task, the data from the sessions were trimmed to 2-min clips (3,600 frames). This resulted in 12 clips overall (three participants, four tasks) for a total of 43,200 frames (minus some lost frames due to blinks).

Because the distances of the fixation points were much greater than the interocular distance, the estimates of vergence were more sensitive to tracker error than estimates of version. Our participants had no measurable eye misalignment according to clinical tests, so we assumed that they correctly converged on surfaces in the direction of gaze (rather than converging in front of or behind the scene). We thus adjusted the vergence estimate for each gaze direction so that the fixation point was on a surface in the scene. These adjusted distances were used as the fixation distances in all of the data analyses.

Representation of defocus blur data

The cameras have limited fields of view, so we selected only the parts of the images within 10° of the fovea (20° diameter) to minimize missing data during analysis. From the estimated 3D position of each visible point in cyclopean-eye coordinates, we calculated the relative distances to fixation. We define *relative distance* as the difference between the radial distance from the cyclopean eye to the fixation point and the radial distance from the cyclopean eye to the scene point where those distances are expressed in diopters (inverse meters). Thus, points with zero relative distance are on the sphere whose center is the cyclopean eye and whose radius is the fixation distance. The cyclopean-centered system yields distances that are often in between distances for left eye– and right eye–centered systems, but the differences are very small, so they are ignored here. We calculated relative distance in each frame for each position in the central visual field. We used Helmholtz coordinates to represent the data with the fovea at the origin (Howard & Rogers, 2002). In these coordinates, azimuths are latitudes and elevations are longitudes. Positive relative distances correspond to scene points farther than fixation and

negative distances to points nearer than fixation. We could not measure focus distance (the distance to which the eyes were accommodated) directly. Instead, we assumed that our young adult subjects accommodated accurately to the distance they were fixating. In support of this assumption, López-Gil et al. (2013) showed that young adults accommodate accurately over a nearly 4D range when they view targets with natural textures binocularly. Specifically, they showed that accommodation is accurate enough over that range to maintain maximum visual acuity.

For each measurement of relative distance, we converted to defocus blur using Equation 2 and the measurements of pupil diameter at the corresponding time point. Thus, our estimates of defocus blur represent the angular diameter of the blur circle created by a point at that distance.

We used a geometric optical model (Figure 1) to determine the magnitude of defocus blur. The human eye has aberrating elements that are not captured by this model (e.g., diffraction, higher-order aberrations, and chromatic aberration), so we wanted to know how these additional effects would affect our estimates of blur in retinal images. In the geometric model, the eye's point-spread function (PSF) is a cylinder whose diameter is described by Equations 1 and 2.

To compare the geometric model to the aberrated eye, we first determined the PSFs associated with one of our participant's eyes by measuring the wavefront aberrations with a Shack-Hartmann wavefront sensor. Pupil diameter during the measurement was 7 mm. Wavelength was 840 nm. The obtained Zernike coefficients were typical of those in healthy young adults (Thibos, Hong, Bradley, & Cheng, 2002). From the wavefront measurements, we calculated the pupil function:

$$P(x, y) = A(x, y)e^{-ikW(x,y)} \quad (3)$$

where A is 1 for points inside the pupil and 0 otherwise, W is the aberrated wavefront, and $k = (2\pi)/\lambda$, where λ is wavelength. The PSF is then the magnitude of the Fourier transform of the pupil function. We modified this calculation to incorporate the effects of chromatic aberration, pupil size, and defocus. We modeled chromatic aberration by assuming an equal-energy white stimulus and, for each wavelength in steps of 1 nm, calculated defocus caused by longitudinal chromatic aberration using the model from Marimont and Wandell (1994):

$$D(\lambda) = 1.7312 - \frac{633.46}{\lambda - 214.10} \quad (4)$$

where λ is wavelength in nanometers and $D(\lambda)$ is diopters as a function of wavelength. Each PSF was normalized and weighted by human photopic spectral sensitivity and recombined to form an overall percep-

tually weighted PSF (Ravikumar, Thibos, & Bradley, 2008). Pupil size was modeled by normalizing the pupil diameter of interest relative to the diameter during the wavefront measurements. Defocus in diopters was converted into the spherical Zernike coefficient via

$$c_2^0 = \frac{Dr^2}{4\sqrt{3}} \quad (5)$$

where c_2^0 is the Zernike coefficient, D is defocus in diopters, and r is pupil radius (Thibos et al., 2002). To compare the PSFs from the aberrated eye and the geometric model, we calculated encircled energy—a measure of the concentration of the PSF—for both functions. We first determined the total energy and centroid of the aberrated PSF. Circles of increasing diameter were created from the centroid, and the energy within each circle was calculated and divided by the total energy. We did the same for the cylinder functions of the geometric model. For both types of PSF, we found the diameter that encircled 50% of the energy. Figure 3 shows the results. Encircled energy diameter in minutes of arc is plotted as a function of relative distance in diopters, where 0 is perfect focus. The solid green and dashed gray lines represent the results for the aberrated and geometric models, respectively. The three sets of contours are the results for pupil diameters of 2, 4, and 7 mm. Most of the observations in our data set had a pupil diameter of ~6 mm. The geometric model is an excellent approximation for all conditions, except where relative distance is 0 or quite close to 0. When the relative distance is less than 0.2D from perfect focus, we slightly underestimated the blur of the retinal image.

Results

We measured defocus blur across the visual field as people performed the four natural tasks shown in Figure 4A. The data from each task were very similar across subjects, so we report across-subject averages.

Figure 4B shows the median relative distances for each task as a function of location in the visual field. Zero indicates distances equal to the fixation distance; negative values are closer and positive ones are farther than fixation. The pattern of relative distance across the visual field differed substantially across tasks. In the *inside walk* task, relative distance tended to be negative (object distance nearer than fixation distance), whereras in the *order coffee* task, it tended to be positive (object distance greater than fixation). In the *make sandwich* and *outside walk* tasks, relative distance was generally positive in the upper visual field and negative in the lower field. There was more variation in median relative distance in the *make sandwich* task than in the other
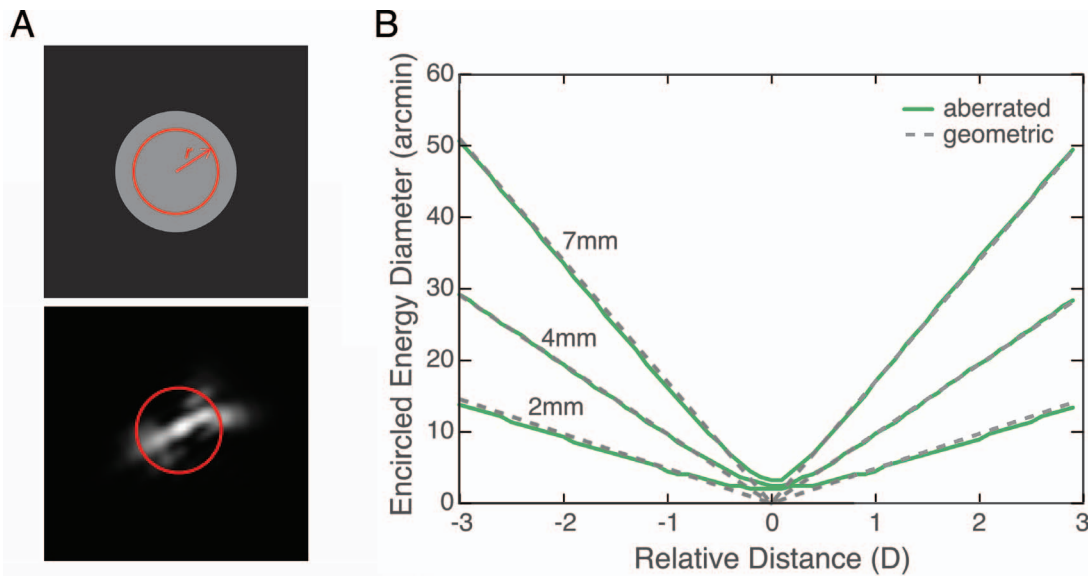
Figure 3. Encircled energy for aberrated eye and for cylinder approximation. (A) The point-spread function (PSF) for the cylinder function we used is shown in the upper panel. The PSF for an eye with typical aberrations is shown in the lower panel. Modeled pupil diameter was 3.5 mm, relative distance was −0.5D, and the light was equal-energy white. The red circles represent the circles used to calculate encircled energy. They have radii of *r*. (B) Diameters of circles that encircle 50% of the energy as a function of relative distance. A relative distance of 0 corresponds to perfect focus. The three sets of curves are for pupil diameters of 7, 4, and 2 mm. (Note that the PSF in the lower left panel is not for a condition on one of the green curves; instead, it lies between the 2- and 4-mm curves.) The solid green lines represent the encircled energy for the aberrated eye. The dashed gray lines represent them for the cylinder approximation.

tasks because the fixation and scene distances were generally shorter in that task.

We next examined the pattern of defocus blur across the central visual field. As we said earlier, defocus blur is unsigned and therefore does not directly signal whether a scene point is farther or nearer than fixation. Figure 4C shows the median blur diameters for the four tasks. Unlike the two-sided relative-distance distributions represented by the medians in 4B, the blur distributions are single sided and bounded at zero, so the median values reflect the overall spread of the distribution away from zero. The pattern of defocus blur across the visual field is much more similar between tasks than the pattern of relative distance. In particular, median blur values tend to increase radially with retinal eccentricity, more so for the *make sandwich* task than for the others.

To get a better gauge on the overall blur statistics, we next examined the results for the weighted combination of data across the four tasks. Figure 5A plots the median relative distances across the central visual field, revealing a distinct pattern. The variation in relative distance is much greater across elevation than across azimuth. Specifically, the medians are positive in the upper visual field because objects tended to be farther than fixation in that part of the field, and the medians are negative in the lower field because objects tended to be nearer than fixation there (Hibbard & Bouzit, 2005; Sprague et al.,

2015). The medians are close to zero along the horizontal meridian, meaning that points left and right of fixation were generally at about the same distance as the fixated point. Figure 5B plots the medians of the corresponding blur-circle diameters for the weighted combination across the tasks. Median blur increases with retinal eccentricity in a mostly radial fashion. We know of course from Figure 5A that the cause of the blur differs across the visual field. We quantified this by calculating for each position in the visual field the percentage of blur observations that are associated with positive relative distances (i.e., caused by scene points farther than fixation). Those percentages are shown in Figure 5C. There is a striking difference between the upper and lower visual fields: Most of the observed blurs in the upper field are associated with positive relative distance and most in the lower field with negative relative distance. We return to this observation later when we ask if viewers attribute blur observed in the upper field to more distant scene points and blur in the lower field to nearer points.

We investigated the degree to which these results depend on the weights assigned to the tasks. We did so by comparing the data in Figure 5A when the weights were determined by ATUS to the data when the same weight (0.25) was assigned to each task. The data (not shown) were qualitatively similar in the two cases, but the change in relative distance as a function of elevation
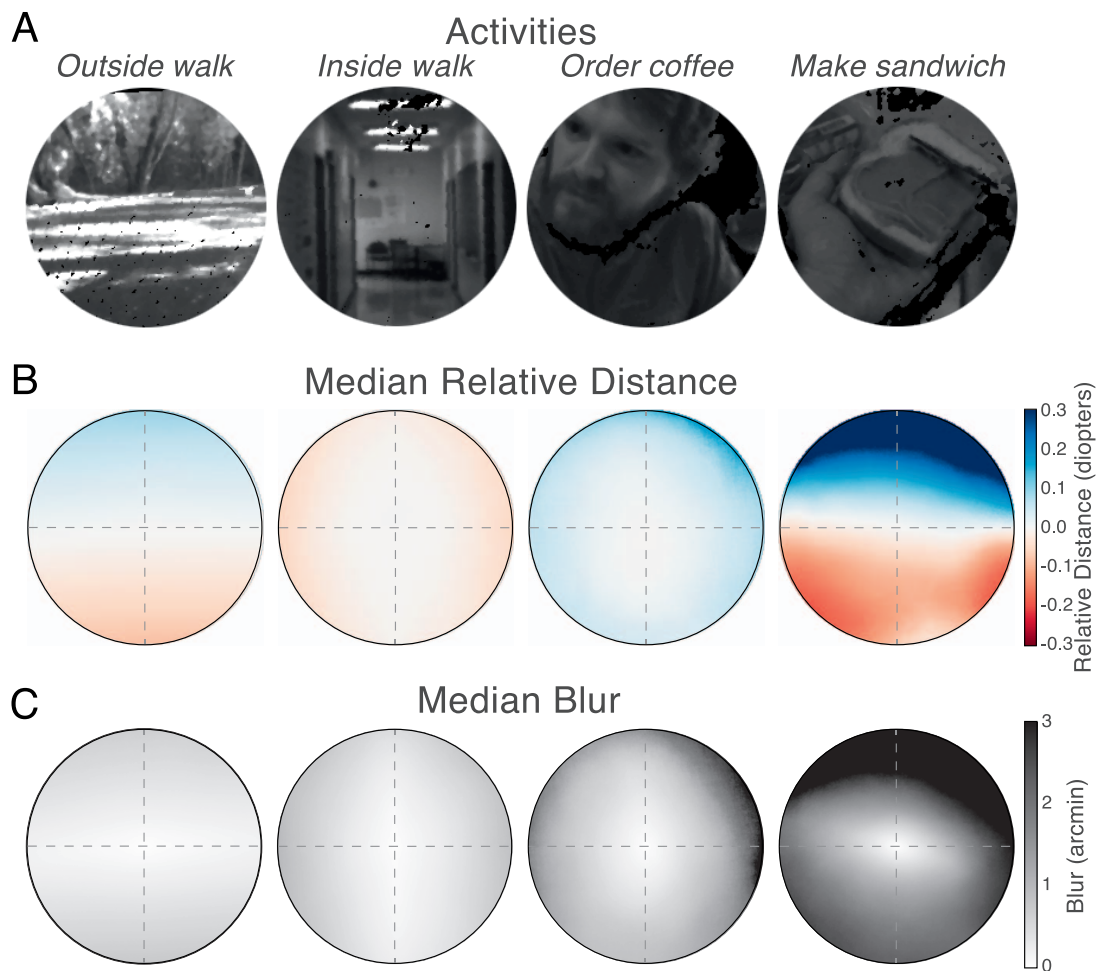
Figure 4. Median relative distances and blur-circle diameters for each task. (A) Icons representing the four natural tasks. From left to right, they are *outside walk*, *inside walk*, *order coffee*, and *make sandwich*. (B) The medians of the relative-distance distributions (i.e., the difference between the distance to a scene point and the distance to the fixation point in diopters) are plotted for each point in the central 20° of the visual field. Upward in the panels is up in the visual field and leftward is left. Each panel shows the medians for the task above it combined across subjects. Blue represents positive relative distance (points farther than fixation), and red represents negative relative distance (nearer than fixation). The values at the top of the right panel are slightly clipped. (C) The medians of defocus blur for the four tasks. The median diameter of the blur circle is plotted for each point in the central visual field. Darker colors represent greater diameters. The values at the top of the right panel are slightly clipped.

was reduced with equal weights compared with ATUS weights.

Videos in Figure 6 show the full relative-distance and defocus blur distributions for the weighted combination across tasks. Figure 6A and C shows the distributions along the vertical meridian and Figure 6B and D the distributions along the horizontal meridian. The relative-distance distributions (A and B) are leptokurtic, meaning they have a dominant central tendency with long tails. The spread of those distributions increases with increasing eccentricity, but the great majority of observations are within 0.5D of best focus. The distributions along the vertical meridian are positively skewed in the upper field (longer tail of positive observations than of negative observations) and negatively skewed in the lower field. The medians along the vertical meridian of course switch from positive in the upper field to negative in the lower. The relative-distance distributions along the horizontal meridian are also leptokurtic and become more spread with greater eccentricity. They do not exhibit left-right asymmetries like the upper-lower asymmetries along the vertical meridian. The blur distributions along the vertical and horizontal meridians (C and D) are similar to one another because positive and negative relative distances contribute equally to blur magnitudes. The spread of the blur distributions increases with eccentricity in a similar fashion along the vertical and horizontal meridians. The great majority of blur observations are smaller than 10 arcmin. We conclude that most observed blurs in the central visual field are close to zero. In other words, large blurs are rare.
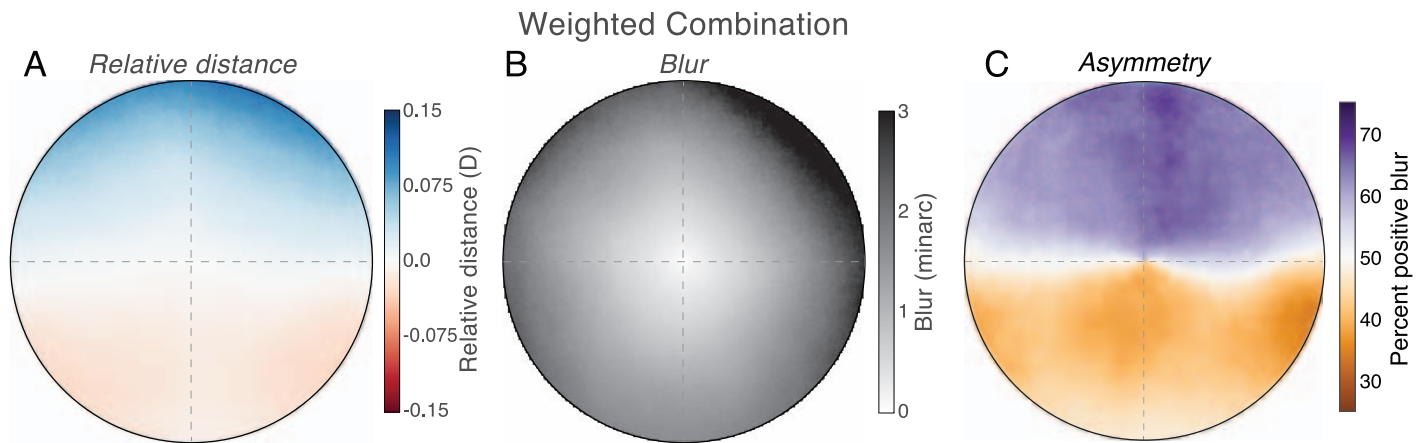
Figure 5. Relative distance, blur, and asymmetry for the weighted combination of data across tasks. (A) The medians of the relative-distance distributions for the weighted combination of the data from the four tasks. Blue and red represent positive and negative relative distances, respectively. (B) Medians of blur for the weighted combination of data. Blur values are the diameters of the blur circle. Darker colors represent larger values. (C) Asymmetry. The percentages of blur observations that are associated with positive relative distance are plotted for each position in the visual field. Purple indicates values greater than 50% and orange values less than 50%.

We have emphasized how the utility of defocus blur as a depth cue is mitigated by its sign ambiguity: The same blur can arise from points farther or nearer than fixation. But our results show that there are systematic asymmetries in the probability that a given blur is created by an object nearer or farther than fixation. These asymmetries provide prior information that could potentially be used to disambiguate the sign. For example, a blurred image in the upper visual field is likely to be caused by an object farther than fixation, and a blurred image in the lower field is likely to be caused by an object closer than fixation.

Are there other patterns in the natural blur distribution that can provide prior information about an object's distance? To examine this, we return to Equation 1. A is the diameter of the viewer's pupil, z_0 is the distance to which the eye is focused, and z_1 is the distance to the object creating the blurred retinal image. From the equation, we can calculate the theoretical probability of z_0 (focal distance) and z_1/z_0 (distance ratio; note that this is different from our definition of relative distance) for different amounts of blur (Held et al., 2010). We set A to 5.8 ± 1.1 mm based on our measurements of pupil diameter in the experimental sessions. The theoretical distributions are shown in Figure 7A. Each color represents a different amount of blur, and the V-shaped lines show the possible combinations of focal distance and distance ratios that could generate this blur. The distributions differ for large and small blurs: Large blurs are consistent with a range of short focal distances (blue), and small blurs are consistent with a larger range of focal distances (green). But for each

blur magnitude, infinite combinations of z_0 and z_1/z_0 are theoretically possible. One clearly cannot estimate focal distance or distance ratio from a given blur observation from these geometric considerations alone.

But natural fixations and scenes are not random, so we expect the empirically observed distributions to differ from the theoretical ones. To investigate this, we used our fixation and scene data to determine which combinations of focal distance and distance ratio are most likely to occur in reality. Figure 7B shows the empirical data in the same coordinates as the theoretical distributions. The empirical distributions are based on the weighted combination of data across tasks. Some properties of the empirical distributions stand out. (a) Focal distances shorter than 0.2 m are essentially never observed, presumably because participants adjusted their views so that very near (and therefore unfocusable) points were not visible in the central visual field. Because of this, large blur magnitudes almost never occur when the object distance is similar to the focal distance (i.e., when the distance ratio is ~1) even though such situations are theoretically possible. (b) Large blur magnitudes (blue points) are generally caused by objects farther than fixation rather than nearer (i.e., large blurs are more likely when the distance ratio is greater rather than less than 1). Thus, unlike the theoretical distributions, large blurs are generally not due to an object being nearer than fixation. (c) Small blur magnitudes (green) are caused by objects being farther or nearer than fixation with roughly equal probability.
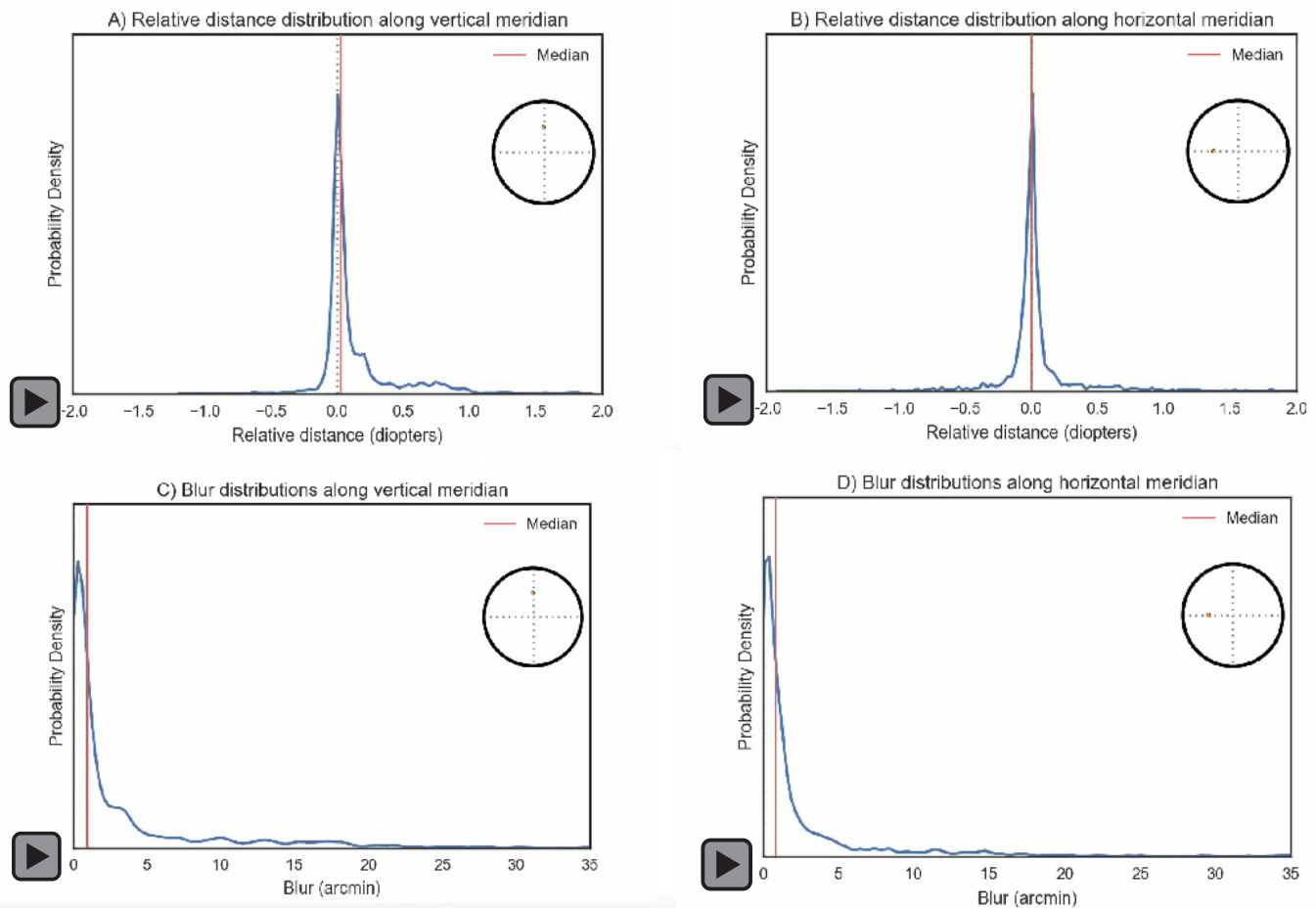
Figure 6. Videos showing relative-distance and blur distributions near the vertical and horizontal meridians of the visual field. (A) Probability of different relative distances for a strip $\pm 2°$ from the vertical meridian. The data are the weighted combination across tasks, averaged across subjects. The vertical dashed line indicates a value of 0. The vertical orange line represents the median relative distance for each elevation. The dot in the circular icon in the upper right indicates the position in the visual field for each distribution. (B) Probability of different relative distances for a strip $\pm 2°$ from the horizontal meridian. The conventions are the same as in A. (C) Probability of different blur values along the vertical meridian. The data are the weighted combination across tasks and subjects. The vertical line represents the median. (D) Probability of different blur values along the horizontal meridian. The vertical line again represents the median.

We can also ask what the probabilities of different combinations of focal distance and distance ratio are without regard to the magnitude of blur. To this end, Figure 8 plots the frequency of different observations of focal distance and distance ratio for the weighted combination of data across tasks. As you can see, distance ratios greater than 1 are much more likely when the focal distance is short, and ratios less than 1 are more likely when the focal distance is long. Thus, the empirically observed combination of focal distance, object distance, and blur is quite different from that expected from geometric considerations alone (Figure 7A). The differences reflect viewers' fixation strategies and the 3D structure of the viewed environment. Similar patterns in binocular disparity as a function of fixation distance have been observed in data with simulated scenes and fixations (Hibbard & Bouzit, 2005; Liu, Bovik, & Cormack, 2008).

In summary,

1. We find that most defocus blur magnitudes in the central visual field are within ± 10 arcmin ($\pm {\sim}0.5D$). Thus, large blur magnitudes are relatively rare.
2. We find that the variation of defocus blur grows in mostly radial fashion with increasing eccentricity. But the causes of the variation with elevation and azimuth differ. In the upper visual field, blur is generally caused by objects that are farther than fixation. In the lower field, it is generally caused by objects nearer than fixation. There is no such asymmetry in the left and right fields.
3. We find that large blur magnitudes are most likely to occur when fixation is near and the object creating the blur is more distant.
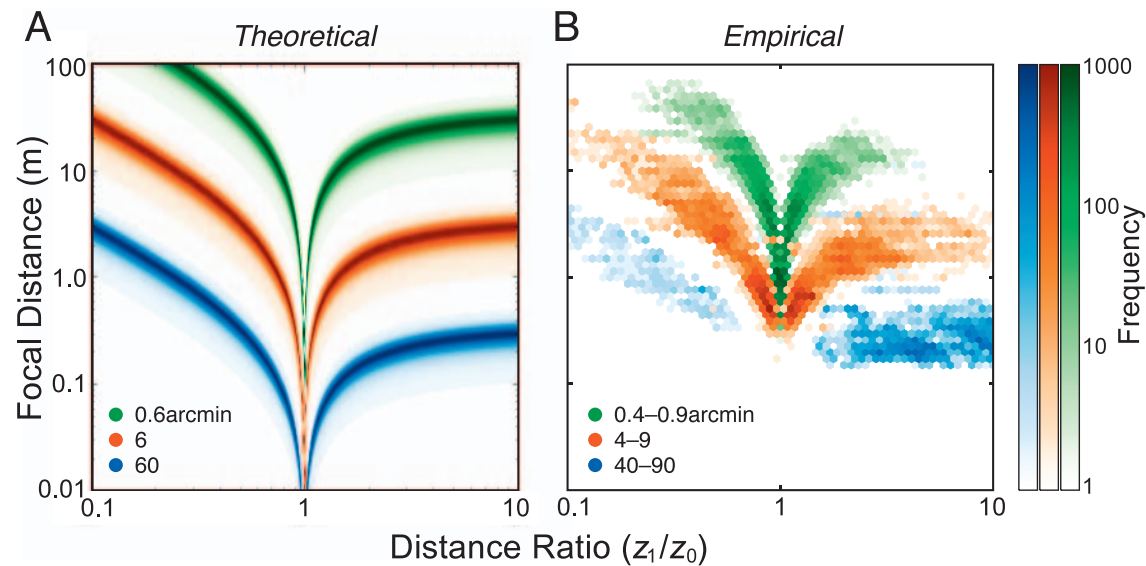
Figure 7. The probability of different amounts of defocus blur given the distance to which the eye is focused (z_0) and the ratio of the object distance divided by the focal distance (z_1/z_0). (A) The theoretical probability distributions for different amounts of blur (green for 0.6 arcmin, orange for 6 arcmin, and blue for 60 arcmin). We assumed a Gaussian distribution of pupil diameters with a mean of 5.8 mm and standard deviation of 1.1 mm; these numbers are consistent with the diameters measured when our subjects were performing the natural tasks. Higher probabilities are indicated by darker colors. As the distance ratio approaches 1, the object moves closer to the focal distance. There is a singularity at a distance ratio of 1 because the object by definition is in focus at that distance. (B) The empirical distributions of focal distance and distance ratio given different amounts of blur. Different ranges of blur are represented by green (0.4–0.9 arcmin), orange (4–9 arcmin), and blue (40–90 arcmin). The number of observations for each combination of focal and relative distance is represented by the darkness of the color, as indicated by the color bars on the right. The data were binned with a range of ~1/15 log unit in focal distance and ~1/30 log unit in distance ratio.

Perceptual experiment

The distribution of blur across the visual field leads to interesting predictions of how blur gradients should be interpreted if the distribution serves as a prior for inferring 3D shape from defocus blur. We next asked whether observers use the natural relationship between relative distance and associated blur across the visual field when interpreting ambiguous stimuli.

Experimental method

Five young adults with normal or corrected-to-normal visual acuity participated. If they normally wear an optical correction, they wore it during the experiment. The stimuli were viewed monocularly on an Iiyama CRT (HM204DT; resolution 1,600 × 1,200; frame rate 60 Hz) from a distance of 115 cm. At that distance, pixels subtended 0.75 arcmin. A circular aperture with a diameter of 14.8° was placed in the optical path from the eye to the stimulus so that the edges of the stimulus on the screen could not be seen. We applied Gaussian blur to white-noise textures on a row-by-row or column-by-column basis. Half of the stimuli were sharp in the center with blur increasing
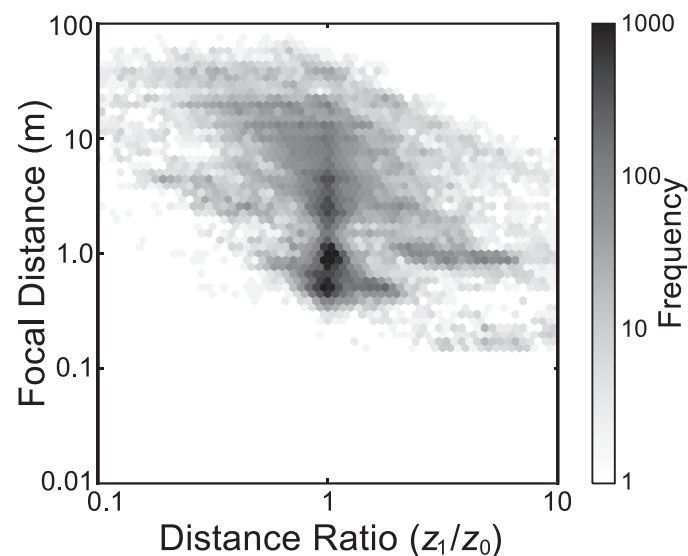

Figure 8. The probability of different focal distances and distance ratios. The number of observations for each combination of focal distance and distance ratio is represented by the darkness of the color, as indicated by the color bar on the right. The data were binned with a range of ~1/15 log unit in focal distance and ~1/30 log unit in distance ratio.

toward the edges (Figure 9A). The other half were blurred in the center with blur decreasing toward the edges (Figure 9B). For the sharp-center stimuli, the standard deviation of the blur kernel in minutes of arc as a function of the number of pixels from center was

$$\sigma = 2 \cdot m \cdot \left(\frac{1}{1 + \exp(-.01 \cdot |x|)} - 0.5 \right) \quad (6)$$

where x is 0 for screen center and 600 (or −600) for stimulus edge, and m is the maximum standard deviation in pixels. When m is greater than zero, this function is nearly a linear function of pixels from −200 to 200 (a third of the screen) and then gradually asymptotes to a value of m (at screen edge). We presented five values of m in pixels: 0, 2, 4, 6, and 8. From Equation 6 and the pixel size, the maximum values for the standard deviation of the blur kernel were 0, 1.5, 3, 4.5, and 6 arcmin. The direction of blur variation was always either vertical or horizontal (Figure 9A, B). When m was 0, there was, of course, no change in blur.

On each trial, a fixation cross first appeared on a uniform gray background for 1 s. The cross was then extinguished and the stimulus presented for 2 s. It had the same mean luminance as the background of the fixation stimulus. Observers were then given nine forced-choice response alternatives (Figures 10A and 11A). Four of the response options were slanted planes (Figure 10A) and four were wedges (Figure 11A). The last option was a frontoparallel plane. Before they began the experiment, observers were briefly trained to make sure they understood the task and response alternatives. No feedback was provided during training or the experiment itself. There were 10 presentations in which the blur gradient was zero and 160 in which it was nonzero. Half of the nonzero gradients were sharp in the middle, and half were sharp on the edges. Of the half that were sharp in the middle, half had vertical gradients (10 for each of the gradient magnitudes) and the other half had horizontal gradients (again 10 for each magnitude).

The stimuli are geometrically consistent with many shapes. Consider the sharp-center stimuli (Figure 9A). Figure 9D shows the 3D surfaces that are geometrically consistent with those stimuli. If the blur gradient is vertical, there are four such surfaces: (a) a surface slanted top-back (blue), (b) a surface slanted top-forward (red), (c) a convex wedge with its nearest point in the middle (green dashed), and (d) a concave wedge (purple dashed). If the gradient is horizontal, there are again four interpretations but rotated 90° in the image plane. Note that the specified shapes in the sharp-edge stimuli would have a plateau in the center. If people use the natural blur statistics, we expect them to perceive stimuli with a vertical blur gradient as slanted top-back because blur above fixation is more likely to be caused

by an increase in relative distance and blur below fixation is more likely to be caused by a decrease in relative distance (Figure 5C). We do not expect an analogous bias for a horizontal gradient because there is no difference in the natural blur distributions in the left and right fields (Figure 5C).

We first examined the pattern of responses when observers selected the planar shapes. Figure 10B shows the results for the sharp-center stimuli; the proportion of "ground," "ceiling," "left," and "right" responses are plotted as a function of the magnitude of the blur gradient when the stimuli were sharp in the center. (We do not plot the results for stimuli that were sharp on the edges here because such stimuli do not have a plausible planar interpretation and observers gave few planar responses with those stimuli.) The proportions are the number of responses of each type divided by the total number of responses (including wedge responses) averaged across subjects. The top and bottom panels show the proportions of responses when the gradient was respectively vertical and horizontal. When the gradient was vertical, the proportion of "ground" responses was much greater than the proportion of other planar responses and was generally greater for larger blur gradients. When the blur gradient was horizontal, the proportions of "left" and "right" responses were very similar. When the magnitude was zero, the great majority of responses was "flat" (not shown). The fact that the subjects reported "flat" when m was zero is important because it shows that the results were not caused by a tendency to see any stimulus as top-back. Thus, a vertical blur gradient was required for subjects to consistently report top-back slant. These results are consistent with the idea that the natural relationship between relative distance and defocus blur across the visual field is used in inferring depth from blurred images.

We next examined the pattern of responses when observers selected the wedge shapes. As we observed previously, the distributions in the natural-scene data (Figure 7B) manifest a tendency for large blurs in nonfoveal locations to be caused by scene points being farther not nearer than the fixation distance. To quantify this tendency, we calculated the proportion of relative distances greater than zero (i.e., farther than fixation) across the central visual field. Excluding the central 0.5° of the data (i.e., the fovea), 55% of the points in the visual field are farther than fixation and 45% are nearer. For short fixation distances, the asymmetry is greater. For long fixation distances, the asymmetry becomes smaller and eventually disappears. At the fixation distance used in our perceptual experiment (1.15 m), the asymmetry is clearly present (Figure 7B). We asked whether the bias toward far relative distance affects 3D percepts. We did this by examining observers' wedge responses. Figure 11A
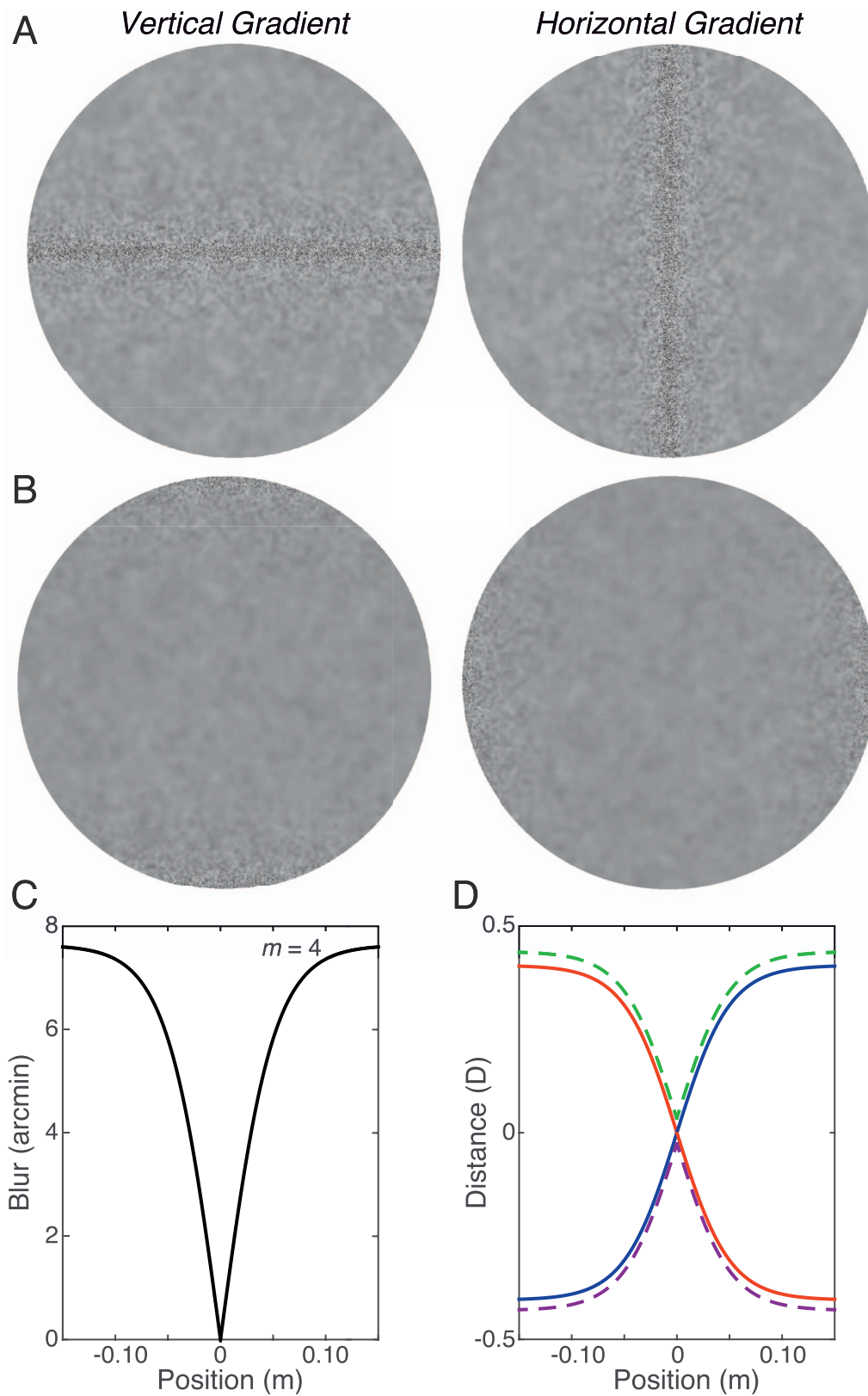
Figure 9. Experimental stimuli and corresponding shapes. (A) White-noise textures with blur gradients. The one on the left has a vertical blur gradient and the one on the right a horizontal blur gradient. In both cases, the blur kernel was zero in the middle of the image and increased monotonically with eccentricity. (B) White-noise textures with blur gradients but with the blur kernel at zero at the edges of the stimuli. (C) Blur as a function of image position in the experimental stimuli with sharp centers. The value of

←

m (Equation 3) is 4. We converted the Gaussian blur kernels with standard deviation σ into cylindrical blur kernels, so the ordinate is now the diameter of the corresponding cylinder. We made this conversion so that we could calculate the corresponding relative distances with Equation 2. (D) The 3D shapes that would create the blur distribution in panel C. Distance from fixation in diopters is plotted as a function of image position. There are four shapes that are consistent with the blur distribution: (1) a surface slanted top-back (blue), (2) a surface slanted top-forward (red), (3) a convex wedge (green dashed), and (4) a concave wedge (purple dashed). The dashed lines have been displaced slightly vertically to aid visibility.
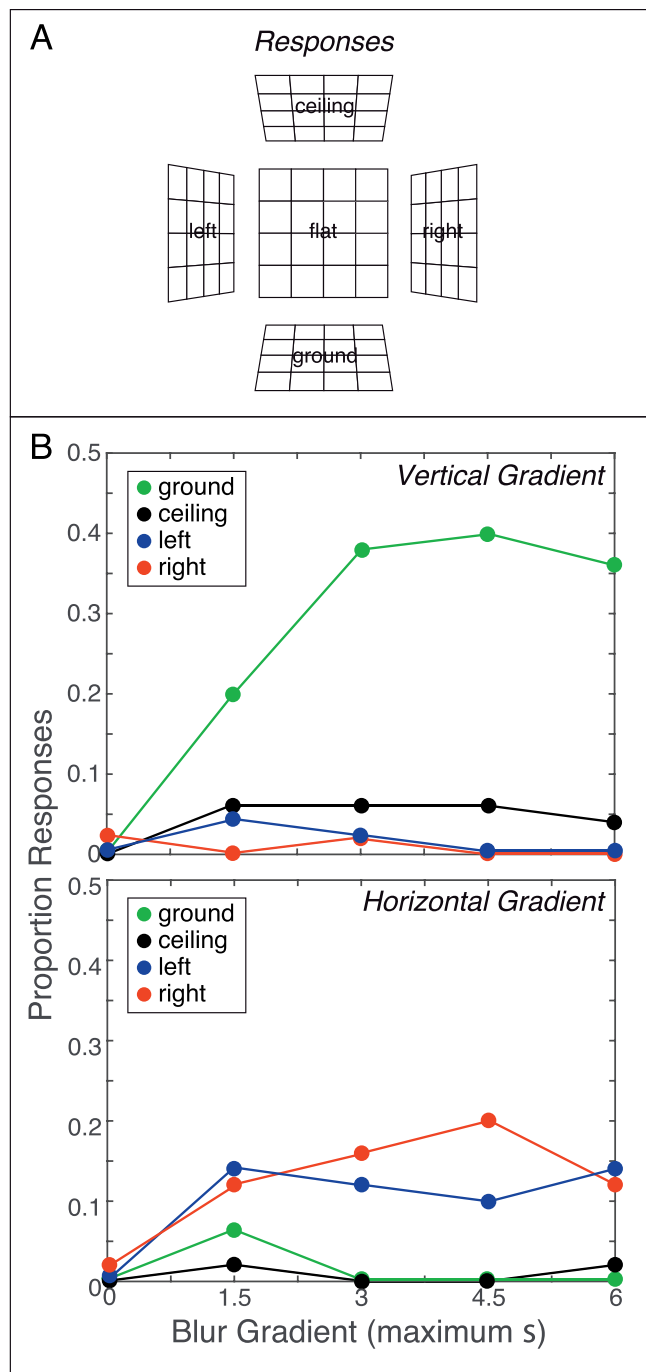

Figure 10. Blur gradients and perception of planar 3D shape. (A) Response icons for five of the nine possible responses. These →

shows the four wedge-like response options. Figure 11B shows the results. When the stimuli were sharp in the center and blurred on the edges (upper panels), observers were most likely to report that they perceived convex wedges, a tendency that increased for larger blur gradients. When the stimuli were sharp on the edges and blurred in the center (lower panels), observers generally reported concave wedges. In each case, observers mostly reported sharp as near and blurred as far, which one expects if using the prior information that most blurred points, particularly very blurred points, are farther than fixation. From our natural statistics, we expect that the observed tendency to see sharp as near will increase at short fixation distances and decrease at long fixation distances.

An analogous perceptual effect is shown in Figure 12. The two panels are photographs of Necker cubes with a pencil running through them. Most people find the upper panel easier to interpret than the lower one. The only difference is that the camera was focused near in the upper panel and far in the lower one. Said another way, the focal plane in the upper panel

←

are the five planar responses. There were also four wedge responses (Figure 11A). (B) Experimental results with vertical and horizontal blur gradients; the upper panel is for vertical gradients and the lower for horizontal gradients. In both panels, the proportion of responses of a particular category (green for ground plane, black for ceiling plane, red for right-side forward, and blue for left-side forward) is plotted as a function of the magnitude of the blur gradient. The abscissa values are the maximum values of σ in Equation 6. The data have been averaged across the five observers. We calculated the departure of the data from equal distribution of responses among the geometrically plausible alternatives for each value of *m*. Specifically, we computed goodness-of-fit χ^2 for the observed responses relative to responses distributed equally among five alternatives. When the blur gradient was vertical, we considered observed responses relative to the alternatives of "flat," "ground," "ceiling," "convex" (horizontal vertex), and "concave" (horizontal vertex). When the gradient was horizontal, we considered responses relative to "flat," "left," "right," "convex" (vertical vertex), and "concave" (vertical vertex). All χ^2 values were statistically significant ($p < 0.01$), except for $m = 1$, horizontal gradient, sharp center. This means that responses were not randomly distributed among the planar alternatives.
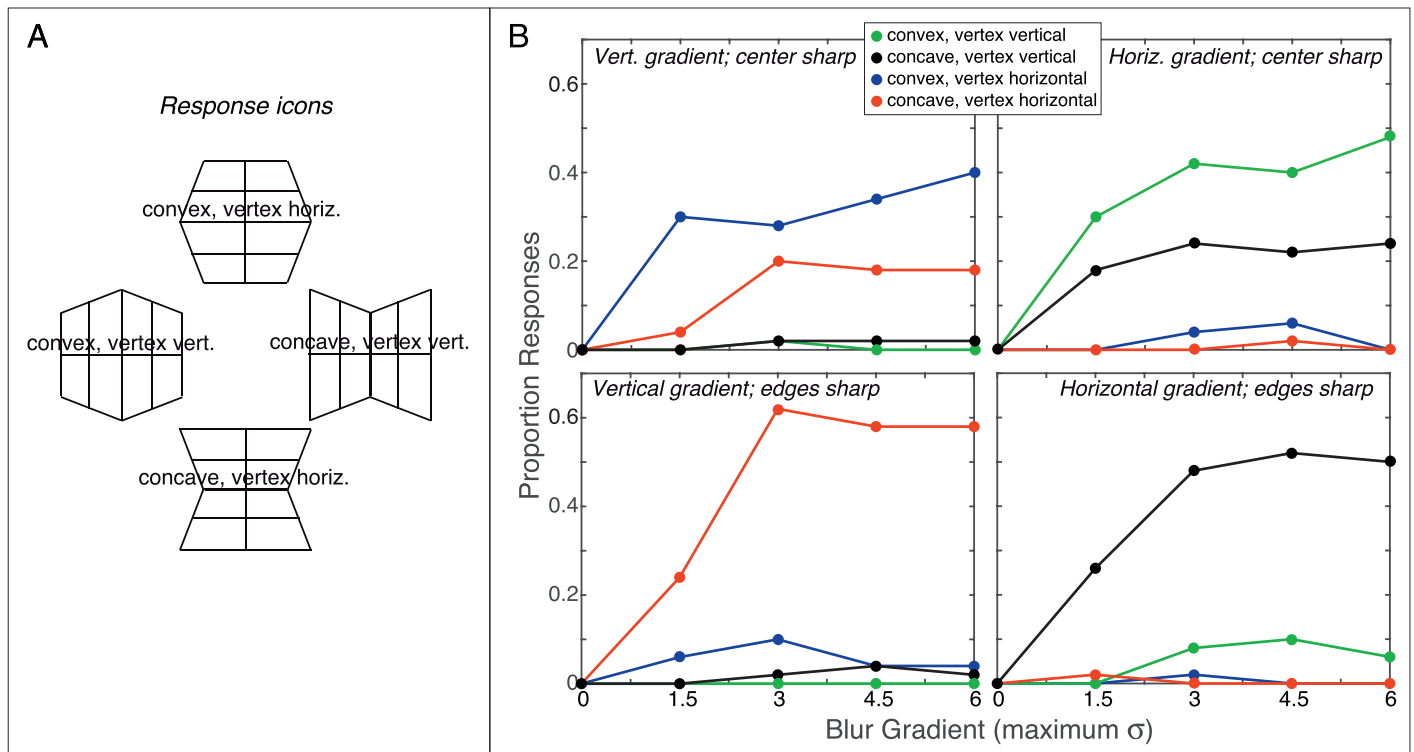
Figure 11. Blur gradients and perception of wedge 3D shape. (A) Response icons for four wedge-like responses (four of nine possible responses). (B) Experimental results. The panels plot the proportion of responses of a particular category as a function of the magnitude of the blur gradient. The abscissa values are the maximum values of σ in Equation 3. The upper panels show the proportions of responses when the center of the stimulus was sharp and the edges blurred. The lower panels show the response proportions when the edges were sharp and the center blurred. The left panels show the responses when the blur gradient was vertical. The right panels show them when the gradient was horizontal. The data have been averaged across the five observers. We calculated the departure of the data from equal distribution of responses among the geometrically plausible alternatives for each value of *m*. Specifically, we computed χ^2 for the observed responses relative to responses distributed equally among three alternatives. When the blur gradient was vertical, we considered observed responses relative to the alternatives of "flat," "convex" (horizontal vertex), and "concave" (horizontal vertex). When the gradient was horizontal, we considered responses relative to "flat," "convex" (vertical vertex), and "concave" (vertical vertex). All χ^2 values were statistically significant ($p < 0.01$), which means that responses were not randomly distributed among the alternatives.

coincides with the near vertex of the cube, so the far parts of the cube are blurred. The focal plane in the lower panel coincides with the far vertex of the cube so the near parts are blurred. Viewers' tendency to see sharp as near is consistent with the upper image and inconsistent with the lower one, and that leads to the difference in how readily the two can be interpreted. Pentland (1985) reported a similar effect. On a display screen, he presented a video of a rotating Necker cube. When the focal distance of the camera was near, the nearest vertex of the cube was sharp and the farthest vertex blurred. In that case, people perceived a rigid cube rotating in the correct direction. But when the camera's focal distance was far (as in the lower panel of Figure 11), people often perceived a nonrigid cube rotating in the opposite direction.

Grosset et al. (2013) observed a related effect. They presented various computer-generated objects (e.g., an image of an aneurysm, a flame, etc.) and asked subjects to indicate which of two points in the stimulus was nearer. When the focal plane was near, subject performance was nearly perfect. When the focal plane was far, performance was generally worse than chance. Subjects' tendency to see sharp as near aided performance when the focal plane was near because the bias was consistent with the correct depth order. The same tendency hurt performance when the plane was far because the bias was then inconsistent with correct depth order. The results from Grosset and colleagues are therefore consistent with our observation that large blurs are more likely to be due to an object being farther rather than nearer than fixation.

The data from our experiment and others show that human observers do in fact use the natural relationship between relative distance and defocus blur to infer the 3D shape of stimuli with ambiguous blur gradients.
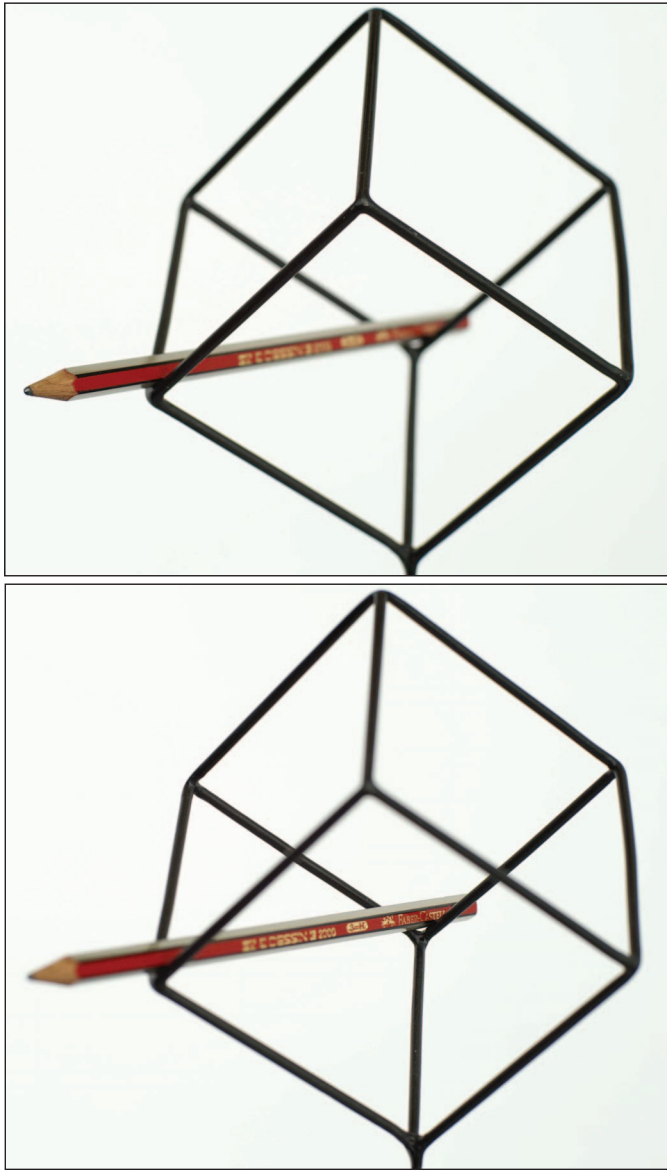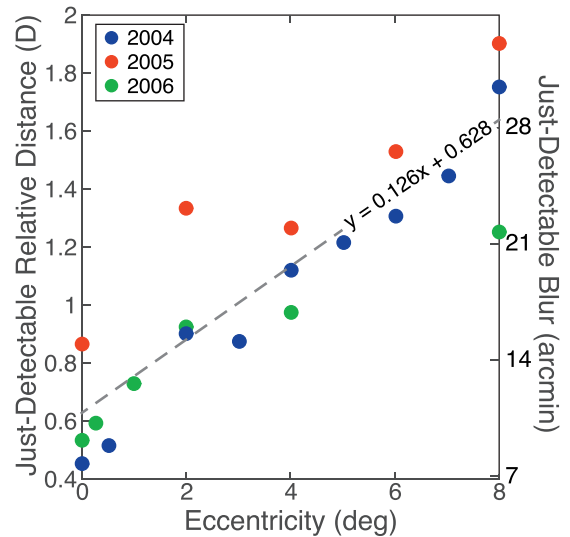
Figure 13. Blur-detection threshold as a function of retinal eccentricity. Detectable changes in defocus are plotted as a function of retinal eccentricity. The left ordinate shows the changes in diopters and the right ordinate the changes in minutes of arc using a 5-mm pupil. The blue points are from Wang and Ciuffreda (2004); they reported the full depth of field, so we divided their thresholds by two. The red points are from Wang and Ciuffreda (2005) and the green points from Wang et al. (2006). The dashed line is the best linear fit to the data.

Frequency of perceptible blur

We now return to the question: Why does the visual world appear predominantly sharp? To investigate, we compared the distribution of blur magnitudes across the central visual field to the appropriate thresholds for distinguishing blurred from sharp.

For blur-detection thresholds at different retinal eccentricities, we used data from three studies by Wang, Ciuffreda, and colleagues (Wang & Ciuffreda, 2004, 2005; Wang, Ciuffreda, & Irish, 2006). To our knowledge, these are the only prior studies that measured blur thresholds at different eccentricities. In all three studies, observers were cyclopleged and thus unable to change accommodative state. Stimuli were viewed monocularly. The focal distance of the fixation target was carefully adjusted to maximize image sharpness. That distance remained fixed. A 5-mm artificial pupil was placed directly in front of the subject's eye. The peripheral stimulus was a high-contrast circular edge centered on fixation. The radius of the edge varied, and those radii defined the retinal eccentricity of the stimulus. To measure thresholds, stimulus distance was slowly increased or decreased until the subject reported that the circular edge appeared blurred. The stimuli were viewed in a Badal lens system so stimulus size at the retina remained constant as focal distance was manipulated. The results are shown in Figure 13. The just-detectable

Figure 12. Blur and the interpretation of 3D shape. The panels are photographs of a Necker cube with a pencil running through it. The camera was focused on the nearest vertex of the cube in the upper panel and on the farthest vertex in the lower panel. Photograph provided by Jan Souman.

Discussion

We observed regularities in the distributions of defocus blur in different parts of the visual field. We showed that human observers use these regularities in interpreting ambiguous blur gradients. We also observed that large blurs are much more likely to be caused by scene points that are farther than fixation than by points that are nearer. Again, human observers seem to have also incorporated this statistical regularity, as evidenced by a tendency to perceive sharp as near and blurred as far.
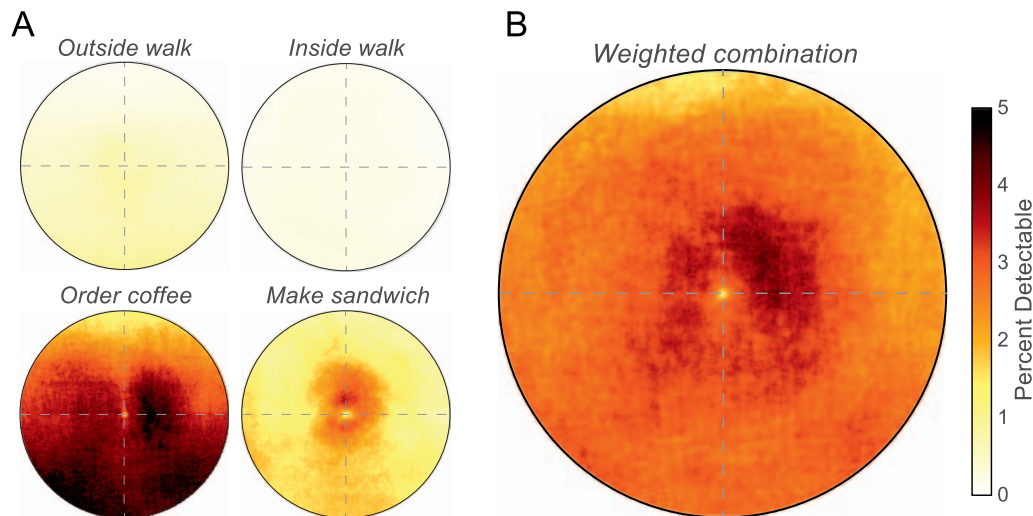
Figure 14. The percentage of detectable blurs across the visual field. (A) Percentage of detectable blur magnitudes in the central visual field for the four tasks. The diameter of the circles is 20° and the fovea is in the center. Darker colors represent higher percentages (see color bar on far right). (B) Percentage of detectable blur magnitudes in the central visual field for the weighted combination across tasks.

change in focal distance increased roughly linearly with retinal eccentricity. We found the best-fitting line to the data using linear regression and then converted the units from diopters to minutes of arc using a pupil diameter of 5 mm and our Equation 2. We could then determine how often blurs in our data set exceed detection threshold.

It is important to note that Wang and Ciuffreda's results are the kind of data required for our purpose. They manipulated the actual focal distance of the stimulus at different retinal eccentricities, so other blurring elements (e.g., diffraction, chromatic aberration) were introduced by the viewer's eye and not rendered into the stimulus. Thus, their data tell us what changes in object distance relative to fixation are detectable. We measured the differences in object and fixation distances, so our measurements are compatible.

The percentages of detectable blurs in the four tasks are shown in Figure 14A. The percentages varied across tasks. For example, the *order coffee* task produced more detectable values than the other tasks. The percentage is higher in that task because ordering coffee is a social task in which participants tended to look at the face of the person to whom they were conversing and then to other objects such as the coffee mug. The face took up a central portion of the visual field, and other objects in the scene tended to be much farther than the face, resulting in greater number of large blur values in the periphery. Fixation of small objects, such as the mug, took up a smaller portion of the visual field, resulting in changes near fixation. By comparison, detectable blurs were quite infrequent in the *outdoor walk* and *indoor walk* tasks because those

scenes were generally more distant than in the *order coffee* and *make sandwich* tasks. Scene distance matters because the magnitude of blur for a given depth interval is roughly proportional to the inverse square of distance, which is evident if one cross-multiplies in Equation 1.

Figure 14B shows the percentages of detectable blurs for the weighted combination of tasks. Detectable blur is surprisingly uncommon. In the weighted combination, blur values exceed threshold less than 4% of the time. Moreover, the percentage does not vary systematically across the central visual field.

The defocus thresholds that Wang and Ciuffreda reported for the fovea were 0.45 to 0.85D, which is somewhat higher than previously reported values (Campbell & Westheimer, 1958; Sebastian, Burge, & Geisler, 2015; Walsh & Charman, 1988). This difference is probably due to subjects employing a stricter criterion in the adjustment method used by Wang and Ciuffreda than in the forced-choice methods in the other studies. To determine how lower thresholds would affect the analysis, we reduced all threshold values by half and recomputed the percentages. The percentages increased roughly twofold throughout the visual field but more in the upper than in the lower field. We plot in Figure 14B the percentages of detectable blurs based on the adjustment data rather than the forced-choice data because the former is more likely to be consistent with everyday perceptual experience.

We conclude that in natural viewing, the defocus blur experienced across the central 20° of the visual field is mostly not noticeably different from no defocus blur. In other words, the visual world may appear
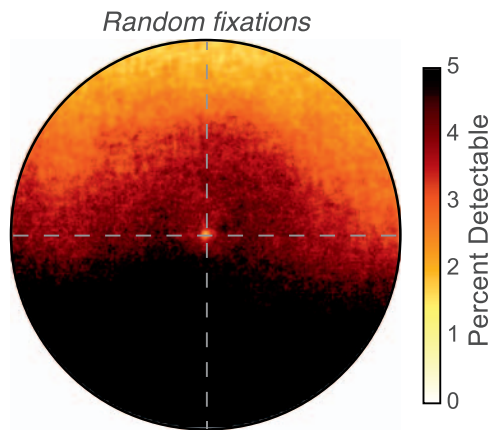
Figure 15. Percentage of detectable blurs with random fixations. Percentage of detectable blur magnitudes for the weighted combination of the four tasks combined across subjects when gaze directions are random. Darker colors represent higher percentages. The data at the bottom of the panel are slightly clipped.

subjectively sharp because blur magnitudes that are large enough to be perceived as blurred are fairly rare. If blur were never detectable in natural vision, it would not be a useful depth cue. But our data show that detectable blurs do occur particularly when fixation distance is short (i.e., in the *order coffee* and *make sandwich* tasks).

It is interesting that the percentage of detectable blurs is fairly constant across the visual field. We wondered if this outcome resulted from how subjects selected scene locations to fixate. We investigated this by comparing blur distributions obtained with real fixations to distributions obtained with random fixations. In the random fixations, gaze directions were chosen from the set of directions observed with real fixations, but the correspondence between gaze directions and video frames was randomly shuffled. For each selected gaze direction, we assigned the fixation distance that corresponded to the scene distance in that direction. (In other words, we did not create fixations in which the viewer would be converged and accommodated in front of or behind surfaces in the scene.) With this procedure, the distributions of gaze directions are the same for real and random fixations. The sets of 3D scenes are also of course the same. The percentages of detectable blurs for random fixations are shown in Figure 15. Detectable blur is much more common than with real fixations, particularly in the lower visual field. This result shows that fixation strategies tend to minimize blur, especially in the lower field. To further examine this observation, we computed the medians and standard deviations of relative distances just above and below the fovea. They are plotted as a function of semicircle radius in Figure 16A and B. The medians in the upper and lower fields are more negative with random fixations and the standard deviations are generally larger. Because real fixations were generally positive in the upper field and negative in the lower, random fixations lead to fewer observations that are
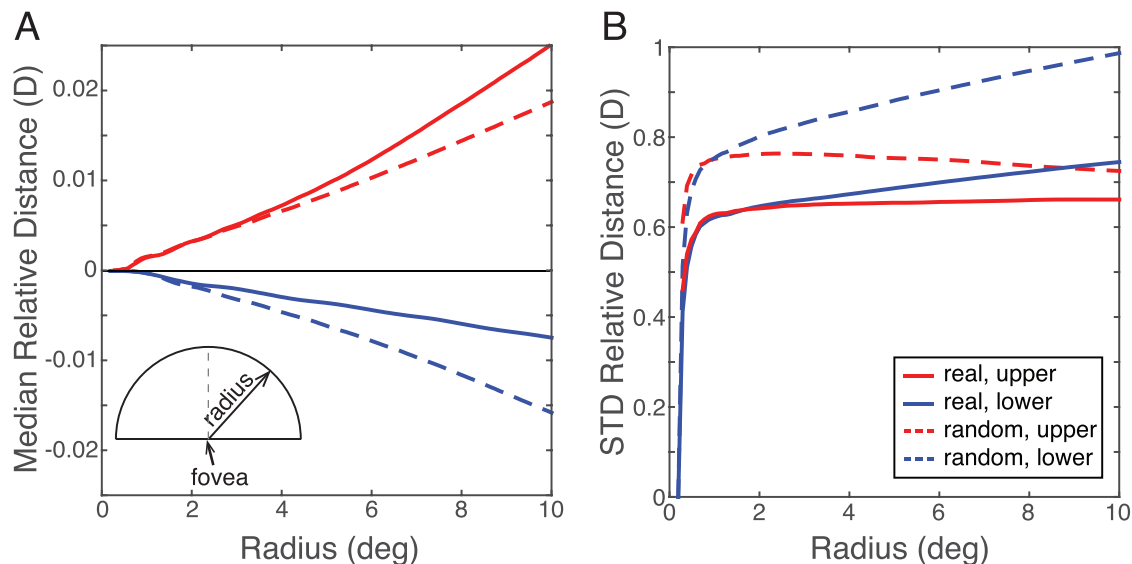

Figure 16. Medians and standard deviations of relative distance with real and random fixations. (A) Median relative distances in the upper and lower fields plotted as a function of radial distance from the fovea. The medians were computed from all relative distances within the semicircular sampling window. The red and blue curves represent the data from the upper and lower fields, respectively. Solid and dashed curves represent the data from real and random fixations, respectively. (B) Standard deviations of relative distances in the upper and lower fields as a function of distance from fovea. The red and blue curves represent data from the upper and lower fields, respectively. Solid and dashed curves represent data from real and random fixations, respectively.
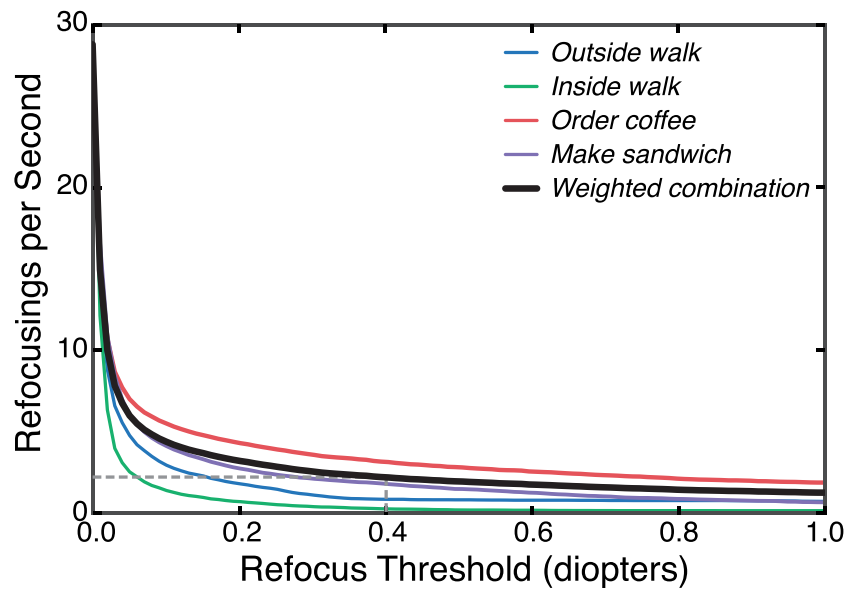
Figure 17. Refocusings required for different thresholds. The estimated number of refocusings per second is plotted as a function of the change in diopters that require a refocus response. The thin colored curves represent those values for each of the individual tasks averaged across subjects. The thick black curve represents those values for the weighted combination of data across the tasks, again averaged across subjects. The dashed lines indicate the number of refocusings needed if the refocus threshold is 0.4 diopters.

substantially different from zero in the upper field and more that are substantially different from zero in the lower field. Thus, the manner in which people select points to fixate in the observed 3D scene causes a reduction of negative relative distance and a reduction in the variation of relative distance. We speculate that people tend to avoid fixation points that would position a depth discontinuity (e.g., an occlusion) in the lower 10° of the visual field. As a consequence, the percentage of detectable blurs is lower in the visual field than a random selection of fixation direction would produce.

Frequency of refocusing

As a viewer fixates one object then another in the environment, the distance to the point of fixation changes. Many of those distance changes will be large enough to require a change in accommodation. Sebastian et al. (2015) claimed that such refocusing occurs ~150,000 times a day. Assuming 16 waking hours, this corresponds to about 2.6 refocusings a second. They did not describe the method by which they reached this conclusion. We used our data set to estimate the number of accommodative changes that should occur during natural behavior.

Figure 17 shows how frequently changes in fixation distance occur that are equal to or greater than a given change in diopters. The thin curves represent those values for the four individual tasks and the thick curve the values for the weighted combination of data across

tasks. The number of estimated refocusings per second decreases quickly as the threshold increases but asymptotes at roughly two per second. (The increase with small thresholds is due in large part to measurement noise.) The frequency of refocusing responses will depend of course on the eye's depth of field (i.e., how much defocus is needed to be detectable and to trigger an accommodative response). Previous work has shown that the eye's depth of field for foveal viewing and typical lighting conditions is 0.25–0.5D (Campbell, 1957; Charman & Whitefoot, 1977; Green & Campbell, 1965; Sebastian et al., 2015). If we adopt a threshold of 0.4D (dashed line), ~ 2.1 refocusings would be required per second for the weighted combination data. Obviously adopting slightly lower or higher thresholds would yield respectively more and fewer refocusings per second, but the differences would be small because of the asymptotic behavior of the estimated curves. People engaged in natural tasks make about three fixations per second (Ibbotson & Krekelberg, 2011), so our analysis suggests that about two-thirds of those fixations create a large enough change in fixation distance to require refocusing. Our estimate of 2.1 refocusings per second projected into 16 waking hours is 120,960 per day, which is quite close to the value stated by Sebastian and colleagues.

Eye shape

We now consider the hypothesis that the eye conforms to a shape that makes the most frequently
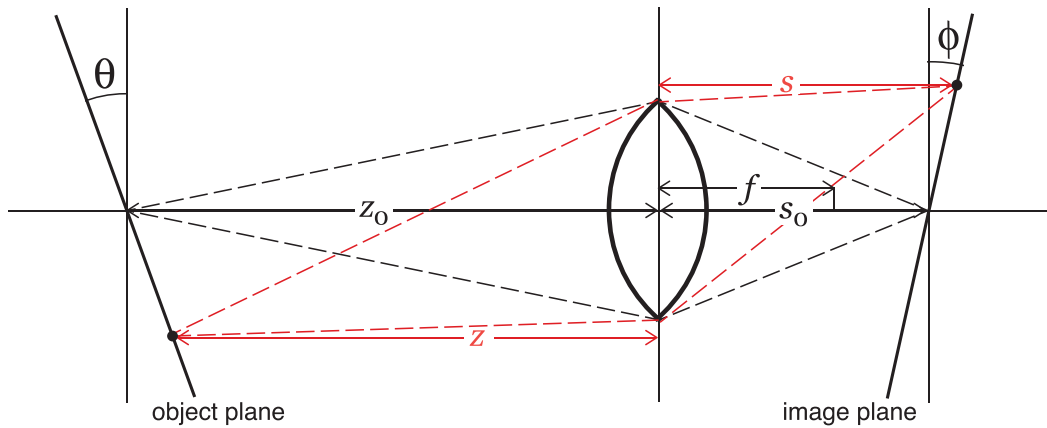
Figure 18. Image formation when the object is a slanted plane. z_0 is the focal distance of the eye measured along the optical axis. f is the focal length and s_0 is the distance from the lens to the plane of best focus along the optical axis. The object plane is slanted by θ relative to frontoparallel. A point on the object plane at distance z forms an image at distance s from the lens. The image is formed in a plane slanted by ϕ relative to frontoparallel.

occurring relative distances in good focus at the retina. If the surface of an object being captured by a conventional imaging device is slanted relative to the optical axis, a slanted imaging surface is required for forming a sharp image along the full extent of the object (Kingslake, 1992). The geometry for a thin-lens system is illustrated in Figure 18.

An object plane at distance z_0 on the optical axis is slanted by θ relative to frontoparallel. Leftward distances from the origin (the center of the lens) are negative; rightward distances are positive. The eye is focused at distance z_0. The object plane is

$$z = -y\tan(\theta) + z_0 \quad (7)$$

where z is distance from the origin and y is distance in the orthogonal direction from the optical axis. The image of the object plane is formed in the plane

$$s = -y\tan(\phi) + s_0 \quad (8)$$

where

$$s_0 = \frac{z_0 f}{z_0 + f} \quad (9)$$

Equation 9 gives the distance along the optical axis from the lens to the image plane when the eye is focused at distance z_0. The image of a point on the object plane at distance z is formed on a plane at distance s:

$$s = \frac{zf}{z+f}. \quad (10)$$

The slant of that image plane is

$$\phi = \tan^{-1}\left[\tan(\theta)\left(\frac{z_0 f}{z_0 + f} - 1\right)\right] \quad (11)$$

We can express the image distances in diopters relative to best focus. In this case, the dioptric distance

of a point on the image plane is

$$\Delta S = \frac{1}{s} - \frac{1}{s_0} = \frac{z+f}{z} - \frac{z_0 + f}{z_0} \quad (12)$$

$$\Delta S = \frac{1}{z} - \frac{1}{z_0} = \Delta Z \quad (13)$$

which means that the variation in image distance in diopters across the image plane is the same as the variation in object distance in diopters across the object plane, provided that the eye is focused at z_0. We can take advantage of this property to determine the image surface that would provide the sharpest image for relative distances at different points in the visual field. We first consider some animal eyes and then the human eye.

Several animals that live on the ground exhibit *lower-field myopia*: The distance corresponding to best focus is nearer for the lower visual field than for the upper field. This has been observed in horse (Sivak & Allen, 1975; Walls, 1942), chicken, quail, crane (Hodos & Erichsen, 1990), pigeon (Fitzke, Hayes, Hodos, Holden, & Low, 1985; García-Sánchez, 2012), turtle (Henze, Schaeffel, & Ott, 2004), salamander, frog (Schaeffel, Hagel, Eikermann, & Collett, 1994), and guinea pig (Zeng, Bowrey, Fang, Qi, & McFadden, 2013). Of course, many animals do not live on the ground. For example, raptors—barn owl, Swainson's hawk, Cooper's hawk, and American kestrel—spend most of their time airborne or perched well above the ground, and they do not exhibit an asymmetry in refraction as a function of elevation (Murphy, Howland, & Howland, 1995).

We investigated whether the change in the refraction of ground-dwelling animals as a function of elevation is well matched for making the ground plane conjugate
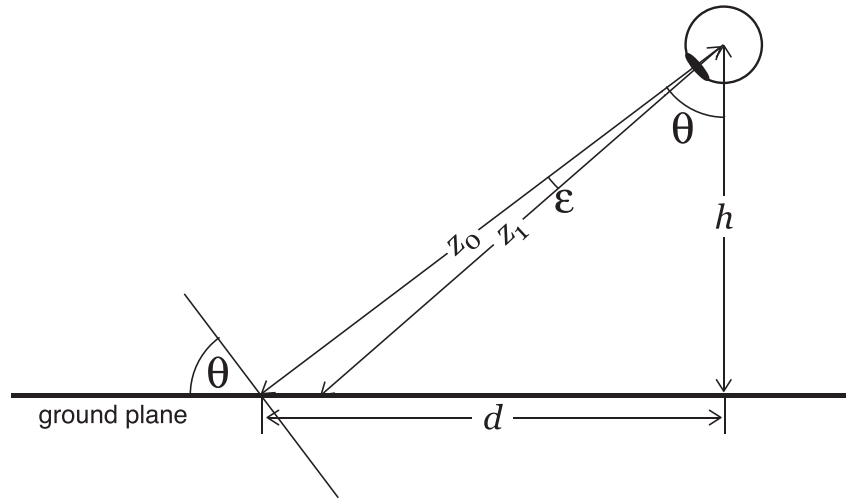
Figure 19. Geometry of object and ground planes for an upright viewer. An eye at height h from the ground views a point on the ground at distance d. The line of sight is at angle θ relative to earth vertical. The ground plane is rotated by θ relative to a frontoparallel plane. The distance of the fixated point along the line of sight is z_0. For a point slightly lower in the visual field, but also on the ground, the distance is z_1. The angular difference between the vectors z_0 and z_1 is ε.

with the retina. The geometry is illustrated in Figure 19. The eye is at height h relative to the ground. The line of sight is rotated by θ relative to earth vertical and intersects the ground at distance d. The distance of that point along the line of sight is z_0. A visual line slightly lower in the visual field intersects the ground at a closer distance, and its distance along a visual line is z_1. If the lower visual line is rotated by ε relative to the line of sight, we have

$$z_0 = \frac{h}{\cos(\theta)}$$
$$z_1 = \frac{h}{\cos(\theta + \varepsilon)} \quad (14)$$

Expressing z_0 and z_1 in diopters and taking the difference

$$\Delta D = \frac{\cos(\theta)}{h} - \frac{\cos(\theta + \varepsilon)}{h} \quad (15)$$

For large values of θ (i.e., gaze away from feet), and $\varepsilon = 1°$:

$$\Delta D \approx \frac{\pi}{180h} \quad (16)$$

This means that the change in diopters for a change in elevation in the visual field is nearly constant for all but near distances (Banks, Sprague, Schmoll, Parnell, & Love, 2015); the change to good approximation depends only on eye height. We refer to the change in diopters across 1° of elevation as the *diopter gradient*.

The diopter gradient as a function of eye height (Equation 13) is represented by the line in Figure 20.

We wanted to see if the diopter gradient of the ground is consistent with asymmetries in the eye shape of various animals. From the measurements of refraction in ground-dwelling animals, we found the best-fitting line to the spherical equivalent of the refraction (the average of the refractions along the two cylinder axes) as a function of elevation for several animals (see caption for details). The ordinate of Figure 20 represents the diopter gradient. The abscissa is the reported height of the animals' eyes above the ground. There is clear correspondence between the magnitude of lower-field myopia and the predictions if the change in refraction with elevation is an adaptation for placing the ground plane in best focus across the retina. Importantly, the correspondence is independent of how far away the animal fixates on the ground, provided that fixation is not too close. Our analysis supports a hypothesis for ground-foraging birds by Hodos and Erichsen (1990):

> The data reported here suggest that lower-field myopia, rather than being an error in the focal length of the eye, represents an adaptive matching of focal length to the eye-to-ground distance. Indeed, for ground-foraging birds, the myopia of the lower visual field would be a refractive error only during flight. (p. 657)

We suggest that many other ground-dwelling animals have a similar adaptation that places the ground into best focus. Given the impressive ability of visual
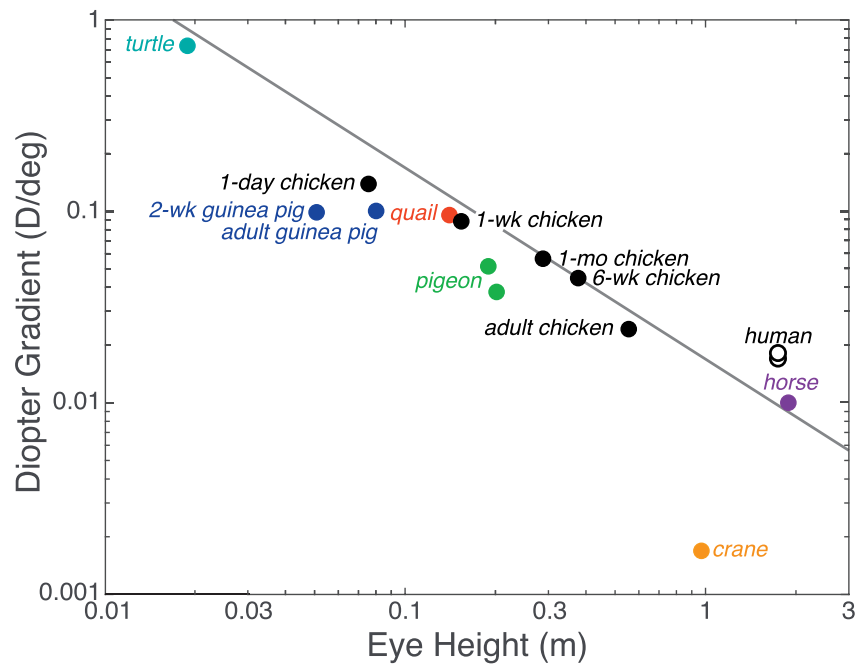
Figure 20. Relationship between eye height and lower-field myopia in ground-dwelling animals. The change in best focus as a function of elevation is plotted as a function of the typical height of the eye above the ground. The change in best focus is diopters per degree of elevation. To compute these values, we found the best-fitting line to data on refraction as a function of elevation. Turtle: figure 3 in Henze et al. (2004); spherical equivalent measured in three animals at 19 eccentricities across 55° of elevation. One-week and 1-month chicken: figure 6 in Schaeffel et al. (1994); refractive state in three 1-week and six 1-month chickens; measured at two eccentricities across 80° of elevation. One-day, 6-week, and adult chicken: table 2 in Hodos and Erichsen (1990); refractive state in three 1-day, three 6-week, and three adult chickens; two eccentricities across 60°. Two-week and adult guinea pig: table 2 in Zeng et al. (2013); spherical equivalent measured in 18 eyes of 2-week-olds and 31 eyes of adults; two eccentricities across 60°. Quail: table 2 in Hodos and Erichsen (1990); refractive state in five adult quails; two eccentricities across 60°. Pigeon: figure 4 in Fitzke et al. (1985); refractive state measured in 72 eyes; numerous eccentricities across 130°. Figure 5.16 in García-Sánchez (2012); spherical equivalent measured in eight eyes at 13 eccentricities across 60° (averaged across three azimuths). Crane: table 2 in Hodos and Erichsen (1990); refractive state in four adult cranes; measured at two eccentricities across 60° of elevation. Horse: table 2 in Sivak and Allen (1975); refractive state measured in four eyes at three eccentricities across 60°. Human: figure 3 from Seidemann et al. (2002); spherical equivalent measured in 62 eyes at nine eccentricities across 55°. Table 2B in Ehsaei et al. (2011); refractive state in 36 emmetropic eyes; measured at seven eccentricities across 60°. Eye height was either provided in the reference on refraction or was estimated from sources on the Internet. The line represents the prediction of the image plane of best focus if it were conjugate with the ground (Equation 16).

feedback to modulate the size and shape of the developing eye (Schaeffel, Glasser, & Howland, 1988; Wallman, Turkel, & Trachtman, 1978), it seems likely that the elevation-dependent change in best focus emerges from visual experience. In other words, the shape of the eye may adapt so as to place the most frequent relative distances into good focus on the retina.

The refraction of the human eye exhibits a qualitatively similar effect. Specifically, the refraction for the lower visual field is more myopic than that for the upper field, whereas the refractions of the nasal and temporal fields are similar (Ehsaei, Mallen, Chisholm, & Pacey, 2011; Seidemann, Schaeffel, Guirao, Lopez-Gil, & Artal, 2002). The diopter gradient across elevation is ~0.018D/°, a small effect. Gradient estimates from the human studies are shown in Figure

20. As can be seen, the observed diopter gradient is roughly consistent with what would be required to place the ground in best focus across the retina when the person is standing.

A related argument has been made for binocular vision. The vertical horopter in humans is slanted top-back. This means that to stimulate corresponding points in the two eyes, objects in the upper and lower visual fields must be respectively farther and nearer than fixation (Cooper, Burge, & Banks, 2011; Helmholtz, 1867; Nakayama, 1977; Schreiber, Hillis, Filippini, Schor, & Banks, 2008). Furthermore, the vertical horopter is more slanted in short animals than in humans (Cooper & Pettigrew, 1979). The principles that underlie lower-field myopia and the slant of the vertical horopter may therefore be similar: They may manifest adaptations to a correlation between elevation

on the retina and the statistical distributions of distances encountered in the world (Sprague et al., 2015).

Conclusions

We measured the statistics of naturally occurring defocus blur and observed some clear regularities. Defocus blur is unsigned, but blur in the upper visual field is most likely to be caused by objects farther than fixation (i.e., positive relative distance), and blur in the lower field is most likely to be caused by objects nearer than fixation (negative relative distance). The fact that human viewers see vertical blur gradients as surfaces slanted top-back shows that they have incorporated and use these statistical regularities when inferring depth from blur. We also found that large blur values are most likely to be caused by objects farther than fixation. The fact that viewers tend to see sharp objects as near and blurred objects as far reflects this statistical regularity.

Resolving the sign ambiguity associated with defocus blur can in principle be accomplished by use of certain optical properties of the human eye. For example, longitudinal chromatic aberration can be used to determine whether the eye needs to accommodate farther or nearer to sharpen a blurred image (Kruger, Mathews, Aggarwala, & Sanchez, 1993). It also affects judgments of the depth order of abutting surfaces (Nguyen, Howard, & Allison, 2005; Zannoli et al., 2016). Some monochromatic higher-order aberrations exhibited by the eye can also in principle provide information about the sign of defocus blur (Wilson, Decker, & Roorda, 2002). These cues in combination with our observations of statistical regularities between blur, distance, and position in the visual field should allow a better understanding of how the visual system can use these signals to infer scene geometry.

Computer vision algorithms that estimate depth from defocus blur deal with the sign ambiguity by focusing the camera such that no scene points are nearer than the focal distance (McCloskey & Langer, 2009; Pentland, 1987). By doing this, the magnitude of observed blur becomes proportional to distance in diopters from the focal plane. Our observations imply that these algorithms do not necessarily have to focus the camera near to avoid the sign ambiguity. Rather, they could rely on the statistical regularities we observed to make probabilistic inferences on whether a given blur observation indicates that the object is nearer or farther than where the camera is focused. In addition, in computer vision applications such as robotics, the designer could use the strategy employed by many ground-dwelling animals: adjust the orientation of the camera sensor planes to be conjugate with the most likely scene geometry and thereby maximize the sharpness of the acquired images.

Finally, we found that the great majority of blur observations in the central visual field are not distinguishable from sharp. Does this imply that blur is not a useful cue for inferring depth or for determining the scale of a scene? It does not. Instead, perceptible blur, although rare, is significant because when one observes it, it means there is something unusual about the current viewing parameters and the scene one is viewing (Grosset et al., 2013; Held et al., 2010; Pentland, 1985; Vishwanath & Blaser, 2010; Zannoli et al., 2016).

References

Banks, M. S., Sprague, W.W., Schmoll, J., Parnell, J.A., & Love, G.D. (2015). Why do animal eyes have pupils of different shapes? *Science Advances*, *1*(7), e1500391.

Brisson, J., Mainville, M., Mailloux, D., Beaulieu, C., Serres, J., & Sirois, S. (2013). Pupil diameter measurement errors as a function of gaze direction in corneal reflection eyetrackers. *Behavior Research Methods*, *45*, 1322–1331.

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

Campbell, F. W. (1957). The depth of field of the human eye. *Optica Acta, 4*, 157–164.

Campbell, F. W., & Westheimer, G. (1958). Sensitivity

of the eye to differences in focus. *Journal of Physiology*, *143*, 18P.

Charman, W. N., & Whitefoot, H. (1977). Pupil diameter and the depth-of-field of the human eye as measured by laser speckle. *Optica Acta*, *24*, 1211–1216.

Choe, K. W., Blake, R., & Lee, S.H. (2016). Pupil size dynamics during fixation impact the accuracy and precision of video-based gaze estimation. *Vision Research*, *118*, 48–59.

Cooper, E. A., Burge, J., & Banks, M.S. (2011). The vertical horopter is not adaptable, but it may be adaptive. *Journal of Vision*, *11*(3):20, 1–19, doi:10.1167/11.3.20. [PubMed] [Article]

Cooper, M. L., & Pettigrew, J.D. (1979). A neuro-physiological determination of the vertical horopter in the cat and owl. *Journal of Comparative Neurology*, *184*, 1–26.

Ehsaei, A., Mallen, E. A., Chisholm, C. M., & Pacey, I. E. (2011). Cross-sectional sample of peripheral refraction in four meridians in myopes and emmetropes. *Investigative Ophthalmology & Visual Science*, *52*, 7574–7585. [PubMed] [Article]

Fitzke, F. W., Hayes, B. P., Hodos, W., Holden, A. L., & Low, J. C. (1985). Refractive sectors in the visual field of the pigeon eye. *Journal of Physiology*, *369*, 33–44.

Gagl, B., Hawelka, S., & Hutzler, F. (2011). Systematic influence of gaze position on pupil size measurement: analysis and correction. *Behavior Research Methods*, *43*, 1171–1181.

Galvin, S. J., O'Shea, R. P., Squire, A. M., & Govan, D. G. (1997). Sharpness overconstancy in peripheral vision. *Vision Research*, *37*, 2035–2039.

García-Sánchez, Y. (2012). *Analysis of retinal image quality for peripheral vision in humans and pigeons (Columba livia)* (Unpublished doctoral dissertation). Cardiff University, Cardiff, UK.

Green, D. G., & Campbell, F. W. (1965). Effect of focus on the visual response to a sinusoidally modulated spatial stimulus. *Journal of the Optical Society of America*, *9*, 1154–1157.

Grosset, A. V., Schott, M., Bonneau, G. P., & Hansen, C. D. (2013). Evaluation of depth of field for depth perception in DVR. *Visualization Symposium (PacificVis)*, 81–88.

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):19, 1–16.

Helmholtz, H. (1867). *Handbuch der Physiologischen Optik*. Rochester, NY: Optical Society of America.

Henze, M. J., Schaeffel, F., & Ott, M. (2004). Variations in the off-axis refractive state in the eye of the Vietnamese leaf turtle (*Geoemyda spengleri*). *Journal of Comparative Physiology A*, *190*, 131–137.

Hibbard, P. B., & Bouzit, S. (2005). Stereoscopic correspondence for ambiguous targets is affected by elevation and fixation distance. *Spatial Vision*, *18*, 399–411.

Hirschmüller, H. (2008). Stereo processing by semi-global matching and mutual information. *Pattern Analysis and Machine Intelligence, IEEE Transactions*, *30*, 328–341.

Hodos, W., & Erichsen, J.T. (1990). Lower-field myopia in birds: an adaptation that keeps the ground in focus. *Vision Research*, *30*, 653–657.

Howard, I. P., & Rogers, B.J. (2002). *Seeing in depth. Vol. 2: Depth perception*. Ontario, Canada: I, Porteous.

Ibbotson, M., & Krekelberg, B. (2011). Visual perception and saccadic eye movements. *Current Opinion in Neurobiology*, *21*, 553–558.

Ivanov, P., & Blanche, T. (2011). Improving gaze accuracy and predicting fixation in real time with video based eye trackers. *Journal of Vision*, *11*(11):505, doi:10.1167/11.11.505. [Abstract]

Kingslake, R. (1992). *Optics in photography*. Bellingham, WA: SPIE Optical Engineering Press.

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

Land, M. F., & Hayhoe, M. (2001). In what ways do eye movements contribute to everyday activities? *Vision Research*, *41*, 3559–3565.

Liu, Y., Bovik, A. C., & Cormack, L. K. (2008). Disparity statistics in natural scenes. *Journal of Vision*, *8*(11):19, 1–14, doi:10.1167/8.11.19. [PubMed] [Article]

López-Gil, N., Martin, J., Liu, T., Bradley, A., Diaz-Muñoz, D., & Thibos, L. N. (2013). Retinal image quality during accommodation. *Ophthalmic & Physiological Optics*, *33*, 497–507.

Maiello, G., Chessa, M., Solari, F., & Bex, P. J. (2015). The (in)effectiveness of simulated blur for depth perception in naturalistic images. *PloS One*, *10*, e0140230.

Marimont, D. H., & Wandell, B.A. (1994). Matching color images: The effects of axial chromatic aberration. *Journal of the Optical Society of America*, *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 of London B: Biological Sciences, 263*, 169–172.

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

Mauderer, M., Conte, S., Nacenta, M. A., & Vishwanath, D. (2014). Depth perception with gaze-dependent depth of field. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 217–226.

McCloskey, S., & Langer, M. (2009). Planar orientation from blur gradients in a single image. *IEEE Conference on Computer Vision and Pattern Recognition, CVPR, 2009*, 2318–2325.

Murphy, C. J., Howland, M., & Howland, H.C. (1995). Raptors lack lower-field myopia. *Vision Research, 35*, 1153–1155.

Nakayama, K. (1977). Geometric and physiological aspects of depth perception. In *21st Annual Technical Symposium* (pp. 2–9). Bellingham, WA: International Society for Optics and Photonics.

Nefs, H. T. (2012). Depth of field affects perceived depth-width ratios in photographs of natural scenes. *Seeing & Perceiving, 25*, 577–595.

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

Okatani, T., & Deguchi, K. (2007). Estimating scale of a scene from a single image based on defocus blur and scene geometry. In *Computer Vision and Pattern Recognition, 2007. CVPR'07* (pp. 1–8).

Palmer, S. E., & Brooks, J. L. (2008). Edge-region grouping in figure-ground organization and depth perception. *Journal of Experimental Psychology: Human Perception and Performance, 34*, 1353–1371.

Pentland, A. P. (1985). The focal gradient: optics ecologically salient. *Investigative Ophthalmology & Visual Science, 26*, 243.

Pentland, A. P. (1987). A new sense for depth of field. *IEEE Transactions on Pattern Analysis and Machine Intelligence, PAMI-9*, 523–531.

Ravikumar, S., Thibos, L. N., & Bradley, A. (2008). Calculation of retinal image quality for polychromatic light. *Journal of the Optical Society of America A, 25*, 2395–2407.

Saarinen, J. (1993). Shifts of visual attention at fixation and away from fixation. *Vision Research, 33*, 1113–1117.

Schaeffel, F., Glasser, A., & Howland, H. C. (1988). Accommodation, refractive error and eye growth in chickens. *Vision Research, 28*, 639–657.

Schaeffel, F., Hagel, G., Eikermann, J., & Collett, T. (1994). Lower-field myopia and astigmatism in amphibians and chickens. *Journal of the Optical Society of America A, 11*, 487–495.

Schreiber, K. M., Hillis, J. M., Filippini, H., Schor, C. M., & Banks, M. S. (2008). The surface of the empirical horopter. *Journal of Vision, 8*(3):7, 1–20, doi:10.1167/8.3.7. [PubMed] [Article]

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]

Seidemann, A., Schaeffel, F., Guirao, A., Lopez-Gil, N., & Artal, P. (2002). Peripheral refractive errors in myopic, emmetropic, and hyperopic young subjects. *Journal of the Optical Society of America A, 19*, 2363–2373.

Shulman, G. L., Sheehy, J. B., & Wilson, J. (1986). Gradients of spatial attention. *Acta Psychologica, 61*, 167–181.

Sivak, J. G., & Allen, D. B. (1975). An evaluation of the "ramp" retina of the horse eye. *Vision Research, 15*, 1353–1356.

Sprague, W. W., Cooper, E. A., Tošić, I., & Banks, M. S. (2015). Stereopsis is adaptive for the natural environment. *Science Advances, 1*(4), 1–17.

Thibos, L. N., Hong, X., Bradley, A., & Cheng, X. (2002). Statistical variation of aberration structure and image quality in a normal population of healthy eyes. *Journal of the Optical Society of America A, 19*, 2329–2348.

Vishwanath, D., & Blaser, E. (2010). Retinal blur and the perception of egocentric distance. *Journal of Vision, 10*(10):26, 1–16, doi:10.1167/10.10.26. [PubMed] [Article]

Wallman, J., Turkel, I. J., & Trachtman, J. N. (1978). Extreme myopia produced by modest change in early visual experience. *Science, 20*, 1249–1251.

Walls, G. L. (1942). *The vertebrate eye and its adaptive radiation*. New York: Hafner.

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.

Wang, B., & Ciuffreda, K.J. (2004). Depth-of-focus of the human eye in the near retinal periphery. *Vision Research, 44*, 1115–1125.

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

Wang, B., Ciuffreda, K. J., & Irish, T. (2006). Equiblur zones at the fovea and near retinal periphery. *Vision Research*, *46*, 3690–3698.

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.

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. [PubMed] [Article]

Zeng, G., Bowrey, H. E., Fang, J., Qi, Y., & McFadden, S. A. (2013). The development of eye shape and the origin of lower field myopia in the guinea pig eye. *Vision Research*, *76*, 77–88.

Zhang, T., O'Hare, L., Hibbard, P. B., Nefs, H. T., & Heynderickx, I. (2014). Depth of field affects perceived depth in stereographs. *ACM Transactions on Applied Perception*, *11*(4), 1–18.