

Dynamic perspective cues enhance depth perception from motion parallax

Athena Buckthought

Psychology Department, Roanoke College,
Salem, VA, USA

Ahmad Yoonessi

McGill Vision Research, Department of Ophthalmology,
McGill University, Montreal, Quebec, Canada

Curtis L. Baker, Jr.

McGill Vision Research, Department of Ophthalmology,
McGill University, Montreal, Quebec, Canada

Motion parallax, the perception of depth resulting from an observer's self-movement, has almost always been studied with random dot textures in simplified orthographic rendering. Here we examine depth from motion parallax in more naturalistic conditions using textures with an overall $1/f$ spectrum and dynamic perspective rendering. We compared depth perception for orthographic and perspective rendering, using textures composed of two types of elements: random dots and Gabor micropatterns. Relative texture motion (shearing) with square wave corrugation patterns was synchronized to horizontal head movement. Four observers performed a two-alternative forced choice depth ordering task with monocular viewing, in which they reported which part of the texture appeared in front of the other. For both textures, depth perception was better with dynamic perspective than with orthographic rendering, particularly at larger depths. Depth ordering performance with naturalistic $1/f$ textures was slightly lower than with the random dots; however, with depth-related size scaling of the micropatterns, performance was comparable to that with random dots. We also examined the effects of removing each of the three cues that distinguish dynamic perspective from orthographic rendering: (a) small vertical displacements, (b) lateral gradients of speed across the corrugations, and (c) speed differences in rendered near versus far surfaces. Removal of any of the three cues impaired performance. In conclusion, depth ordering performance is enhanced by all of the dynamic perspective cues but not by using more naturalistic $1/f$ textures.

Introduction

Motion parallax, the differential motion of retinal images of objects at different distances resulting from natural observer movement, is a powerful depth cue. As our vantage point moves, objects that are closer to us move faster across our field of view than objects that are farther away (Gibson, Gibson, Smith, & Flock, 1959; Helmholtz, 1925; Howard & Rogers, 2012). For example, as an observer fixates on a point in space while making translational head movements, objects nearer than fixation move in one direction on the retina, and farther objects move in the opposite direction. This relative motion of objects or of texture elements lying on surfaces, is a highly effective cue for depth perception, particularly for judging the relative depths of nearby surfaces (e.g., Rogers & Rogers, 1992). The visual stimuli employed in most previous studies of motion parallax were simplified in ways that made it more technically feasible to produce them, but they differ from those in natural motion parallax in two potentially important ways.

First, previous studies have been limited in using random dot textures, which differ greatly from textures in natural scenes, for example, in their luminance distribution and power spectrum. Visual cortex neurons have receptive fields that are optimally efficient for encoding the rich spatial properties of natural images (Olshausen & Field, 1997), suggesting that these properties are important for visual perception. Previous studies suggest that naturalistic information may enhance depth perception (Cooper & Norcia, 2014; Lee & Saunders, 2011; Saunders & Chen, 2015) and act as more effective stimuli in other ways, e.g., noise masking

Citation: Buckthought, A., Yoonessi, A., & Baker, C. L., Jr. (2017). Dynamic perspective cues enhance depth perception from motion parallax. *Journal of Vision*, 17(1):10, 1–19, doi:10.1167/17.1.10.

doi: 10.1167/17.1.10

Received August 9, 2016; published January 11, 2017

ISSN 1534-7362



(Hansen & Hess, 2012) or binocular rivalry (Baker & Graf, 2009).

Second, most previous motion parallax studies used simulated surfaces whose texture motions are synchronized to observer movement in a simple proportional manner, i.e., orthographic rendering (e.g., Bradshaw, Hibbard, Parton, Rose, & Langley, 2006; Nawrot & Joyce, 2006; Ono & Ujike, 2005; Rogers & Rogers, 1992; Yoonessi & Baker, 2011). These stimuli are ideal for examining the relative motion cue in isolation. However, in naturally occurring motion parallax, the visual motion contains several other “dynamic perspective” cues related to eccentricity, distance, and viewing angle. The possible contribution of these cues has only been explored to a limited extent with motion parallax (Rogers, 2012; Rogers & Rogers, 1992). However, demonstrations of improved depth perception from analogous cues in stereopsis (Backus, Banks, van Ee, & Crowell, 1999; Bradshaw, Glennerster, & Rogers, 1996; Rogers & Bradshaw, 1993) suggest that dynamic perspective cues might enhance depth from motion parallax and make it less ambiguous (Rogers, 2012). (Note that dynamic perspective cues refer only to the optic flow or motion of texture elements and not pictorial or static perspective cues, e.g., texture gradients or changes of object size with distance, which could potentially give rise to depth in a static image.)

In the present study, we examine the effects of using motion parallax stimuli that are more naturalistic with regard to both of the above concerns. First of all, in addition to random dots, we also use a more naturalistic type of texture, Gabor micropatterns with different sizes and orientations on a midgray background (Figure 1A). The relative numbers of different spatial scales of Gabors are adjusted so as to produce an overall Fourier spectrum falling off in proportion to spatial frequency ($1/f$) as with natural images (Field, 1987; Kingdom, Hayes, & Field, 2001).

Here we also use more naturalistic dynamic perspective rendering and compare the results to orthographic rendering. A comparison of orthographic and perspective rendering is depicted in Figure 2, exaggerating the differences for the purpose of illustration. Each panel shows a rectangular gridded surface in the frontoparallel plane as viewed from left and right positions when an observer is performing right-to-left head movements while fixating at the center. Orthographic rendering uses a simplified method for calculating the optic flow or motion of points on an image at different depths. If the rectangular surface lies in the fixation plane (i.e., has zero depth), then no motion of the surface occurs with head movements (Figure 2A). If the rectangular surface is farther than the fixation point, it will move in the same direction as the head movement (Figure 2C), but if it has near depth, it

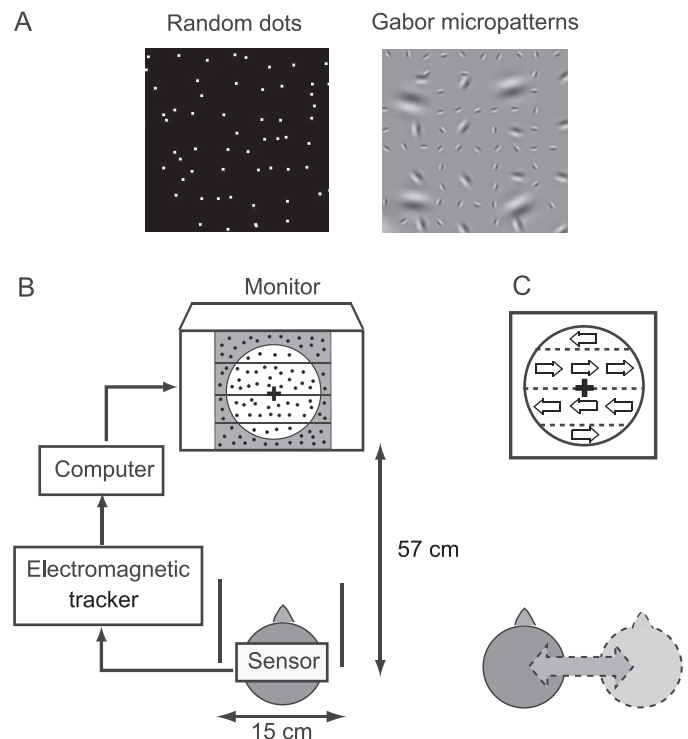


Figure 1. Stimuli and setup for measurement of depth from motion parallax. (A) Examples of random dot patterns (left) and Gabor micropattern textures (right) used in these experiments. (B) Schematic of the experimental setup used to measure head position and synchronize visual stimulus to head movement. Observers moved their head freely from side to side within a 15-cm span between two vertical bars acting as guides for the range of motion. An electromagnetic sensor registered head position. The computer updated stimulus motions in real time, in synchrony with head movement data, without any noticeable time lag. (C) When the observer performs side-to-side head movements while fixating at the center of the screen, the rendered depth is proportional to the ratio of stimulus motion to head movement (“syncing gain”). This is true for orthographic rendering, while for perspective rendering the stimulus motion is more complex to simulate the surface in depth as it appears from different eye positions.

moves in the opposite direction with equal speed (Figure 2E).

Dynamic perspective rendering calculates more precisely the transformation of the rectangular surface as it should appear from different positions. First consider the rectangular surface that lies in the fixation plane (i.e., has zero depth; Figure 2B). As the surface is viewed from the right side, its right edge is closer to the eye and appears slightly larger than the left edge. The opposite is true when the rectangular surface is viewed from a position to the left side. Thus, as we observe a frontoparallel rectangular object from one side, its retinal image is actually trapezoidal in shape. Due to shape constancy, this subtle change in shape may not

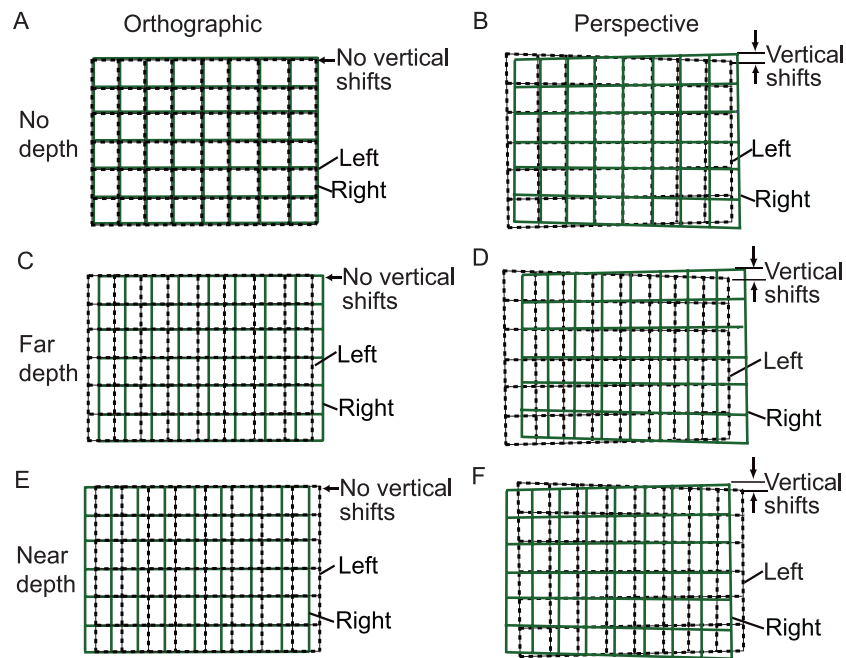


Figure 2. Conceptual illustration of the differences between orthographic (left column, A, C, E) and perspective (right column, B, D, F) rendering of a single (gridded) frontoparallel surface for three cases: (A, B) No depth. (C, D) Far depth. (E, F) Near depth. The grid image is shown as viewed from left and right end points of lateral head excursion with the appropriate transformations exaggerated for the purpose of illustration.

usually be consciously noticeable, but it is present in the retinal image and therefore available as a depth cue. Now we can consider what happens if the rectangular surface has depth in front of or behind fixation. As the observer makes head movements, the surface will move in the same direction for far depth (Figure 2D) or the opposite direction for near depth (Figure 2F). These changes with depth are similar to orthographic rendering, but the shifts are larger for near than far surfaces. Again, as we view the rectangular surface in depth from one side, either the left or right edge will be closer to the eye and will be larger so that the retinal image is actually trapezoidal in shape.

The differences between orthographic and dynamic perspective rendering can be further illustrated with a schematic diagram showing the optic flow pattern with arrows indicating direction and speed of moving texture elements (Figure 3, top left). The optic flow pattern is shown for a near surface above a far surface, which are above and below a central fixation point, respectively. In the case of orthographic rendering, the texture elements on far and near surfaces move in opposite directions with equal speeds. For perspective rendering, the texture elements on far and near surfaces also move in opposite directions, but elements on the near surface move faster, and there are other subtle differences in the overall optic flow pattern as we saw with the rectangular surfaces.

Thus, dynamic perspective rendering has three additional depth cues that are not present in ortho-

graphic rendering, which we explain in turn. The first cue is the presence of small vertical displacements, which are present in addition to the larger horizontal displacements and will occur for any image features not in the median plane of the image. This cue is analogous to vertical disparity in stereopsis (Howard & Rogers, 2012; Read & Cumming, 2006; Rogers & Bradshaw, 1993).

The second additional cue in dynamic perspective rendering is the presence of lateral gradients in the speeds of points on the rendered surface across the extent of the corrugations in depth. As shown in Figure 3 (top left), for a frontoparallel surface at far depths, the displacements are slightly larger at the outer edges of the display, which are closer to one eye than the other. For a near surface, this gradient reverses; speeds are slightly lower at the outer edges than the center. Typically, there are gradients in both vertical and horizontal speeds (Read & Cumming, 2006; Rogers & Bradshaw, 1993).

The third cue is the presence of speed differences between points on near and far surfaces. This is the familiar depth cue mentioned earlier in which close objects move faster than distant objects when an observer makes head movements: The retinal image velocity of separate objects in the visual field is inversely proportional to the square of the distance from the observer's fixation plane (Helmholtz, 1925; Howard & Rogers, 2012).

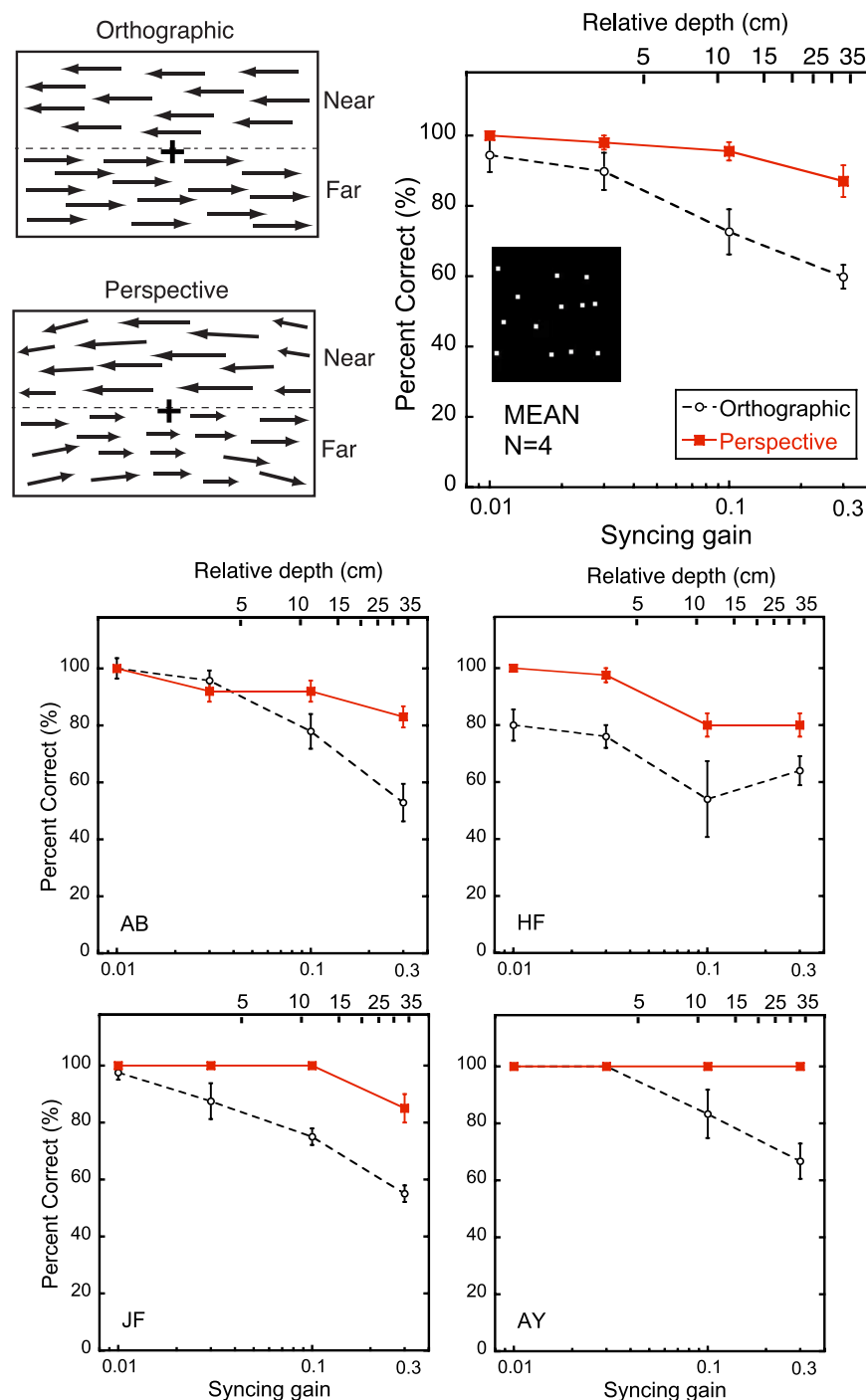


Figure 3. Depth ordering performance in motion parallax with random dot textures, using orthographic versus perspective rendering. Top left shows schematic illustration comparing orthographic and perspective rendering of dot trajectories. Results are shown for the mean of all observers (large graph, upper right) and for four individual observers (smaller graphs, bottom). Each graph shows performance as a function of syncing gain, equivalent to rendered depths double-labeled across the top of each graph. Depth perception was better with perspective (red) than orthographic (black) rendering, particularly at large rendered depths. Error bars in this and all subsequent figures indicate standard errors.

In the first part of this study, we demonstrate enhanced depth perception with perspective compared to orthographic rendering of dynamic element motion, for both random dot and Gabor textures, with better

performance for random dots. Then, for random dot stimuli, we examine the effects of selectively removing each of the three cues that distinguish dynamic perspective from orthographic rendering and find that

all of them contribute substantially to the perception of depth from motion parallax. Finally, we show that by adding perspective cues to individual micropatterns, thereby removing an inadvertent “flatness cue,” performance for Gabor textures becomes comparable to that for random dots.

General materials and methods

We only briefly summarize the hardware and software setup because they have been described in detail in previous papers (Yoonessi & Baker, 2011, 2013). The observer made lateral head movements, and texture motion on the monitor was synchronized to these movements in order to simulate a real three-dimensional surface in depth. An overall schematic of the experimental setup is shown in Figure 1B, and its details are described in the following sections.

Visual stimuli

Visual stimuli were generated with a Macintosh computer (Mac Pro, 2×2.8 GHz, 4-GB RAM, OSX v10.5) using Matlab code (version 2007b, Mathworks, Natick, MA) and Psychophysics Toolbox Version 3 (Brainard, 1997). The stimuli were presented on a CRT monitor (Sony Trinitron A7217A, 1024×768 pixels, 75-Hz refresh rate or 13.3 ms per frame), which was gamma-corrected for the Gabor texture stimuli with a mean luminance of 40 cd/m^2 . The stimuli were viewed monocularly to avoid a cue conflict with stereopsis at a distance of 57 cm.

Random dot textures

The stimulus patterns consisted of white (80.1 cd/m^2) dots on a black (0.08 cd/m^2) background. Each dot was circular in shape, 0.3° in diameter, rendered with high-quality antialiasing using the “DrawDots” function of Psychophysics Toolbox. Pilot studies showed that depth perception was good over a wide range of densities of the dots (or Gabors); thus, the exact density to be used was not critical. Here we used a dot density of 0.83 dots/deg^2 .

Gabor micropattern textures

The micropattern textures consisted of sine-phase Gabors with four spatial frequencies (1, 2, 4, and 8 cycles/deg) and random orientations. OpenGL 4.0 Graphics Shading Language was used for fast rendering of the Gabor elements (with the exception of Experiment 4). The relative numbers of micropatterns

at the four spatial frequencies were adjusted to produce an approximation to a $1/f$ power spectrum for the texture as a whole (Kingdom et al., 2001). The individual Gabor elements were randomly scattered to create each texture but were not allowed to overlap with one another. The RMS contrast of each texture was adjusted to 15.8%, which was well above threshold for all observers. The density of texture elements was $0.83 \text{ elements/deg}^2$.

Stimulus display

The displacements of the random dot or Gabor micropattern elements were modulated using square wave profiles (0.05 cycles/deg) to create shearing motion patterns (Figure 1C). The stimuli were presented within a circular mask of 28° , which resulted in about 1.4 cycles/image of visible square wave modulation. A fixation point was presented prior to and throughout each stimulus presentation at the center of the circular mask. The fixation point was always set at the transition point (Figure 1C) between the oppositely moving peaks and troughs of the bidirectional square wave modulation waveform. This texture motion simulated adjacent surfaces that were behind (half cycles moving in the same direction as head movement) and in front (half cycles moving oppositely to head movement) of the monitor screen, respectively.

To achieve precise real-time synchronization of visual stimuli to head position, we used a digital position measurement system in conjunction with exploiting the graphics card GPU capabilities for drawing. To produce the correct retinal motion giving rise to depth from motion parallax, the dot or Gabor motions were synchronized to measured changes in head position. On each frame update, the difference between the current and previous head position was multiplied by a gain parameter and applied to the one-dimensional modulation profile to modulate the texture element positions and thereby generate a shearing pattern. The display system did not produce “dropped frames,” so the stimulus motion appeared very smooth and systematically proportionate to head movement. The delay between head movement and stimulus update was approximately 20 ms, which did not produce any noticeable sensorimotor lag.

We use the ratio between head movement and image motion, which we call “syncing gain,” as an important parameter in our experiments (Longuet-Higgins & Prazdny, 1980; Ono & Ujike, 2005; Yoonessi & Baker, 2011). The syncing gain is linearly proportional to the rendered depth, and therefore, our graphs of depth ordering performance versus syncing gain are double-labeled for relative depth. The stimulus can also be described in terms of equivalent disparity; for example, 1 min of disparity equivalence means that a stimulus

element has translated 1 min for a head translation of the 6.5 cm interocular distance (Nawrot & Joyce, 2006). Similarly to binocular stereopsis, larger disparity equivalence, corresponding to larger element translations, generates larger magnitudes of simulated depth.

Orthographic and dynamic perspective rendering

For orthographic rendering, there is a simple relationship between the head movement and the speeds of the random dot or Gabor elements with speed linearly proportional to depth in front of or behind fixation (also directly proportional to syncing gain). For dynamic perspective rendering, the computations of element speeds are more complex, involving motion with different magnitudes and directions, in order to accurately render the depth specified by a given syncing gain value. The image transformations for perspective rendering were calculated using equations 5 and 6 in Read and Cumming (2006), appendix A (p. 1346), which they used for stereopsis:

$$x_L = f \frac{Z \sin H_L - (X - D) \cos H_L}{[(X - D) \sin H_L + Z \cos H_L]} \quad (1)$$

$$x_R = f \frac{Z \sin H_R - (X + D) \cos H_R}{[(X + D) \sin H_R + Z \cos H_R]} \quad (2)$$

$$y_L = - \frac{fY}{[(X - D) \sin H_L + Z \cos H_L]} \quad (3)$$

$$y_R = - \frac{fY}{[(X + D) \sin H_R + Z \cos H_R]} \quad (4)$$

The equations calculate the projection in retinal coordinates of the left and right eyes (horizontal coordinates: x_L , x_R ; vertical coordinates: y_L , y_R) of an object that is at a particular location in external space (X , Y , Z in a head-centered coordinate system). The rotation angles of the left and right eyes are specified by H_L and H_R , respectively. The focal length of the eye is f . These equations can easily be adapted for motion parallax if we assume that just one eye moves between the indicated positions of the left and right eyes. The lateral head movement or distance from one viewpoint to the next is D . Thus, we used these equations to calculate the position of moving texture elements on the screen as the observer makes head movements. We calculated the observer's changing eye position in retinal coordinate space and used these equations to give the appropriate transformations to calculate the moving texture element positions in external space. These equations are appropriate for perspective rendering because all the vertical and horizontal displacements are calculated correctly for all texture element positions, allowing for the effects of eccentricity. This also takes into consideration the lateral

movement and rotation of the eye as the observer makes head movements and fixates in the center (Read & Cumming, 2006). Translational vestibulo-ocular reflex and ocular following response eye movements likely occur during presentation and might be imperfect, which may add small inaccuracies in rendering (Adeyemo & Angelaki, 2005; Miles, 1998; Quaia, Sheliga, Fitzgibbon, & Optican, 2012).

Head movement recording

The head position and orientation for six degrees of freedom (6-DOF) were measured (0.5 mm and 0.1° resolution, respectively) using an electromagnetic position-tracking device (Flock of Birds, Ascension Technologies, Shelburne, VT) with a medium range transmitter. The sensor was fastened to the observer's forehead using a headband. The head movement was sampled at 100 Hz and transferred to the host computer using a serial port/USB connection. The change in lateral head position was used for real-time modulation of the stimulus motion as described previously, and the complete 6-DOF position/orientation was recorded to hard disk for subsequent analysis. Observers freely moved their head laterally back and forth while viewing the stimulus during each trial, traversing a path between two vertical bars with a spacing of 15 cm about once every second. The head position data for every trial was monitored and recorded, and observers were asked to make adjustments if necessary. The observers' lateral head movements were consistent in displacement and speed from trial to trial. The average velocities ($\pm SD$) across all trials for the four observers were as follows: 14.79 ± 1.42 cm/s, 15.85 ± 1.55 cm/s, 14.63 ± 1.47 cm/s, and 15.14 ± 1.62 cm/s. Head movements in other translational or rotational directions were minimal and not systematic (Yoonessi & Baker, 2011).

Psychophysical task and observers

Observers performed a two-alternative forced choice depth ordering task, in which they reported which modulation half cycle of the texture adjacent to the fixation point appeared in front of the other. No feedback as to correctness of response was provided. A total of 60 trials were run for each stimulus condition, presented in pseudorandomized order. The stimulus presentation time was 5 s.

Four observers (AB, AY, HF, and JF) participated in all of the experiments with the exception of Experiment 4, in which only three observers participated (AB, JF, and HF). Two of the observers (JF and HF) were naive to the purpose of the experiment, and

the other two (AB and AY) were authors. All observers had normal or corrected-to-normal vision. These experiments conformed with McGill University's ethical guidelines for human experimentation as well as the Declaration of Helsinki. All observers gave prior informed written consent for participation.

Experiment 1: Comparison of orthographic and perspective rendering

In Experiment 1, we compared orthographic and dynamic perspective rendering for two types of textures (random dots, Gabor micropatterns).

Results

The graphs in Figure 3 show depth ordering performance with random dot stimuli at a series of syncing gain values, which correspond to increasing amounts of simulated relative depth (top axes). The large, upper right graph is the mean for four observers with individual data shown in the smaller, bottom graphs. Depth ordering performance was better for dynamic perspective (red) than orthographic (black) rendering, which is consistent with our hypothesis that dynamic perspective cues improve depth perception. Furthermore, performance systematically declined as the rendered depth was increased; this is consistent with the predictions of neural models (Fernandez & Farell, 2008) that there is an upper limit on the depth that can be perceived as found in previous studies (Ono, Rivest, & Ono, 1986; Yoonessi & Baker, 2011). Results for the same measurements using Gabor micropattern stimuli are shown in a similar format in Figure 4. Again, performance is systematically better for dynamic perspective than for orthographic rendering with a general decline at larger syncing gains or depths. Comparison of the results in Figures 3 and 4 also showed that depth ordering performance was somewhat lower with the Gabor micropatterns than the random dots, for both dynamic perspective and orthographic rendering. A repeated-measures ANOVA confirmed that the main effect of type of texture (Gabors vs. dots), $F(1, 3) = 312.5$, $p = 0.0001$; main effect of rendering, $F(1, 3) = 77.1$, $p = 0.003$; and main effect of syncing gain, $F(3, 9) = 37.5$, $p = 0.001$, were all significant. The interactions were not significant—rendering by texture type: $F(1, 3) = 0.387$, $p = 0.578$; rendering by syncing gain: $F(3, 9) = 2.60$, $p = 0.12$; texture type by syncing gain: $F(3, 9) = 2.54$, $p = 0.12$, with the exception of the three-way interaction of rendering by texture type by syncing gain:

$F(3, 9) = 5.26$, $p = 0.02$, which was significant. Although caution should be used in drawing conclusions from a study with a small number of participants, these results confirm the overall trend that depth ordering performance is better for perspective than orthographic rendering with both types of stimuli.

This difference between random dot and Gabor micropattern textures was somewhat surprising because using a stimulus that is more naturalistic, at least in having a $1/f$ spectrum, actually lowered depth ordering performance. A possible explanation for this difference is that perspective rendering was applied only to the optic flow of texture elements and not to the individual elements themselves (see General materials and methods); this would have a negligible effect on the small dots but could create a significant cue conflict with the Gabors especially for larger elements. We return to this point in Experiment 4.

Experiment 2: Removal of individual dynamic perspective cues

Experiment 1 indicated that depth ordering performance was better with dynamic perspective than orthographic rendering for both types of textures. As discussed earlier, there are three dynamic perspective cues that could potentially contribute to this superior performance: vertical displacements, speed differences between near and far surfaces, and lateral gradients in speed (both horizontal and vertical). In this section, we selectively removed each of these three cues from the perspective rendered stimulus to assess their contributions to depth ordering performance.

Methods

The stimulus display, experimental setup, depth ordering task, and general procedures were the same as in Experiment 1 but using versions of the stimulus with selective removal of individual dynamic perspective cues. This experiment was carried out using random dots, which the previous experiment demonstrated to give better performance than the Gabor textures.

Stimulus conditions

No vertical displacements

This stimulus was produced by setting all vertical displacements to zero in the dynamic perspective rendered stimulus as shown schematically in Figure 5 (upper left, bottom-most stimulus schematic). As shown, the displacements in other directions as well as

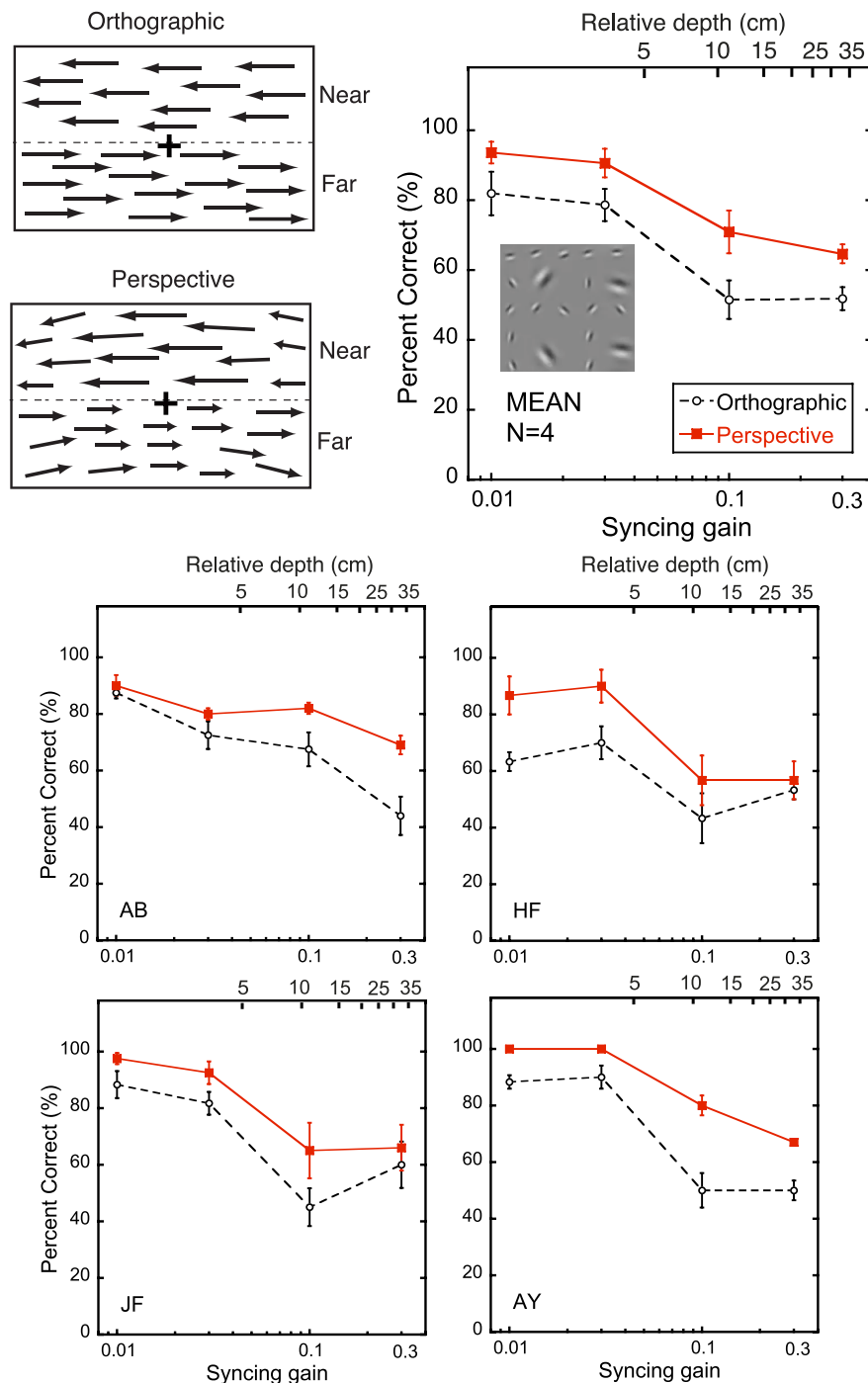


Figure 4. Same as Figure 3 but for Gabor micropattern textures. As before, depth ordering performance was better for perspective than orthographic rendering.

lateral gradients and speed differences between near and far surfaces were still present.

No speed differences between near and far surfaces

This stimulus was produced by eliminating the speed differences between near and far surfaces in the perspective rendered stimulus. Thus, the moving dots at an equal distance in front of or behind the fixation

point had the same speed (see Figure 6, upper left), but the other cues (vertical displacements, lateral gradients in speed) were still present.

No lateral gradients in speed

This stimulus was produced by eliminating the lateral gradients in speed within each of the surfaces in the perspective rendered stimulus (see Figure 7, upper

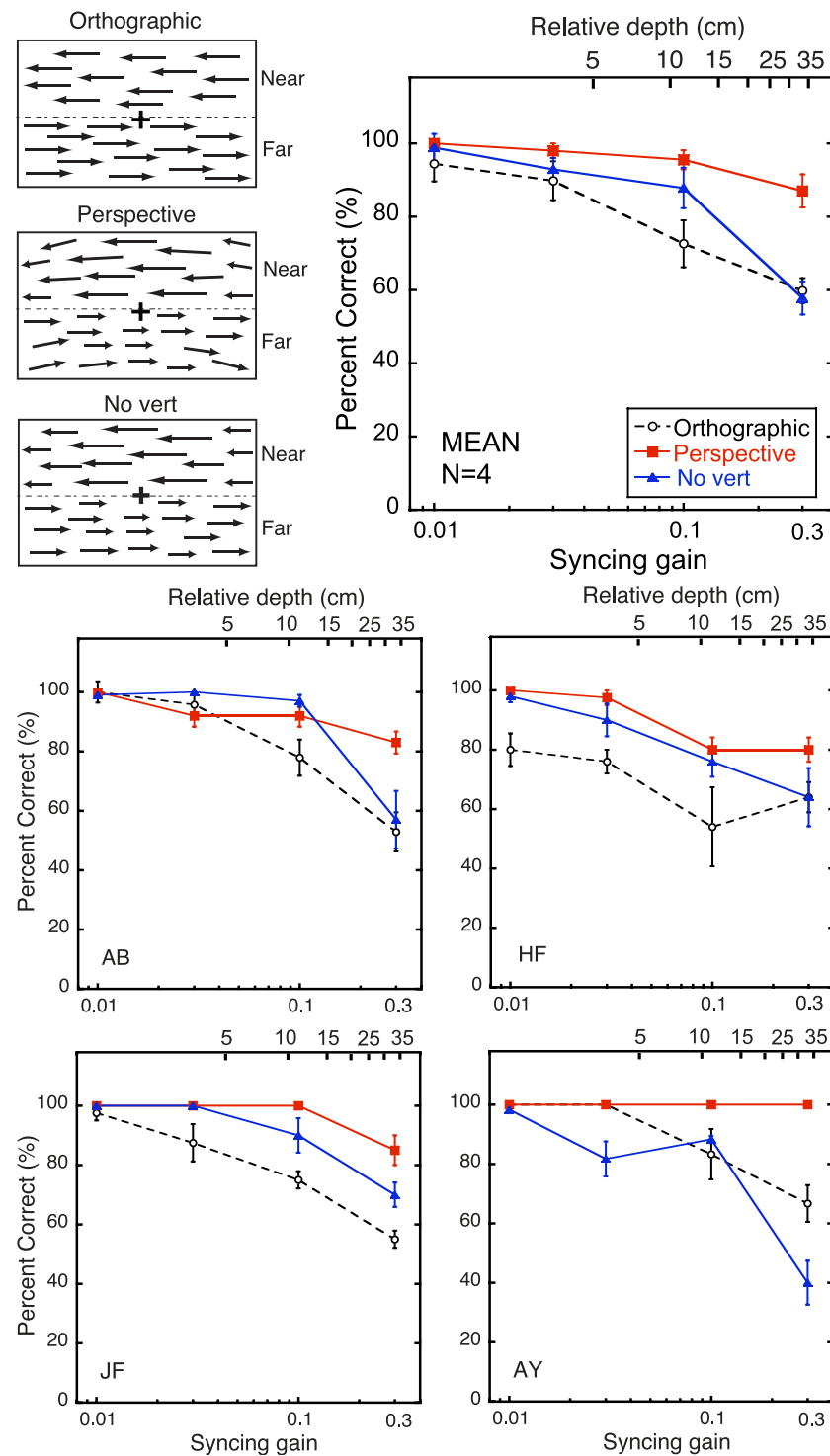


Figure 5. Effect of removing vertical displacements from dynamic perspective rendering. Upper left shows schematic illustrations of random dot trajectories, comparing orthographic and perspective rendering and perspective with vertical displacements removed. Graph at upper right shows depth ordering performance (mean of four observers) with the random dot stimulus for perspective rendering with all vertical displacements removed (blue). The results for orthographic (black) and perspective (red) rendering are also reproduced from Figure 3 for reference. Bottom four graphs show results for the individual observers. Depth ordering performance with vertical displacements removed is poorer than for full dynamic perspective rendering, particularly at larger rendered depths.

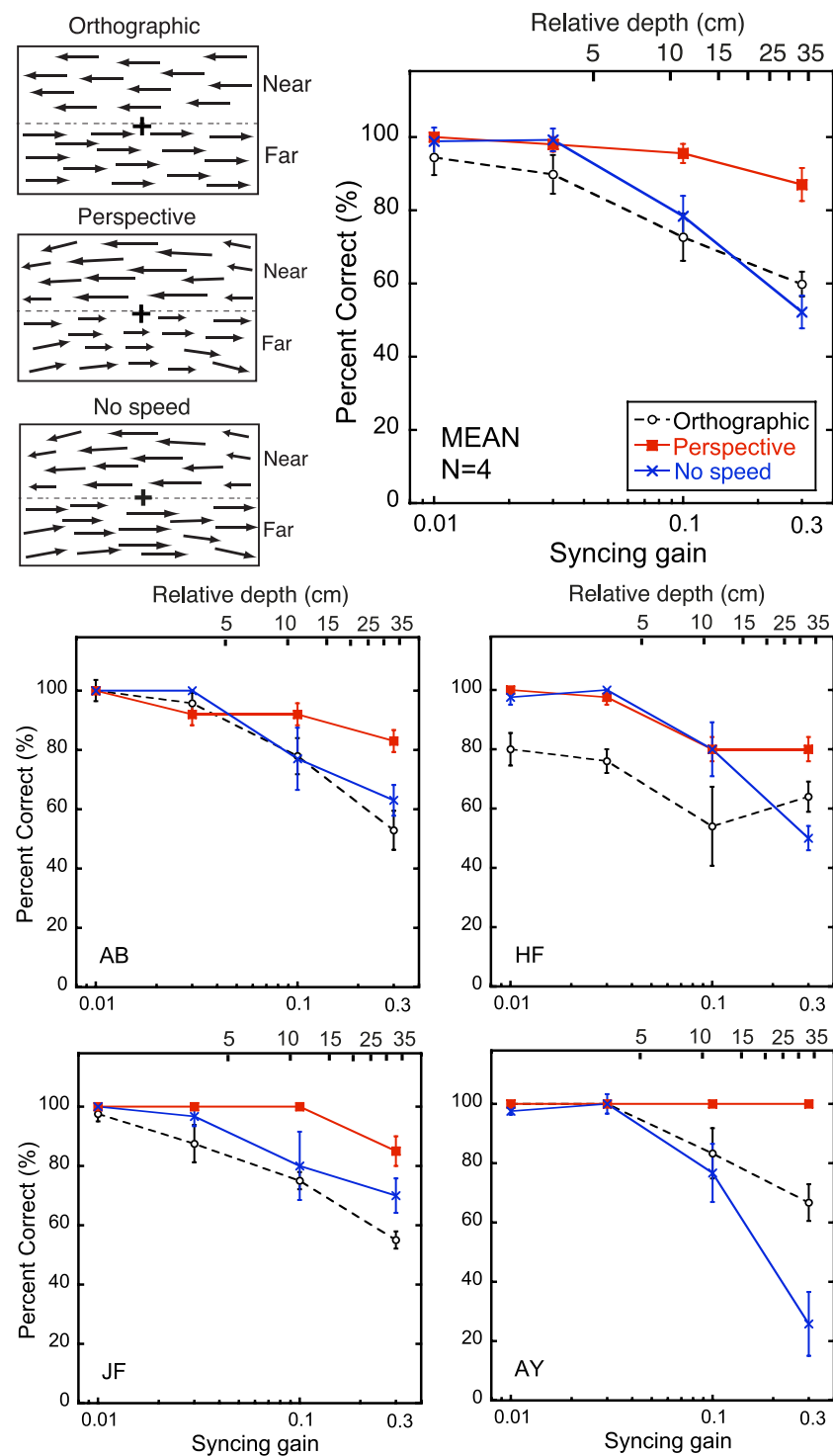


Figure 6. Same as Figure 5 but for perspective rendering with all speed differences between near and far surfaces removed. Depth ordering performance was poorer than that for perspective rendering, particularly at large rendered depths.

left). These gradients could include components in both horizontal and vertical directions, and both were removed. As shown in the schematic of dot motions, the differences in speed between near and far surfaces

were still present. Vertical displacements were also still present but are not prominent in the illustration because the variation of this cue (gradient) in the horizontal direction was removed.

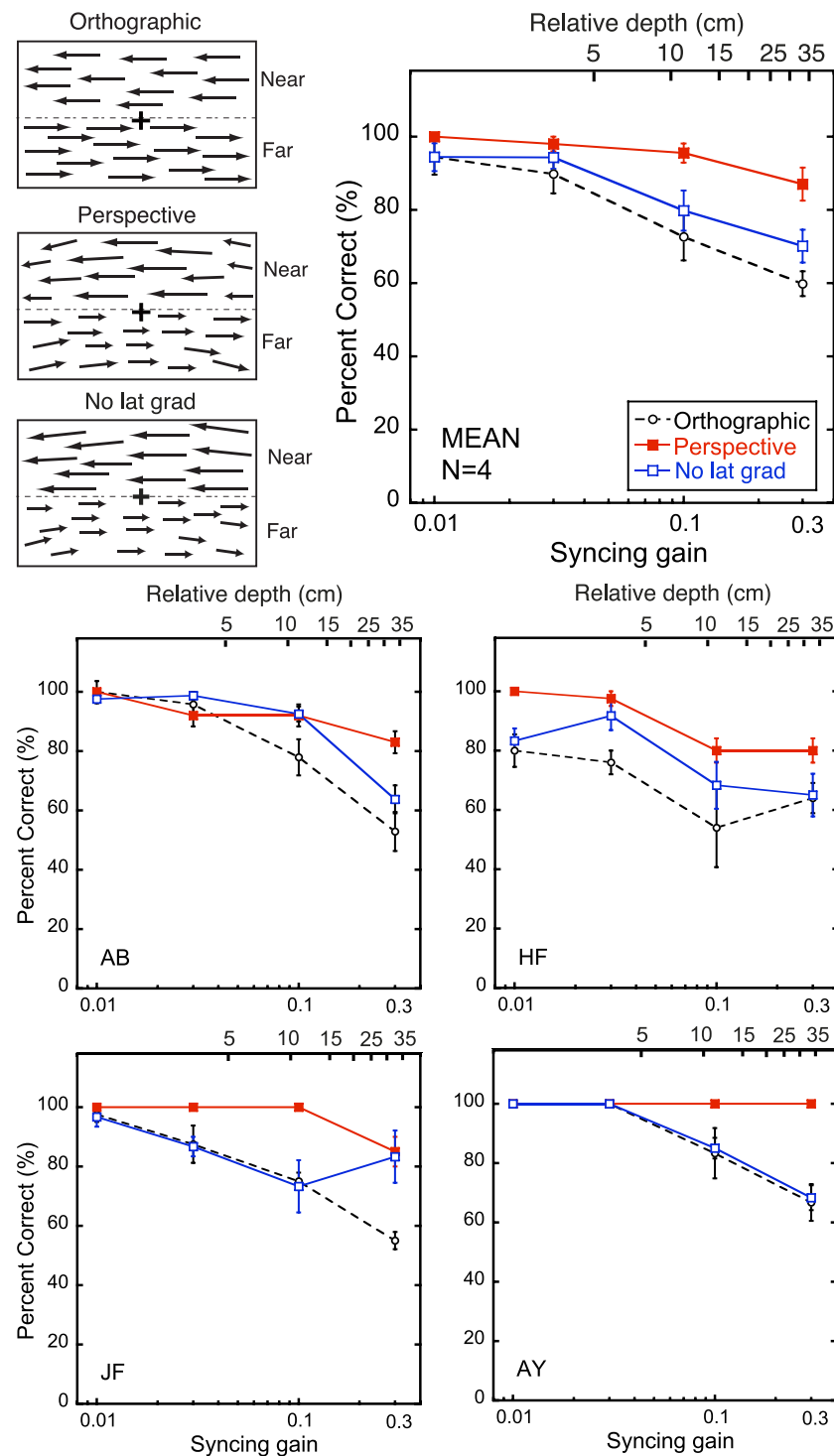


Figure 7. Same as Figure 5 but for perspective rendering with all lateral gradients in speed removed. Depth ordering performance was poorer than for perspective rendering, particularly at large rendered depths.

Results

No vertical displacements

Depth ordering performance with the perspective rendered stimulus in which the vertical displacements have been removed is shown in Figure 5 (blue) with comparison to the earlier results with orthographic

(black) and dynamic perspective (red) rendering. The results suggest that removing the vertical displacements impaired depth ordering performance. As before, we also found that, in general, depth ordering performance declined as the rendered depth was increased. A repeated-measures ANOVA showed that the main

effect of cue condition (full perspective rendering versus that with vertical displacements removed) was not significant, $F(1, 3) = 6.32$, $p = 0.09$, but the main effect of syncing gain, $F(3, 9) = 49.6$, $p = 0.0001$, and the cue by syncing gain interaction were both significant, $F(3, 9) = 5.39$, $p = 0.022$. The interaction occurred because depth ordering performance declined with syncing gain to a greater extent when the vertical displacements were not present. The effect of removing the vertical displacements was more pronounced for some observers (e.g., AY) than others.

We noticed that this removal of vertical motions could potentially cause the stimulus to be perceived as slightly convex, which might also have an effect on the difficulty of the depth ordering task. For this reason, we also tested an alternative condition in which the vertical displacements were randomized instead of eliminated. In this case, the surfaces were perceived to be frontoparallel. Depth ordering results with this control condition (Figure S1, Supplementary Materials) were comparable to those obtained with the vertical displacements removed. Thus, these results reinforce the idea that vertical displacements contribute to depth ordering performance.

No speed differences

Depth ordering performance for the perspective rendered stimulus in which the speed differences between the near and far surfaces were removed is shown in Figure 6 (blue). Removing these speed differences had the effect of lowering depth ordering performance. A repeated-measures ANOVA confirmed that the main effect of cue condition (i.e., perspective rendering versus that with speed differences removed), $F(1, 3) = 10.1$, $p = 0.045$, and syncing gain, $F(3, 9) = 54.5$, $p = 0.0001$, were both significant as well as the cue by syncing gain interaction, $F(3, 9) = 6.38$, $p = 0.013$. The significant interaction indicated that, when the speed differences were removed, depth ordering performance dropped off more steeply as the rendered depth was increased. We note that there were individual differences in the effects of removing speed differences for different observers (e.g., the effect was quite pronounced for observer AY and less so for observer JF).

No lateral gradients in speed

Depth ordering performance for the perspective rendered stimulus in which the lateral gradients in speed have been removed is shown in Figure 7 (blue). Removing these speed gradients lowered depth ordering performance. A repeated-measures ANOVA revealed that the main effect of cue condition (i.e., perspective rendering vs. perspective rendering with

lateral gradients removed) was significant, $F(1, 3) = 17.9$, $p = 0.024$, as well as syncing gain, $F(3, 9) = 20.8$, $p = 0.0001$, but with no significant interaction, $F(3, 9) = 1.67$, $p = 0.24$. Again, there were some individual differences in the extent to which removing lateral gradients in speed affected depth ordering performance.

The lateral speed gradients in normal dynamic perspective include both horizontal and vertical components, and we questioned what their respective contributions could be. Therefore, we tested stimuli in which we eliminated the lateral gradients either in the horizontal or in the vertical speeds along the corrugations in the perspective rendered stimulus while retaining the respective orthogonal speed gradients. The results confirmed that depth ordering performance was also impaired by removing the lateral gradients in either horizontal speeds (Figure S2, Supplementary Materials) or in vertical speeds (Figure S3, Supplementary Materials).

In summary, the removal of any of the three perspective cues (speed differences between near and far surfaces, vertical displacements, or lateral gradients in speed along the corrugations) impaired depth ordering performance. However, the effect of removing these cues differed somewhat across subjects. It was also a general finding across all conditions that the depth ordering performance dropped off as the rendered depth was increased, consistent with previous studies (Ono et al., 1986; Yoonessi & Baker, 2011, 2013).

Experiment 3: Noise coherence thresholds

In the experiments so far, we only measured percentage correct in depth ordering rather than titrating the level of difficulty to obtain thresholds because the long trial durations (5 s each) together with multiple syncing gain values would otherwise necessitate excessive time donations from our volunteer observers. However, it is possible that the results could have been affected by a floor or ceiling effect, and the task could either be too difficult or easy to reveal any significant variation with the cue manipulations. To assess this possibility using a more sensitive measure, we also tested the observers on the same stimulus conditions and task but with additive depth noise to degrade performance (Yoonessi & Baker, 2011) at single values of syncing gain. These coherence noise thresholds were obtained from measuring depth ordering performance at a range of different percentages of added noise.

As before, observers performed a two-alternative forced choice depth ordering task in which they reported which modulation half cycle of the texture appeared in front of the other. A percentage of “signal dots” were presented at the near and far depths, and the remaining “noise dots” were assigned random depths (ranging between the near and far surfaces). The percentage of the signal dots (depth coherence) was manipulated to control the difficulty level of the task so that the depth ordering performance increased monotonically with depth coherence. No feedback was given as to correctness of responses. The stimulus presentation time was 5 s. Sixty trials were run at each of eight levels defining the percentage of noise dots (25%, 37.5%, 50%, 62.5%, 75%, 81.25%, 87.5%, and 100%) in pseudorandomized order. Three observers were tested at a syncing gain of 0.1, and one observer (HF) was tested at a syncing gain of 0.01 because she had difficulty with the task at 0.1.

To quantify psychophysical sensitivity, a cumulative Gaussian function was fit to the proportion correct data for a total of 640 trials, and the percentage of noise dots corresponding to 75% correct performance was taken as the threshold. Curve fitting and calculation of standard errors was performed using the Prism statistics software package (GraphPad Software, Inc., La Jolla, CA).

Results

Coherence noise thresholds are shown for the different stimulus conditions in Figure 8, which include orthographic (gray) and perspective (orange) rendering as well as perspective rendering with each of the three cues selectively removed (blue). The results consistently showed better tolerance to added noise dots for perspective than for orthographic rendering. Furthermore, the removal of any of the three perspective cues lowered the percentage of noise for performance at the criterion level although to different extents across observers.

Experiment 4: Higher precision perspective rendering removing flatness cue

The results of Experiment 1 were somewhat unexpected because depth ordering performance was lower for the Gabor micropattern stimuli than the random dots even though the Gabors are more naturalistic texture stimuli. Based on the importance of perspective information demonstrated in the above experiments, we wondered whether the poorer performance with Gabor

textures might be due to the simplified manner in which perspective rendering was done. The dynamic perspective rendering captured the optic flow (pattern of motion vectors) of texture elements but did not transform each texture element in size, orientation, or aspect ratio in accordance with its image when viewed from different head positions. (This simplification in rendering has been used previously in applied video studies and is referred to as “billboarding”—Kaiser, Montegut, & Proffitt, 1995). This, in effect, produced a “flatness” cue, a potential cue conflict that could diminish the perceived depth and lower depth ordering performance.

Hence, in Experiment 4, we used a more precise rendering that incorporated both the correct optic flow of elements and also the rendering of each individual texture element with the appropriate perspective transformations. Note that the Gabors were still at the same sizes regardless of the rendered depth; i.e., the typical variation of texture element size with distance that occurs in a naturalistic setting was not incorporated (Howard & Rogers, 2012). This was the case because the micropattern textures had Gabors that were always the same four sizes regardless of the syncing gain or depth being rendered. But the individual Gabor micropatterns now had subtle changes in aspect ratio and orientation as they were viewed from different vantage points.

Methods

The experimental setup and task were the same as in Experiments 1, 2, and 3. However, the OpenGL Graphics Shading Language, used in the previous experiments, could not render the texture elements fast enough for adequate real-time performance. Instead, the rendered images were precalculated and stored prior to the experiment for each of 180 head positions equally spaced over a 15-cm excursion. At the start of each stimulus trial, all of the precalculated images were loaded into host machine memory, and then within each trial, appropriate images were selected and displayed as each new head position was registered by the Flock of Birds device. This did not add any appreciable sensorimotor time lag to the stimulus display, and the resultant image motion appeared smooth and seamless. Different sets of randomized images were generated for each stimulus condition, observer, and testing session so that the same set of images was never shown twice.

Results

Depth ordering performance with this more precise perspective rendering, which not only captured the

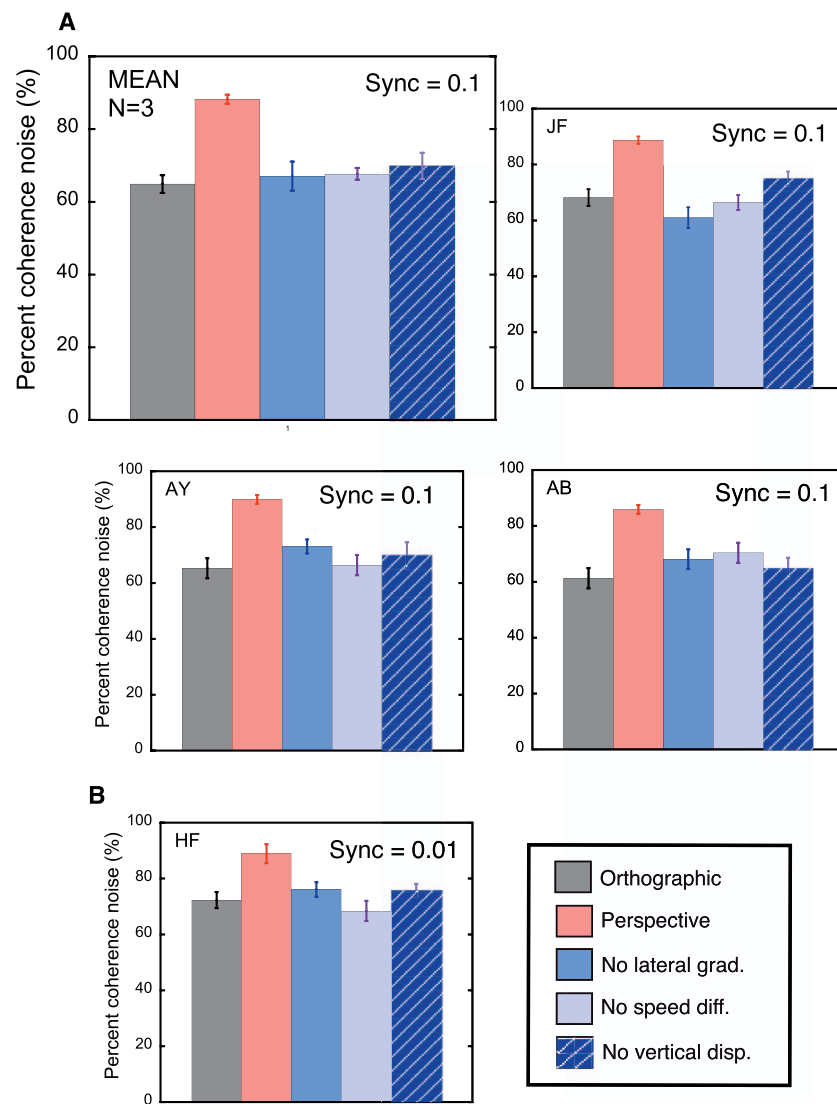


Figure 8. Depth coherence thresholds with the random dot stimulus for orthographic and perspective rendering as well as the cases in which each of the three perspective cues have been selectively removed: lateral gradients in speed, speed differences between near and far surfaces, and vertical displacements. Three observers were tested at a syncing gain of 0.1, and one observer (HF) was tested at a syncing gain of 0.01. (A) Coherence thresholds for the mean of the three observers (large graph) and for individual observers (small graphs), all at the syncing gain of 0.1. (B) Thresholds for an individual observer (HF) at the syncing gain of 0.01. Depth ordering performance was better for perspective than orthographic rendering, and removal of any of the three cues lowered depth ordering performance.

optic flow of texture elements, but also rendered the perspective distortions of each individual Gabor element, is shown in Figure 9 (green). The results indicated that this more precise perspective rendering yielded an improvement in depth ordering performance over the simple dynamic perspective rendering (Figure 9, red) that we had used in Experiment 1. A repeated-measures ANOVA confirmed this with a significant main effect of rendering, $F(1, 6) = 22.0$, $p = 0.043$, and syncing gain, $F(3, 6) = 13.5$, $p = 0.004$; the rendering condition by syncing gain interaction was not significant, $F(3, 6) = 1.69$, $p = 0.27$.

The depth ordering performance with this more precise perspective rendering was comparable to the performance we had observed earlier with the perspective rendered random dots (Figure 9, black). A repeated-measures ANOVA confirmed that the main effect of texture type (improved perspective rendering with Gabors vs. perspective rendered dots) was not significant, $F(1, 6) = 2.09$, $p = 0.29$; the main effect of syncing gain was significant, $F(3, 6) = 24.2$, $p = 0.001$; and the interaction was not significant, $F(3, 6) = 1.33$, $p = 0.35$. Thus, this newer method of rendering with the Gabor micropattern texture improved depth ordering performance to the level that was obtained earlier with

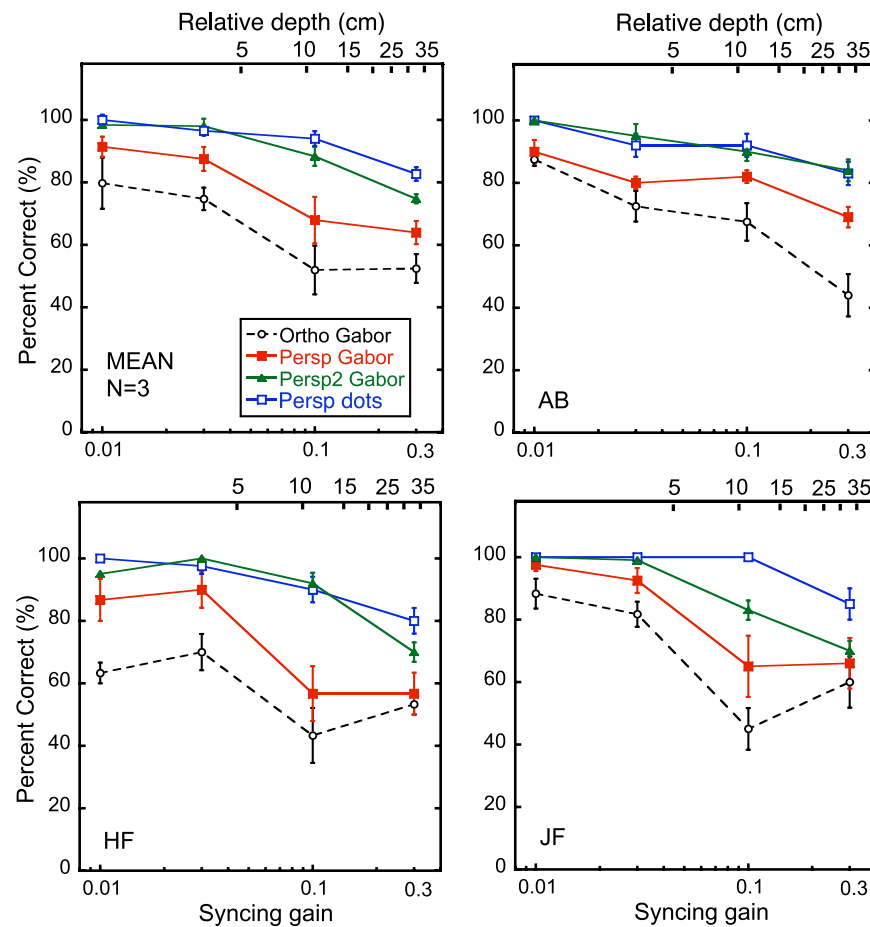


Figure 9. Depth ordering performance with dynamic perspective rendering and Gabor micropatterns (“Persp Gabor,” red, same as in Experiment 1) and a more precise version, which also rendered individual texture elements with appropriate perspective transformations (“Persp2 Gabor,” green). Depth ordering performance for Gabors with orthographic rendering (black) and dots with perspective rendering (blue) is also shown for comparison. Depth perception was better with this more precise perspective rendering and became comparable to that with perspective rendered random dots.

random dots. This result supports the idea that the earlier reduced performance with Gabor micropattern textures was due to a “flatness” cue conflict.

Discussion

The results showed that depth ordering performance was better with perspective than with orthographic rendering for both the random dot and Gabor micropattern textures. Furthermore, each of the three perspective cues (vertical displacements, lateral gradients of speed across the corrugations, and speed differences between near and far surfaces) contributed to improving depth ordering performance. The lateral gradients of speed included gradients in both horizontal and vertical speed components, both of which turned out to be important. Although we might have imagined that just one of these cues could have accounted for the difference between orthographic and perspective ren-

dering, we found that all of these cues were important but to different extents across observers. The results using coherence noise thresholds also supported the same conclusions. Our results do not allow us to distinguish between different models of cue combination, for example, strong or weak fusion (Johnston, Cumming, & Landy, 1994; Landy & Kojima, 2001; Landy, Maloney, Johnston, & Young, 1995; Parker, Cumming, Johnston, Hurlbert, & Gazzaniga, 1995; Young, Landy, & Maloney, 1993). One possible interpretation is that removal of a cue resulted in cue conflicts with a consequent impairment of depth perception. Usually perspective cues occur together, and the removal of one cue at a time provides conflicting information, which could be interpreted in different ways. The implication is that previous studies using orthographic rendering (e.g., Bradshaw et al., 2006; Nawrot & Joyce, 2006; Ono & Ujike, 2005; Rogers & Rogers, 1992; Yoonessi & Baker, 2011) may have underestimated how well we perceive depth from

motion parallax. The results also suggest that models of optic flow in motion parallax (e.g., Fernandez & Farell, 2008; Koenderink, 1986; Longuet-Higgins & Prazdny, 1980; Mayhew & Longuet-Higgins, 1982) should incorporate the effects of all of these dynamic perspective cues. Some observers also reported that the displays with orthographic rendering were somewhat nonrigid and appeared to rotate although this did not occur with perspective rendering. In general, these observations suggest that perspective rendering comes closer to capturing the depth cues that are most relevant to the visual system.

In the first experiment, depth ordering performance was somewhat lower overall for the Gabor micropattern textures than the random dots even when perspective rendering was used. This result was somewhat unexpected because these were more naturalistic stimuli. In Experiment 4, we explored one possible reason for this difference, which is the presence of a potential “flatness cue” conflict arising from the lack of small perspective transformations of individual elements in size, aspect ratio, and orientation as they should appear from different head positions (Kaiser et al., 1995). With the more precise rendering that provided these transformations in Experiment 4, depth perception with the Gabor micropatterns became comparable to that with the random dots. It is important to note that we still kept the same sizes of Gabor micropatterns when they appeared in different depth planes rather than systematically varying their sizes with distance as might appear in a naturalistic setting. Presumably, we would obtain even better depth ordering performance if we included this extra depth cue (Howard & Rogers, 2012). The present work makes an important contribution in being one of only a few studies to investigate the effect of texture image characteristics in depth from motion parallax, which has important theoretical and practical applications, such as in virtual reality and computer games.

Relation to previous studies

The results can be compared with previous studies of perspective cues on depth from motion parallax or structure from motion (i.e., with a stationary observer). Rogers and Rogers (1992) manipulated perspective information in three different structure from motion conditions within a 17° aperture in which the observer was stationary: (a) The display monitor was rotated (perspective, structure from motion), (b) the stimulus was altered so that one of the vertical edges was shorter than the other (simulated perspective with vertical displacements only), and (c) both the vertical edges and the overall width of the stimulus were altered (simulated perspective with vertical displacements plus width

change). The presence of the perspective information in which all cues were present resulted in the best and most unambiguous perception of depth, and depth was more ambiguous in the conditions in which only partial perspective cues were presented. They also compared these results to motion parallax with orthographic rendering, in which nonvisual information was present because of the observer’s movement. Depth perception with this motion parallax stimulus was almost as good as that for the perspective rendered structure from motion stimulus. Thus, either visual perspective cues (structure from motion) or nonvisual information (motion parallax) could improve depth perception and make it less ambiguous. However, they did not include a motion parallax condition with perspective cues as in the present study.

In a more recent motion parallax study, one perspective cue, vertical displacements, was shown to improve depth perception and make it less ambiguous (Rogers, 2012). However, another study failed to find any effect of vertical displacements in motion parallax although this was the only perspective cue that was used in isolation, so this may have been less than optimal and does not correspond to any condition used in the present study (George, Johnson, & Nawrot, 2013). This result might have been due to their relatively small stimulus size ($8.9^\circ \times 8.9^\circ$). The effects of the other dynamic perspective cues were not tested (George et al., 2013).

The results of the present study are also consistent with neurophysiological studies because neurons in area MT involved in processing depth from motion parallax are sensitive to dynamic perspective rendering cues (Kim, Angelaki, & DeAngelis, 2015). Because the processing of dynamic perspective cues requires mechanisms that integrate motion signals over large regions of the visual field, these cues are likely not analyzed in the middle temporal area (MT) but in brain areas that process large-field motion, such as the caudal intraparietal sulcus or ventral intraparietal area of the parietal lobe, or medial superior temporal area, and the signals are fed back to MT (Schindler & Bartels, 2016; Sereno, Trinath, Augath, & Logothetis, 2002).

Comparison to stereopsis

Although stereoscopic depth arises from the slightly different vantage points of the two eyes, depth from motion parallax arises from the change in vantage from one eye position to another. Hence, the dynamic perspective cues in motion parallax are comparable to cues that are also present in stereoscopic depth perception (Bradshaw et al., 1996; Mayhew & Longuet-Higgins, 1982; Read & Cumming, 2006; Rogers & Bradshaw, 1993). In stereopsis, retinal disparities

decrease with viewing distance in an analogous manner to the variation of object speed with distance in motion parallax (Ono et al., 1986).

Vertical displacements in motion parallax are analogous to vertical disparities in binocular vision, which occur because any feature not in the median plane will be closer to one eye than the other and consequently will project to different vertical positions and will be of different vertical sizes in the two eyes (e.g., Koenderink, 1986; Mayhew & Longuet-Higgins, 1982; Read & Cumming, 2006). In addition, there are also lateral gradients of disparities across a binocularly viewed image (e.g., Backus et al., 1999; Rogers & Bradshaw, 1993), which are important in the perception of surface slant (Backus et al., 1999). The presence of all of these perspective cues in stereopsis has been shown to be effective in scaling depth from horizontal disparities and influencing the perception of shape (Bradshaw et al., 1996; Rogers & Bradshaw, 1993).

Conclusions

Depth perception from motion parallax is enhanced by each of the three dynamic perspective cues. Previous studies using orthographic rendering may have underestimated the depth perception that is possible, and this should motivate the use of more naturalistic perspective rendering in the future. This suggests that an important goal would be to revise models of motion parallax to incorporate the effects of all of these perspective cues. Depth perception with more naturalistic texture stimuli was comparable to that with random dots, provided that the textures were rendered with correct perspective transformations of the micropattern shapes. This has the practical implication that the use of simplified rendering may have a detrimental effect on depth perception and that there are advantages to using the more precisely calculated rendering.

Keywords: *optic flow, cue combination, human psychophysics, structure from motion, motion perception*

Acknowledgments

This work was funded by a grant from the Natural Sciences and Engineering Research Council of Canada (OPG0001978) to C. L. Baker, Jr. We thank our observers for their participation. A preliminary report of these findings was presented at the Vision Sciences Society conference (Buckthought, Yoonessi, & Baker, 2014).

Commercial relationships: none.

Corresponding author: Athena Buckthought.

Email: athenabuck1@gmail.com.

Address: Psychology Department, Roanoke College, Salem, VA, USA.

References

- Adeyemo, B., & Angelaki, D. E. (2005). Similar kinematic properties for ocular following and smooth pursuit eye movements. *Journal of Neurophysiology*, 93(3), 1710–1717. [PubMed]
- Backus, B. T., Banks, M. S., van Ee, R., & Crowell, J. A. (1999). Horizontal and vertical disparity, eye position, and stereoscopic slant perception. *Vision Research*, 39(6), 1143–1170. [PubMed]
- Baker, D. H., & Graf, E. W. (2009). Natural images dominate in binocular rivalry. *Proceedings of the National Academy of Sciences, USA*, 106(13), 5436–5441. [PubMed]
- Bradshaw, M. F., Glennerster, A., & Rogers, B. J. (1996). The effect of display size on disparity scaling from differential perspective and vergence cues. *Vision Research*, 36(9), 1255–1264. [PubMed]
- Bradshaw, M. F., Hibbard, P. B., Parton, A. D., Rose, D., & Langley, K. (2006). Surface orientation, modulation frequency and the detection and perception of depth defined by binocular disparity and motion parallax. *Vision Research*, 46(17), 2636–2644. [PubMed]
- Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, 10(4), 433–436. [PubMed]
- Buckthought, A., Yoonessi, A., & Baker, C. L. (2014). Dynamic perspective cues enhance depth from motion parallax. *Journal of Vision*, 14(10): 734, doi: 10.1167/14.10.734. [Abstract]
- Cooper, E. A., & Norcia, A. M. (2014). Perceived depth in natural images reflects encoding of low-level luminance statistics. *Journal of Neuroscience*, 34(35), 11761–11768. [PubMed]
- Fernandez, J. M., & Farell, B. (2008). A neural model for the integration of stereopsis and motion parallax in structure-from-motion. *Neurocomputing*, 71(7–9), 1629–1641. [PubMed]
- Field, D. J. (1987). Relations between the statistics of natural images and the response properties of cortical cells. *Journal of the Optical Society of America, A*, 4, 2379–2394. [PubMed]
- George, J. M., Johnson, J. I., & Nawrot, M. (2013). In pursuit of perspective: Does vertical perspective disambiguate depth from motion parallax? *Perception*, 42(6), 631–641. [PubMed]
- Gibson, E. J., Gibson, J. J., Smith, O. W., & Flock, H.

- (1959). Motion parallax as a determinant of perceived depth. *Journal of Experimental Psychology*, 58(1), 40–51. [PubMed]
- Hansen, B. C., & Hess, R. F. (2012). On the effectiveness of noise masks: Naturalistic vs. unnaturalistic image statistics. *Vision Research*, 60, 101–113. [PubMed]
- Helmholtz, H. V. (1925). Physiological optics. *Optical Society of America*, 3, 318.
- Howard, I. P., & Rogers, B. J. (2012). *Perceiving in depth* (Vol. 2). Oxford, UK: Oxford University Press.
- Johnston, E. B., Cumming, B. G., & Landy, M. S. (1994). Integration of stereopsis and motion shape cues. *Vision Research*, 34(17), 2259–2275. [PubMed]
- Kaiser, M. K., Montegut, M. J., & Proffitt, D. R. (1995). Rotational and translational components of motion parallax: Observers' sensitivity and implications for three-dimensional computer graphics. *Journal of Experimental Psychology: Applied*, 1(4), 321–331. [PubMed]
- Kim, H. R., Angelaki, D. E., & DeAngelis, G. C. (2015). A novel role for visual perspective cues in the neural computation of depth. *Nature Neuroscience*, 18(1), 129–137. [PubMed]
- Kingdom, F. A., Hayes, A., & Field, D. J. (2001). Sensitivity to contrast histogram differences in synthetic wavelet-textures. *Vision Research*, 41(5), 585–598. [PubMed]
- Koenderink, J. J. (1986). Optic flow. *Vision Research*, 26(1), 161–179. [PubMed]
- Landy, M. S., & Kojima, H. (2001). Ideal cue combination for localizing texture-defined edges. *Journal of the Optical Society of America A*, 18(9), 2307–2320. [PubMed]
- Landy, M. S., Maloney, L. T., Johnston, E. B., & Young, M. (1995). Measurement and modeling of depth cue combination: In defense of weak fusion. *Vision Research*, 35(3), 389–412. [PubMed]
- Lee, Y. L., & Saunders, J. A. (2011). Stereo improves 3D shape discrimination even when rich monocular shape cues are available. *Journal of Vision*, 11(9):6, 1–12, doi:10.1167/11.9.6. [PubMed] [Article]
- Longuet-Higgins, H., & Prazdny, K. (1980). The interpretation of a moving retinal image. *Proceedings of the Royal Society of London B: Biological Sciences*, 208(1173), 385–397. [PubMed]
- Mayhew, J. E. W., & Longuet-Higgins, H. C. (1982). A computational model of binocular depth perception. *Nature*, 297(5865), 376–378. [PubMed]
- Miles, F. A. (1998). The neural processing of 3-D information: Evidence from eye movements. *European Journal of Neuroscience*, 10(3), 811–822. [PubMed]
- Nawrot, M., & Joyce, L. (2006). The pursuit theory of motion parallax. *Vision Research*, 46(28), 4709–4725. [PubMed]
- Olshausen, B. A., & Field, D. J. (1997). Sparse coding with an overcomplete basis set: A strategy employed by V1? *Vision Research*, 37(23), 3311–3325. [PubMed]
- Ono, H., & Ujike, H. (2005). Motion parallax driven by head movements: Conditions for visual stability, perceived depth, and perceived concomitant motion. *Perception*, 34, 477–490. [PubMed]
- Ono, M. E., Rivest, J., & Ono, H. (1986). Depth perception as a function of motion parallax and absolute-distance information. *Journal of Experimental Psychology: Human Perception and Performance*, 12(3), 331–337. [PubMed]
- Parker, A. J., Cumming, B. G., Johnston, E. B., Hurlbert, A. C., & Gazzaniga, M. S. (1995). Multiple cues for three-dimensional shape. In M. S. Gazzaniga (Ed.), *The cognitive neurosciences* (pp. 351–364). Cambridge, MA: MIT Press.
- Quaia, C., Sheliga, B. M., Fitzgibbon, E. J., & Optican, L. M. (2012). Ocular following in humans: Spatial properties. *Journal of Vision*, 12(4):13, 1–29, doi:10.1167/12.4.13. [PubMed] [Article]
- Read, J. C., & Cumming, B. G. (2006). Does depth perception require vertical-disparity detectors? *Journal of Vision*, 6(12):1, 1323–1355, doi:10.1167/6.12.1. [PubMed] [Article]
- Rogers, B. (2012). Motion parallax, pursuit eye movements and the assumption of stationarity. *Journal of Vision*, 12(9): 251, doi:10.1167/12.9.251. [Abstract]
- Rogers, B. J., & Bradshaw, M. F. (1993). Vertical disparities, differential perspective and binocular stereopsis. *Nature*, 361(6409), 253–255. [PubMed]
- Rogers, S., & Rogers, B. J. (1992). Visual and nonvisual information disambiguate surfaces specified by motion parallax. *Perception & Psychophysics*, 52(4), 446–452. [PubMed]
- Saunders, J. A., & Chen, Z. (2015). Perceptual biases and cue weighting in perception of 3D slant from texture and stereo information. *Journal of Vision*, 15(2):14, 1–24, doi:10.1167/15.2.14. [PubMed] [Article]
- Schindler, A., & Bartels, A. (2016). Motion parallax links visual motion areas and scene regions. *Neuroimage*, 125, 803–812. [PubMed]
- Sereno, M. E., Trinath, T., Augath, M., & Logothetis, N. K. (2002). Three dimensional shape representa-

- tion in monkey cortex. *Neuron*, 33(4), 635–652. [PubMed]
- Yoonessi, A., & Baker, C. L., Jr. (2011). Contribution of motion parallax to segmentation and depth perception. *Journal of Vision*, 11(9):13, 1–21, doi: 10.1167/11.9.13. [PubMed] [Article]
- Yoonessi, A., & Baker, C. L., Jr. (2013). Depth perception from dynamic occlusion in motion parallax: Roles of expansion-compression versus accretion-deletion. *Journal of Vision*, 13(12):10, 1–16, doi:10.1167/13.12.10. [PubMed] [Article]
- Young, M. J., Landy, M. S., & Maloney, L. T. (1993). A perturbation analysis of depth perception from combinations of texture and motion cues. *Vision Research*, 33(18), 2685–2696. [PubMed]