

# Contribution of motion parallax to segmentation and depth perception

Ahmad Yoonessi

McGill Vision Research, McGill University, Montreal,  
QC, Canada



Curtis L. Baker Jr.

McGill Vision Research, McGill University, Montreal,  
QC, Canada



Relative image motion resulting from active movement of the observer could potentially serve as a powerful perceptual cue, both for segmentation of object boundaries and for depth perception. To examine the perceptual role of motion parallax from shearing motion, we measured human performance in three psychophysical tasks: segmentation, depth ordering, and depth magnitude estimation. Stimuli consisted of random dot textures that were synchronized to head movement with sine- or square-wave modulation patterns. Segmentation was assessed with a 2AFC orientation judgment of a motion-defined boundary. In the depth-ordering task, observers reported which modulation half-cycle appeared in front of the other. Perceived depth magnitude was matched to that of a 3D rendered image with multiple static cues. The results indicate that head movement might not be important for segmentation, even though it is crucial for obtaining depth from motion parallax—thus, concomitant depth perception does not appear to facilitate segmentation. Our findings suggest that segmentation works best for abrupt, sharply defined motion boundaries, whereas smooth gradients are more powerful for obtaining depth from motion parallax. Thus, motion parallax may contribute in a different manner to segmentation and to depth perception and suggests that their underlying mechanisms might be distinct.

Keywords: 3D surface and shape perception, active vision, motion—3D, segmentation, motion parallax, shape and contour

Citation: Yoonessi, A., & Baker, C. L., Jr. (2011). Contribution of motion parallax to segmentation and depth perception. *Journal of Vision*, 11(9):13, 1–21, <http://www.journalofvision.org/content/11/9/13>, doi:10.1167/11.9.13.

## Introduction

In order to perceive the 3D layout of a scene from 2D retinal images, the visual system exploits a number of cues that provide information about distances of different surfaces from the observer. The visual system is faced with several computational problems: parsing the image into surfaces that belong to different objects (“segmentation”), determining the relative distances of these surfaces (“depth ordering”), and obtaining information about how far apart these surfaces are from one another (“depth magnitude”). Such problems are ill-posed, and therefore, the visual system needs to combine different sources of information to make perceptual inferences. Each of the available depth cues (e.g., stereopsis, shape from shading, interposition, etc.) is limited in its effectiveness and working range (Cutting & Vishton, 1995). One of the more powerful sources of information is motion parallax, i.e., relative retinal image motion resulting from active movement<sup>1</sup> of the observer (Gibson, Gibson, Smith, & Flock, 1959; Helmholtz, 1925). Its effectiveness, particularly for depth perception, lies in the sensorimotor relationship between observer movement and consequent retinal image motion, which is dependent on observer movement, scene layout, and point of fixation. This information can be used to segment object boundaries (i.e., figure–ground segmentation) and

also to perceive depth relationships between different objects and within individual objects. However, the relative contribution of motion parallax to segmentation and to depth perception has not been studied systematically. Furthermore, while the same visual information is simultaneously available for both kinds of percept, it is unclear to what extent they share a common neural mechanism.

In the simplest case of lateral translation, the pattern of optic flow formed from observer movement is depicted in Figure 1. When an observer fixates a point at optical infinity (Figure 1a), different parts of the retinal image will move oppositely to head movement with an inverse relationship to depth. In a more typical case when the fixation point is at an intermediate distance on an object or texture marking (Figure 1b), translational vestibuloocular reflex (TVOR) and visually driven eye movements will compensate for observer movement to keep the object of interest in the fovea. In this situation, a bidirectional optic flow pattern is formed: For surface elements in front of the fixation point, there is an inverse relationship between flow and distance; for surface elements behind the fixation point, there is a proportional relationship. At a more local level of objects within a scene, there are two important special cases of boundaries to consider (Figure 1c). When the boundary is parallel to the observer movement, there will be a shearing motion between the opposite sides of the boundary, and when the boundary is orthogonal to the

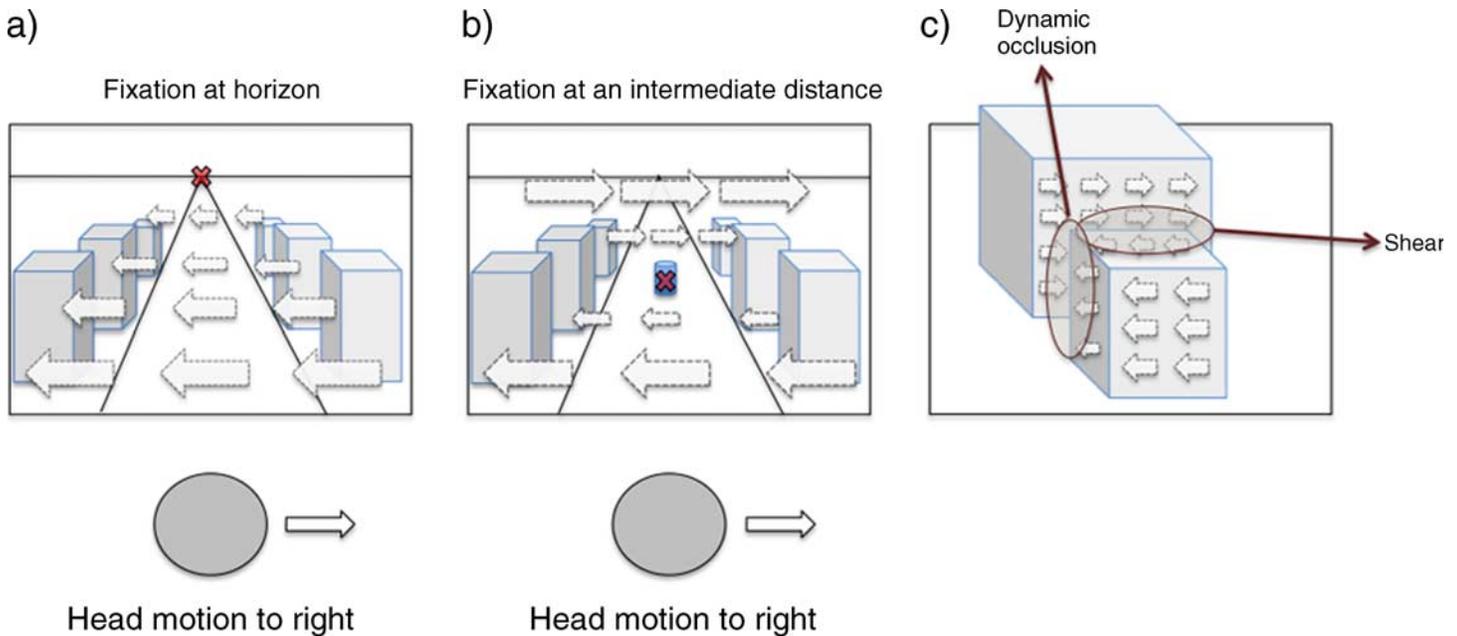


Figure 1. Patterns of retinal image motion created from rightward lateral observer translation. (a) Fixation at horizon—retinal motion will be opposite to head movement with an inverse relationship to depth. (b) Fixation at an intermediate distance—retinal image motion will be opposite to head movement for objects nearer than fixation point and proportional to head movement at farther depths. (c) Pattern of shear and dynamic occlusion at object boundaries—boundaries parallel to direction of observer movement will create shearing motion, while orthogonal boundaries will give rise to expansion–compression and accretion–deletion.

direction of observer movement, the nearer surface will occlude the farther surface without any accompanying shear. However, the occlusion case is more complex since it consists of two phenomena occurring simultaneously: expansion–compression motion and accretion–deletion (Andersen & Braunsten, 1983; Ono, Rogers, Ohmi, & Ono, 1988; Yonas, Craton, & Thompson, 1987). (Note that when observer movement is neither parallel nor orthogonal to an occlusion boundary, the flow pattern will be a mixture of shear and dynamic occlusion.) In this paper, we will defer consideration of occlusion and explore perception of shear boundaries from motion parallax that are free of accretion–deletion cues.

Natural scenes may contain two distinct types of boundaries defined by retinal image motion: abrupt, sharply defined motion boundaries between regions of relatively uniform optic flow, which are typically produced from occlusion boundaries, and gradual gradients of optic flow produced from a receding ground plane or by differences in depth of points on a slanted or curved surface within an object. Gibson et al. distinguished these two cases, referring to them as “two-velocity motion” and “motion perspective” (or “flow field”) motion, respectively (Gibson et al., 1959; Gibson & Carel, 1952). Rogers and Graham (1979) demonstrated that motion parallax using simulated depth profiles with either abrupt transitions between regions of uniform image velocity or with smooth gradients of motion could give compelling depth percepts, but they did not compare quantitative psychophysical performance systematically.

Information in the optic flow from motion parallax can be mathematically expressed as a mixture of a translational component of image motion in the opposite direction to the observer’s movement and a rotational component that provides motion in the same direction as the observer (Longuet-Higgins & Prazdny, 1980; Trucco & Verri, 1998). Simple lateral translation of a camera would create pure translational image motion with no rotational component. However, in the case of a human or animal observer, eye movements (TVOR and visually driven eye movements) occur during observer translation that keep the object of interest in the fovea. These fixational eye movements add a rotational component to the optic flow. Therefore, during a translational movement of the human observer, the optic flow is the vector sum of a translational and a rotational component arising from head and eye movements, respectively (thus, pure translation of the retinal image rarely occurs). At the fixation point, these two components are of equal magnitude but opposite sign and, therefore, cancel each other locally, resulting in net local motion at the fixation point being zero. It has been shown mathematically that the rotational component of the optic flow by itself is not dependent on depth, and therefore, it is impossible to obtain depth information from the rotational component alone (Longuet-Higgins & Prazdny, 1980; Trucco & Verri, 1998). Thus, eye movements alone (without head translation) cannot disambiguate depth from retinal image motion—e.g., eye movements in a stationary observer, or rotational VOR from simple head rotations, do not provide sufficient information for depth perception.

It is desirable to specify motion parallax as a function of the relative depth between near vs. far surfaces, since there may be a maximal limit (Ono, Rivest, & Ono, 1986) analogous to the disparity limit in stereopsis. To produce a stimulus having systematically different values of simulated depth difference, we vary the proportionality factor between head movement and stimulus motion (shear), which we term the “syncing gain” (Ono & Ujike, 2005; Ujike & Ono, 2001). Based on the analogy between motion parallax and stereopsis (both of which entail integrating retinal images from distinct views of the visual scene), the motion parallax stimuli can be expressed in terms of “equivalent disparity” (Rogers & Graham, 1979). However, due to rendering inaccuracies and eye movements, the analogy to disparity becomes progressively more inaccurate at larger values of relative depth (see [Discussion](#) section), so here we will primarily represent our results in terms of syncing gain.

Segmentation from motion, often called “form from motion” or “motion-defined form,” has been widely studied (e.g., Baker & Braddick, 1982; Braddick, 1993; Nakayama, Silverman, Macleod, & Mulligan, 1985; Stoner & Albright, 1993). The majority of these studies found that a small amount of motion (as little as two frames) can be sufficient for observers to segment surfaces. Orientation and vernier discrimination thresholds for luminance- and motion-defined forms are similar, at least at high contrasts, but the sensitivity to motion-defined form deteriorates rapidly when stimulus parameters such as speed and contrast are non-optimal (Regan, 1986, 1989). Evidence from neurophysiology (Allman, Miezen, & McGuinness, 1985; Cao & Schiller, 2003; Frost & Nakayama, 1983; Von Grunau & Frost, 1983), fMRI (Orban et al., 2003; Vanduffel et al., 2001), and psychophysics (Baker & Braddick, 1982; Braddick, 1993; Nakayama et al., 1985; Stoner & Albright, 1993) has suggested specialized mechanisms for motion segmentation in the early visual pathways. Observer movement is a major source of retinal image motion, but segmentation from motion parallax is often ignored, and to our knowledge, it has not been studied psychophysically. In addition, previous studies of segmentation from motion have focused mainly on frontoparallel surfaces with sharp boundaries, whereas most surfaces in the natural world are curved and thus contain smooth gradients of motion.

Motion parallax has historically been associated with depth perception rather than segmentation. Previous experiments regarding depth from motion parallax have often reported inconsistent results, i.e., motion parallax was deemed to be ineffective (Epstein & Park, 1964; Farber & McConkie, 1979), effective but not dependent on head/eye movement (Braunstein & Tittle, 1988), dependent on head movement alone (Rogers & Graham, 1979), or dependent on eye movement alone (Nawrot, 2003; Nawrot & Joyce, 2006). However, it should be noted that the term “motion parallax” has been used inconsistently, sometimes to describe motion as seen by a

stationary observer (Braunstein & Tittle, 1988), which is now usually called structure from motion or kinetic depth, or to denote image motion resulting from observer movement (Rogers & Graham, 1979), which we call motion parallax in this document.

To examine motion parallax generated from an active observer, the stimulus must be synchronized with observer movement. This requires measurement of position and/or orientation and a fast computing system to provide real-time updating of the stimulus in proportion to the head movement. Such a system to study shear-based motion parallax was first developed by Rogers and Graham (1979) using random dot patterns whose motions were modulated by analog measurements of head position to produce periodic one-dimensional profiles of relative image motion. This profile or “envelope” was a 1D waveform, which determined the amount of dot motion relative to the other dots on every horizontal line on the screen. Here, we implement such a setup digitally, using an accurate head-free position/orientation tracking system connected to hardware capable of creating visual stimuli in real time ([Figure 2](#)). The most common approach to head tracking for motion parallax has utilized a chin rest on a linear rail, in some cases with a metronome to encourage regular, constant-speed movement (Ono et al., 1986; Rogers & Graham, 1979). Such a setup uncomfortably constrains the observer’s movement and is unnatural compared to real-life conditions. For example, experiments designed to encourage constant-speed head movement, such as those using computer-driven rails (Ono & Ujike, 2005; Ujike & Ono, 2001), would be lacking in acceleration, which is an especially potent stimulus for vestibular sensors of linear motion (Angelaki & Hess, 2005). Therefore, we employ a setup in which observers can move their head freely in a relatively unconstrained manner more similar to natural head movements and incorporating more naturalistic acceleration components.

The optic flow resulting from motion parallax has long been thought to provide rich information for depth perception and perceptual segmentation (Gibson et al., 1959; Helmholtz, 1925), and it is an appealing idea that there might be common neural substrates for encoding the similar optic flow information underlying these perceptual abilities (Allman et al., 1985; Nakayama & Loomis, 1974). However, while the optic flow cues for segmentation and depth arise from similar ecological circumstances, they could differ substantially in the nature of their low-level encoding: For example, boundary segmentation might work best for abrupt, sharply defined boundaries between regions of uniform optic flow, as produced by object occlusion, whereas theoretical analyses of depth perception (Longuet-Higgins & Prazdny, 1980; Ullman, 1979) suggest that it would profit better from more gradual gradients of optic flow, which are more rich in deformation components of image motion (Koenderink & Van Doorn, 1975). In general, it is unclear whether segmentation and depth behave similarly across stimulus

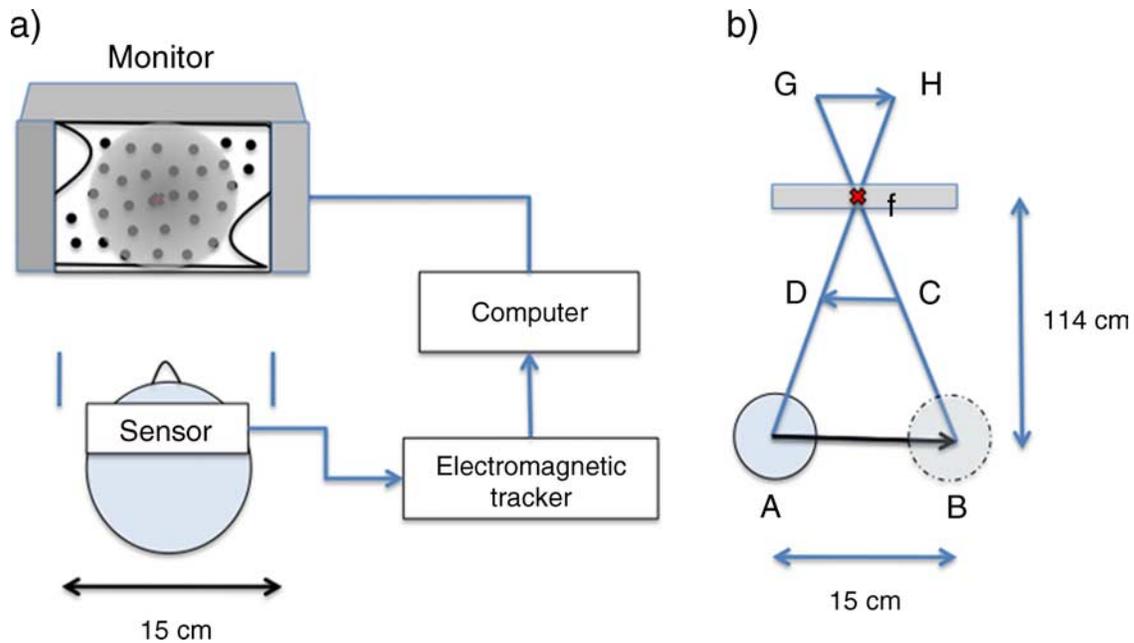


Figure 2. (a) Schematic diagram of the system used to measure head position and synchronize visual stimulus to head movement. Observers moved their head freely within a 15-cm span between two vertical bars acting as guides for the excursion. An electromagnetic sensor placed on the observer's forehead registered the head position/orientation. Stimulus motions were updated in real time, in synchrony with head movement data, without noticeable latency. (b) Geometry defining the virtual depth in motion parallax with an intermediate-distance fixation point,  $f$ . When the observer performs a lateral translation ( $A \rightarrow B$ ) while fixating on the monitor screen, the virtual stimulus depth is proportional to the ratio of stimulus motion to head movement ("syncing gain") =  $CD/AB$ . The arrows ( $C \rightarrow D$ ,  $G \rightarrow H$ ) indicate the projected motion of the object on the fixated monitor plane but is not identical to the retinal image motion due to rendering inaccuracies. Note that syncing gains higher than unity will produce non-ecological conditions.

conditions, suggesting a common neural substrate, or if they behave differently, suggesting distinct mechanisms. To address this question, here, we assess psychophysical performance on depth and segmentation tasks using a common type of stimulus display and head movement-monitoring system.

Signals from the vestibular system and from the visual/proprioceptive consequences of eye movements provide important sources of information about body movement. It is clear that such "extraretinal" information arising from self-movement can potentially disambiguate the depth sign (Naji & Freeman, 2004; Nawrot & Joyce, 2006; Rogers & Graham, 1979; Wexler, Panerai, Lamouret, & Droulez, 2001; Wexler & Van Boxtel, 2005), but its role in segmentation is unclear. Therefore, it is reasonable to ask how and to what extent this extraretinal information is used for segmentation from motion. To address these questions, in this study, we employ an additional "playback" condition (Nadler, Angelaki, & DeAngelis, 2008; Nadler, Nawrot, Angelaki, & DeAngelis, 2009; Wexler et al., 2001) in which identical visual information is provided when observers are stationary.

In this study, we explore the situation depicted in Figure 1b, i.e., shear-based motion parallax, using a bidirectional image motion stimulus with a central fixation point. Our pilot experiments as well as previous studies (Baker & Braddick, 1982; Ono et al., 1986; Rogers &

Graham, 1982; Ujike & Ono, 2001) suggested that such a bidirectional optic flow pattern provides better segmentation and depth perception. This arrangement is particularly advantageous in that it allows significantly more realistic rendering of naturalistic retinal image motions from motion parallax (see Discussion section). A defined fixation point not only enables more accurate depth rendering but also reduces potential variability in the retinal image motion across conditions, trials, and observers.

In our first experiment, we employ a psychophysical task in which performance is contingent on motion-based segmentation to make a 2AFC judgment of the orientation of a motion-defined boundary. In order to assess whether active head movement contributes to segmentation, we compare performance during active head movement, with visual stimulus motion synchronized to it, and during passive viewing of stimuli whose motions are a "playback" recreated from the head movement data collected previously in the active condition. Motion-based segmentation can clearly occur in the absence of adequate cues for depth perception, e.g., in "form-from-motion" experiments (Regan, 1986; Sachtler & Zaidi, 1995; Sary, Vogels, & Orban, 1993). However, in the naturally common context of motion parallax, does concomitant depth perception facilitate segmentation? In this experiment, we also compare psychophysical performance for both square- and sine-wave modulating functions, since the accompanying

depth perception from motion parallax might be more robust for a smoothly varying sinusoidal modulating function than for an abruptly changing square-wave profile.

In the second and third experiments, we utilize two psychophysical measures of depth perception, with stimuli constructed in the same way as those used in the segmentation experiment. First, we measure performance in depth ordering from motion parallax, in which observers perform a 2AFC judgment of the perceived relative depth between two surfaces. Previous reports of the importance of head movement in obtaining depth from motion parallax (Ono et al., 1986; Rogers & Graham, 1979, 1982) did not employ quantitative measurements in a well-defined 2AFC task. In addition, we assess the magnitude of perceived depth, as a function of syncing gain, using a task in which observers match the depth seen in each stimulus to one of a series of rendered 3D surfaces. In both sets of measurements, we compare performance for both square- and sine-wave modulating functions, since our preliminary experiments (Yoonessi & Baker, 2009) suggested that the accompanying depth perception from motion parallax might be more robust for a smoothly varying sinusoidal modulating function than for an abruptly changing square-wave profile.

## General materials and methods

In order to achieve quality real-time synchronization of visual stimuli to head position, we employed a digital position measurement system in conjunction with exploiting the graphics card GPU capabilities for drawing. The observers were instructed to freely translate their head laterally back and forth while viewing the stimulus during each trial, traversing a path limited by a pair of vertical bars with a spacing corresponding to a distance of about 15 cm. The head position data for every trial was recorded on hard disk for later analysis. The overall schematic of the system is shown in Figure 2a, and its details will be 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 written with Psychophysics Toolbox (Brainard, 1997) and presented on a CRT monitor (Trinitron A7217A, 1024 × 768 pixels, 75 Hz), which was gamma-corrected with a mean luminance of 40 cd/m<sup>2</sup>. The stimuli were viewed from a distance of 114 cm, with monocular viewing to avoid cue conflict with stereopsis. The stimulus patterns consisted of white (80.31 cd/m<sup>2</sup>) dots on a black (0.07 cd/m<sup>2</sup>) background, with a dot density of 1.04 dots/deg<sup>2</sup>. Each dot was of circular shape,

0.2 deg (2 pixels) in diameter, rendered with high-quality anti-aliasing. The dots' displacements were modulated using sine- or square-wave profiles to create shearing motion patterns (Figure 1b) similar to those described by Rogers and Graham (1979).

To emulate a motion parallax situation and provide potential depth percepts, the dot motions were synchronized to measured changes in head position (see below). On each frame update, the difference between current and previous head position was multiplied by a gain parameter (see below) and applied to the 1D modulation profile to modulate the dot positions and thereby generate a shearing pattern. The accuracy of syncing gain values was verified by measuring the amount of stimulus motion on the screen. In order to obtain good real-time performance, the “dentsync” settings for Psychtoolbox drawing were employed so that the stimulus drawing was synced to the vertical retrace but without pausing execution of the program until the next vertical retrace. This option resulted in smoother real-time performance in exchange for a small jitter in presentation time. We measured the actual presentation time on each trial and verified that, in practice, the variance was negligible. Using these measures, 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 noticeable sensorimotor lags in these experiments.

We employ the ratio between head movement and image motion, which we call “syncing gain,” as the principal parameter that is varied in our experiments. This parameter has an important relationship to the geometry of motion parallax (Figure 2b) and has been employed previously (Ono & Ujike, 2005; Ujike & Ono, 2001). In addition, describing motion parallax in this way has been frequently used in describing optic flow (Longuet-Higgins & Prazdny, 1980) and in computer vision (Trucco & Verri, 1998). Ideally, the syncing gain is linearly proportional to relative depth (Figure 2b), and therefore, our graphs of performance vs. syncing gain will be double-labeled for calculated relative depth. Representing the stimulus as a function of syncing gain is more accurate than describing it in terms of equivalent disparity (Rogers & Graham, 1979), which would be dependent on fixation point and eye movements that can be complex during head movement (see Discussion section).

The spatial frequency of the modulation was 0.1 cpd, which seemed to provide the best depth percepts on our setup, and is close to the optimal value reported by Rogers and Graham (1982). The stimuli were presented within a circular mask of 18 deg of visual angle, which resulted in about 1.5 cycles/image of visible modulation.

A fixation point was presented prior to and throughout each stimulus presentation, at the center of the circular mask. The stimulus pattern was modulated with sine phase modulation, i.e., the transition point was always at the center of the screen. The fixation point was always set at

the transition point (Figure 1b) between the oppositely moving peaks and troughs of the bidirectional modulation waveform. Such dot motion simulated surfaces that were behind (half-cycle moving the same direction as head movement) and in front (half-cycle moving in the opposite direction of head movement) of the monitor screen, respectively. The fixation point therefore served to maintain the same pattern of retinal image motion across conditions, observers, and trials.

## Head movement recording

The head position and orientation (0.5 mm and 0.1 deg resolution, respectively, for 6 DOF) were measured using an electromagnetic position-tracking device (Flock of Birds, Ascension Technologies, VT, USA) with a medium-range transmitter. The sensor was secured on the observer's forehead using an elastic band. The head movement data were 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 above, and the complete 6-DOF position/orientation was recorded to hard disk for subsequent analysis.

The observers were instructed to perform simple back-and-forth lateral head translations, using two vertical bars as guides (Figure 2a) to encourage reproducible extents of excursion. Since the observer's head movement was not physically constrained to a 1D path, as in previous motion parallax experiments (Rogers & Graham, 1979, 1982), there was some potential for serious variance in psychophysical results due to differences in head movement behavior between observers and/or across trials. To assess this question, we quantitatively analyzed the recordings of head movement data obtained during the psychophysical experiments. Analyses of the raw head position signal, average span, velocity, acceleration, and Fourier amplitude spectrum of head movements are described in the [Supplementary materials](#) (Figures S3–S6 and Tables S1 and S2). Observers typically performed lateral head movements in a quite stereotyped manner (e.g., the average velocity across trials for the four observers was  $16.38 \pm 1.68$  cm/s,  $18.51 \pm 1.94$  cm/s,  $15.00 \pm 2.95$  cm/s, and  $16.79 \pm 3.76$  cm/s). This consistency should not be surprising, due to the instructions given to the observers to make lateral translational movements within defined limits, the limited duration of each trial, and the biomechanical constraints of comfortable head movement.

## Observers

Four observers (YA, MA, BC, and BJ) participated in the depth ordering and depth magnitude tasks, and three (YA, MA, and BJ) participated in the segmentation experiments. Two of the observers (MA and BJ) were

naive to the purpose of the experiment and the other two (YA and BC) were the authors. All observers had normal or corrected-to-normal vision. Each observer gave prior consent for participation according to university guidelines.

## Experiment 1: Segmentation

**Experiment 1** is aimed at studying segmentation from motion parallax. We compared performance under two conditions: “head-synced,” in which the stimulus motion is synchronized to the head movement, and “playback,” in which the observer's head is stationary while viewing the same stimulus motions as in the head-synced conditions.

## Materials and methods

To assess segmentation performance, we chose a task that depended on seeing the boundary between adjacent regions of differently moving textures. Observers were instructed to attend to the orientation of the motion-defined boundary intersecting the fixation point, using a square-wave modulation pattern (Figure 3a). On each trial, the 1D modulation pattern was tilted slightly left or right oblique (Figure 3b). Note that in this as well as the following experiments, the depth variations in the modulation pattern were always rendered as parallel to the plane of the screen, and the orientation of the modulation pattern was around a frontoparallel axis from the center of the screen. The observer's task was to press a button to indicate a 2AFC judgment of the orientation of the motion-defined boundary. In order to perform the task correctly, the observer had to be able to segregate the oppositely moving regions of dot textures to see the oriented boundary. If the stimuli were rendered as having zero depth variations, then the task would be impossible. The simulated depth order was randomly selected on each trial. In order to preclude task performance based on motion of any single subregion of the stimulus, the texture dots always moved horizontally.

For each observer, the head-synced condition was first tested—i.e., the stimulus motion was synchronized to the head movement. Then, in the subsequent playback condition, the observer remained stationary while the stimulus motion was controlled by the head position data previously recorded for that observer with head syncing. Therefore, the difference between the two conditions was purely non-visual, since the visual stimulus information was identical between the two conditions. We also made the same measurements using a 1D sine-wave modulation pattern (Figures 4a and 4b). Note that in the playback condition the same visual information was recreated on the screen. Therefore, what was generated was the same

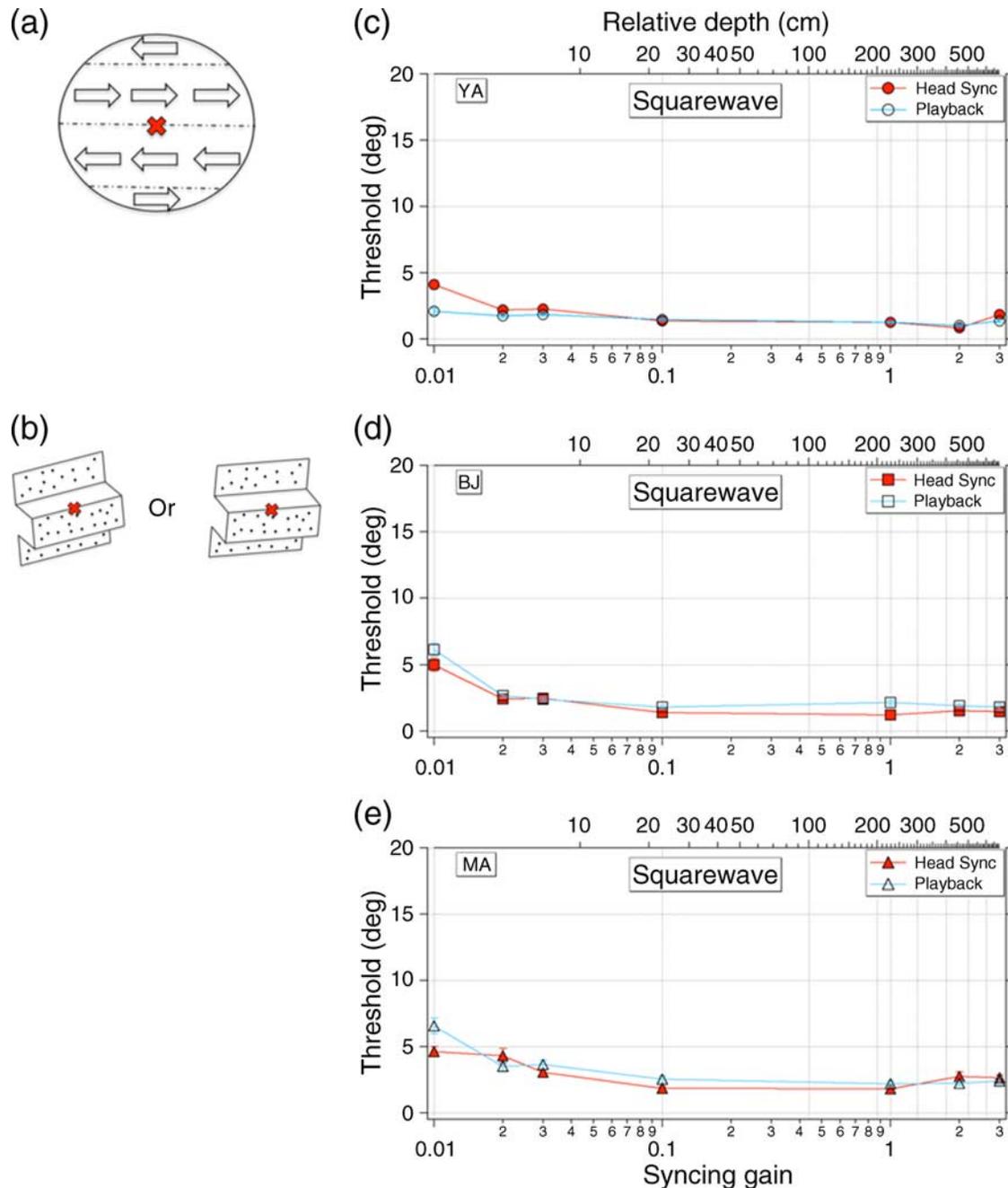


Figure 3. Segmentation performance measured with orientation discrimination. (a) Square-wave modulation pattern of relative motion of random dot textures, with fixation mark (red) at sine phase central boundary. (b) Cartoon depictions of 2AFC orientation judgment, left vs. right oblique. (c–e) Threshold values of envelope orientation discrimination for three observers plotted as functions of syncing gain (ratio between stimulus motion and head movement). Increasing values of syncing gain correspond to greater degrees of stimulated relative depth (top axis). Error bars indicate  $\pm SE$ .

pattern of visual motion on the display—so if the observer maintained fixation as instructed, the retinal image motion should be approximately the same. This stimulus was not inherently ambiguous for the segmentation task, because as noted above the judgment was of the orientation of the rendered depth boundaries, not the plane in which they lie.

Following a button press to initiate each trial, stimuli were presented for 1 s. In pilot experiments, this duration

was demonstrated to be sufficient for observers to comfortably and naturally perform a full cycle of head movement, while being more than sufficient for good task performance. To maintain comparability, the playback conditions were also presented for the same duration.

For each condition and modulation pattern, a series of values of syncing gain were tested using a method of constant stimuli. Five level values of orientation,

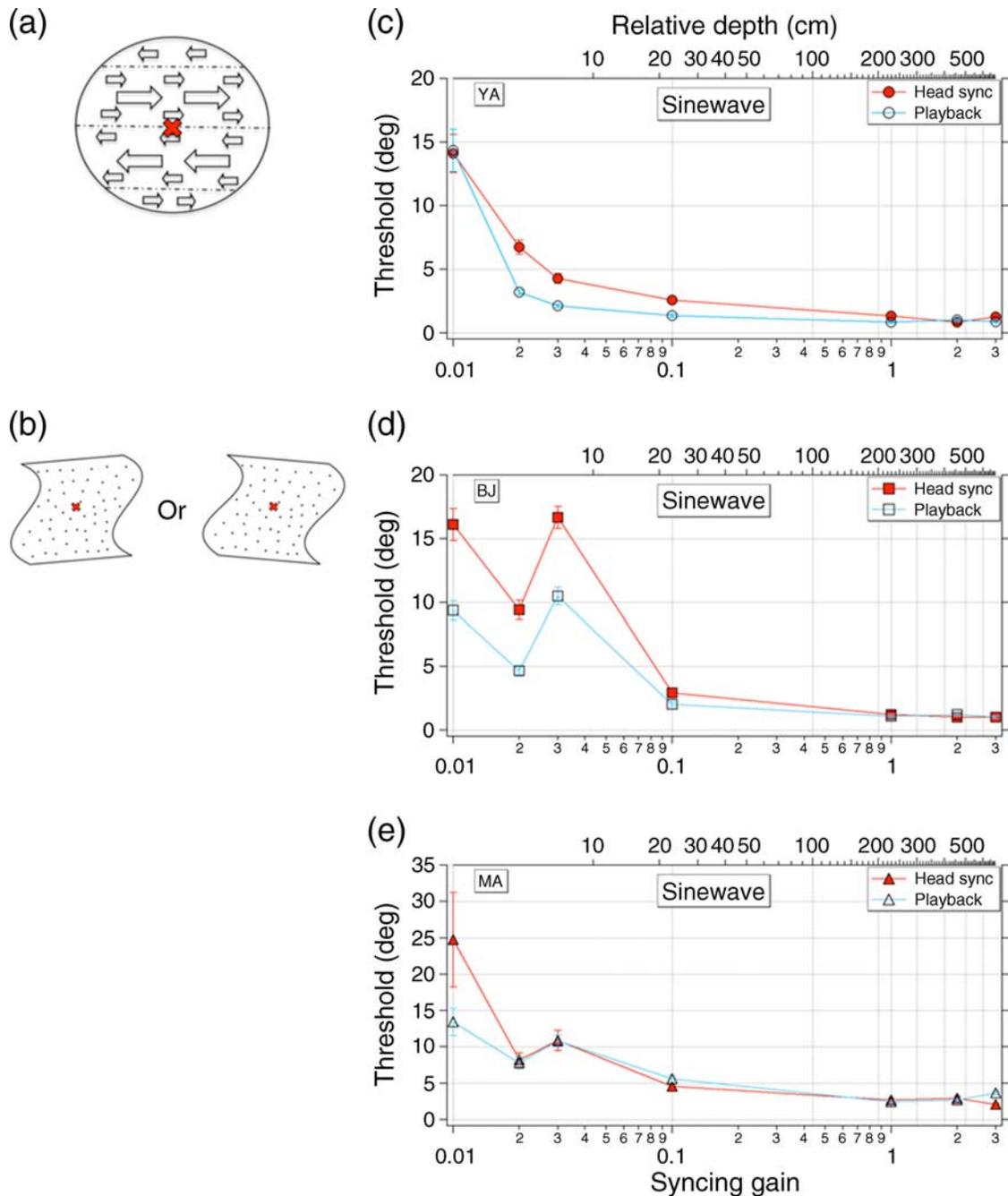


Figure 4. Same as Figure 3 but for sine-wave modulation pattern.

determined from pilot measurements, were tested in pseudorandom order, in blocks containing 20 trials per level value. Blocks were accumulated to provide at least 60 trials per level value. A psychometric function of percent correct vs. boundary orientation was constructed, and the width parameter (standard deviation) of the best-fitting cumulative Gaussian was taken as the JND threshold. Bootstrap estimates of standard deviations of estimated parameters were obtained using Prism software (GraphPad, CA, USA).

## Results

The results for the segmentation task in the case of square-wave modulation are shown in the right-hand panels of Figure 3 for three observers. Each graph shows the orientation discrimination threshold as a function of syncing gain, with the head movement condition in filled symbols and the playback condition in open symbols. The pattern of results is similar across observers: The threshold is very low and approximately constant for syncing gains

down to about 0.03 and becoming somewhat higher for the lower syncing gains. For the most part, thresholds for the head movement and playback conditions are surprisingly similar; however, a two-way independent measures ANOVA test suggests there is a significant difference between head sync and playback thresholds for two of the three observers (YA:  $F(1, 28) = 21.81$ ,  $P < 0.0001$ ; BJ:  $F(1, 28) = 605.24$ ,  $P < 0.0001$ ; MA:  $F(1, 28) = 2.7$ ,  $P = 0.1118$ ).

Figure 4 shows results of the same experiment but using sine-wave modulation, illustrated in the same manner as in Figure 3. The best performance is for higher syncing gains, above about 0.10, corresponding to greater depth differences. At syncing gains below 0.10, thresholds rise substantially. Again the thresholds for the head movement and playback conditions are quite similar for most of the range of syncing gains tested, and a two-way independent measures ANOVA test suggests there is a significant difference between head sync and playback conditions for two of the observers (YA:  $F(1, 28) = 9.48$ ,  $P = 0.0046$ ; BJ:  $F(1, 28) = 80.13$ ,  $P < 0.0001$ ; MA:  $F(1, 28) = 1.89$ ,  $P = 0.1803$ ). At very low syncing gains (below 0.10), head movement seems to interfere with the orientation judgment. This is not what one might naturally expect, since in the head movement condition the brain has access to more information about the structure, and ideally, the extra-retinal cues should aid performance. However, these results suggest that in this task the visual system for the most part ignores the head movement information or is even impeded by it at the lowest syncing gains tested.

Comparing the data plotted in Figures 3 and 4, we see that segmentation performance is somewhat better for square-wave than for sine-wave modulation at syncing gains below 0.10, and a two-way independent measures ANOVA test suggests that head sync and playback data are significantly different between the two modulation waveforms for both head sync and playback conditions (YA, head sync:  $F(1, 28) = 89.99$ ,  $P < 0.0001$ ; YA, playback:  $F(1, 28) = 61.23$ ,  $P < 0.0001$ ; BJ, head sync:  $F(1, 28) = 358.3$ ,  $P < 0.0001$ ; BJ, playback:  $F(1, 28) = 93.37$ ,  $P < 0.0001$ ; MA, head sync:  $F(1, 28) = 26.57$ ,  $P < 0.0001$ ; MA, playback:  $F(1, 28) = 88.43$ ,  $P < 0.0001$ ).

## Experiment 2: Depth ordering

Experiment 2 is aimed at studying the conditions that create optimal depth ordering from motion parallax, using a similar stimulus as in Experiment 1.

### Materials and methods

The setup was similar to that in Experiment 1 except that the modulation pattern was always horizontal, and the observer's task was to judge whether the half-cycle above

or below the fixation point appeared to be in front of the other in a 2AFC task (Figure 5a). The correct condition corresponds to the half-cycle moving opposite to head movement as being in front. Within each block of trials, three values of syncing gain were randomly interleaved, with 20 trials per value—sets of these three values were chosen to be separated by a factor of 10 (e.g., 0.02, 0.20, 2.0). Such blocks were accumulated to provide at least 60 trials per value of syncing gain.

A playback condition was not systematically tested, since without head or eye movements there would be no logical source of information for disambiguating depth and therefore no correct or incorrect answer. The stimulus presentation time for each trial was 5 s, to allow sufficient time for the depth percept to reliably form (Rogers & Graham, 1982).

### Results

The results for the depth-ordering task with square-wave modulation are shown in Figure 5b for two observers and in Figure 5c for another two observers. In each graph, the percentage of trials for correctly reported depth ordering is plotted as a function of syncing gain. The vertical dashed line shows the point beyond which the rendered depth would exceed the viewing distance. The results indicate good performance only at very low syncing gains (0.03 or less), corresponding to very small differences in rendered depth (less than 10 cm). With increase in syncing gain (corresponding to larger relative depths), performance gradually falls toward chance.

The same experiment but using sine-wave modulation yielded results shown in Figure 6b for two observers and Figure 6c for another two observers. Performance is very good across a range of syncing gains up to about 0.30, and then falls off at higher values. Note that in comparison to the square-wave modulation (Figure 5), reliable depth ordering now extends to much higher syncing gains, corresponding to much greater rendered depths—chance performance is only reached at rendered depths corresponding to the viewing distance itself. A two-way independent measures ANOVA test showed a significant difference between sine- and square-wave modulations for syncing gains greater than 0.10 (YA:  $F(1, 24) = 70.90$ ,  $P < 0.0001$ ; BJ:  $F(1, 24) = 6.98$ ,  $P = 0.0143$ ; MA:  $F(1, 24) = 35.89$ ,  $P < 0.0001$ ; BC:  $F(1, 16) = 42.16$ ,  $P < 0.0001$ ). In addition, observers reported seeing more clearly defined depth in the sine-wave modulation condition than when viewing the square-wave modulations.

Since the playback condition would not logically disambiguate depth ordering (at least with eye movements suppressed with a static fixation mark), we did not systematically test it. However, we did make exploratory measurements using the playback condition (i.e., without head movement) and verified that depth-ordering performance was indeed at chance levels. However, somewhat

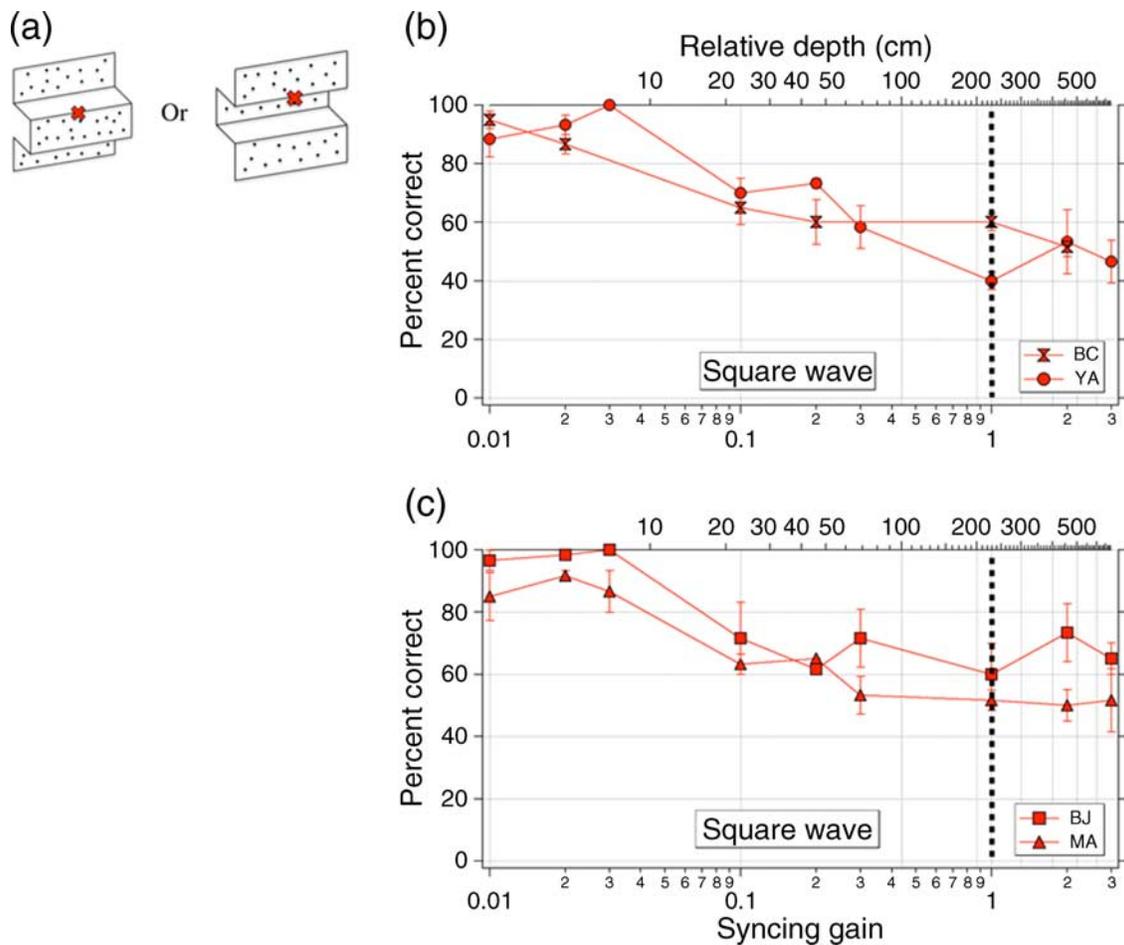


Figure 5. Depth-ordering performance for square-wave modulation patterns. (a) Cartoon depictions of perceived 3D stimuli for 2AFC task to report which surfaces to either side of fixation mark (red) appeared in front of or behind the other. (b) Percent correct performance for two observers, plotted as functions of syncing gain (ratio between head movement and stimulus movement). Dashed vertical line indicates point beyond which simulated relative depth exceeds the physical viewing distance. (c) Same as (b) for another two observers. Error bars indicate  $\pm SE$ .

interestingly, there was a marked bias for judging the lower half of the stimulus as being nearer, even though the surface rendered as nearer was actually in the lower hemifield in only half of the trials (see [Discussion](#) section).

### Experiment 3: Depth magnitude

[Experiment 2](#) demonstrated that observers could use motion parallax information to correctly judge depth order. To measure the importance of motion parallax for depth magnitude (i.e., how far apart perceived surfaces appear to be from one another), we conducted a depth-matching task using the same stimuli.

#### Methods

These measurements were similar to those of [Experiment 2](#), except that observers matched the apparent

magnitude of depth seen in the stimulus to one of a series of 3D rendered surfaces ([Figures 7a](#) and [8a](#)) that were presented on a secondary display screen. These images portrayed a similar 1D pattern of depth modulation but with multiple depth cues (excluding motion parallax) to provide a rich, robust percept of depth. Perspective views of the surfaces were rendered using conventional ray tracing from an oblique viewpoint. The surfaces were covered with noise texture and rendered with ray tracing to have correct shadow and shading cues. The amount of depth was varied parametrically so that the step size between two subsequent indices was constant, and therefore, there was a linear relationship between the amount of rendered depth and the rating. The rendered images were assigned ratings from 0 to 10 ([Figures 7a](#) and [8a](#)).

Observers were allowed unlimited time to look back and forth between the two screens, before then entering a corresponding index to signal their decision as to which surface had the most similar apparent depth. Each observer performed this task only for the values of

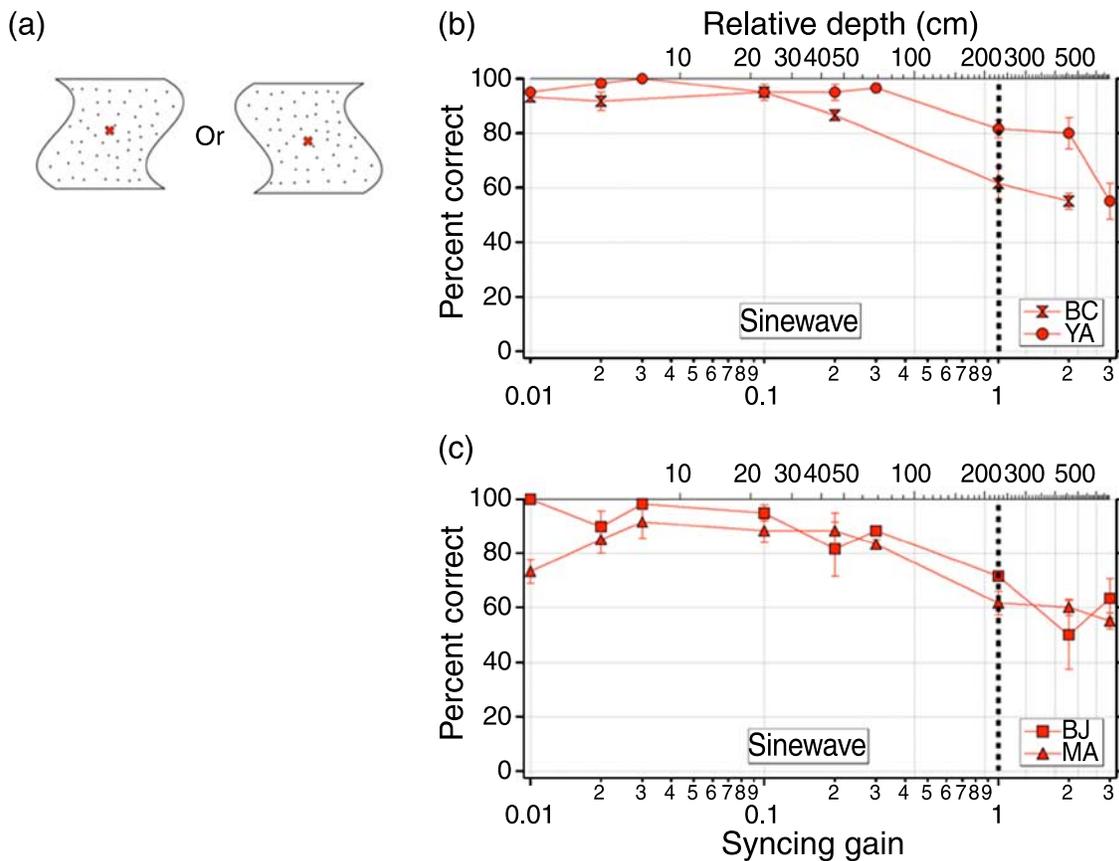


Figure 6. Same as Figure 5 but for sine-wave modulation pattern.

syncing gain that had provided good depth ordering in the previous experiment. Within each block of trials, a series of syncing gains were pseudorandomly interleaved, with five instances of each. Trial blocks were accumulated to provide at least 15 trials per syncing gain.

## Results

The results of this task for square-wave modulation are shown in Figure 7b for three observers, with the index of matched depth graphed as a function of syncing gain. The reported values of perceived depth followed an approximately linear relationship between syncing gain and depth. The black dashed line is a reference to indicate the slope of a linear relationship between the matching index and syncing gain. Note that since the matching images are metrically ambiguous due to the lack of viewpoint information, the important parameter is the slope and not the offset of this relationship. The depth matches for one of the observers (BJ) in Figure 7b did not show a clear relationship, though a substantial amount of depth was reported for all of the conditions tested. However, the matched depths for the other two observers increased systematically with syncing gain in an approximately linear relationship with slopes comparable to that of the reference line, indicating orderly depth magnitude

judgments (reference line: 28.0; YA:  $48.8 \pm 3.9$ ; BJ:  $12.3 \pm 4.1$ ; MA:  $29.3 \pm 4.0$ ).

The results for sine-wave modulation are shown in Figure 8b for three observers, now over a broader range of syncing gains in accord with their better depth-ordering performance for sine-wave patterns (Figure 6). The results show that the reported values of perceived depth followed an approximately linear relationship for syncing gains up to about 0.10 (reference line: 28.0; YA:  $60.8 \pm 3.5$ ; BJ:  $26.9 \pm 2.8$ ; MA:  $26.6 \pm 1.7$ ), with subsequent saturation.

## Discussion

Our results have demonstrated that segmentation and depth from shear motion produced by head movement depend on stimulus parameters quite differently. Segmentation performance was relatively robust across a wide range of syncing gains, whereas good depth perception proved to be more fragile. In addition, square-wave modulations produced better segmentation performance, while sinusoidal patterns yielded good depth only at low syncing gains. These results suggest that segmentation from motion not only does not depend on, but also does not benefit from, accompanying perceived depth in the same visual stimulus.

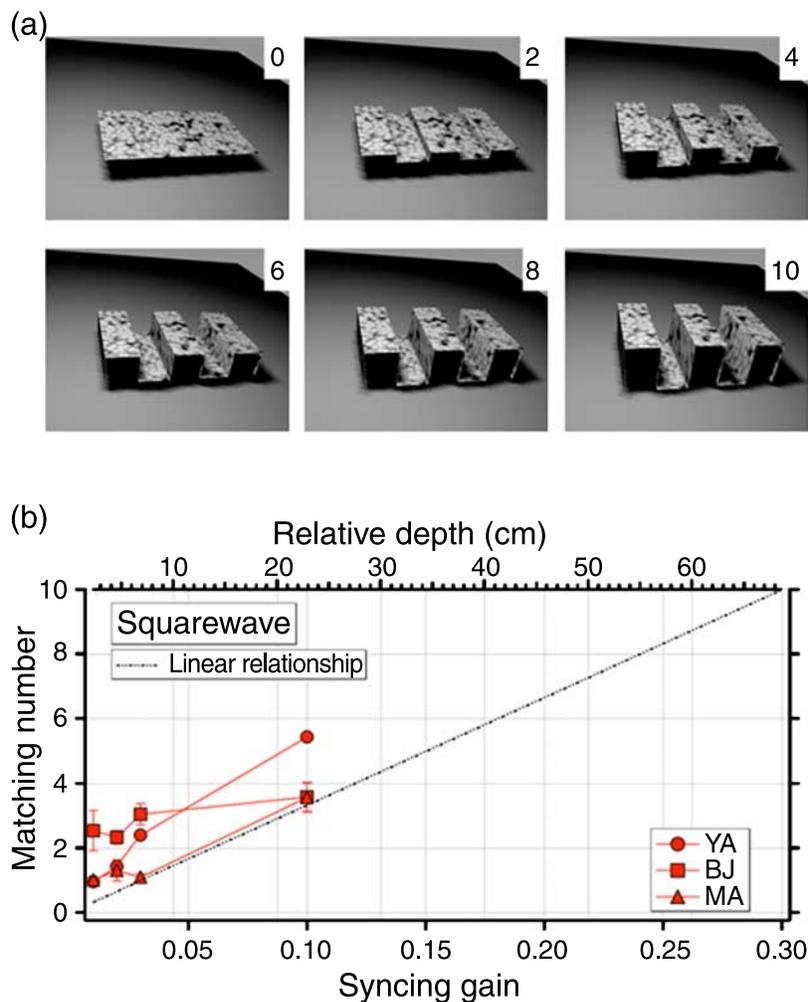


Figure 7. Depth-matching judgments for square-wave modulation. (a) Observers matched perceived depth magnitude in motion parallax stimulus to one of an ordered series of static, multicue renderings. (b) Average depth matches plotted as functions of syncing gain. Black dashed reference line depicts linear relationship between numbers reported and syncing gain, which corresponds to relative depth rendered by motion parallax stimulus. Error bars indicate  $\pm SE$ .

The three tasks chosen in this study were somewhat different in nature—these differences were in part due to practical considerations and in part to the fundamentally different nature of the underlying abilities being assessed. However, as discussed earlier, the motivation for this study was to measure the performance and interaction in these seemingly different abilities that in ecological conditions are based on the same visual information and take place concurrently. The second task (depth ordering) requires a long presentation time (5 s) to provide time for active head movement during each stimulus condition. Therefore, systematic collection of data for full psychometric functions in which the level of difficulty is titrated (e.g., with added noise) would entail excessively long periods of data collection. We chose instead to measure only percentage errors at a fixed level of difficulty for our primary data set and to do secondary measurements of full psychometric functions with added noise at a few fixed values of syncing gain (see [Supplementary materials](#)). The

third task (depth magnitude) is fundamentally different from the others and inherently not amenable to objective 2AFC designs. Such magnitude estimation is a Type 2 judgment (Sperling, Doshier, & Landy, 1990), which does not possess a correct or incorrect answer on every trial (see Kingdom & Prins, 2010). Though the task could be done within a 2AFC design, the nature of the task would still be subjective.

### Contribution of motion parallax to segmentation

We found very good performance in the segmentation task over most or all of the 300-fold range of syncing gain that we explored ([Figures 3 and 4](#)). It would be interesting to find the limits, if an even larger range were examined. Performance must necessarily deteriorate at very low gains due to minimum motion displacement limits (Baker

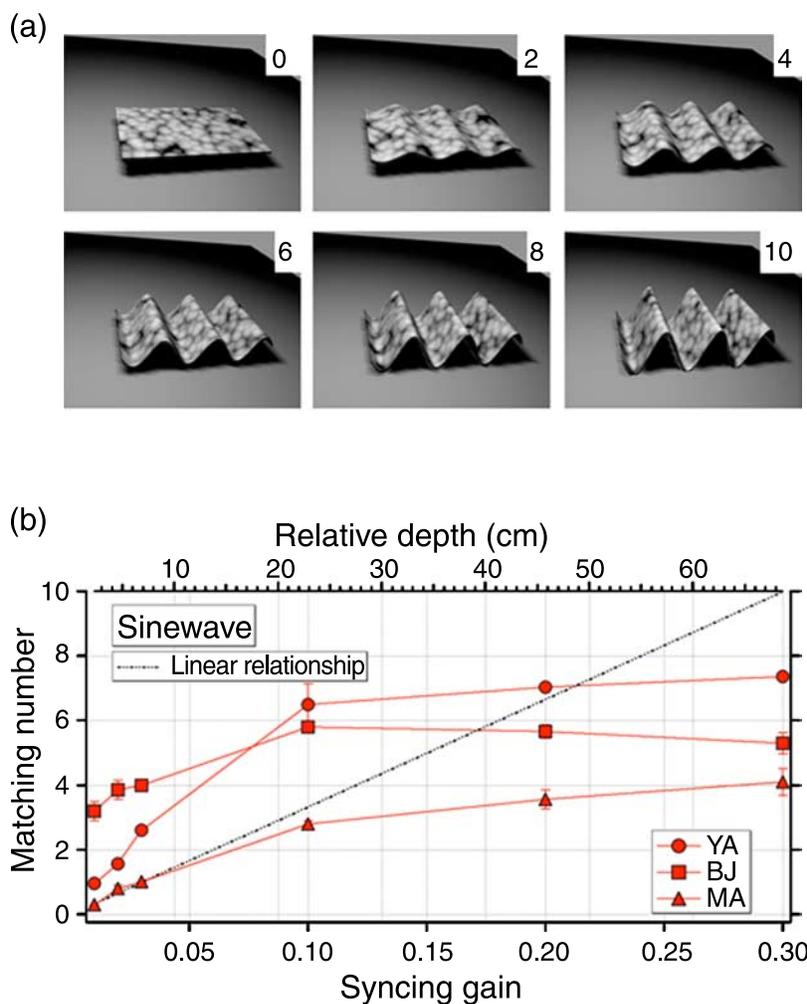


Figure 8. Same as Figure 7 but for sine-wave modulation pattern.

& Braddick, 1985; Westheimer & McKee, 1978). On our setup, it was not feasible to examine syncing gains below 0.01, because quantization error effects due to the pixel resolution resulted in an absence of motion. At very high gains, performance might also degrade due to the resultant large displacements (Baker & Braddick, 1985) or high velocities (Westheimer & McKee, 1975). However, a practical issue is that higher syncing gains start to introduce border artifacts (accretion–deletion), since we keep the texture motion always horizontal. For most of the syncing gain range that we did employ, accretion–deletion artifacts were negligible due to the low dot density and were not detectable even with close scrutiny.

The segmentation task was presented with 1-s duration, much less than that used for the depth task (5 s), which required a longer time for the depth percept to form. In selected examples, we verified that increasing the presentation time from 1 to 5 s did not produce any significant improvement in the segmentation thresholds (see Supplementary materials).

Sine- and square-wave modulations both produced good segmentation performance across a broad range of syncing

gains. However, comparing Figures 3 and 4, they differed substantially at low syncing gains, where square waves yielded better performance. This finding is consistent with psychophysical results using shear motion stimuli with stationary observers (Sachtler & Zaidi, 1995; Watson & Eckert, 1994). This difference between sine- and square-wave envelopes might be explained by higher local motion energy in the case of square-wave modulation—i.e., the lack of effect at higher syncing gain could be due to a minimum required threshold for relative motion energy, beyond which the measured thresholds are limited by other factors. Alternatively, if the segmentation mechanism is relatively more “edge-based” than “region-based,” then the abrupt boundaries of the square-wave condition might afford an important advantage.

It seemed possible that the remarkably similar segmentation results for head movement and playback conditions might be due to a floor or ceiling effect. To assess this possibility, we explored two possible ways of adding coherence noise to degrade performance. One was to randomize the values of 1D motion envelope modulation for a percentage of the dots. We tested this type of noise

on a subset of conditions and found that it had little impact on the depth task, perhaps because the noise appeared as particles superimposed transparently on a 3D structure—in any case, it lowered performance for the segmentation task similarly for head movement and playback conditions. The other noise degradation approach that we tested was application of a random jitter to each dot's motion, proportional to the magnitude of its displacement. This type of noise breaks down the depth percept more effectively, progressively degrading the apparent 3D solidity of the structure. We tested the depth and segmentation tasks on one observer with the addition of this kind of noise (see [Supplementary materials](#)). In the segmentation task, thresholds rose similarly for the head movement and playback conditions, starting at around 60% noise and becoming impossible at about 80% (Figure S1a). In the depth task, performance started to drop with the addition of about 60% noise and went to chance performance with about 80% noise (Figures S1b and S1c). Therefore, these results indicate that the lack of effect of synchronization to head movement over most of the syncing gain range is not a floor or ceiling effect.

The lack of effect of head movement on segmentation seems surprising since the accompanying depth perception (at least at lower syncing gains) should provide an additional cue that could make the task easier, as suggested by cue combination phenomena (e.g., Landy, Maloney, Johnston, & Young, 1995). However, with motion parallax-mediated segmentation, head movement does not augment segmentation and may even degrade performance at lower syncing gains (Figures 3 and 4, filled vs. open lines/symbols). The actual interference by head movement could be due to imperfections in rendering real-world optic flow (see below), causing cue conflicts between eye/head movement and visual information that could hinder performance. One might argue that our psychophysical task for segmentation does not require a 3D judgment, and therefore, it should not be surprising that the depth percept accompanying head movement does not improve performance. However, our goal was not to test a 3D segmentation task but to see whether the availability of extraretinal information such as head movement, and the consequent perception of depth, would *facilitate* segmentation judgments—however, our results show that this is not the case.

## Contribution of motion parallax to depth perception

We were concerned about three potential sources of cue conflict in these experiments, all of which we evaluated in pilot tests in which the display monitor was viewed through a matte black cardboard tunnel. The first concern was that visibility of the surrounding room, in the observer's periphery, could provide a cue to the visual system that rendering inaccuracies (see below), particularly

at higher syncing gains, might demonstrate the stimulus to be “fake,” thereby interfering with performance. Second, the visibility of the display monitor surface could provide a “flatness cue” to the observer, interfering with the percept of depth. However, using the matte black cardboard tunnel to prevent observers from seeing anything besides the display screen did not change the performance on the depth-ordering task, and therefore, we did not use it for systematic data collection. A third potential cue conflict arises because the (stationary) circular mask gives rise to accretion–deletion of the moving texture that it encloses, which is non-ecological for the half-cycle of the stimulus that moves in a way to simulate a surface that is in front of the monitor. That is, the half-cycle whose motion is rendered as nearer than the monitor should logically occlude the mask, which is on the display screen—but instead, the mask occludes the stimulus. To prevent such a problem, we inserted a physical mask halfway along the depth of the black tunnel. However, in pilot experiments, such a modification did not change the pattern of depth-ordering results and, therefore, was not employed in collecting data for our main experiments.

Our observers were able to achieve good depth-ordering performance with shear-defined motion parallax but only at the lower syncing gains (Figures 5 and 6). This failure of depth perception at higher gains cannot be attributed to a simple deterioration of motion detection at higher image motion velocities, since good segmentation performance was obtainable throughout this range (Figures 3 and 4). Note that this failure occurred for syncing gains approaching and exceeding unity, corresponding to rendered relative depths on the order of the simulated viewing distance itself. Such relative depths would be increasingly rare or even impossible in reality, and therefore, it should not be surprising that the visual system might lack mechanisms for representing them. In addition, as will be discussed below, the 3D rendering accuracy progressively deteriorates with syncing gain, which might also contribute to loss of good depth perception.

This critical dependence on syncing gain might go some ways toward reconciling earlier, seemingly contradictory results of Braunstein and Andersen (1981) and Farber and McConkie (1979) vs. those of Rogers and Graham (1979) regarding the quality and nature of perceived depth from motion parallax or kinetic depth, since most of these studies examined only single values or very limited ranges of the amount of relative motion. This range was at very low syncing gains in Rogers and Graham's study and their reports of good depth percepts are in agreement with our results. Ono et al. (1986) varied viewing distance over an approximately 4-fold range in a motion parallax stimulus and found that the quality of depth perception decreased as the magnitude of perceived depth increased, consistent with our results.

Good depth-ordering performance was obtained over a much wider range of syncing gains for sine-wave than for

square-wave modulation patterns (Figure 5 vs. Figure 6). This result suggests a difference between sharp boundaries and surfaces slanted in depth, reminiscent of Gibson's distinction between "two-velocity" vs. "flow field" optic flow (Gibson et al., 1959): "Although no clear line can be drawn between them, the two-velocity type of motion parallax applies to the problem of perceiving a group of objects in otherwise empty space, while the flow-velocity type of motion parallax applies to the perceiving of a background surface such as a wall (or substratum)." The difference between depth perception for sine- and square-wave modulations suggests that shear-based motion parallax is more important for determining depth in surfaces that are slanted or curved in depth, when optic flow contains a gradient of different velocities, rather than for flat surfaces in the observer's frontoparallel plane. In everyday life, this difference could mean that shear-based motion parallax information plays a particularly useful role in activities such as walking, when it is important to correctly estimate the slant of the ground plane.

Sharp boundaries like those in the square-wave modulation patterns occur naturally in motion parallax when a closer opaque surface partially occludes a farther surface, i.e., an object boundary. Such dynamic occlusion boundaries are frequently accompanied by accretion–deletion cues as surface texture is covered or uncovered. In that situation, dynamic occlusion often might dominate shear-based motion parallax information, and it might then make sense for the visual system to utilize the more reliable occlusion cue. It is rare for the visual system to encounter an abrupt motion boundary that is purely shear, without any dynamic occlusion, in ecological conditions. Even though an ideal observer model would predict that the visual system should fully utilize the shear information in the square-wave modulation to get depth across all syncing gains, this evidently does not occur in human observers. Evolutionary constraints might play an important role in the visual system's lack of response to the information available in a stimulus that would rarely occur in natural world. However, in the sine-wave case, the visual system does not have any other source of motion information to disambiguate the depth order. Furthermore, since most of the surfaces in the real world are slanted with regard to the observer, gradients of visual motion are frequent in ecological conditions. Therefore, it might make sense for the visual system to rely relatively more on shear-based motion parallax information to perceive depth in the case of sine-wave modulation.

Relative image motion that is not synchronized to head movement, as in our "playback" condition, contains insufficient information to disambiguate depth. Consequently, unlike in the segmentation experiments, we did not systematically test the playback condition for the depth-ordering task. Nevertheless, we did run exploratory tests of a few playback conditions without head movement—not surprisingly, the results showed chance performance on depth ordering, but interestingly, there was a pronounced

bias for observers to report the envelope half-cycle below the fixation point as the nearer surface. This phenomenon is consistent with much earlier reports of a bias for the lower hemifield to appear nearer (Bourdon, 1902). Such a bias might be explained by the comparative scarcity of motion parallax in the upper hemifield under ecological conditions—objects in the field of view are most often positioned on the ground and will therefore usually appear in the lower hemifield.

For the depth-matching task, it is important to realize that the numerical indices provided for the matching images (Figures 7a and 8a) are not metrically calibrated, because the observer has no information regarding the position of the virtual camera employed for the rendering. Therefore, the vertical offset of the matching data (Figures 7b and 8b) is not meaningful—e.g., the corresponding number for the lowest syncing gain value is ambiguous from the numbered images, and therefore, observers might report different values for this syncing gain. However, the depth difference between two *consecutive* numbered images (i.e., slope) is independent of the viewing distance. Therefore, the important parameter in these data is the slope of the line and not its vertical offset.

The depth-matching results (Figures 7 and 8) show that for the low syncing gains (below 0.1) at which observers report good depth ordering, the perceived depth magnitude increases almost linearly with syncing gain. This suggests that magnitude estimation is most accurate when the depth-ordering task is easy—however, with increased syncing gain, magnitude estimation becomes less accurate.

## Role of head and eye movements

During lateral head translation (or indeed any observer movement), fixational eye movements act reflexively to keep the object of interest in the fovea and thereby minimize motion blur (Angelaki & Hess, 2005; Carpenter, 1988). Without such eye movements, the pattern of optic flow during head translation would almost always be unidirectional, in the opposite direction of head movement, and would cause substantial motion blur and degradation of perception. In the presence of fixational eye movements, this pattern will be bidirectional around the fixation point: Points in front of fixation will move opposite to the head movement and farther points will move in the same direction (Figure 1b). These compensatory eye movements may, in general, be a combination of translational vestibuloocular reflex (TVOR) and visually driven eye movements. Unlike the rotational VOR in which eye rotation ideally has a gain of unity, the amplitude of eye movements in TVOR is dependent on the fixation distance (Angelaki, McHenry, Dickman, Newlands, & Hess, 1999; Busetini, Miles, Schwarz, & Carl, 1994; Ramat & Zee, 2003; Schwarz, Busetini, & Miles, 1989). Therefore, the vestibular system needs access to information regarding the current fixation

distance, in order to generate proportional eye movement amplitude. Several sources such as vergence, accommodation, and vertical disparities have been suggested as cues to fixation distance (see Angelaki & Hess, 2005 for a review). Therefore, when an observer looks at a stationary monitor and performs a lateral head translation, the TVOR eye movement should be proportional to the actual distance of the monitor and not the simulated distance, causing possibly significant inaccuracies or cue conflicts if the observer's gaze were tracking parts of a moving texture. We sought to minimize this problem by utilizing a bidirectional pattern that simulates the optic flow from an object at an intermediate distance and setting the simulated fixation point on the monitor plane (Figure 1b). We placed the fixation point on the zero-crossing of this modulation pattern, which corresponds to being rendered in the same depth plane as the display screen, with the stimulus motion being symmetrical on both sides (nearer and farther).

Nevertheless, the eye movement behavior during our experiments might deviate from that in natural viewing. It has been proposed (Angelaki & Hess, 2005) that the TVOR by itself is not sufficient to keep the object of interest completely in the fovea and that visual signals may often contribute to eliminating the residual retinal slip. However, the conflict between depth cues from accommodation and from visual optic flow could amplify the residual retinal slip of the TVOR. Thus, in the simulated conditions rendered on a monitor, this retinal slip might be greater than in natural viewing—if so then visual inputs to the eye movements might play a more important role and, thereby, require a greater than expected exposure time for depth perception compared to the natural situation.

Recent research has suggested the possible importance of eye movements as an extraretinal source of information to disambiguate depth from the optic flow information during motion parallax (Nadler et al., 2009; Nawrot, 2003). These eye movements function to eliminate the retinal slip from an imperfect TVOR and are distinct from voluntary eye movements (such as pursuit). Furthermore, our experiments would seem to provide a counterexample to the idea that eye movement signals are essential to disambiguate depth in motion parallax. The translational head movement with fixation on a static monitor gives rise to TVOR eye movements. The TVOR is thought to be “calibrated” for viewing distance, using cues such as accommodation or disparity (e.g., Angelaki & Hess, 2005), and so in our experiments with an explicit, static fixation point, the TVOR alone should be able to minimize retinal slip, without any need for postulating a covert eye movement signal to cancel it out. Yet, in our experiments, there is very good depth perception without any pursuit eye movements. It is possible that, in different viewing situations, multiple sources of information, including eye and head movement signals, may contribute to disambiguating depth (Rogers & Rogers, 1992). The roles of eye and head movements in motion parallax need to be clarified by further experiments.

## Effects of limitations in rendering motion parallax stimuli

It is important to note that without eye tracking it is impossible to accurately simulate on a 2D screen the same optic flow pattern created from observer movement in the natural world. In real life, the pattern that is created on the retina will be a function of both translational and rotational optic flow components, and the accompanying fixational eye movements will be a complex mixture of TVOR and visually driven eye movements. In an attempt to simulate this pattern, a one-dimensional modulating function has typically been applied—as done previously and also in our experiment—to head position measurements, which has been of different wave shapes, e.g., sine or square wave. However, this modulating function is not accurately representative of real-world optic flow: In a realistic rendering, the image motion of the more distant surface should be faster than that of the nearer surface (necessitating an asymmetric modulation waveform). In addition, the image speeds of the dots would have to vary systematically with distance from the fixation point and traverse curved paths (necessitating a 2D modulation function). Note that the magnitude of the latter error will increase with syncing gain and, as noted earlier, may be a contributing factor to the decline of depth ordering at higher syncing gains (Figures 5 and 6). Finally, the retinal image motion will be perturbed by inaccuracies in fixational eye movements, which are implicitly assumed to be perfect in our experiments. Therefore, the resultant percept might not be identical to that in a natural situation.

Due to the above-mentioned stimulus imperfections, the observed optic flow during a lateral head translation may not match the optic flow that would be expected from a head movement. This discrepancy might be the reason that observers in motion parallax experiments often report perceiving a solid 3D structure that has translational and/or rotational object motion. In this situation, retinal image motion might be divided into two parts: a part matching the observer's self-motion that corresponds to static 3D structure and a residual part that is interpreted as object motion. In the conditions with high syncing gains, observers reported seeing the surface as rotating, consistent with earlier reports (Rogers & Graham, 1982). It should also be noted that observer motion with respect to a stationary object or surface produces the same optic flow as that from rotation of the object about its central axis. Therefore, any unmatched optic flow can logically be attributed to object rotation. Another reason for this percept might be that parts of the stimulus that are moving in the same direction as the observer are simulating an eye rotation. At higher syncing gains, this motion is bigger than the real rotation that would naturally occur because of head translation, and thus, the residual motion is interpreted as object rotation.

In the conditions with very small syncing gain, nearly all observers perform depth ordering almost perfectly.

However, at least for low syncing gains, most observers reported not consciously seeing any stimulus motion and reported the stimulus as static, but still they could clearly perform the depth-ordering task. This suggests that in those conditions the movement of the dots nearly matches the visual system's "expected" optic flow, and therefore, the surface is seen as stationary. In this condition, the dot motion is attributed only to depth and not to object motion. These observations suggest that it is not necessary to perceive optic flow to obtain depth information from motion parallax, since in these conditions observers do not perceive any motion of the stimulus, but they can still perform the task.

Since naturally occurring optic flow fills the entire visual field, it would seem advantageous to employ a much larger stimulus. However, with increase in the viewable screen size, the discrepancies between simulated optic flow pattern and the optic flow from the natural world become greater. This problem becomes more significant with increasing distance from the fixation point, where rotational eye movements cause the optic flow pattern to be curved.

In pilot experiments, we found that larger dot sizes caused a degradation of performance on depth tasks. This should not be surprising, since larger dot sizes would make cue conflicts with size and perspective cues more apparent, particularly for the sine-wave modulation case. (Note that this cue conflict would similarly be an issue with any micropatterns, not just dots.) Therefore, in the final experiment, the dots were chosen to be very small (i.e., 2 pixels) to minimize such cue conflicts.

## Possible neural mechanisms

Studies of neural correlates of motion parallax have been limited, due to the difficulties of neurophysiological recording in a moving animal. However, segmentation and depth from motion in the case of a stationary animal have been studied extensively and have suggested various candidate neural mechanisms that could be relevant to the analysis of motion parallax information.

Theoretical models for detection of motion discontinuities typically involve an early fine-scale stage consisting of filters or neurons that are selectively responsive to uniform motion within small regions (e.g., moving texture elements or small patches of moving texture) and a later stage that detects discontinuities across a coarser scale in the outputs of the first-stage filters. The early stage filters that detect local motion are typically modeled as quasi-linear spatiotemporal filters that mimic direction-selective neurons in early visual cortex such as motion energy models built from summation of local spatiotemporal filters (Adelson & Bergen, 1985). The late stage operation might compute discontinuities in responses from the early filters—for example, to estimate optic flow velocity vectors of a two-dimensional velocity field (Bülthoff,

Little, & Poggio, 1989; Hildreth, 1984). An early proposal that was couched in neurophysiological terms was that of a center-surround spatially antagonistic receptive field, selective for opposite directions of motion in the center and surround (Nakayama & Loomis, 1974), i.e., "motion opponency."

The simplest form for such a motion-opponent receptive field would be "single opponent" in which a neuron has a particular preferred direction of motion within the center of its receptive field, with suppression by the same direction of motion (and/or enhancement by the opposite direction) in the surround. Neurons whose responses are consistent with single-opponent motion signals have been described primarily in extrastriate cortical areas. In recent neurophysiological studies using an animal that was translated on a moveable platform, Nadler et al. (2008, 2009) demonstrated neurons in area MT that are tuned to near and far depth defined by monocular motion parallax stimuli and could potentially be employed in forming perceptual depth information.

Neurons that combine motion and disparity information would be potentially promising candidates for depth from motion parallax. Such integration has been well documented in area MT neurons (Bradley, Qian, & Andersen, 1995; DeAngelis & Uka, 2003; Pack, Born, & Livingstone, 2003), although these experiments were not specifically designed to address motion parallax. Some neurons in the higher tier area MST, which receives much of its input from MT, prefer one direction of motion when the disparity corresponds to the stimulus being closer than the fixation plane and the opposite direction of motion when the disparity corresponds to a farther stimulus (Roy, Komatsu, & Wurtz, 1992; Roy & Wurtz, 1990).

A more complex form of motion opponency would be a "double-opponent" receptive field for which the neuron responds to relative anti-phase motion between center and surround regardless of the absolute directions of motion in either. Such motion double opponency has been found in many of the neurons in primate area MT (Allman et al., 1985; Born, 2000; Pack, Hunter, & Born, 2005) and its homologue in the cat, area PMLS (Von Grunau & Frost, 1983), as well as the primate superior colliculus (Bender & Davidson, 1986) and the optic tectum of the pigeon (Frost & Nakayama, 1983). Note that both single- and double-opponent receptive fields will give preferential responses to motion discontinuities, and so could provide useful information for segmentation of motion-defined boundaries. However, for a neuron to selectively respond to the *sign* of relative depth, its receptive field must preserve information about the directions of motion—thus, it must possess a *single*-opponent, not double-opponent, receptive field organization in order to provide depth-ordering information.

Boundaries in a natural scene are often oriented and defined by more than one cue. Thus, a plausible first step to segment such boundaries would be to signal their orientation. Furthermore, such orientation selectivity

should be similar for boundaries defined by more than one cue such as texture, luminance, or motion—i.e., “cue-invariant” (Albright, 1992). Neurons responding selectively to the orientation of motion-defined shear boundaries have been described in primate areas MT (Marcar, Xiao, Raiguel, & Orban, 1995; Xiao, Raiguel, Marcar, & Orban, 1997) and V4 (Mysore, Vogels, Raiguel, & Orban, 2006), but in those studies, it was only the V4 neurons that exhibited cue invariance for both luminance- and motion-defined boundaries. Weaker cue-invariant orientation selectivity in a minority of neurons has also been found in V2 and to a lesser degree in V1 (Marcar et al., 1995). These studies report different patterns of responses, depending on the relationship between the direction of texture motion and the boundary orientation. Such results might suggest that psychophysical performance for segmentation and depth tasks would differ for shear vs. occlusion.

Parietal cortex in primates contains neurons that combine visual and motor information (Mountcastle, Lynch, Georgopoulos, Sakata, & Acuna, 1975; Taira, Mine, Georgopoulos, Murata, & Sakata, 1990). In particular, the intraparietal sulcus contains three adjacent areas, whose neurons respond preferentially during visuomotor behavior within “grasp-related space” in AIP, “immediate extrapersonal space” in MIP, and “perioral space” in VIP (Colby & Duhamel, 1991; Colby & Goldberg, 1999). Possibly, such a segregation of function might be related to the different kinds of motion parallax performance that we find for varying ranges of relative depth. To integrate this information, a coordinate transformation is needed from observer to object coordinates. Such a role might be played by parietal cortex neurons such as those in area LIP, which shift the position of their receptive field with an eye movement in such a way as to always correspond to the same portion of the visual field (Duhamel, Colby, & Goldberg, 1992)—thus, these neurons encode the world in a gaze-centric coordinate system. This type of coding would be particularly advantageous for motion parallax stimuli, since it would facilitate discounting changes in the optic flow due to eye movements.

## General conclusion

This work has examined the importance of shear-based motion parallax in segmentation and depth perception. Our findings suggest that the visual system utilizes different mechanisms to obtain depth and segmentation from the same visual information. Thus, the utilization of motion parallax is dependent not only on the available information in the stimulus but also on the computational problem it faces.

The differences in depth task performance between sine- and square-wave modulations lend support to Gibson’s distinction between “two-velocity” and “flow field” stimuli.

Our findings suggest that shear-based motion parallax more effectively signals small depth differences across depth gradients within an object but supports segmentation of abrupt boundaries arising from larger depth differences between separate objects or an object and a background. Therefore, processing of visual motion information might be categorically different for abrupt motion boundaries and for smooth gradients of motion.

## Acknowledgments

This work was funded by a grant from the Natural Sciences and Engineering Research Council of Canada (OPG-0001978) to CB. We would like to thank Michael Langer, Chris Pack, and Fredrick Kingdom for their comments and discussion during the experiments and writing of this manuscript. Likewise, we are grateful to our observers for their participation. Early reports of these findings were presented at the Annual Meeting of the Visual Sciences Society (Yoonessi & Baker, 2009, 2010).

Commercial relationships: none.

Corresponding author: Ahmad Yoonessi.

Email: ahmad.yoonessi@mail.mcgill.ca.

Address: McGill Vision Research, 687 Pine Ave W, H4-14, Montreal, Quebec H3A 1A1, Canada.

## Footnote

<sup>1</sup>In the interests of clarity, we shall maintain a convention of using the term “movement” to refer to displacement of the observer and “motion” for changes in the visual stimulus (i.e., retinal image).

## References

- Adelson, E. H., & Bergen, J. R. (1985). Spatiotemporal energy models for the perception of motion. *Journal of the Optical Society of America*, 2, 284–299.
- Albright, T. D. (1992). Form-cue invariant motion processing in primate visual cortex. *Science*, 255, 1141–1143.
- Allman, J., Miezen, F., & McGuinness, E. (1985). Stimulus specific responses from beyond the classical receptive field: Neurophysiological mechanisms for local–global comparisons in visual neurons. *Annual Review of Neuroscience*, 8, 407–430.
- Andersen, G. J., & Braunstein, M. L. (1983). Dynamic occlusion in the perception of rotation in depth. *Perception & Psychophysics*, 34, 356–362.

- Angelaki, D. E., & Hess, B. J. M. (2005). Self-motion-induced eye movements: Effects on visual acuity and navigation. *Nature Reviews Neuroscience*, 6, 966–976.
- Angelaki, D. E., McHenry, M. Q., Dickman, J. D., Newlands, S. D., & Hess, B. J. M. (1999). Computation of inertial motion: Neural strategies to resolve ambiguous otolith information. *Journal of Neuroscience*, 19, 316.
- Baker, C. L., Jr., & Braddick, O. J. (1982). Does segregation of differently moving areas depend on relative or absolute displacement? *Vision Research*, 22, 851–856.
- Baker, C. L., Jr., & Braddick, O. J. (1985). Eccentricity-dependent scaling of the limits for short-range apparent motion perception. *Vision Research*, 25, 803–812.
- Bender, D. B., & Davidson, R. M. (1986). Global visual processing in the monkey superior colliculus. *Brain Research*, 381, 372–375.
- Born, R. T. (2000). Center-surround interactions in the middle temporal visual area of the owl monkey. *Journal of Neurophysiology*, 84, 2658.
- Bourdon, B. (1902). *La perception visuelle de l'espace*. Paris: Librairie Schleincher Frères.
- Braddick, O. J. (1993). Segmentation versus integration in visual motion processing. *Trends in Neuroscience*, 16, 263–268.
- Bradley, D., Qian, N., & Anderson, R. A. (1995). Integration of motion and stereopsis in middle temporal cortical area of macaques. *Nature*, 373, 609–611.
- Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision*, 10, 433–436.
- Braunstein, M. L., & Andersen, G. J. (1981). Velocity gradients and relative depth perception. *Perception & Psychophysics*, 29, 145–155.
- Braunstein, M. L., & Tittle, J. S. (1988). The observer-relative velocity field as the basis for effective motion parallax. *Journal of Experimental Psychology: Human Perception and Performance*, 14, 582–590.
- Bülthoff, H., Little, J., & Poggio, T. (1989). A parallel algorithm for real-time computation of optical flow. *Nature*, 337, 549–553.
- Busetini, C., Miles, F. A., Schwarz, U., & Carl, J. R. (1994). Human ocular responses to translation of the observer and of the scene: Dependence on viewing distance. *Experimental Brain Research*, 79, 484–494.
- Cao, A., & Schiller, P. H. (2003). Neural responses to relative speed in the primary visual cortex of rhesus monkey. *Visual Neuroscience*, 20, 77–84.
- Carpenter, R. H. S. (1988). *Movements of the eyes* (2nd ed.). London: Pion.
- Colby, C. L., & Duhamel, J. R. (1991). Heterogeneity of extrastriate visual areas and multiple parietal areas in the macaque monkey. *Neuropsychologia*, 29, 517–537.
- Colby, C. L., & Goldberg, M. (1999). Space and attention in parietal cortex. *Neuroscience*, 22, 319–349.
- Cutting, J. E., & Vishton, P. M. (1995). Perceiving layout and knowing distances: The integration, relative potency, and contextual use of different information about depth. In *Handbook of perception and cognition: Perception of space and motion* (vol. 5, pp. 69–117). San Diego, CA: Academic Press.
- DeAngelis, G. C., & Uka, T. (2003). Coding of horizontal disparity and velocity by MT neurons in the alert macaque. *Journal of Neurophysiology*, 89, 1094.
- Duhamel, J. R., Colby, C. L., & Goldberg, M. E. (1992). The updating of the representation of visual space in parietal cortex by intended eye movements. *Science*, 255, 90.
- Epstein, W., & Park, J. (1964). Examination of Gibson's psychophysical hypothesis. *Psychological Bulletin*, 62, 180–196.
- Farber, J. M., & McConkie, A. B. (1979). Optical motions as information for unsigned depth. *Journal of Experimental Psychology: Human Perception and Performance*, 5, 494–500.
- Frost, B. J., & Nakayama, K. (1983). Single visual neurons code opposing motion independent of direction. *Science*, 220, 744–745.
- 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, 4051.
- Gibson, J. J., & Carel, W. (1952). Does motion perspective independently produce the impression of a receding surface. *Journal of Experimental Psychology*, 44, 16–18.
- Helmholtz, H. V. (1925). Physiological optics. *Optical Society of America*, 3, 318.
- Hildreth, E. C. (1984). Computations underlying the measurement of visual motion. *Artificial Intelligence*, 23, 309–354.
- Kingdom, F. A. A., & Prins, N. (2010). *Psychophysics: A practical introduction*. London: Academic Press, Elsevier.
- Koenderink, J. J., & Van Doorn, A. J. (1975). Invariant properties of the motion parallax field due to the movement of rigid bodies relative to an observer. *Optica Acta*, 22, 773–791.
- 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, 389–412.
- Longuet-Higgins, H., & Prazdny, K. (1980). The interpretation of a moving retinal image. *Proceedings of the Royal Society of London B: Biological Sciences*, 208, 385–397.
- Marcas, V. L., Xiao, D., Raiguel, S. E., & Orban, G. A. (1995). Processing of kinetically defined boundaries in the cortical motion area MT of the macaque monkey. *Journal of Neurophysiology*, 74, 1258.
- Mountcastle, V. B., Lynch, J. C., Georgopoulos, A., Sakata, H., & Acuna, C. (1975). Posterior parietal association cortex of the monkey: Command functions for operations within extrapersonal space. *American Physiological Society*, 38, 871–908.
- Mysore, S. G., Vogels, R., Raiguel, S. E., & Orban, G. A. (2006). Processing of kinetic boundaries in macaque V4. *Journal of Neurophysiology*, 95, 1864.
- Nadler, J. W., Angelaki, D. E., & DeAngelis, G. C. (2008). A neural representation of depth from motion parallax in macaque visual cortex. *Nature*, 452, 642–645.
- Nadler, J. W., Nawrot, M., Angelaki, D. E., & DeAngelis, G. C. (2009). MT neurons combine visual motion with a smooth eye movement signal to code depth-sign from motion parallax. *Neuron*, 63, 523–532.
- Naji, J. J., & Freeman, T. C. A. (2004). Perceiving depth order during pursuit eye movement. *Vision Research*, 44, 3025–3034.
- Nakayama, K., & Loomis, J. (1974). Optical velocity patterns, velocity-sensitive neurons, and space perception: A hypothesis. *Perception*, 3, 63–80.
- Nakayama, K., Silverman, G. H., Macleod, D. I. A., & Mulligan, J. (1985). Sensitivity to shearing and compressive motion in random dots. *Perception*, 14, 225–238.
- Nawrot, M. (2003). Eye movements provide the extra-retinal signal required for the perception of depth from motion parallax. *Vision Research*, 43, 1553–1562.
- Nawrot, M., & Joyce, L. (2006). The pursuit theory of motion parallax. *Vision Research*, 46, 4709–4725.
- Ono, H., Rogers, B. J., Ohmi, M., & Ono, M. E. (1988). Dynamic occlusion and motion parallax in depth perception. *Perception*, 17, 255–266.
- 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.
- 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, 331–337.
- Orban, G. A., Fize, D., Peuskens, H., Denys, K., Nelissen, K., Sunaert, S., et al. (2003). Similarities and differences in motion processing between the human and macaque brain: Evidence from fMRI. *Neuropsychologia*, 41, 1757–1768.
- Pack, C. C., Born, R. T., & Livingston, M. S. (2003). Two-dimensional substructure of stereo and motion interactions in macaque visual cortex. *Neuron*, 37, 525–535.
- Pack, C. C., Hunter, J. N., & Born, R. T. (2005). Contrast dependence of suppressive influences in cortical area MT of alert macaque. *Journal of Neurophysiology*, 93, 1809.
- Ramat, S., & Zee, D. (2003). Ocular motor responses to abrupt interaural head translation in normal humans. *Journal of Neurophysiology*, 90, 887.
- Regan, D. (1986). Form from motion parallax and form from luminance contrast: Vernier discrimination. *Spatial Vision*, 1, 305–318.
- Regan, D. (1989). Orientation discrimination for objects defined by relative motion and objects defined by luminance contrast. *Vision Research*, 29, 1389–1400.
- Rogers, B. J., & Graham, M. (1979). Motion parallax as an independent cue for depth perception. *Perception*, 8, 125–134.
- Rogers, B. J., & Graham, M. (1982). Similarities between motion parallax and stereopsis in human depth perception. *Vision Research*, 22, 261.
- Rogers, S., & Rogers, B. J. (1992). Visual and nonvisual information disambiguate surfaces specified by motion parallax. *Perception & Psychophysics*, 52, 446–452.
- Roy, J. P., Komatsu, H., & Wurtz, R. H. (1992). Disparity sensitivity of neurons in monkey extrastriate area MST. *Journal of Neuroscience*, 12, 2478.
- Roy, J. P., & Wurtz, R. H. (1990). The role of disparity-sensitive cortical neurons in signalling the direction of self-motion. *Nature*, 348, 160–162.
- Sachtler, W., & Zaidi, Q. (1995). Visual processing of motion boundaries. *Vision Research*, 35, 807–826.
- Sary, G., Vogels, R., & Orban, G. A. (1993). Cue-invariant shape selectivity of macaque inferior temporal neurons. *Science*, 260, 995–995.
- Schwarz, U., Busetini, C., & Miles, F. A. (1989). Ocular responses to linear motion are inversely proportional to viewing distance. *Science*, 245, 1394.
- Sperling, G. B., Doshier, B. A., & Landy, M. S. (1990). How to study the kinetic depth experimentally. *Journal of Experimental Psychology: Human Perception and Performance*, 16, 445–450.

- Stoner, G., & Albright, T. D. (1993). Image segmentation cues in motion processing: Implications for modularity in vision. *Journal of Cognitive Neuroscience*, *5*, 129–149.
- Taira, M., Mine, S., Georgopoulos, A. P., Murata, A., & Sakata, H. (1990). Parietal cortex neurons of the monkey related to the visual guidance of hand movement. *Experimental Brain Research*, *83*, 29–36.
- Trucco, E., & Verri, A. (1998). *Introductory techniques for 3-D computer vision*. Upper Saddle River, NJ: Prentice Hall.
- Ujike, H., & Ono, H. (2001). Depth thresholds of motion parallax as a function of head movement velocity. *Vision Research*, *41*, 2835–2843.
- Ullman, S. (1979). *The interpretation of visual motion*. Cambridge, MA: MIT Press.
- Vanduffel, W., Fize, D., Mandeville, J. B., Nelissen, K., Van Hecke, P., Rosen, B. R., et al. (2001). Visual motion processing investigated using contrast agent-enhanced fMRI in awake behaving monkeys. *Neuron*, *32*, 565–577.
- Von Grunau, M., & Frost, B. J. (1983). Double-opponent-process mechanism underlying RF-structure of directionally specific cells of cat lateral suprasylvian visual area. *Experimental Brain Research*, *49*, 84–92.
- Watson, A. B., & Eckert, M. P. (1994). Motion-contrast sensitivity: Visibility of motion gradients of various spatial frequencies. *Journal of the Optical Society of America*, *11*, 496–505.
- Westheimer, G., & McKee, S. P. (1975). Visual acuity in the presence of retinal-image motion. *Journal of the Optical Society of America*, *65*, 847–850.
- Westheimer, G., & McKee, S. P. (1978). Stereoscopic acuity for moving retinal images. *Journal of the Optical Society of America*, *68*, 450–455.
- Wexler, M., Panerai, F., Lamouret, I., & Droulez, J. (2001). Self motion and perception of stationary objects. *Nature*, *409*, 85–88.
- Wexler, M., & van Boxtel, J. J. A. (2005). Depth perception by the active observer. *Trends in Cognitive Sciences*, *9*, 431–438.
- Xiao, D. K., Raiguel, S., Marcar, V., & Orban, G. A. (1997). The spatial distribution of the antagonistic surround of MT/V5 neurons. *Cerebral Cortex*, *7*, 662.
- Yonas, A., Craton, L. G., & Thompson, W. B. (1987). Relative motion: Kinetic information for the order of depth at an edge. *Perception & Psychophysics*, *41*, 53–59.
- Yoonessi, A., & Baker, C. L., Jr. (2009). Is segmentation from motion parallax influenced by perceived depth? [Abstract]. *Journal of Vision*, *9*(8):935, 935a, <http://www.journalofvision.org/content/9/8/935>, doi:10.1167/9.8.935.
- Yoonessi, A., & Baker, C. L., Jr. (2010). Contribution of motion parallax to depth ordering, depth magnitude and segmentation [Abstract]. *Journal of Vision*, *10*(7):1194, 1194a, <http://www.journalofvision.org/content/10/7/1194>, doi:10.1167/10.7.1194.