

Boundary segmentation from dynamic occlusion-based motion parallax

Ahmad Yoonessi

McGill Vision Research Unit, Department of
Ophthalmology, McGill University, Montreal, QC, Canada



Curtis L. Baker, Jr.

McGill Vision Research Unit, Department of
Ophthalmology, McGill University, Montreal, QC, Canada



Active observer movement results in retinal image motion that is highly dependent on the scene layout. This retinal motion, often called motion parallax, can yield significant information about the boundaries between objects and their relative depth differences. Previously we examined segmentation from shear-based motion parallax, which consists of only relative motion information. Here, we examine segmentation from dynamic occlusion-based motion parallax, which contains both relative motion and accretion-deletion. We utilized random dots whose motion was modulated with vertical low spatial frequency envelopes and synchronized to head movements (Head Sync), or recreated using previously recorded head movement data for the same stationary observer (Playback). Observers judged the orientation of a boundary between regions of oppositely moving dots in a 2AFC task. The results demonstrate that observers perform poorer when the stimulus motion is synchronized to head movement, particularly at smaller relative depths, even though that head movement provides significant information about depth. Both expansion-compression and accretion-deletion in isolation could support segmentation, albeit with reduced performance. Therefore, unlike our previous results for depth ordering, expansion-compression and accretion-deletion contribute similarly to segmentation. Furthermore, human observers do not appear to utilize depth information to improve segmentation performance.

Introduction

When a human observer moves about in a natural (three-dimensional [3-D]) static environment, a complex pattern of retinal image motion is formed. This pattern, often called motion parallax, is dependent on the 3-D scene layout and can provide information

about the boundaries between surfaces and their respective locations in depth. When the resulting motion boundaries between surfaces are parallel to the direction of observer movement, shear boundaries are produced, consisting of relative textural motion. However, boundaries orthogonal to the direction of observer movement produce dynamic occlusion boundaries, which are more complicated than shear boundaries, because they contain not only relative motion (in this case, expansion-compression), but also the appearance and disappearance of texture elements along the moving contour. The latter phenomenon (accretion-deletion) can provide significant information for detection of boundaries between surfaces and their relative depths (Yonas, Craton, & Thompson, 1987).

Segmentation from dynamic occlusion has been previously studied, however, only for stationary observers. Studies of motion-defined form (Baker & Braddick, 1982; Regan, 1986, 1989) demonstrated that stationary human observers are remarkably good in obtaining figure-ground segmentation from a mixture of shear and dynamic occlusion. Detection of motion-defined boundaries has been investigated with stationary observers for expansion-compression as well as shear, but not with accretion-deletion (Nakayama, Silverman, Macleod, & Mulligan, 1985; Shimojo, Silverman, & Nakayama, 1989; Watson & Eckert, 1994; Sachtler & Zaidi, 1995). Many of these studies found that a small amount of motion, in some cases as little as two frames, is sufficient for segmenting motion boundaries. Furthermore, accretion-deletion has been shown to be a powerful cue for depth in the case of a stationary observer (Gibson, Kaplan, Reynolds, & Wheeler, 1969; Kaplan, 1969; Thompson, Mutch, & Berzins, 1985; Yonas et al., 1987; Craton & Yonas, 1990; Hegdé, Albright, & Stoner, 2004; Kromrey, Bart, & Hegdé, 2011).

Citation: Yoonessi, A., & Baker, C. L., Jr. (2014). Boundary segmentation from dynamic occlusion-based motion parallax. *Journal of Vision*, 14(4):15, 1–15, <http://www.journalofvision.org/content/14/4/15>, doi:10.1167/14.4.15.

Segmentation from motion parallax has not been systematically investigated, even though most naturally occurring retinal image motion occurs due to the observer's own movement in the environment. Previous psychophysical studies of motion parallax have concentrated on its role in depth perception (e.g., Rogers & Graham, 1979, 1982, 1983; Ono, Rogers, Ohmi, & Ono, 1988), and its contribution to boundary segmentation has been quantitatively studied only for simple shear patterns (Yoonessi & Baker, 2011a). However, pure shear is relatively rare in everyday life—most naturally occurring motion parallax entails a substantial degree of dynamic occlusion as well as shear.

In general, active self-movement of the observer generates a pattern of retinal image motion that depends on the characteristics of head and eye movements and the visual scene. If there are no compensatory eye movements during a lateral translation, only objects at infinity will remain stable and the visual scene will experience significant blur. However, if compensatory eye movements are made, the objects at the foveal depth plane will remain stabilized, whereas acuity will be decreased for the rest of the visual scene (Angelaki & Hess, 2005). In the context of motion parallax, self-movement results in not only a pattern of optic flow that is dependent on surface distances, but also extraretinal information that can disambiguate depth relationships (Wexler & Van Boxtel, 2005). From a theoretical point of view, it might be reasonable to expect that the depth information obtained during motion parallax should logically improve segmentation performance, since a depth difference alone provides evidence for an occlusion boundary. Furthermore, unlike other segmentation cues such as luminance or contrast, abrupt differences in optic flow in motion parallax can only arise from an occlusion boundary.

Many theoretical studies demonstrated that employing depth information can improve the performance of a computer algorithm for segmentation (e.g., Sun, Sudderth, & Black, 2012), but the psychophysical data investigating whether human observers can improve segmentation by incorporating depth information is limited. Previous psychophysical findings demonstrate that good depth perception can be obtained from dynamic occlusion stimuli in motion parallax (e.g., Yoonessi & Baker, 2013). A simple cue-summation model might suggest that with the increase in number of reliable information sources about an object, psychophysical performance should logically improve. Thus the visual system might incorporate these different cues in a simple weighted sum manner to infer the location and orientation of the occlusion boundaries, but the results of our earlier shear study did not show a facilitation of segmentation performance by incorporating concomitant depth information (Yoonessi & Baker, 2011a). However, a different

outcome might be expected for dynamic occlusion, because accretion-deletion can contribute to depth from dynamic occlusion even in the absence of head movements (Yoonessi & Baker, 2013). In order to examine the importance of head movement and the resulting extraretinal depth information, here we compare two conditions. In the Head Sync condition, in which stimulus motion was synchronized to the head movement so as to mimic natural motion parallax, observers voluntarily initiated head movement excursion in a free and unconstrained manner, to simulate natural viewing conditions. We varied the ratio between head movements and the stimulus motion, which we call “syncing gain,” as the primary variable—more details about this parameter can be found in our earlier studies (Yoonessi & Baker, 2011a, 2013). In the Playback condition, previously recorded head movement data was used to recreate the same visual information on the screen for a stationary observer (Wexler, Panerai, Lamouret, & Droulez, 2001; Nadler, Nawrot, Angelaki, & DeAngelis, 2009; Yoonessi & Baker, 2011a). Thus in this condition there should be little or no input from extraretinal sources, such as vestibular sensors (otoliths) or eye movements, so the difference between the Head Sync and Playback conditions should be primarily nonvisual. We previously made this comparison for pure shear, which contains only relative motion. However, dynamic occlusion is more complicated than shear, since it entails accretion-deletion as well as relative motion. These cues rely differently on head movements, in that expansion-compression requires head movements to be depth-unambiguous, whereas accretion-deletion does not require extraretinal information to provide valid depth ordering (Yoonessi & Baker, 2013). Thus, one might expect the relative motion component of dynamic occlusion to behave similarly as in pure shear, but the accretion-deletion cue's dependence on head movement might be very different (Yoonessi & Baker, 2013).

An important question about dynamic occlusion is how information from the two cues, expansion-compression and accretion-deletion, are combined. In our recent study of depth perception from dynamic occlusion-based motion parallax (Yoonessi & Baker, 2013), we found that observers exhibited very good depth ordering when both cues were available; however, when the cues were presented in isolation, observers could obtain some degree of depth from expansion-compression, but not at all from accretion-deletion. These results were markedly incompatible with a simple cue summation rule, in which cues contribute additively in proportion to their reliability (Landy, Maloney, Johnston, & Young, 1995). If the same cue combination rules apply to segmentation as found for depth, a similar asymmetry might be

expected in which accretion-deletion in isolation would not provide segmentation. In our previous study, accretion-deletion appeared to provide particularly powerful facilitation of depth at high values of syncing gain when co-occurring motion information was present—if segmentation behaves in a similar manner, we would expect better segmentation performance at these gain values compared to that from pure shear. On the other hand, a different outcome would suggest that the two tasks are preformed by distinct mechanisms.

Here we assess the relative contribution of the two cues at different values of syncing gain by placing them in conflict, where expansion-compression and accretion-deletion signal opposing depth signs, a condition that cannot exist in ecological settings. Even though the conflict between the two cues only involves depth, any change in performance might suggest dependence of segmentation on concomitant depth information. On the other hand, similar results would suggest that depth information from relative motion does not affect segmentation. In order to separate the contributions of the two cues we compare the segmentation performance for sine wave modulation, which consists only of expansion-compression information, and a condition in which only accretion-deletion information is present without relative motion.

Our results demonstrate that the head movement does not improve segmentation performance, suggesting that concomitant depth information is not utilized. In marked contrast to previous depth-ordering results (Yoonessi & Baker, 2013), accretion-deletion as well as expansion-compression in isolation can provide good segmentation performance and act in a manner consistent with a simple cue summation.

General materials and methods

Here we briefly summarize the hardware and software setup. A more detailed description can be found in our previous papers (Yoonessi & Baker, 2011a, 2013).

Task and visual stimuli

The stimuli were generated with a Macintosh (Mac Pro, 2×2.8 GHz, 4 GB Ram, OSX v10.5) using Matlab (Mathworks, MA, USA) code written with Psychophysics Toolbox (Brainard, 1997; Pelli, 1997; Kleiner, Brainard, & Pelli, 2007), 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^2 . The stimuli were viewed from a

distance of 114 cm with monocular viewing, using an eye patch, to avoid cue conflict from stereopsis.

Stimulus patterns consisted of white (80.31 cd/m^2) dots on a black (0.07 cd/m^2) background. Each dot was 0.2° and the density of dots was 1.04 dots/deg^2 . Each dot was of circular shape with high quality anti-aliasing. The dots' displacements were modulated using sine or square wave profiles to create motion patterns (Figure 1a).

We utilized OpenGL texture mapping and OpenGL shading language to render the stimulus in the same manner described previously (Yoonessi & Baker, 2013). To emulate a motion parallax situation and provide potential depth percepts, the motions of the dots were synchronized to measured changes in head position (see below). On each frame update, the difference between current and previous head position was calculated, and then multiplied by a scale factor, the “syncing gain” (see below)—the one-dimensional (1-D) modulation profile was multiplied by this number, and used to modulate the dot horizontal displacement. 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 is proportional to rendered depth and might be a better representation of information obtained from motion parallax than absolute values of velocity or equivalent disparity. Note that a syncing gain of unity produces a simulated depth corresponding to the magnitude of the viewing distance itself.

The spatial frequency of the modulation was always 0.1 cpd. The stimuli were presented within a circular mask of 18° of visual angle, which resulted in about 1.5 cycles/image of visible modulation. The modulation waveform was a 1-D sine or square wave, and was bidirectional, which corresponds to peaks and troughs moving in opposite directions. 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. A fixation point was presented before and during each stimulus presentation, at the center of the circular mask. The stimulus was modulated with sine phase modulation, which corresponds to the modulator zero-crossing point at the center of the screen. The fixation point was always set at this zero-crossing (Figure 1b). The approximate simulated depth for each condition was calculated based on triangulation between the fixation point, the syncing gain, and the amount of head movements, and is shown as an additional axis along the top of each data graph.

In order to aid comparison between the results of this study and previous results for segmentation from shear and depth from dynamic occlusion (Yoonessi & Baker, 2011a, 2013), we kept all stimulus parameters

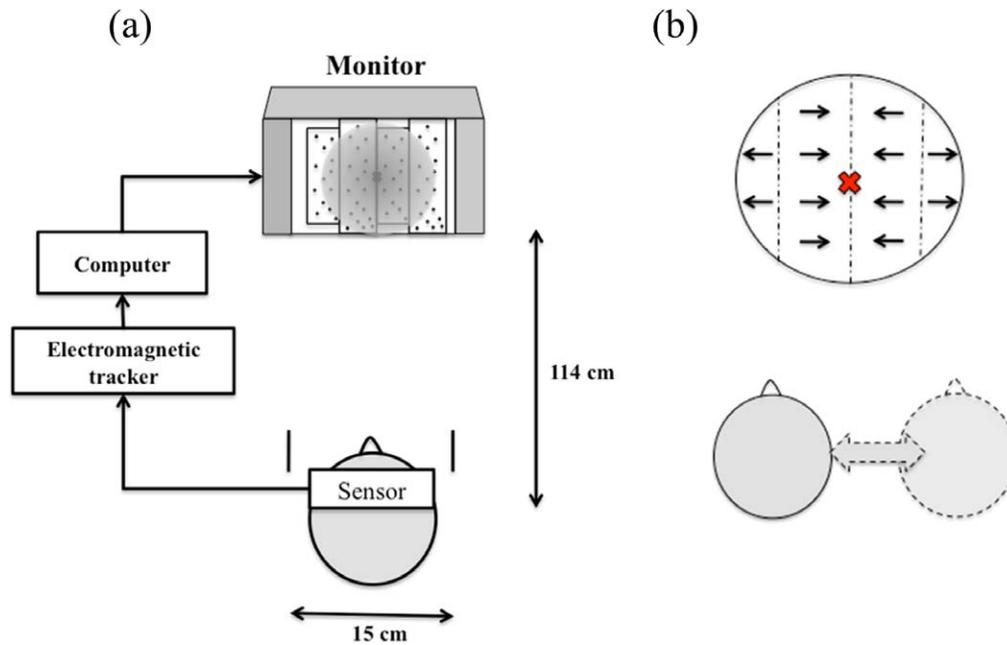


Figure 1. Experimental setup for motion parallax experiments. (a) As the human observer moves laterally, the computer updates visual stimuli on the monitor in synchrony with head position provided by the electromagnetic tracking of a sensor placed on the observer's forehead. (b) Visual stimulus as seen by observers, consisting of regions of dots moving in opposite directions to one another. The fixation point was always at the center of the screen, even though the boundary could move past it.

such as density, spatial frequency, etc., identical to those in the previous studies.

Stimuli were presented with 1-s duration, which was sufficient for good performance on the segmentation task. We verified formally that segmentation performance did not appreciably benefit from increasing the presentation time from 1–5 s (see Supplementary materials).

Head movement recording

Observers were instructed to freely translate their head laterally back and forth while viewing the stimulus during each trial, traversing a path corresponding to a distance of about 15 cm. Head position and orientation data were recorded with 6-DOF using an electromagnetic position-tracking device (Flock of Birds, Ascension Technologies, Shelburne, VT) with a medium-range transmitter. The sensor, secured on the observer's forehead using an elastic band, recorded head position and orientation data with 0.5 mm and 0.1° resolution, respectively. The head movement data was sampled at 100 Hz and transferred to the computer using a serial port / USB connection. Observers were instructed to perform only a lateral head translation, and used two vertical bars as guides for end points of the movement (Figure 1a). Only the x -direction data was employed for the visual stimulus, and any movement in the y - or z -direction was excluded from

use in the rendering. However, the full 6-DOF position and orientation data was recorded on the hard disk for subsequent analysis.

Segmentation performance was measured using a 2AFC orientation judgment. The task was to judge the orientation of the perceived boundary between adjacent regions of dots moving in opposite directions. The boundary was not detectable in any single frame and therefore the task was only possible by spatiotemporal integration. The boundary was slightly tilted, left- or right-oblique, in the monitor plane (around the z -axis) as indicated in Figure 2b, and observers pressed one of two possible keys to report perceived boundary orientation. Even though the boundary was slightly tilted, the motion of each dot was always horizontal, and therefore task performance was not possible using only local motion cues. On different trials the syncing gain parameter was varied, which corresponded to different rendered depth differences between the two surfaces. This depth difference was always parallel to the monitor plane—note that if the depth had been rendered as zero, there would be no motion in the stimulus and the task would have been impossible. Observers were instructed to maintain gaze on a static fixation mark at the center of the screen, positioned halfway between the rendered depths of the adjacent surfaces, and always on the same depth plane as the monitor screen. When the near-rendered surface covered parts of the far surface around the center of the display, the fixation point still remained visible to

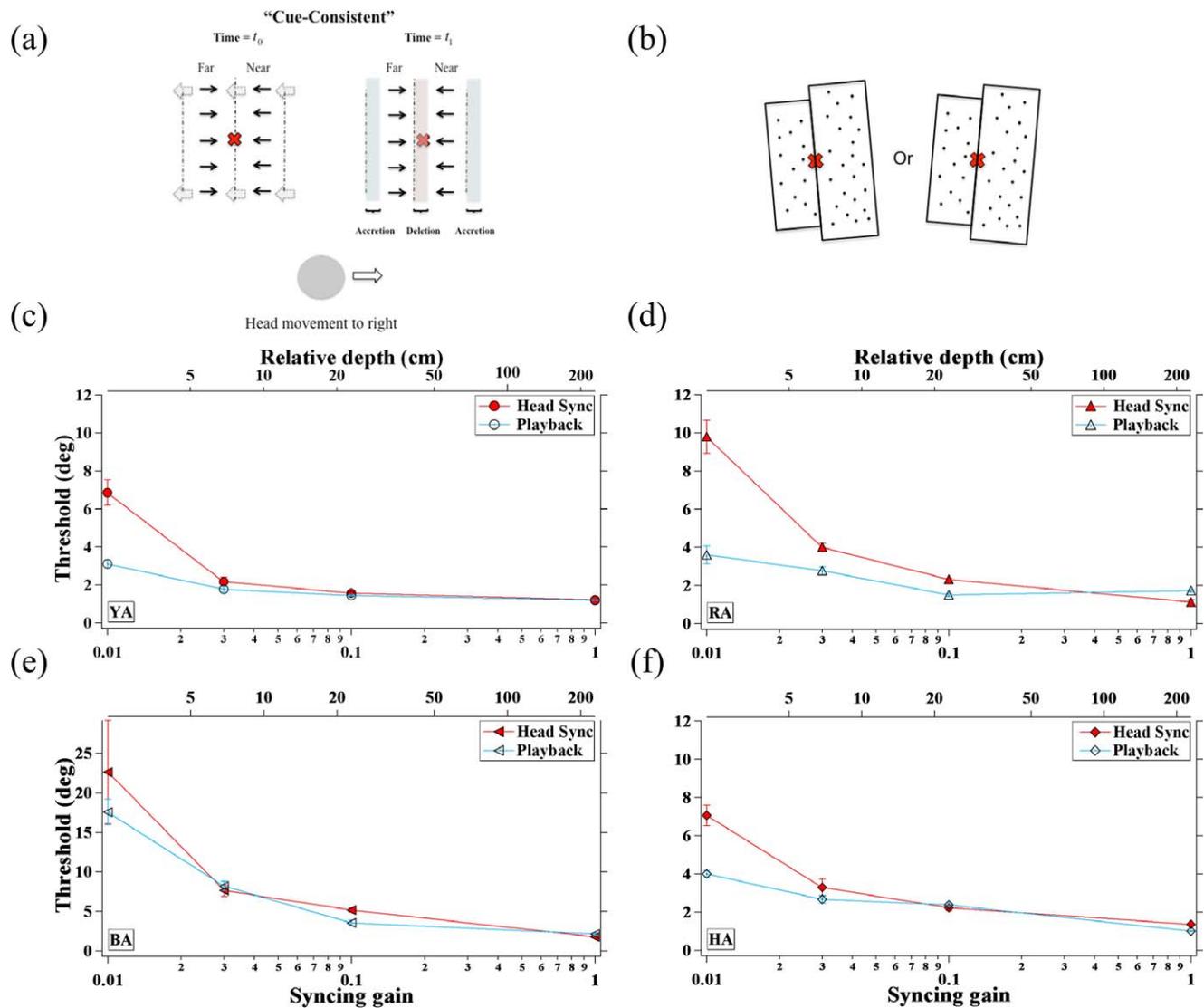


Figure 2. Segmentation experiment for square wave modulation. (a) Schematic depiction of Cue-Consistent combination of the expansion-compression and accretion-deletion cues. Smaller filled arrows represent motion of the random dot textures (expansion-compression), whereas larger open arrows represent motion of the boundary along which accretion-deletion occurs. Surfaces are labeled Near and Far, as signaled by both the expansion-compression and accretion-deletion cues. As the observer moves, the leading edge of the near surface causes deletion of texture on the far surface, and the trailing edge of the near surface gives rise to accretion of the far surface texture. (b) Cartoon drawing of the 2AFC orientation judgment task. (c–f) Results for four observers plotted as just noticeable difference threshold versus syncing gain. Filled red symbols indicate data for the Head Sync condition, and open blue symbols the Playback condition in which the observer is stationary. Error bars, here and in subsequent figures, indicate \pm SE.

ensure consistency in gaze direction during and across trials. The head-synced condition was tested first, and in the subsequent block of experiments the previously recorded head movement data was used to recreate the same visual information on the screen (Playback) for the same observer with their head held stationary. Therefore the difference between the two conditions (Head Sync and Playback) should be predominantly nonvisual. The stimulus was not inherently ambiguous, and depth in the stimulus was not essential to the judgment being made; however, we were interested in

examining the influence of head movements and of concomitant depth information on the segmentation judgment.

Each value of syncing gain, modulation pattern (sine or square wave) and Cue-Consistent/Cue-Conflict condition (see below) was tested in separate blocks using a method of constant stimuli. In each block five values of orientation were presented in a random order, with 20 repetitions per level value. Trial blocks were accumulated, such that each level value was tested at least 60 times. A cumulative Gaussian function was

then fit to the percent correct versus orientation results to obtain a just noticeable difference threshold. Curve fits and bootstrap estimates of the curve fit threshold parameters were obtained using Prism software (Graphpad, CA, USA).

Observers

Four observers (YA, RA, BA, and HA) participated in these experiments, three of whom (RA, BA, and HA) were naive to the purpose of the experiment. All observers had normal or corrected-to-normal vision. All experiments were conducted in accordance with the university's ethical guidelines, and observers gave prior consent to their participation in the experiment. All experimental procedures adhered to the Declaration of Helsinki.

Results

First we assessed segmentation performance with both cues present in a relatively natural manner, i.e., with square wave depth modulation. Figure 2a shows a schematic depiction of this Cue-Consistent condition, in which the smaller filled arrows indicate texture motion, and the large open arrows indicate boundary motion. In this condition, the boundary moved synchronously with the surface that was moving oppositely to the head movement. Therefore the surface rendered as nearer to the observer by the relative motion (expansion-compression) cue occluded the far surface. The Near and Far labels above each region in Figure 2a indicate the near and far surfaces as rendered from the relative motion cue. Figure 2b depicts a cartoon drawing of the 2AFC orientation judgment in which observers chose one of the two possible responses accordingly. Figure 2c through f shows the results for four observers, with each graph showing the measured orientation thresholds plotted against the syncing gain parameter. The top axis indicates the corresponding relative depths rendered in the stimuli. These depth values are computed by geometric triangulation between head movement and viewing distance, with the monitor plane having zero depth—they are only approximations, because of rendering inaccuracies and head movement variability (see Yoonessi & Baker, 2011a, for more details). The filled red symbols indicate thresholds for the Head Sync condition in which the stimulus motion was synchronized to the head movement. The thresholds are generally very low at the high syncing gains, but they gradually increase at lower gains. The absolute orientation thresholds are different among the observ-

ers, but the pattern of gradual increase at lower syncing gains remains fairly consistent.

In the Playback condition, depicted with open blue symbols, the same stimulus motion was recreated while the observer was stationary—thus the difference between the two conditions should be primarily nonvisual. The results show similar thresholds for Head Sync and Playback conditions at higher syncing gains, whereas at lower gains (below 0.10) the thresholds for the Head Sync condition are significantly higher than for the Playback condition. A statistical two-way independent measures ANOVA test confirms a significant difference between the Head Sync and Playback conditions, YA: $F(1, 16) = 30.87, p < 0.0001$; RA: $F(1, 16) = 53.09, p < 0.0001$; HA: $F(1, 16) = 24.95, p = 0.0001$; BA: $F(1, 16) = 13.76, p = 0.0019$.

To examine the interaction of the expansion-compression and accretion-deletion cues in segmentation we utilized a Cue-Conflict condition, in which the boundary moved in synchrony with the surface that was moving in the same direction as the observer (Figure 3a). Therefore the surface rendered as farther from the observer (by the relative motion cue) covered or uncovered texture elements of the near surface—a situation that would not occur in ecological settings. Note that the two cues produce conflicting depth cues, but consistent segmentation information. The boundary between the adjacent surfaces is still defined by two compatible information sources, whereas in the depth-ordering task these two cues are incompatible. Here we aimed to examine whether this conflicting depth information would influence segmentation performance.

Figure 3a depicts a schematic drawing of the Cue-Conflict condition. The labels Near and Far indicate near and far surfaces according to the expansion-compression information. Figure 3b shows a cartoon drawing of the 2AFC judgment. Figure 3c through f shows the results for four observers, plotted as thresholds versus syncing gain (bottom axis) or corresponding rendered depth (top axis), with Head Sync and Playback conditions shown as filled red and open blue symbols respectively. The results show good performance across the full range of syncing gains for two observers (YA and RA) and slight decline at low syncing gains for two others (HA and BA). The Head Sync and Playback conditions have fairly similar thresholds. A statistical two-way ANOVA test shows a significant difference between Head Sync and Playback conditions for only one of the observers in the Cue-Conflict condition, YA: $F(1, 16) = 41.98, p < 0.0001$; RA: $F(1, 16) = 0.26, p = 0.6140$; HA: $F(1, 16) = 0.09, p = 0.7669$; BA: $F(1, 16) = 1.92, p = 0.1853$.

Performance for the Cue-Conflict condition (Figure 3) is overall slightly better than for the Cue-Consistent condition (Figure 2). A statistical two-way ANOVA

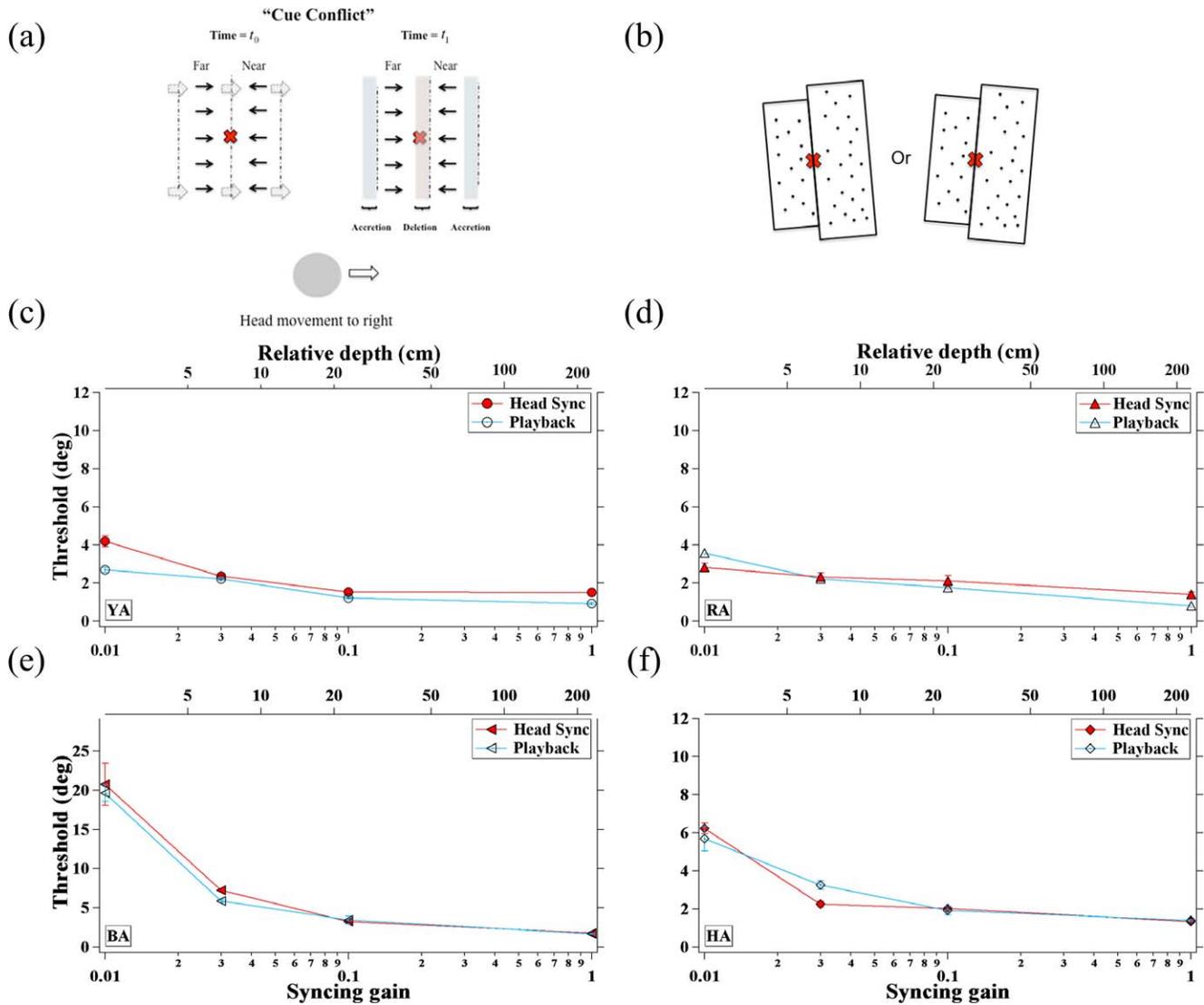


Figure 3. Same as Figure 2 but for Cue-Conflict condition, in which surfaces that were rendered as far by the relative motion cue would occlude the surface that was rendered as near by the expansion-compression cue.

test shows a significant difference between most comparisons of the Cue-Consistent and Cue-Conflict for the Head Sync condition [YA, Head Sync: $F(1, 16) = 7.60, p = 0.0141$; YA, Playback: $F(1, 16) = 2.67, p = 0.1217$; RA, Head Sync: $F(1, 16) = 71.04, p < 0.0001$; RA, Playback: $F(1, 16) = 3.91, p = 0.0655$; HA, Head Sync: $F(1, 16) = 7.08, p = 0.0171$; HA, Playback: $F(1, 16) = 6.68, p = 0.0200$; BA, Head Sync: $F(1, 16) = 16.85, p = 0.0008$; BA, Playback: $F(1, 16) = 52.67, p < 0.0001$]. Note that most of the improvement for the Cue-Conflict condition is due to better performance for the Head Sync condition—consequently, the difference between the Head Sync and Playback conditions is now smaller in the Cue-Conflict condition, especially at low syncing gains. This improvement in performance for the Cue-Conflict condition is contradictory to what one might expect, since the visual system is widely assumed

to be optimized for stimuli encountered in the natural world. However, in this case the performance is better for the nonecological condition.

Note that even though the Cue-Conflict condition in Playback experiments (i.e., without head movement) would seem to be identical to that for the Cue-Consistent stimuli, it is conceivable that observers might have used different head movements in the Head Sync versions of the two conditions. To help preclude this possibility, the Playback stimuli for the Cue-Consistent and Cue-Conflict conditions utilized each observer’s head movement recordings obtained from their respective trials with the head-synchronized stimuli. There still might have been some differences in the vestibular input, contributing differently for Head Sync and Playback conditions. However, the recorded head movements for Cue-Consistent and Cue-Conflict

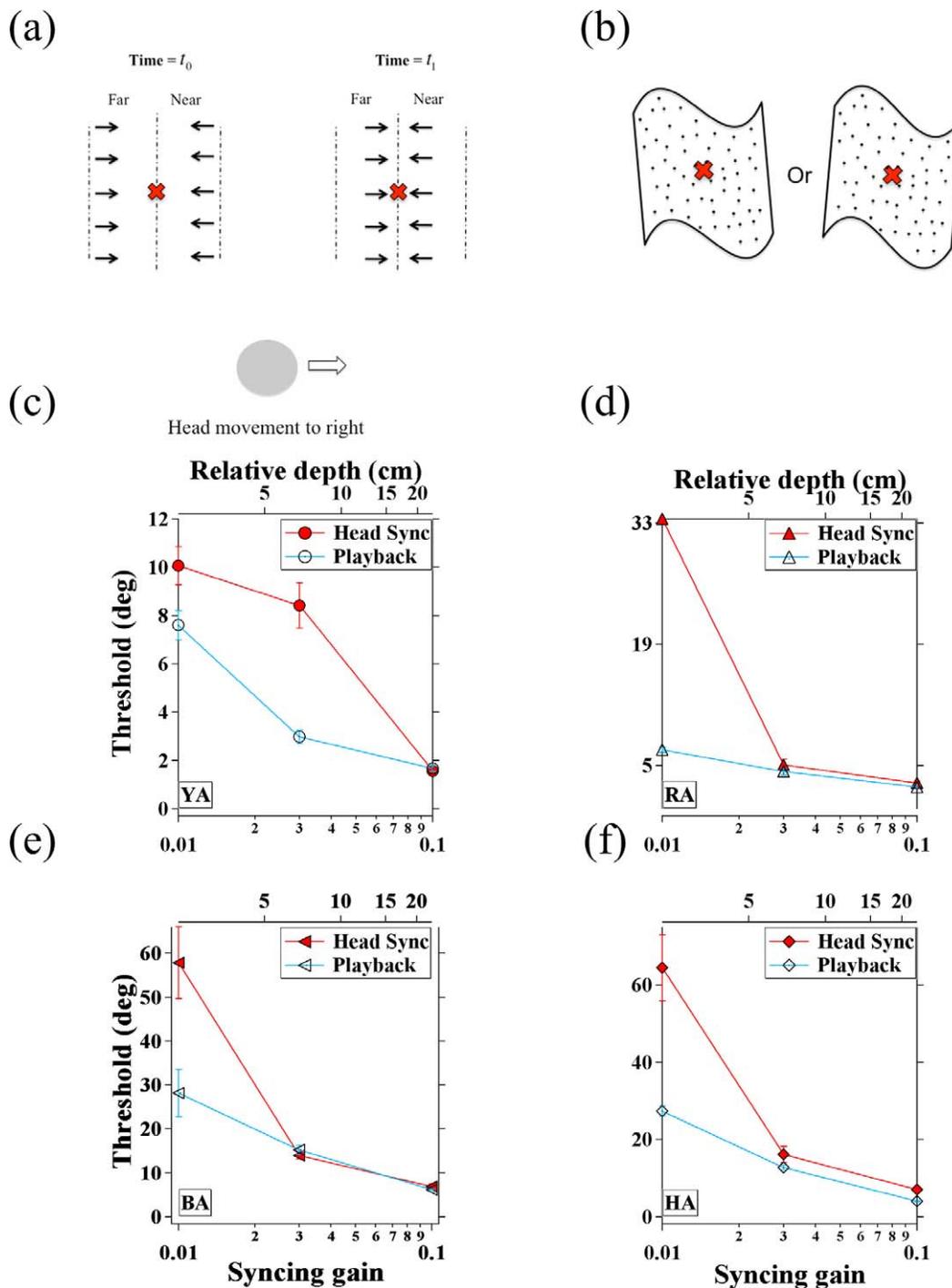


Figure 4. Same as Figure 2, but for sine wave modulation. The conditions that contain accretion-deletion were avoided, resulting in a smaller range of possible syncing gains.

conditions showed no significant differences, so this possibility appears unlikely.

The results from the Cue-Conflict condition suggest not only that accompanying depth does not facilitate segmentation, but also that it might interfere with psychophysical performance, since the condition with ecologically invalid depth demonstrates better segmentation. Comparing to previous results for shear (see

Yoonessi & Baker, 2011a; figure 3c through e), the performance is similar, which suggests a negative role for depth. Furthermore, in the high syncing gains where accretion-deletion contributes primarily to depth (Yoonessi & Baker, 2013; Figures 3b and 4b), segmentation performance is only slightly better than from shear based motion parallax, which would seem to suggest a minimal role for accretion-deletion.

To examine whether segmentation can be obtained by expansion-compression without accompanying accretion-deletion, we tested performance using a sinusoidal modulation envelope (Figure 4a). This stimulus brings about an odd situation, in that for smaller relative depths the nearer parts of the rendered surface (hills) will never occlude the farther parts (valleys), with occlusion (and therefore accretion-deletion) only occurring at the larger syncing gains. Correct handling of such accretion-deletion for a sinusoidal modulation requires use of a different approach to applying the 1-D modulation function, which would compromise comparisons with the other experiments. To avoid this problem we tested only a restricted range of smaller syncing gains. Figure 4b shows a cartoon drawing of the orientation judgment. Figure 4c through f shows the results, in which the performance is worse than for square wave modulation (Figure 2) at the smallest syncing gains. But similarly to the square wave case, performance is better in the Playback condition. A statistical two-way ANOVA confirms a significant difference between Head Sync and Playback conditions, YA: $F(1, 12) = 30.78, p = 0.0001$; RA: $F(1, 12) = 102.53, p < 0.0001$; HA: $F(1, 12) = 23.51, p = 0.0004$; BA: $F(1, 12) = 8.56, p = 0.0127$. These results demonstrate that expansion-compression alone, i.e., in the absence of accretion-deletion, is able to support segmentation, albeit not quantitatively as well as for square wave modulation (Figure 2), which also contains accretion-deletion. The segmentation results for sine wave modulation are similar to those we previously obtained for shear motion parallax (see Yoonessi & Baker, 2011a; figure 4c through e), which suggests that expansion-compression component contributes similarly to both shear and dynamic occlusion.

To assess segmentation performance in the absence of expansion-compression, we employed an Accretion-Deletion Only condition (Figure 5a). In this condition the textures were static and only the boundary between them moved across subsequent frames. With the motion of the boundary, dots from one surface were progressively deleted as dots from the other surface were revealed. However, since the textures do not move with the boundary, each boundary exhibits accretion and deletion simultaneously, as texture from the far surface is covered up and that from the near surface is exposed. In this condition, there is boundary motion but no texture motion; therefore, only accretion-deletion information is available. Figure 5c through f shows the results for this accretion-deletion stimulus plotted as thresholds versus syncing gain for each of the observers. The results show a substantial deterioration at low syncing gains for three out of four observers, whereas at high gains the performance is similar to that for the Cue-Consistent condition (Figure 2). The difference between Head Sync and Playback conditions

is similar to that seen earlier for the Cue-Consistent and Cue-Conflict conditions (Figures 2 and 3). A statistical two-way ANOVA test shows a significant difference between Head Sync and Playback conditions in the Accretion-Deletion Only condition for two of the observers, YA: $F(1, 16) = 15.58, p = 0.0012$; RA: $F(1, 16) = 4.20, p = 0.0572$; HA: $F(1, 16) = 8.03, p = 0.0120$; BA: $F(1, 16) = .06, p = 0.8105$. Furthermore, for the higher syncing gains of 0.10–1.0 the thresholds are similar to those for the Cue-Consistent condition (Figure 2). The fact that observers are able to do this task at all demonstrates that relative texture motion (expansion-compression) is not necessary to obtain segmentation from the dynamic occlusion stimulus. This result seems surprising in view of our previous demonstration that the same Accretion-Deletion Only stimulus did not provide depth at all, even with head synchronization (Yoonessi & Baker, 2013, figure 7).

Discussion

Segmentation from dynamic occlusion based motion parallax

Relative image motion of adjacent surfaces can be sufficient for segmentation (Baker & Braddick, 1982; Regan, 1986, 1989), but in naturally occurring motion parallax there is accompanying head movement. Thus from a theoretical point of view it is conceivable that the depth perception resulting from head movements in a motion parallax situation might facilitate segmentation performance, since the head movements can provide added information about an object's 3-D shape. Furthermore, previous results demonstrate that human observers are able to perceive reliable depth information from these visual stimuli (Yoonessi & Baker, 2013). However, we found that head movement did not improve segmentation performance compared to similar Playback measurements, across a wide range of syncing gains (Figures 2 through 4). This result is contrary to what one might expect, since in the head-synchronized conditions the observer has more information about the surface depth order and shape, which might facilitate segmentation. At smaller syncing gains the head movement actually appears to interfere with segmentation judgments, more so in the Cue-Consistent than in the Cue-Conflict condition, even though these conditions provide robust depth ordering (Yoonessi & Baker, 2013). Since for the small syncing gains the amount of texture motion is very small, such decrease in performance might be due to imperfections in fixation and subsequent loss of high spatial frequency information in the retinal image. Measure-

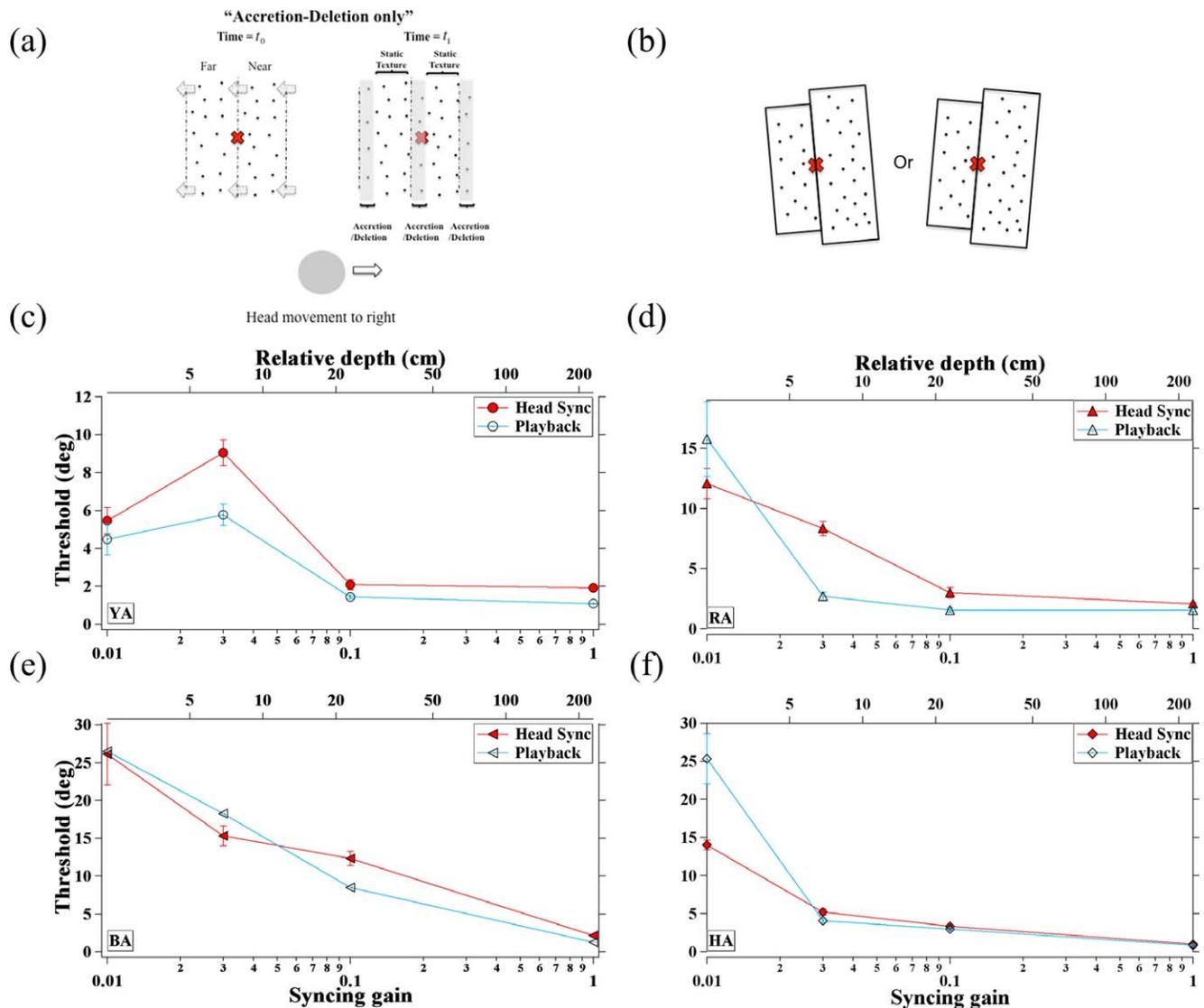


Figure 5. Segmentation performance for Accretion-Deletion Only condition. (a) Schematic depiction of the stimulus. The textures were static, but the boundary between them moved. (b) Cartoon drawing of the 2AFC orientation judgment task. (c–f) Results for four observers in the Head Sync and Playback conditions.

ments of eye movements in these conditions would be pertinent to resolving this question.

One might expect from the cue combination literature (e.g., Landy et al., 1995) that performance in the Cue-Conflict situation should be less than in the Cue-Consistent because of the lower number of reliable information sources, even if the depth only provides additional information about the surface shape and is not necessarily related to the judgment being made. However, the results (Figures 2 and 3) indicate that, in fact, the segmentation performance is somewhat better in the Cue-Conflict condition for the smaller syncing gains. This might be explained in terms of the relationship between the relative motion of the boundary and the observer. In the Cue-Consistent condition, the boundary motion is in the opposite

direction to the observer’s translational movements (Figure 2b). If the translational vestibulo-ocular reflex (TVOR) was perfect, retinal image motion of the boundary should be equal in speed (though opposite in direction) for Cue-Consistent versus Cue-Conflict conditions. However, if the TVOR is imperfect (TVOR gain less than unity; Ramat & Zee, 2003), then the boundary speed in the retinal image will be faster for the Cue-Consistent condition (even though in screen coordinates the speeds are identical), and this faster speed might degrade segmentation performance. In any case, it is interesting that in comparing Cue-Conflict with Cue-Consistent stimuli, with and without synchronization to head movement, the best performance in both cases is obtained with the most nonecological conditions.

Since the head movements were not physically constrained to a 1-D path, observers could have made small head movements in y - and z -directions, that were not incorporated in the rendering. Our analysis of the recorded head movement data suggested that the amount of movement in the z -direction is negligible, but there is a small amount of movement in the y -direction that was not utilized in creating the visual stimulus. However, it seems unlikely that the segmentation performance for the Head Sync conditions would significantly improve by adding a small amount of motion in the y -direction. In addition, since we employed orthographic rendering and there is no vertical shortening of distant edges (perspective cues), the surface is perceived as slightly rotating during observer movement. However, this small perceived rotation is identical in both Head Sync and Playback conditions, and cannot explain the differences in performance between these conditions.

Comparison to shear-based motion parallax

Dynamic occlusion is different from shear in that the stimulus contains accretion-deletion and also a moving boundary, whereas in shear, the location of the envelope boundary is stationary. Segmentation performance for a stationary observer has been demonstrated to be poorer in the case of moving envelopes rather than stationary envelopes for a stimulus containing expansion-compression only (Watson & Eckert, 1994). Based on this finding, one might expect the segmentation performance to be poorer for dynamic occlusion-based motion parallax than for pure shear.

In the segmentation results for dynamic occlusion (Figures 2 and 4) and for shear (figures 3 and 4, Yoonessi & Baker, 2011a), we found thresholds are very low for mid-range and higher syncing gains, with a progressive increase at the lower gains which is less pronounced for playback than head-synchronized motion. However, the thresholds at low syncing gains are slightly higher than those previously found for the shear, whereas at high gains the thresholds are similar or slightly lower (figures 3 and 4; Yoonessi & Baker, 2011a). Thus, in agreement with Watson and Eckert (1994), the segmentation from relative motion information might be weaker in the case of a moving envelope boundary, which is compensated for by the addition of accretion-deletion information at higher syncing gains (Yoonessi & Baker, 2013).

The two cues in dynamic occlusion exploit distinct parts of the motion flow field. Expansion-compression is a more global, region-based cue comprising motion vectors across extended regions, whereas accretion-deletion only occurs at the boundaries and therefore is a

more local, edge-based cue. There has been extensive debate on whether segmentation in human vision is edge-based or region-based (e.g., Biederman & Ju, 1988; Wolfson & Landy, 1998; Ben-Shahar & Zucker, 2003; Motoyoshi & Kingdom, 2010). An edge-based processing segregates two surfaces based on sharp differences between adjacent image attributes that are very close to a putative boundary, whereas a region-based approach relies on whether nearby extensive regions are similar or different. Therefore a region-based segmentation suggests a substantial similarity for shear and dynamic occlusion, whereas edge-based segmentation could utilize accretion-deletion information and thereby produce significant differences between the two. On the other hand, a system based solely on edge-detection would be unable to detect gradual gradients across surfaces. Our previous work with shear-based motion parallax (Yoonessi & Baker, 2011a) showed that segmentation performance was impaired for a sinusoidal depth modulation pattern compared to that for a square wave, particularly at small syncing gains. Though for technical reasons we were only able to test a limited range of syncing gains for dynamic occlusion, we found a similar pattern of results (Figure 4 compared to Figure 2): Thresholds decline markedly with decreasing syncing gains, with performance for the sine wave modulation at the smallest syncing gains becoming substantially lower than for square wave patterns. This result is consistent with findings for stationary observers (Watson & Eckert, 1994; Sachtler & Zaidi, 1995). This difference might be explained by the higher motion energy difference in the case of square wave modulation. Furthermore, if segmentation utilizes edge-based processing, then performance for square wave envelopes might benefit from the sharp edges in the texture motion at the boundaries. The difference in performance for sine and square wave modulations might also be due to accretion-deletion, as this cue was absent in the sine wave case. However, in light of the similarity between these results and the previous results for shear (Yoonessi & Baker, 2011a), this idea appears unlikely.

Combination of accretion-deletion and relative motion cues

Boundaries defined by dynamic occlusion differ importantly from those in shear motion, in that they contain accretion-deletion information in addition to relative texture motion. These information sources are presumably combined to perform a boundary-related psychophysical task. A simple weak fusion model of cue combination would suggest a weighted average of expansion-compression and accretion-deletion signals, with the weights adjusted according to informational content and reliability of the cue (Landy et al., 1995).

However, a strong fusion model would entail nonlinear summation, with interaction between the different information sources prior to their summation. We previously found that expansion-compression and accretion-deletion contribute in a quite asymmetric manner for depth from dynamic occlusion (Yoonessi & Baker, 2013). Our results showed that accretion-deletion can significantly facilitate the perception of depth, but it could not provide a depth percept in isolation. Therefore since our earlier results employing the same stimuli for a depth-ordering task were clearly incompatible with a weak fusion model, it might be expected that the same cue combination rules would govern segmentation (Yoonessi & Baker, 2013). However, if depth and segmentation are processed independently, the cue combination rules could be very different for segmentation.

In a stimulus containing only accretion-deletion (Figure 5), the texture on both sides of the boundary was static and only the boundary moved across successive frames. However, the results show that observers are still able to perform the segmentation task, albeit with reduced performance at low syncing gain, with performance very similar to that for the Cue-Consistent condition at higher syncing gains. Thus in marked contrast to previous depth results (Yoonessi & Baker, 2013), accretion-deletion alone can support segmentation. That being said, the converse is also true: The results for a sine wave modulation (Figure 4) demonstrate that expansion-compression alone can also support segmentation, albeit not quite as well as when both cues are present.

The number of accretion-deletion events is dependent on syncing gain, dot density and the span of the head movements, and increases in proportion to the magnitude of rendered depth (i.e., syncing gain). If only the accretion-deletion rate were determining the psychophysical performance, identical results for the Accretion-Deletion Only and the Head Sync conditions would have been expected, whereas the difference in performance for these conditions suggests that the accretion-deletion events do not fully account for improvement of segmentation performance at higher syncing gains. In addition, in Figures 2 and 3, note that the psychophysical performance remains relatively unchanged across a 10–50-fold increase in syncing gain, and consequently a proportionate increase in the number of accretion-deletion events. Thus the improvement in performance at high syncing gains is not simply correlated with the increase in accretion-deletion rate. However, results from the Accretion-Deletion Only condition (Figure 5) suggest that accretion and deletion by itself can contribute to segmentation for these values of syncing gain.

Taken together, these findings suggest that expansion-compression and accretion-deletion have similar

operating ranges for contribution to segmentation, unlike depth ordering in which the expansion-compression mainly contributed at low and accretion-deletion at high syncing gains (Yoonessi & Baker, 2013). At low syncing gains, expansion-compression primarily supports segmentation, though accretion-deletion alone can yield segmentation with reduced performance. At high syncing gains, pure accretion-deletion can provide performance almost as good as in the Cue-Consistent condition; however, our experiments cannot rule out a role for expansion-compression in this range of gains. It would make sense that accretion-deletion would play a greater role at higher syncing gains, due to the greater number of texture elements covered and uncovered (greater signal-to-noise ratio). These findings are compatible with a weak fusion model (Landy et al., 1995), in which each cue is able to provide segmentation in isolation, and the performance in the presence of both cues is better than for either cue alone. Thus the nature of cue summation in motion parallax is qualitatively different for segmentation and depth-ordering tasks.

Computer vision

In computer vision, segmentation is often improved by employing co-occurring depth information. A common approach is layered segmentation, in which depth is utilized to group motion into layers belonging to objects (Jepsen, Fleet, & Black, 2002; Sun, Sudderth, & Black, 2010, 2012). Depth order and occlusion cues can provide reliable information about image regions belonging to the same objects, and thereby improve segmentation. This approach can extract occlusion boundaries reliably, and often requires more than two frames, similar to motion parallax (Gruber & Weiss, 2006; Sun et al., 2010). Thus a computer algorithm can benefit from camera/observer movement since it adds an extra source of information about surface shape that can be used to improve segmentation performance. However, our results are contrary to this—we found that human observers have poorer performance in the Cue-Consistent conditions, in which extra information about depth is richly available and reliable (Figure 2). From a computer vision standpoint, the Cue-Conflict and Playback conditions should have yielded poorer psychophysical performance, because of the smaller number of information sources or their unreliability.

Possible neural mechanisms

Single- and double-motion opponent receptive field organization in visual neurons (Frost & Nakayama, 1983; Von Grunau & Frost, 1983; Allman, Miezin, &

McGuinness, 1985; Born, 2000; Pack, Hunter, & Born, 2005) could potentially contribute to both segmentation and depth from motion parallax. Selective neuronal responses to the orientation of motion-defined boundaries have been described in visual cortex of primates (Marcar, Xiao, Raigel, & Orban, 1995; Xiao, Raigel, Marcar, & Orban, 1997) and cats (Gharat & Baker, 2012). Single unit recordings from macaque monkeys in response to a shear-type stimulus synchronized to the animal's translation revealed responses in areas middle temporal (MT) and medial-superior temporal (MST) that were correlated with rendered depth order (e.g., Nadler, Angelaki, & DeAngelis, 2008; Nadler et al., 2009; Nadler et al., 2013). However, only single-opponent receptive field neurons provide useful information for depth, while both single- and double-opponent receptive fields could contribute to segmentation. This discrepancy in the availability of neurons might be an underlying reason for segmentation in motion parallax being more robust than depth (Yoonessi & Baker, 2011a, 2013). Investigations of neural responses to motion defined boundaries have so far neglected accretion-deletion, though it has been hypothesized that visual cortex neurons that are selectively responsive to second-order stimuli might be able to detect accretion-deletion information (Hegd e et al., 2004).

Neuronal responses are often described as “form-cue invariant” if their tuning properties do not depend on the cue defining the stimulus (Albright, 1992; Baker, 1999). Neuronal responses for boundaries defined by first or second order information have been found to be form-cue invariant in cats (Gharat & Baker, 2012) and in primates (Albright, 1992; Sary, Vogels et al., 1993). Our segmentation results obtained here are consistent with this idea, as the segmentation from expansion-compression or accretion-deletion is very similar and is not dependent on the cues defining the stimulus. This suggests that the same neurons might be able to detect motion-defined boundaries, irrespective of whether they are defined by expansion-compression or by accretion-deletion.

Conclusions

Our results have demonstrated that expansion-compression and accretion-deletion in isolation can provide reliable boundary segmentation and therefore, markedly unlike in depth perception, their contribution appears to obey a relatively simple summation. This difference suggests that segmentation and depth perception from dynamic occlusion might be served by distinct underlying mechanisms. In addition, taken together with our previous findings, it appears that movement of the observer can improve depth percep-

tion, but at the expense of weakening the ability to segment boundaries.

Keywords: motion parallax, segmentation, cue combination, dynamic occlusion

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 and Chris Pack for their useful comments during the course of this study. We also would like to thank our observers for their participation. An earlier report of findings discussed in this paper was presented at the annual meeting of the Vision Sciences Society (Yoonessi & Baker, 2011b).

Commercial relationships: none.

Corresponding author: Ahmad Yoonessi.

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

Address: McGill Vision Research, McGill University, Montreal, QC, Canada.

References

- Albright, T. D. (1992). Form-cue invariant motion processing in primate visual cortex. *Science*, 255(5048), 1141–1141.
- 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.
- Angelaki, D. E., & Hess, B. J. (2005). Self-motion-induced eye movements: Effects on visual acuity and navigation. *Nature Reviews Neuroscience*, 6(12), 966–976.
- Baker, C. L., Jr. (1999). Central neural mechanisms for detecting second-order motion. *Current Opinion in Neurobiology*, 9(4), 461–466.
- Baker, C. L., Jr., & Braddick, O. J. (1982). Does segregation of differently moving areas depend on relative or absolute displacement? *Vision Research*, 22(7), 851–856.
- Ben-Shahar, O., & Zucker, S. W. (2003). The perceptual organization of texture flow: A contextual inference approach. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 25(4), 401–417.
- Biederman, I., & Ju, G. (1988). Surface versus edge-

- based determinants of visual recognition. *Cognitive Psychology*, 20(1), 38–64.
- Born, R. T. (2000). Center-surround interactions in the middle temporal visual area of the owl monkey. *Journal of Neurophysiology*, 84(5), 2658–2669.
- Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision*, 10, 433–436.
- Craton, L. G., & Yonas, A. (1990). Kinetic occlusion: Further studies of the boundary flow cue. *Perception & Psychophysics*, 47(2), 169–179.
- Frost, B. J., & Nakayama, K. (1983). Single visual neurons code opposing motion independent of direction. *Science*, 220(4598), 744–745.
- Gharat, A. M., & Baker, C. L., Jr. (2012). Motion-defined contour processing in early visual cortex. *Journal of Neurophysiology*, 108, 1228–1243.
- Gibson, J. J., Kaplan, G. A., Reynolds, H. N., & Wheeler, K. (1969). The change from visible to invisible: A study of optical transitions. *Perception & Psychophysics*, 5, 113–116.
- Gruber, A., & Weiss, Y. (2006). Incorporating non-motion cues into 3D motion segmentation. *Proceedings of the 9th European Conference on Computer Vision, Part III* (pp. 84–97). Springer: Berlin Heidelberg.
- Hegd e, J., Albright, T. D., & Stoner, G. R. (2004). Second-order motion conveys depth order information. *Journal of Vision*, 4(10):1, 838–842, <http://www.journalofvision.org/content/4/10/1>, doi:10.1167/4.10.1. [PubMed] [Article]
- Jepsen, A. D., Fleet, D. J., & Black, M. J. (2002). A layered motion representation with occlusion and compact spatial support. In *Computer Vision—ECCV 2002* (pp. 692–706). Springer: Berlin Heidelberg.
- Kaplan, G. A. (1969). Kinetic disruption of optical texture: The perception of depth at an edge. *Attention, Perception, & Psychophysics*, 6(4), 487–492.
- Kleiner, M., Brainard, D., Pelli, D., Ingling, A., Murray, R., & Broussard, C. (2007). What’s new in Psychtoolbox-3. *Perception*, 36(14), 1.
- Kromrey, S., Bart, E., & Hegd e, J. (2011). What the “moonwalk” illusion reveals about the perception of relative depth from motion. *PLoS ONE*, 6(6), e20951.
- 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.
- Marcar, 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(3), 1258.
- Motoyoshi, I., & Kingdom, F. A. A. (2010). The role of co-circularity of local elements in texture perception. *Journal of Vision*, 10(1):3, 1–8, <http://www.journalofvision.org/content/10/1/3>, doi:10.1167/10.1.3. [PubMed] [Article]
- Nadler, J. W., Angelaki, D. E., & DeAngelis, G. C. (2008). A neural representation of depth from motion parallax in macaque visual cortex. *Nature*, 452(7187), 642–645.
- Nadler, J. W., Barbash, D., Kim, H. R., Shimpf, S., Angelaki, D. E., & DeAngelis, G. C. (2013). Joint representation of depth from motion parallax and binocular disparity cues in macaque area MT. *Journal of Neuroscience*, 33(35), 14061–14074.
- 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(4), 523–532.
- Nakayama, K., Silverman, G. H., Macleod, D. I. A., & Mulligan, J. (1985). Sensitivity to shearing and compressive motion in random dots. *Perception*, 14(2), 225–238.
- Ono, H., Rogers, B. J., Ohmi, M., & Ono, M. E. (1988). Dynamic occlusion and motion parallax in depth perception. *Perception*, 17(2), 255–266.
- 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(3), 1809–1815.
- Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, 10(4), 437–442.
- Ramat, S., & Zee, D. (2003). Ocular motor responses to abrupt interaural head translation in normal humans. *Journal of Neurophysiology*, 90(2), 887–902.
- 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(10), 1389–1400.
- Rogers, B. J., & Graham, M. (1979). Motion parallax as an independent cue for depth perception. *Perception*, 8(2), 125–134.
- Rogers, B. J., & Graham, M. (1982). Similarities

- between motion parallax and stereopsis in human depth perception. *Vision Research*, 22(2), 261–270.
- Rogers, B. J., & Graham, M. (1983). Anisotropies in the perception of three-dimensional surfaces. *Science*, 221(4618), 1409–1411.
- Sachtler, W., & Zaidi, Q. (1995). Visual processing of motion boundaries. *Vision Research*, 35(6), 807–826.
- Sary, G., Vogels, R., & Orban, G. A. (1993). Cue-invariant shape selectivity of macaque inferior temporal neurons. *Science*, 260(5110), 995–997.
- Shimojo, K., Silverman, G. H., & Nakayama, K. (1989). Occlusion and the solution to the aperture problem for motion. *Vision Research*, 29(5), 619–626.
- Sun, D., Sudderth, E. B., & Black, M. J. (2010). Layered image motion with explicit occlusions, temporal consistency, and depth ordering. *Advances in Neural Information Processing Systems*, 23, 2226–2234.
- Sun, D., Sudderth, E. B., & Black, M. J. (2012). Layered segmentation and optical flow estimation over time. In *Computer Vision and Pattern Recognition (CVPR)*, 2012 IEEE Conference on (pp. 1768–1775). IEEE.
- Thompson, W. B., Mutch, K. M., & Berzins, V. A. (1985). Dynamic occlusion analysis in optic flow fields. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 7(4), 374–383.
- 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(1), 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(2), 496–505.
- Wexler, M., Panerai, F., Lamouret, I., & Droulez, J. (2001). Self-motion and the perception of stationary objects. *Nature*, 409, 85–88.
- Wexler, M., & Van Boxtel, J. J. (2005). Depth perception by the active observer. *Trends in Cognitive Sciences*, 9(9), 431–438.
- Wolfson, S. S., & Landy, M. S. (1998). Examining edge- and region-based texture analysis mechanisms. *Vision Research*, 38(3), 439–446.
- 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(7), 662–677.
- Yonas, A., Craton, L. G., & Thompson, W. B. (1987). Relative motion: Kinetic information for the order of depth at an edge. *Perception & Psychophysics*, 41(1), 53–59.
- Yoonessi, A., & Baker, C. L., Jr. (2011a). 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. [PubMed] [Article]
- Yoonessi A., & Baker, C. L., Jr. (2011b). Segmentation and depth from motion parallax-induced dynamic occlusion. *Journal of Vision*, 11(11):62, <http://www.journalofvision.org/content/11/11/62>, doi:10.1167/11.11.62. [Abstract]
- Yoonessi A., & Baker, C. L., Jr. (2013). Depth perception from dynamic occlusion in motion parallax: Roles of expansion-compression vs. accretion-deletion. *Journal of Vision*, 13(12):10, 1–16, <http://www.journalofvision.org/content/13/12/10>, doi:10.1167/13.12.10. [PubMed] [Article]