

Depth perception from dynamic occlusion in motion parallax: Roles of expansion-compression versus accretion-deletion

Ahmad Yoonessi

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



Curtis L. Baker, Jr.

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



Motion parallax, or differential retinal image motion from observer movement, provides important information for depth perception. We previously measured the contribution of shear motion parallax to depth, which is only composed of relative motion information. Here, we examine the roles of relative motion and accretion-deletion information in dynamic occlusion motion parallax. Observers performed two-alternative forced choice depth-ordering tasks in response to low spatial frequency patterns of horizontal random dot motion that were synchronized to the observer's head movements. We examined conditions that isolated or combined expansion-compression and accretion-deletion across a range of simulated relative depths. At small depths, expansion-compression provided reliable depth perception while accretion-deletion had a minor contribution: When the two were in conflict, the perceived depth was dominated by expansion-compression. At larger depths in the cue-conflict experiment, accretion-deletion determined the depth-ordering performance. Accretion-deletion in isolation did not yield any percept of depth even though, in theory, it provided sufficient information for depth ordering. Thus, accretion-deletion can substantially enhance depth perception at larger depths but only in the presence of relative motion. The results indicate that expansion-compression contributes to depth from motion parallax across a broad range of depths whereas accretion-deletion contributes primarily at larger depths.

motion is highly dependent on the scene layout (Helmholtz, 1925; Gibson, Gibson, Smith, & Flock, 1959). This visual cue, often called motion parallax, provides powerful information about the boundaries between objects and their relative depth differences; it can reliably provide 3-D scene layout and help enable navigation in the environment (Helmholtz, 1925). Two distinct types of motion boundaries are formed as a result of this retinal image motion. Boundaries that are parallel to the direction of observer movement provide a shearing motion, in which the only source of information regarding depth differences is relative motion. Boundaries that are orthogonal to the direction of observer movement provide dynamic occlusion, in which the front surface dynamically covers and uncovers the far surface. In this situation, two sources of depth information are available: relative motion of texture elements, more specifically, the optic flow component “expansion-compression,” which is comparable to that of the shearing motion, and, additionally, the covering and uncovering of parts of the farther texture, i.e., “accretion-deletion,” which can provide powerful information for depth sign (Yonas, Craton, & Thompson, 1987). Thus accretion-deletion and expansion-compression are two distinct components of the dynamic occlusion phenomenon.¹

Most previous studies of motion parallax examined the simpler case of shearing motion (Rogers & Graham, 1979, 1982; Ujike & Ono, 2001; Yoonessi & Baker, 2011a), demonstrating that it could provide reliable depth ordering and magnitude. Controlling a shear motion stimulus is simpler, and the results are easier to interpret due to the absence of multiple cues. But while dynamic occlusion provides significantly

Introduction

During active observer movement in the natural environment, the resultant pattern of retinal image

Citation: Yoonessi, A., & Baker, C. L., Jr. (2013). Depth perception from dynamic occlusion in motion parallax: Roles of expansion-compression versus accretion-deletion. *Journal of Vision*, 13(12):10, 1–16, <http://www.journalofvision.org/content/13/12/10>, doi:10.1167/13.12.10.

more powerful information for depth (Cutting & Vishton, 1995), its contribution to motion parallax has only begun to be explored (Rogers & Graham, 1983; Ono, Rogers, Ohmi, & Ono, 1988). The relative motion information in dynamic occlusion is comparable to that of shear motion parallax whereas accretion and deletion of texture on partially occluded surfaces provides additional, particularly powerful information for depth ordering. In this paper, we will examine how these two kinds of information contribute to depth from motion parallax with dynamic occlusion.

Accretion-deletion and expansion-compression are fundamentally different in their dependence on head movements for effective contribution to depth. Dynamic occlusion stimuli containing accretion and deletion of texture elements can provide reliable depth percepts in the absence of accompanying head movement (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). Depth relationships from relative motion (either expansion-compression or shear), however, are ambiguous without head movement information, i.e., relative motion information by itself cannot disambiguate depth sign. Thus, in the absence of accretion-deletion, the depth order from relative motion can only be disambiguated with extraretinal information, which normally arises from synchronous head movement. It is an open question whether the extraretinal information accompanying head movement might also enhance the utility of accretion-deletion for depth perception.

In natural motion parallax, accretion and deletion of textures occurs in association with synchronous boundary motion. Hence, accompanying boundary motion might play an important role in how accretion-deletion contributes to depth. On the other hand, utilizing relative motion information does not particularly depend on the motion of the boundary. If, in the absence of the boundary motion, depth perception disappeared, it would suggest a greater role for accretion-deletion. But if, in the absence of relative motion information, depth perception changed significantly, a greater role for expansion-compression might be suggested.

Here, we devised experimental conditions to examine the contributions of relative motion (expansion-compression) and accretion-deletion to depth perception. Depth perception was assessed with a psychophysical task in which observers performed two-alternative forced choice (2AFC) judgments of the perceived relative depth order of two surfaces. The stimulus consisted of alternating strips of random dots, whose relative motion was synchronized to the observer's

head movement in such a way as to simulate a range of relative depths. In order to assess how dynamic occlusion supports depth perception across a broad range of simulated depth, we devised a “Cue-Consistent” condition in which the expansion-compression and accretion-deletion cues signaled similar depth signs, mimicking ecological conditions. Furthermore, we tested the relative strengths of expansion-compression and accretion-deletion cues by creating a “Cue-Conflict” condition (which would not occur in the natural world), in which the two cues were placed in conflict with one another and signaled opposing depth signs.

To further analyze the contribution of each of these sources of information to psychophysical performance and assess the nature of cue combination, we designed experimental conditions with different levels of contribution for expansion-compression and accretion-deletion. In order to minimize the contribution of accretion-deletion, we employed a “Fixed-Boundary” condition, which would greatly reduce the contribution of accretion-deletion. To measure how accretion-deletion supports depth perception when accompanying expansion-compression is ambiguous, we tested a “Playback” condition in which observers were stationary, and only accretion-deletion information could disambiguate the depth. We further investigated to what extent depth is obtainable by expansion-compression or accretion-deletion in isolation by creating “Transparent” and “Accretion-Deletion-Only” conditions, in which the only cue to disambiguate the depth was expansion-compression or accretion-deletion, respectively. The results indicate that expansion-compression contributes to depth from motion parallax across a broad range of depths although more so at smaller depths. Accretion-deletion alone is unable to provide any depth perception; however, it acts powerfully at larger depths to facilitate co-occurring relative motion, regardless of its informational validity.

Materials and methods

Our hardware setup for producing head-synchronized relative motion stimuli was described in detail in our previous study (Yoonessi & Baker, 2011a), and will be only briefly summarized here. 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 while using a pair of vertical bars as guides for the range of movement. The head position/orientation data for every trial was recorded for later analysis.

Visual stimuli

The stimuli were generated with a Macintosh (Mac Pro, 2 × 2.8 GHz, 4 GB Ram, OSX v10.5) using Matlab code and the Psychophysics Toolbox (Brainard, 1997). Stimuli were presented on a CRT monitor (Trinitron A7217A, 1024 × 768 pixels, 75 Hz), which was gamma-corrected with a mean luminance of 40 cd/m² and viewed from a distance of 114 cm.

Texture patterns consisted of white (80.31 cd/m²) dots on a black (0.07 cd/m²) background (density 1.04 dots/deg²). Each dot was of circular shape (0.2°), produced with high-quality anti-aliasing; the dots remained the same shape and size in all conditions. We utilized two methods for drawing dots. In the first method, which was used for the conditions without accretion-deletion, we controlled individual dot positions on each frame. In the conditions with accretion-deletion, we employed OpenGL texture mapping to change the location of a texture region between consecutive frames. In order to keep the conditions comparable, the textures used in the second method were image-captured from drawings produced by the first method. In order to ensure dot textures were identical, we measured their physical envelope and dot size on the screen to ensure that texture patterns produced by the two methods were indistinguishable. For the Accretion-Deletion-Only condition, we employed OpenGL Shading Language to process the stimulus on the graphics card rather than the host computer. This enabled us to move the transparency layer (alpha layer) separately from the image texture map quickly enough for it to be synchronized to the head movement without any noticeable lag.

To simulate a motion parallax situation, the motions of the dots were synchronized to measured changes in head position (Figure 1a, see below). On each frame update, the difference between current and previous head position on the horizontal axis was calculated and then multiplied by a scaling factor (i.e., syncing gain); the 1-D modulation profile (square wave) was multiplied by this number and used to modulate the dots' horizontal displacement. In order to obtain good real-time performance, the “dentsync” settings in Psychtoolbox were used for drawing. Therefore, the stimulus drawing was synced to the vertical retrace, but execution of the program was not paused until the vertical retrace had been reached. This option results 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 movements 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 might be a better representation of information obtained from motion parallax than velocity values or the equivalent disparity because it matches the geometry of movement and is proportional to rendered depth. Furthermore, in the case of motion parallax, absolute velocity values do not possess significant information as the same visual scene could give rise to different retinal image velocities depending on the observer's own velocity of movement. More details about this parameter can be found in our previous study (Yoonessi & Baker, 2011a).

The spatial frequency of the modulation of horizontal dot displacement 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° of visual angle, which resulted in about 1.5 cycles/image of visible modulation.

The sine-phase modulation waveform produced bidirectional texture motion, which corresponds to peaks and troughs of rendered depth, moving in opposite directions. Such 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 at the zero-crossing point of the modulation waveform (Figure 1b). The fixation point served to maintain the same pattern of retinal image motion across conditions, observers, and trials.

Head movement recording

Head position and orientation data in six degrees of freedom were recorded using an electromagnetic position-tracking device (Flock of Birds, Ascension Technologies, Shelburne, VT) with a medium-range transmitter, which provided position (0.5 mm resolution) and orientation (0.1° resolution) of a sensor secured on the observer's forehead. The head movement data were sampled at 100 Hz and transferred to the computer using a serial port/USB connection. The change in horizontal (X) position was used for real-time modulation of the stimulus motion as described above, and the complete position and orientation data were recorded to hard disk for subsequent analysis.

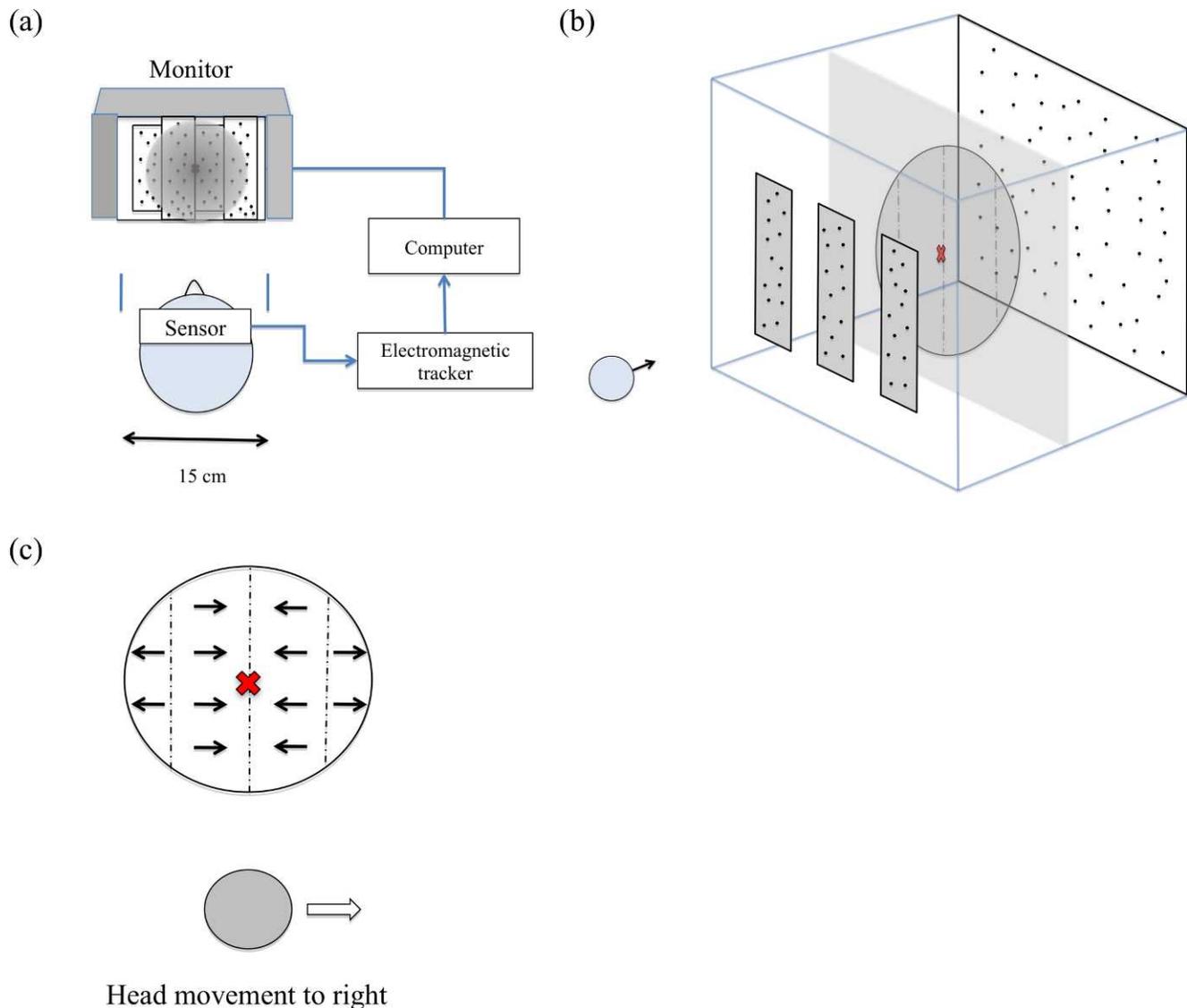


Figure 1. Experimental setup for motion parallax experiments. (a) As human observer moves laterally, computer updates visual stimuli on monitor in synchrony with head position provided by electromagnetic tracking of sensor placed on observer's forehead. (b) The 3-D volume that was simulated, in which near opaque surfaces consisting of vertical strips of random dot textures occlude a far surface of random dot texture. The projection of these surfaces was rendered on a virtual plane located halfway between the near and far surfaces, indicated in gray in the cartoon drawing. A circular aperture was imposed, leaving about 1.5 cycles visible to the observer. (c) 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.

Observers were instructed to perform only a lateral head translation, using two vertical bars as guides for end points of the movement (Figure 1a). The observer's head movement was not physically constrained to a 1-D path as in previous motion parallax experiments (Rogers & Graham, 1979, 1982), and there was no auditory signal provided for synchronization of head movements; consequently, there was potential for variance in psychophysical results due to differences in head movement between observers and even from one trial to the next for a given observer. Furthermore,

translational vestibulo-ocular reflex eye movements could depend on the head acceleration values during translation. However, our analysis of the recorded head movement data showed that the temporal frequency and the velocity of head movements were very similar across trials (average velocity for three observers: RA: 20.76 ± 1.89 , YA: 22.27 ± 2.20 , IVI: 16.83 ± 1.62 , GE: 18.65 ± 1.35 cm/s). We previously (Yoonessi & Baker, 2011a) analyzed head movement span, velocity, acceleration, and Fourier power spectrum for our shear experiment using this same hardware setup. The results

of head movement analysis in this study show strong similarity to our previously analyzed head movement data. Observers, on average, performed lateral head movements with similar velocity and temporal frequency. 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 movements.

Psychophysical task and procedures

Depth-ordering performance was measured using a 2AFC depth-ordering task. The observer's task was to judge which side of the vertical boundary edge (left or right) appeared nearer. The fixation point was always visible on the screen even when the near-rendered surface occluded the area around it. This creates a conflict because the fixation point is on the monitor plane and should logically be occluded in ecological conditions. However, without this compromise, the possibility of uncontrolled eye movements could introduce significant variance in the retinal image motion within and across trials.

Each value of syncing gain (0.01–0.3), modulation pattern, and cue condition was tested using a method of constant stimuli. In each block of trials, two values of syncing gains (i.e., 0.01 and 0.1) were presented in a random order with 20 repetitions of each. Only two gain values were tested together to prevent excessively long trial blocks. Trial blocks were accumulated such that each syncing gain value was tested at least 60 times.

The maximum syncing gain tested in this study was constrained due to consequences of the boundary motion: For larger syncing gains, the amount of image motion for each texture region became a significant fraction of the display size, such that the consequent “wraparound” would make the depth-ordering task ill-defined. To avoid this problem, we did not test syncing gains above 0.30. For the viewing distance of 114 cm used here, the range of tested syncing gains (0.01 to 0.30) resulted in simulated relative depths of about 1–65 cm.

The stimuli were viewed monocularly to avoid a cue conflict with stereopsis. The psychophysical tasks were performed in normal room illumination, and observers were able to see other parts of the room. This could be a source of cue conflict because other objects in the room might provide information about the flatness of the rendering. We investigated this issue in preliminary trials using a cardboard black tunnel to prevent observers from seeing any other parts of the visual field but found no effects on the depth results.

Observers

Five observers (YA, RA, IVI, GE, and CS) with normal or corrected-to-normal vision participated—four of whom were naive to the purpose of the experiment. Two of the observers (YA, RA) participated in all of the experiments, and the others were only employed in a subset of the conditions so as to accumulate at least four observers for each experiment. 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 depth perception for our dynamic occlusion stimulus with both expansion-compression and accretion-deletion cues present in a manner like that which occurs during natural motion parallax. This Cue-Consistent stimulus condition is depicted schematically in Figure 2a; in this and subsequent diagrams for other stimulus conditions, small solid arrows indicate random dot texture motion within the envelope half cycles closest to the fixation point (red “X”). The texture motion used here is always orthogonal to the boundary and in opposite directions on adjacent sides of each boundary, i.e., expansion-compression. The large open arrows indicate envelope boundary motion while the gray and red stippled areas indicate regions where accretion and deletion, respectively, occur. For this Cue-Consistent condition, the textured surface rendered as nearer to the observer moves oppositely to the observer movement. The far-rendered surface moves in the same direction as the observer movement, and its texture is covered and uncovered by that of the nearer surface. Consequently, texture deletion occurs along the leading edge of the near surface, and accretion occurs along its trailing edge.

All observers experienced good perception of adjacent surfaces in distinct depth planes. In agreement with the subjective perception, the depth-ordering performance was very good across the full range of simulated relative depths for all the observers (Figure 2b).

We were concerned that possible effects of syncing gain here might have been masked by a ceiling effect. To address this question, we added coherence noise to the stimulus for representative low and high values of the syncing gain (0.02 and 0.20) by changing the amount of motion for a fraction of the dots to a random amount proportional to the observer's head

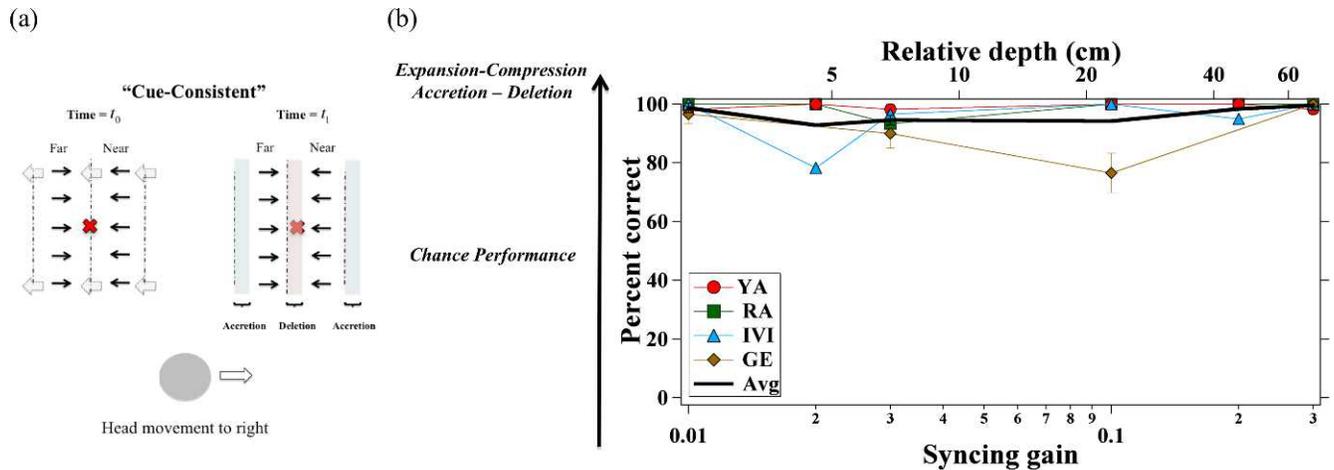


Figure 2. Depth ordering for Cue-Consistent condition. (a) Schematic depiction of stimulus in which both expansion-compression and accretion-deletion cues are synchronized to head movement in an ecologically valid manner. In this and subsequent stimulus schematics, small arrows indicate motion of random dot textures (expansion-compression), large open arrows depict boundary motion, and blue/red stippling indicates accretion-deletion. Fixation point at red “X.” (b) Performance in 2AFC depth-ordering task, in which observers judged which side appeared nearer (scored correct if the side moving opposite to head movement was judged as near) versus syncing gain for four observers (average indicated by thick black line). Error bars, here and in subsequent figures, indicate \pm SE.

movement (for details, see Yoonessi & Baker, 2011a). The results (see Supplementary Materials) showed that noise degraded performance more at a low syncing gain whereas it had remarkably little effect on the performance at a high syncing gain (Supplementary Materials, Figure S1). A possible reason for this result might be that at low syncing gains, the information from accretion-deletion is less because fewer texture elements are covered and uncovered, thereby making this weaker information more vulnerable to noise. However, at high syncing gains the information from accretion-deletion is progressively greater so that a given amount of coherence noise would degrade performance much less.

In the above stimulus, as in naturally occurring motion parallax, the expansion-compression information is ambiguous as to depth-sign unless it is accompanied by synchronous head movements whereas the accretion-deletion cue is unambiguous even for a stationary observer. To assess the role of head movement, we tested the Playback condition (Figure 3a), in which the visual stimulus was recreated from previously recorded head movement data (see Yoonessi & Baker, 2011a, for more details). Therefore, between the two conditions (Cue-Consistent and Playback), the visual stimulus is identical, and the only difference is in the extra retinal information from head movement.

Depth-ordering performance for four observers in the Playback condition is shown in Figure 3b. The responses were scored as correct when the judged depth ordering was consistent with the accretion-deletion cue because only that cue contained unambiguous depth information. The performance was near chance at low

syncing gains but almost perfect at high syncing gains. This result suggests that observers could effectively utilize the accretion-deletion cue at larger but not at smaller relative depths.

The results so far suggest a greater role of accretion-deletion at higher syncing gains and expansion-compression at lower syncing gains but provide no indication of how the cues are combined. To examine the interaction of these two cues in depth ordering, we measured performance in a Cue-Conflict condition, in which the relative motion and accretion-deletion signaled opposite depth signs (Figure 4a). In this condition, the stimulus region moving opposite to head movements, which corresponds to the near surface as rendered by the relative motion cue, was covered by the region moving in the same direction as head movement; thus, the far surface dynamically covers and uncovers the near surface. Because there were two conflicting sources of depth-order information, there was no logically correct manner in which to score the correctness of performance. We scored a given response as correct if the side moving oppositely to head movement was chosen as appearing nearer, which was consistent with the relative motion information. The results (Figure 4b) show that for low syncing gains, the perceived depth sign corresponded to the depth provided by expansion-compression information whereas at high syncing gains (above 0.10) the perceived depth corresponded to that provided by the accretion-deletion cue. At low syncing gains, the stimulus appeared perceptually similar to the Cue-Consistent conditions, and observers were unaware of the unrealistic nature of the stimulus. However, at

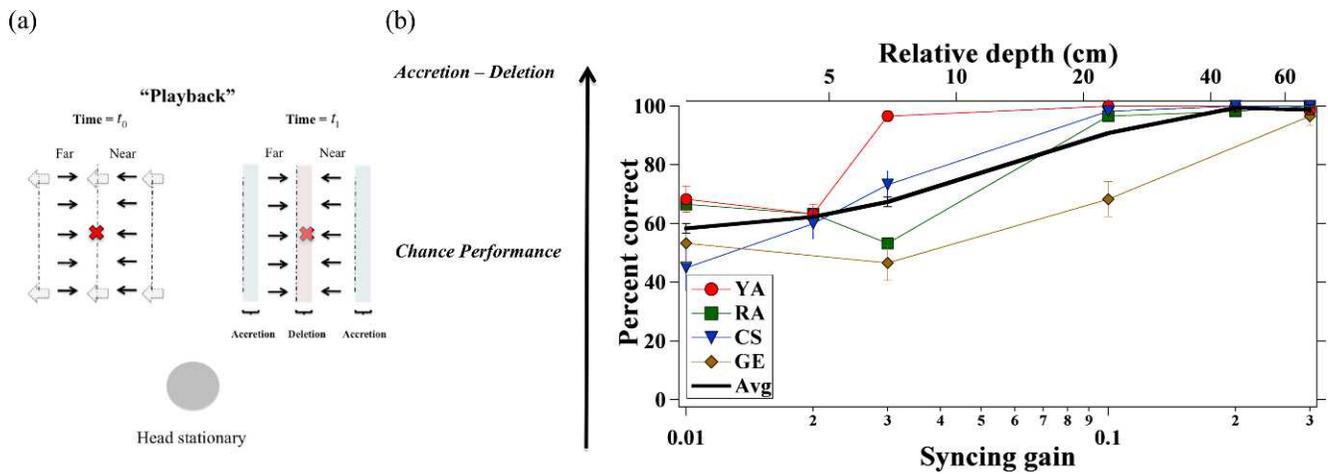


Figure 3. Depth ordering for Playback condition, shown in same format as Figure 2. (a) Visual stimuli, same as Cue-Consistent condition, but observer’s head is stationary, and stimulus motion is controlled by previously recorded head movements of the same observer in the Cue-Consistent condition. Note that, in this condition, the only unambiguous cue to depth is the accretion-deletion. (b) Performance for four observers, scored as correct if consistent with the depth ordering rendered by accretion-deletion.

higher syncing gains observers reported the stimulus appeared unnatural although they were unaware of what caused this phenomenon. Thus, in agreement with Ono et al. (1988), expansion-compression contributes substantially to depth ordering—so much so that it can dominate conflicting accretion-deletion at small relative depths. Conversely, at larger simulated depths, the accretion-deletion cue dominates conflicting information from relative motion.

In the stimuli used above, the boundary between adjacent texture regions moved back and forth with the observer’s head movement; normally, boundary motion is inextricably linked to both accretion-deletion and expansion-compression. To examine the role of boundary motion, we created a Fixed-Boundary

condition in which the texture surfaces were moving, but their boundary was stationary (Figure 5a). The texture within each region moved consistently with the observer’s head movement, but each dot disappeared as it reached the border and reappeared at a random vertical position along the other side of the boundary with the total number of dots in each region always being constant. Thus, in this condition, there was no boundary motion, only relative texture motion (expansion-compression) with accretion-deletion along the stationary boundary. The Fixed-Boundary condition possesses the same amount of texture motion energy as the Cue-Consistent condition (and all other previous conditions) but without boundary motion. Note that in this condition the accretion-deletion information is

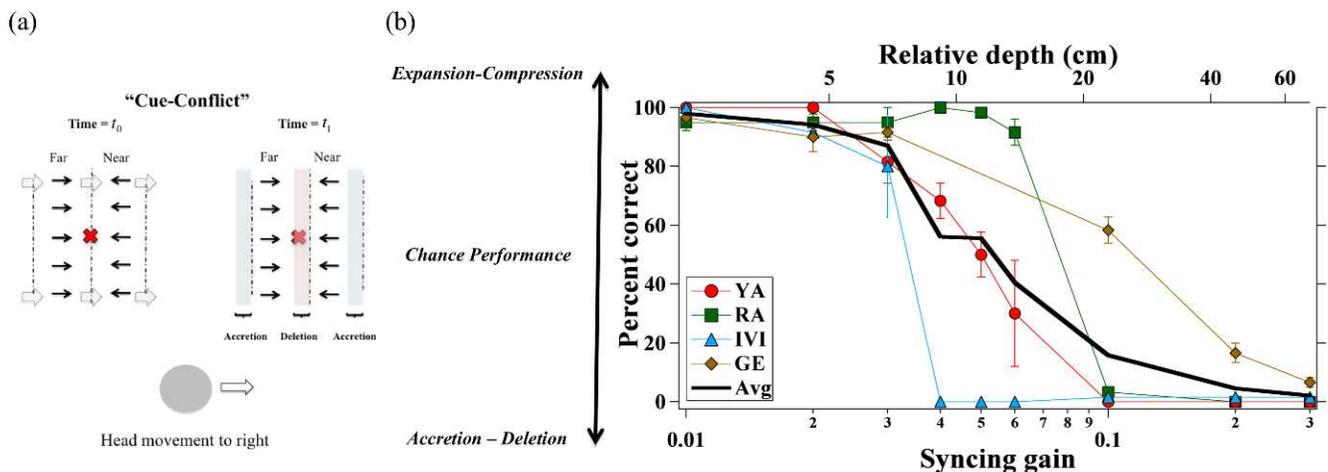


Figure 4. Depth ordering for Cue-Conflict condition. (a) Schematic depiction of the stimulus in which expansion-compression and accretion-deletion signal opposing depth signs. Thus, the texture rendered as near by the relative motion (expansion-compression) cue is covered and uncovered by the far-rendered surface, a condition that cannot exist in ecological conditions. (b) Performance for four observers, scored as correct if consistent with the expansion-compression cue.

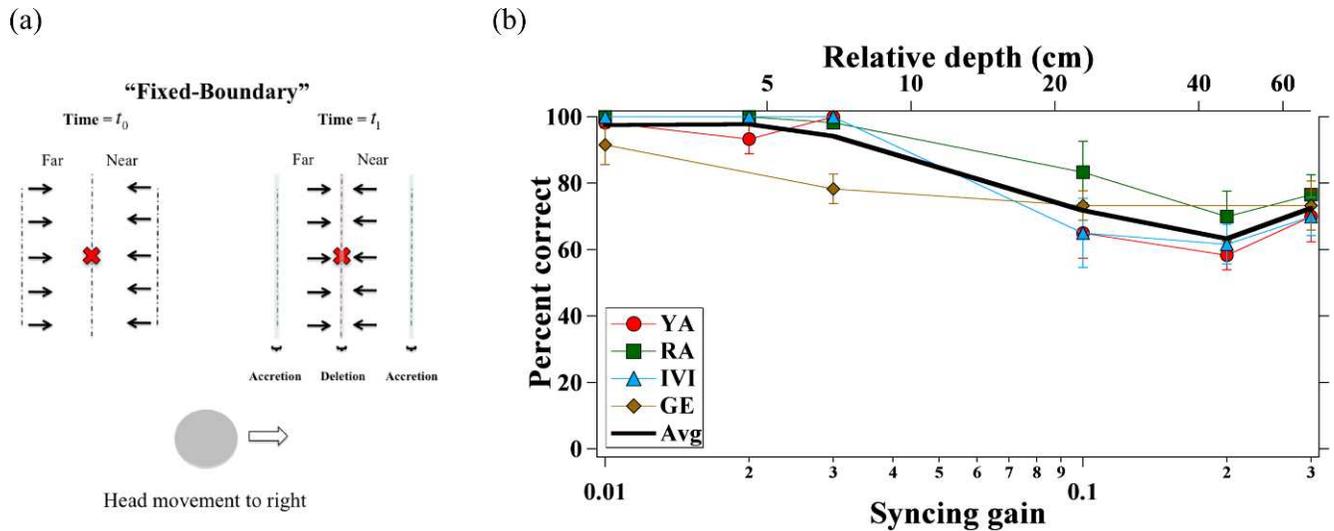


Figure 5. Depth ordering for Fixed-Boundary condition. (a) Schematic depiction of stimulus in which boundaries are stationary, but the texture in each region moves consistently with head movement. Note that accretion-deletion at boundaries in this condition does not convey depth-ordering information. (b) Performance for four observers.

depth-ambiguous because there is always an equal amount of deletion (for compressive motion) or accretion (for expansive motion) on both sides of the boundary whereas in ecological conditions the amount of accretion and deletion on two sides of the moving boundary is very different and therefore unambiguous for depth ordering.

The psychophysical performance for this condition (Figure 5b) was very good for low syncing gains. However, with an increase in syncing gain, the performance gradually dropped although never quite to chance levels. Therefore, the performance is similar to that for the Cue-Consistent condition (Figure 2) at small but not at high syncing gains. A statistical two-way ANOVA test shows that the results for the Fixed-Boundary condition are significantly different from those for the Cue-Consistent (Figure 2b) condition, YA: $F(1, 24) = 72.81, p < 0.0001$; RA: $F(1, 24) = 21.85, p < 0.0001$; GE: $F(1, 16) = 10.49, p < 0.0051$; IVI: $F(1, 24) = 24.85, p < 0.0001$. These results suggest that expansion-compression can support depth ordering, albeit only at small relative depths, even when the accompanying accretion-deletion information is depth-ambiguous and even in the absence of a moving boundary.

The dependence of depth-ordering performance on relative depth in the Fixed-Boundary condition is complementary to that for the Playback condition; however, both results are consistent with a common principle that when one of the information sources is depth-ambiguous the other source can effectively contribute at syncing gains for which it performs best. These results, taken together, reinforce the idea suggested from the Cue-Conflict experiment and the earlier report of Ono et al. (1988), that accretion-

deletion contributes more to depth perception at larger rendered depths, and relative motion plays a greater role at smaller rendered depths.

All the stimuli so far contained both accretion-deletion and relative motion (expansion-compression) information. We wondered to what extent depth perception could be obtained solely from the expansion-compression cue without any accretion-deletion; the previous result suggests that depth ordering might still be good at low syncing gains. In the “Transparent” condition (Figure 6a), each texture region was rendered as if it were transparent rather than opaque as in all our previous stimuli, thereby avoiding any accretion-deletion. The relative texture motion was in sync with the boundary motion (similar to the Cue-Consistent condition), but the textures could transparently superimpose upon or pull apart from one another; thus, regions with zero and doubled dot densities were created alternately as the observer moved back and forth. While this variation in texture density does not occur in natural dynamic occlusion, note that it does not provide any cue to depth ordering.

The results (Figure 6b) show that all observers performed depth ordering very well at low syncing gains, and the performance remained well above chance at higher syncing gains for three of the four observers. This condition further demonstrates that the good performance at high syncing gains in the Cue-Consistent condition (Figure 2) is not solely due to the presence of accretion-deletion as there was no accretion-deletion in this stimulus. However the performance at high syncing gains was not as good as in the Cue-Consistent condition (Figure 2). A statistical two-way ANOVA test shows that the results for the Transparent condition are significantly different from

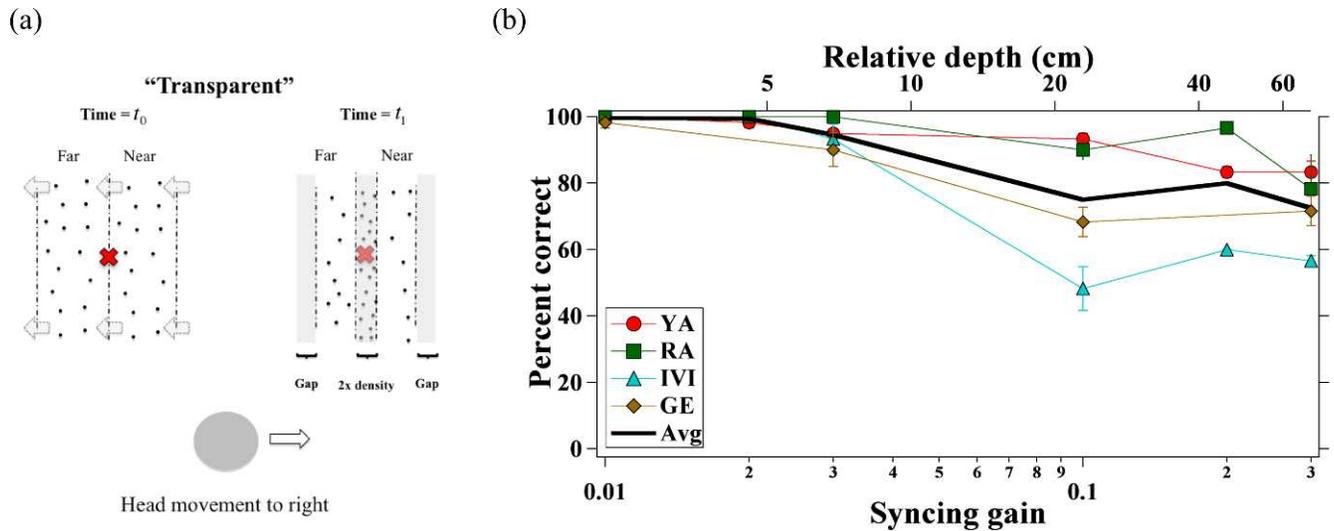


Figure 6. Depth ordering for Transparent condition. (a) Schematic depiction of Transparent condition in which the textures were treated as if they were transparent and without any accretion-deletion. As motion progresses, regions with zero and doubled dot densities are created alternately (gray areas) as the observer moves back and forth. (b) Performance for four observers.

those in the Cue-Consistent condition, YA: $F(1, 24) = 48.41$, $p < 0.0001$; RA: $F(1, 24) = 6.47$, $p = 0.0179$; GE: $F(1, 16) = 8.34$, $p = 0.0107$; IVI: $F(1, 24) = 119.71$, $p < 0.0001$. In summary, the results from this condition show that expansion-compression, in the absence of accretion-deletion, is sufficient to provide good depth ordering at smaller relative depths and in at least some observers a limited degree of performance at larger relative depths.

The results from the preceding experiments all support the idea that accretion-deletion is relatively more useful at higher syncing gains. This idea is consistent with the greater amount of information provided by this cue at higher syncing gain values, i.e., for each back-and-forth motion, a greater amount of texture is covered and uncovered. The results so far are also consistent with a quasilinear cue summation (e.g., Landy & Kojima, 2001), in which the relative weights of the two cues vary with syncing gain. This idea would predict that the accretion-deletion cue in isolation from expansion-compression should support good depth-ordering performance at high but not low syncing gains, i.e., as for the Playback condition (Figure 3). To test this idea, we next examined an “Accretion-Deletion-Only” condition (Figure 7a) in which the textures were static, and only the boundary between the two textures was moving in synchrony with the head movement. Along the leading edge of the moving boundary, texture elements from the far-rendered surface disappeared while dots from the near surface became visible and conversely for the trailing edge (Figure 7a). Thus, in principle, the stimulus contains sufficient information to unambiguously specify depth order. Therefore, the only cue for depth was the relationship of the motion of this boundary to the head

movement, i.e., the head movement is necessary for this stimulus to be depth-unambiguous.

However, it was immediately very clear that observers reported seeing no depth in the stimulus and found the depth-ordering task unnatural even though the boundary motion was identical to that in the Cue-Consistent condition. Consequently, we formally tested psychophysical performance only at four representative values of syncing gain. The results (Figure 7b) demonstrate that observers were at chance performance; they were unable to utilize this information for depth perception. This result demonstrates that the accretion-deletion cue by itself is not sufficient to provide a percept of depth notwithstanding its demonstrated potency in the Cue-Conflict and Playback conditions.

Discussion

Contributions of expansion-compression versus accretion-deletion cues

We investigated the roles of expansion-compression and accretion-deletion to depth ordering in a series of experiments, either placing them in conflict with one another or testing conditions that would rely more on one or the other. When both cues were present but in conflict with one another (Figure 4b), the perceived depth ordering was consistent with the information provided by expansion-compression at lower syncing gains and with accretion-deletion at higher gains in agreement with earlier results of Ono et al. (1988).

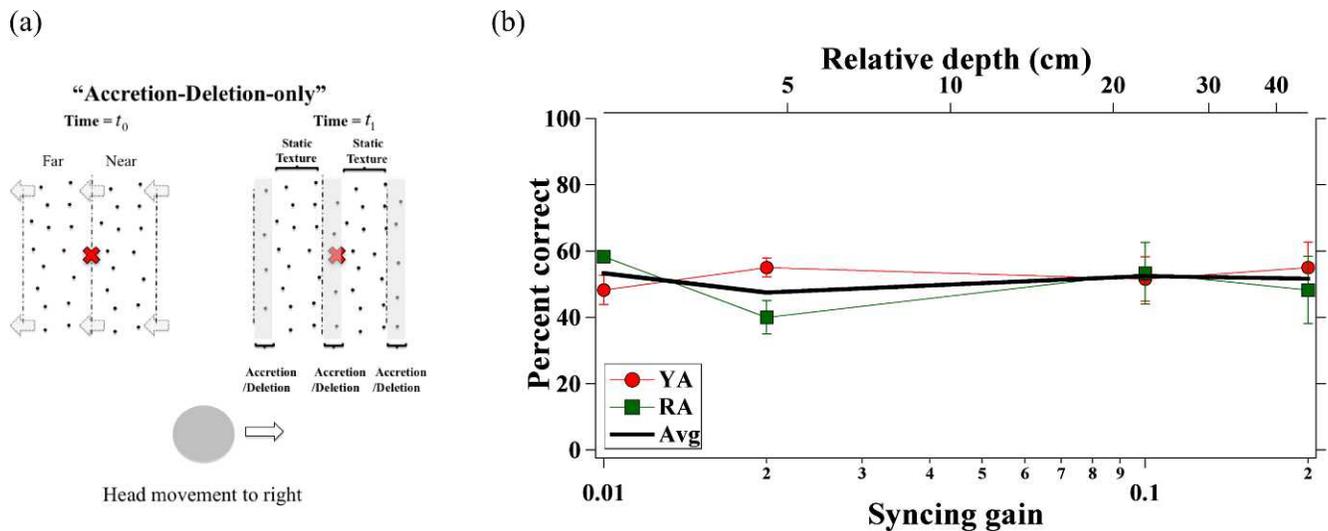


Figure 7. Depth ordering for Accretion-Deletion-Only condition. (a) Schematic depiction of the stimulus in which textures are stationary and only the boundary moves in synchrony with the head movement. Note that the accretion-deletion information contains unambiguous depth-order information. (b) Performance for two observers—note that observers reported no perceived depth in these stimuli.

Because the lower syncing gains will result in fewer accretion-deletion events (i.e., texture elements covered or uncovered) due to the smaller boundary displacement, the point of transition in depth order from expansion-compression to accretion-deletion dominance might be expected to depend on texture properties such as dot density (see Supplementary Materials, Figure S2). The results for the Transparent condition (Figure 6b) demonstrate that expansion-compression alone, i.e., without any accretion-deletion, supports depth ordering better at lower than at higher syncing gains. Conversely, the Playback experiment (Figure 3b) provided effective depth information only from accretion-deletion (because any expansion-compression would be depth-ambiguous due to the lack of head movement), and its results were complementary to those for the Transparent condition, i.e., best performance at the higher syncing gains. Thus all these results indicate that both expansion-compression and accretion-deletion contribute importantly to depth ordering with predominant roles at smaller versus larger relative rendered depths. This idea is consistent with the informational content of accretion-deletion, i.e., the greater number of texture elements covered/uncovered at higher syncing gains.

Relative motion (shear or expansion-compression) is depth-ambiguous unless it is synchronized appropriately with head movement; however, in the case of dynamic occlusion, accretion-deletion along a moving boundary does contain sufficient information for depth ordering even in the absence of observer movement. Our results (Figure 3) using the Playback condition demonstrate that human observers can indeed perform correct depth ordering without synchronized head

movement although only at the higher syncing gains at which accretion-deletion is best able to contribute.

Cue combination

Depth information can be obtained by the visual system from multiple visual cues. A simple “weak fusion” model of cue combination (Landy, Maloney, Johnston, & Young, 1995; Landy & Kojima, 2001) would suggest that these cues are used to estimate separate depth maps in isolation, which are then combined to form the scene layout. This idea is partly compatible with computational schemes (Marr, 1982) in which different cues are utilized to create a 2.5-D sketch of the visual scene. Evidence from psychophysical studies has demonstrated that some visual sources are combined in a manner that is compatible with a weak fusion model, e.g., texture and stereo (Young, Landy, & Maloney, 1993), motion and stereo (Johnston, Cumming, & Landy, 1994). In contrast, a strong fusion model would entail a nonlinear interaction between the different sources of information with these cues not necessarily represented in a separable manner. A weak fusion model would predict minimal interaction between expansion-compression and accretion-deletion and that each should be able to provide some degree of depth information in isolation. In such a model, information from each of the cues is combined as a linear weighted sum in which the weights are proportional to the informativeness or reliability of the cues. This is at least qualitatively consistent with the results for the Cue-Conflict condition, in which the apparent weights of the cues vary according to their

reliability. At lower syncing gains, the accretion-deletion information will be impoverished due to the smaller number of dots that are covered or uncovered, so it should have a lower signal-to-noise ratio. At higher syncing gains, the reliability of accretion-deletion is substantially increased due to the greater number of accretion-deletion events, and thus it makes sense that its relative weight would become proportionately larger.

However, another experimental condition demonstrates properties of cue combination that are more compatible with a strong fusion model. In the Accretion-Deletion-Only condition (Figure 7), the textures are static, and only the boundary between them moves, generating accretion-deletion along its path. In principle, this stimulus contains sufficient information to infer depth because it is synchronized to the head movement. However, the total failure of observers to see depth in this stimulus (Figure 7) suggests that even though accretion-deletion can dominate depth ordering at high syncing gains when other cues are present, human observers are not able to exploit it in the absence of other information. This behavior is not merely due to correlation between cues as the reliability of the accretion-deletion cue is not diminished in absence of texture motion. If the combination of cues were optimal as an “ideal observer” model would suggest (Geisler, 2003), the observers should have been able to utilize accretion-deletion information to disambiguate depth order. Taken together, all these results suggest that accretion-deletion can support psychophysical performance in depth ordering only when accompanied by relative texture motion. Interestingly, the accompanying texture motion need not occur in an ecologically valid, cue-consistent manner in order to facilitate the use of the accretion-deletion information to perform the task.

The asymmetrical roles of accretion-deletion and expansion-compression cues in depth from motion parallax might arise from the differing nature of their occurrence in natural scenes. Depth-related expansion-compression often occurs naturally in the absence of accompanying accretion-deletion, e.g., across surfaces that are gradually curved in depth (i.e., having smooth depth gradients) or in optic flow related to the ground plane (as discussed below). But accretion-deletion effects occur rarely, if ever, in natural situations except when accompanied by expansion-compression. Thus from an evolutionary perspective, it should not be surprising that the visual system lacks a mechanism to exploit the depth information in the Accretion-Deletion-Only stimulus but does exploit its rich informativeness when accompanied by expansion-compression. This behavior might be analogous to the combination of luminance and color: Pure luminance boundaries, like relative motion, are frequent in nature, but a pure

color boundary, like pure accretion-deletion, is extremely rare. Probably as a consequence, color in isolation does not produce depth, but it can enhance the depth provided by luminance (Kingdom, 2003). This is analogous to our finding that accretion-deletion alone fails to provide depth but that it can enhance the depth provided by expansion-compression.

Furthermore, in ecological settings in which a foreground surface frequently possesses a substantial area, accretion of texture elements and their deletion occur at different spatial positions with significant distance between them, rather than along a common boundary as in our Accretion-Deletion-Only stimulus. It is possible that the integration of accretion and deletion operates in a way that requires spatial separation between the locations of accretion and deletion of texture elements. In such a case, accretion and deletion that occur in the same spatial location might cancel each other out, resulting in no net response.

Comparison to shear-based motion parallax

We kept the depth-ordering task and stimuli as close as possible to those in our previous study on shear (Yoonessi & Baker, 2011a) in order for the two studies to be comparable. All stimulus parameters, such as spatial frequency, dot density, dot size, etc., were identical to those previously employed.

Motion parallax from both shear and dynamic occlusion could provide good depth ordering but over different ranges of syncing gains. Depth from shear was only good at low syncing gains (Figure 5, Yoonessi & Baker, 2011a), but for dynamic occlusion, it was almost perfect across a wide range (Figure 2). In fact, depth ordering for dynamic occlusion was most vulnerable to coherence noise at the lower syncing gains (Supplementary Materials, Figure S1). Thus, for depth ordering, shear yields best performance at smaller relative depths, and dynamic occlusion is best at larger depths, presumably due to a stronger contribution from accretion-deletion.

In this regard, an interesting comparison is between the Transparent condition (Figure 6), which does not contain accretion-deletion information, and previous shear results for square-wave modulation. Both of these conditions contain only relative motion information because there is no accretion-deletion. The pattern of data for both conditions is similar, but for at least three of the four observers, the performance is somewhat better for the Transparent condition than for shear at relatively higher syncing gains. This might suggest that relative motion information provides better depth ordering for dynamic occlusion than for pure shear although the comparison may be compli-

cated by the moving boundary in dynamic occlusion. Furthermore, in the Fixed-Boundary condition (Figure 5) the performance is still relatively better than that in the previous shear results, which is also consistent with this idea. Note that the Fixed-Boundary condition does have accretion-deletion, but it occurs in a manner that cannot disambiguate the depth order.

These comparisons suggest that shear and dynamic occlusion behave quite differently for depth ordering. Depth from dynamic occlusion is more robust across a wider range of syncing gains in comparison to shear, probably due, to some extent, to the additional information provided by accretion-deletion. The superior depth perception from dynamic occlusion compared to shear is consistent with the general idea that occlusion is thought to be the most powerful among the various depth cues (Cutting & Vishton, 1995).

One hypothesis emerging from results of this and our previous study of shear is that the most powerful role of shear motion parallax could be for perception of the ground plane during observer movement whereas dynamic occlusion-based motion parallax might contribute relatively more to perception of object boundaries. In a pure shear stimulus, depth ordering was much better when there was a gradient of depth difference (sine-wave modulation) rather than a sharp boundary (square-wave modulation). Such gradients are most prominently on the ground plane because it is almost always slanted in depth in relation to the observer's vantage point. However, we found that expansion-compression motion by itself was quite good in supporting depth order (Figure 6), and when accretion-deletion was present (Figure 2), the performance was almost perfect. This suggests that a human observer may rely relatively more on dynamic occlusion than on shear boundaries to detect depth differences between objects of finite size at intermediate viewing distances. Thus, shear and dynamic occlusion might play complementary roles in the perception of a scene whereby shear motion would be more informative about the overall scene layout with dynamic occlusion contributing more to object perception.

Similarities to stereopsis

In stereopsis, the images formed in the two eyes are combined to form a single final percept. This combination is relatively straightforward when every image region in one eye corresponds to a matching region in the other eye. However, in ecological conditions, some regions of the image will only be visible in one of the eyes and therefore will be unpaired. This problem, often called “DaVinci stereopsis” (Shimojo & Nakayama, 1990), poses a similar computational problem as dynamic occlusion. Both dynamic occlusion and

monocular occlusion give rise to points that are unmatched spatiotemporally in dynamic occlusion and spatially in monocular occlusion. The disparity in the two images is analogous to expansion-compression information as both are related to relative depth, and the unmatched areas in monocular occlusion correspond to accretion-deletion information in motion parallax.

Shimojo and Nakayama (1990) demonstrated that “valid” paired regions in monocular occlusion (i.e., resulting from geometry that might occur in real life) give rise to a depth percept whereas invalid regions (which would not occur naturally) result in binocular rivalry or suppression. In order for the monocular region to be valid, both its position relative to the boundary and the eye to which it is presented must be consistent with ecological conditions. The valid presentation is analogous to our Cue-Consistent condition (Figure 2), in which both expansion-compression and accretion-deletion occur in an ecologically valid manner. However, in the stereo-invalid regions, the two sources of information contradict one another, e.g., the monocular region is presented in the wrong eye for its location relative to the boundary. This is analogous to our Cue-Conflict condition (Figure 4), in which two information sources contradict one another. However, unlike our Cue-Conflict experiment, in monocular occlusion, a reversal of the depth percept has not been reported. At intermediate values of syncing gain, in which the performance is at chance levels (Figure 3), we do see depth on every trial but a randomly different depth order. Thus, there is a switching between alternate percepts that may be comparable to binocular rivalry, but in this case, the percept could alternate between depth provided by the two cues of expansion-compression and accretion-deletion. Suppression is analogous to performance at low or high syncing gain values in which one of the cues is systematically ignored.

In DaVinci stereopsis, the distance of the monocular region from the occlusion boundary is proportional to the amount of depth difference between the occluding surface and the monocular region (Shimojo & Nakayama, 1990; Gillam & Nakayama, 1999; Cook & Gillam, 2004). Furthermore, there is a maximum amount of distance between the monocular region and the boundary in which the monocular region is perceived as depth (Nakayama & Shimojo, 1990; Tsirlin, Wilcox, & Allison, 2012). This distance is analogous to the syncing gain parameter in our experiment, which is also proportional to rendered depth. However, the depth from dynamic occlusion seems to be very robust with syncing gain whereas monocular occlusion is unable to provide depth in some conditions even with an increase in the monocular

region's distance to the boundary (Gillam, Cook, & Blackburn, 2003), but see Tsirlin et al. (2012).

Computational algorithms

Detection of motion boundaries has been one of the principal methods in computer vision to detect occlusion boundaries. Luminance edges in a static image are simple to detect, but they may arise from causes other than occlusion edges, e.g., shadows, shading, highlights, texture, etc. Analogous edge-finding approaches may be used to find sharp transitions in optic flow information. Employing optic flow is more reliable because such discontinuities always correspond to a depth difference either within or between objects (Thompson et al., 1985).

Most traditional computer vision algorithms for motion field estimation rely on an assumption that the motion field is smooth almost everywhere but, notably, not at occlusion boundaries. Therefore, it becomes relevant to have an algorithm that looks for discontinuities in optic flow to segment boundaries and help find relative depth order. Nakayama and Loomis (1974) suggested a motion “convexity function” approach in which velocities integrated in local center and surround regions are subtracted to enhance motion discontinuities. Vaillant and Faugeras (1992) used extremal edges to show that at least three different views of a scene are necessary to reliably extract occlusion boundaries, which suggests dynamic occlusion could be more powerful than stereopsis for occlusion detection. More biologically plausible models rely on responses to spatiotemporal filters (Niyogi, 1995b), whose responses are used to detect occlusion boundaries. More recently, Bayesian approaches have been utilized, in which a cost function derived from the responses of spatiotemporal filters is minimized to provide boundaries with reasonable similarities to occlusion boundaries perceived by the visual system (Black & Fleet, 2000).

In most computer vision algorithms described so far, a segmentation step precedes depth ordering. However, to find the complete solution, it is necessary to infer the “border ownership,” i.e., the adjacent texture to which the boundary belongs. In a retinal image motion created from a simple observer translation, this can be achieved by comparing the magnitude and relative direction of optic flow to the direction of the observer movement (Thompson et al., 1985). The side of a boundary with the larger amplitude motion vector in the direction opposite to the moving observer movement (i.e., the “Near” surface in Figure 2b) will “own” the boundary. The discontinuity in image motion will always move in synchrony to the side that owns it (as in our Cue-Consistent condition, Figure 2b). Thus,

another approach is to specify border ownership by comparing the motion of the boundary to that of the regions on either side of it (Darrell & Fleet, 1995). Most of these algorithms do not utilize accretion-deletion of surface texture and would give equal performance for shear and dynamic occlusion.

Some algorithms have been proposed to detect accretion-deletion (Mutch & Thompson, 1985; Thompson et al., 1985; Niyogi, 1995a; Sun, Sudderth, & Black, 2010; Humayun, Aodha, & Brostow, 2011), but incorporating accretion-deletion does not seem to have yielded significant improvements in performance. A serious problem is that it is often very difficult to distinguish between accretion-deletion events and points that are not visible in some frames due to other reasons, such as sudden changes in luminance (e.g., specular reflections). Thus, for a computer algorithm to relate unmatched points to depth is not straightforward whereas a human observer can perform such a computation with impressive robustness to noise (see Supplementary Materials, Figure S1). Our findings suggest that for computer algorithms to better approach human performance, emphasis should be given to more robustly exploiting the accretion-deletion cue.

Possible neural mechanisms

The differential image motion in motion parallax can be thought of as a type of “second-order” stimulus because it entails boundaries defined by a difference other than luminance or color. Neurons that respond to second-order stimuli have been found in the early visual cortex of cats (e.g., Mareschal & Baker, 1999; Tanaka & Ohzawa, 2006; Rosenberg & Issa, 2011) and macaque monkeys (Peterhans & Von der Heydt, 1991; Li et al., 2012). Hegdé et al. (2004) suggested that these neurons might be suitable candidates to detect dynamic occlusion boundaries. Consistent with this idea, neuronal second-order responses often extend to surprisingly high carrier temporal frequencies (Rosenberg & Issa, 2011; Gharat & Baker, 2012), which would make them particularly useful for encoding accretion-deletion information in dynamic occlusion boundaries. Cue-invariant second-order responses of these neurons have been demonstrated for shear-type motion-defined boundaries (Gharat & Baker, 2012), but this question has not yet been examined for dynamic occlusion stimuli.

In order to fully utilize dynamic occlusion information for depth ordering, the visual system needs to specify which side “owns” the boundary. Mechanisms for such border ownership would be of potential importance for a depth-ordering task such as ours because the difference between the Cue-Consistent and Cue-Conflict conditions essentially involves which side

the boundary belongs to. Neurons that code for border ownership of simple contrast or texture boundaries have been found in macaque areas V2 and V4 (Zhou, Friedman, & Von der Heydt, 2000). These neurons' responses to a luminance-defined edge is contingent upon the "owner" of the boundary between the adjacent regions as cued by accompanying stimulus information outside the classical receptive field. Evidence from human fMRI also implicates area V2 in border ownership processing (Fang, Boyaci, & Kersten, 2009). Single-unit recordings from area V2 of macaque monkeys (Von der Heydt, Krieger, & He, 2003) demonstrate responses to differential motion stimuli in stationary animals that were contingent on border ownership although such studies have not yet explored border ownership in motion parallax.

Due to a visual neuron's limited receptive field size, its response cannot disambiguate the overall direction of an object's motion from the motion of a single contour of that object (known as the "aperture problem") (Marr, 1982). Accretion and deletion of contours within a receptive field's end zone could potentially provide additional information to resolve such ambiguities (Anderson & Sinha, 1997). A minority of neurons in primate V1 and a majority in MT exhibit end-stopping (Pack, Gartland, & Born, 2004), i.e., their response to an optimally oriented contour is inhibited if that contour extends beyond the neuron's classical receptive field. Consequently, such end-stopped neurons respond much better to contours with terminators. Furthermore, these neurons' responses are correlated with whether an occluder is perceived as belonging to a moving stimulus or arising extrinsically from an occluding object (Duncan, Albright, & Stoner, 2000; Pack et al., 2004). These findings suggest that such neurons might provide accretion-deletion and depth-order information in the context of motion parallax.

Conclusions

Our results suggest that expansion-compression and accretion-deletion cues play complementary roles in obtaining depth from dynamic occlusion motion parallax with relative motion being prominently utilized at smaller depth differences and accretion-deletion contributing more when there are larger depth differences. We have demonstrated that accretion-deletion can only enhance depth perception when it occurs in the presence of relative motion but cannot create a perception of depth in isolation even though it contains, in principle, sufficient information to do so. Dynamic occlusion appears to provide overall superior depth perception compared to shear, probably at least in part due to accretion-deletion.

Keywords: motion parallax, depth-ordering, 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 Unit, Department of Ophthalmology, McGill University, Montreal, Quebec, Canada.

Footnote

¹Note we do not use dynamic occlusion and accretion-deletion synonymously.

References

- Anderson, B. L., & Sinha, P. (1997). Reciprocal interactions between occlusion and motion computations. *Proceedings of the National Academy of Sciences, USA*, *94*, 3477–3480.
- Black, M. J., & Fleet, D. J. (2000). Probabilistic detection and tracking of motion boundaries. *International Journal on Computer Vision*, *38*(3), 231–245.
- Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision*, *10*, 433–436.
- Cook, M., & Gillam, B. (2004). Depth of monocular elements in a binocular scene: The conditions for da Vinci stereopsis. *Journal of Experimental Psychology*, *30*, 92–103.
- Craton, L. G., & Yonas, A. (1990). Kinetic occlusion: Further studies of the boundary flow cue. *Perception and Psychophysics*, *47*(2), 169–179.
- 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 W. Epstein & S. Rogers (Eds.), *Handbook of perception and cognition*, Vol. 5 (pp. 69–117). San Diego, CA: Academic Press.
- Darrell, T., & Fleet, D. J. (1995). Second-order method for occlusion relationships in motion layers. *MIT Media Lab TR*, 314.
- Duncan, R. O., Albright, T. D., & Stoner, G. R. (2000). Occlusion and the interpretation of visual motion: Perceptual and neuronal effects of context. *Journal of Neuroscience*, 20, 5885–5897.
- Fang, F., Boyaci, H., & Kersten, D. (2009). Border ownership selectivity in human early visual cortex and its modulation by attention. *Journal of Neuroscience*, 29(2), 460–465.
- Geisler, W. S. (2003). Ideal observer analysis. In L. Chalupa & J. Werner (Eds.) *The visual neurosciences* (pp. 825–837). Boston, MA: MIT Press.
- Gharat, A. M., & Baker, C. L., Jr. (2012). Motion-defined contour processing in early visual cortex. *Journal of Neurophysiology*, 108, 1228–1243.
- Gibson, E. J., Gibson, J. J., Smith, O. W., & Flock, H. (1959). Motion parallax as a determinant of perceived depth. *Journal of Experimental Psychology*, 58(1), 40–51.
- Gibson, J. J., Kaplan, G. A., Reynolds, H. N., & Wheeler, K. (1969). The change from visible to invisible: A study of optical transitions. *Perception and Psychophysics*, 5, 113–116.
- Gillam, B., Cook, M., & Blackburn, S. (2003). Monocular discs in the occlusion zones of binocular surfaces do not have quantitative depth: A comparison with Panum's limiting case. *Perception*, 32, 1009–1019.
- Gillam, B., & Nakayama, K. (1999). Quantitative depth for a phantom surface can be based on cyclopean occlusion cues alone. *Vision Research*, 39, 109–112.
- Hegd , 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]
- Helmholtz, H. V. (1925). Physiological optics. *Optical Society of America*, 3, 318.
- Humayun, A., Aodha, O. M., & Brostow, G. J. (2011). Learning to find occlusion regions. In *Proceedings of conference on computer vision and pattern recognition (CVPR)*, June 21–23, 2011, Colorado Springs, CO, USA.
- Johnston, E. B., Cumming, B. G., & Landy, M. S. (1994). Integration of stereopsis and motion shape cues. *Vision Research*, 34(17), 2259–2275.
- Kaplan, G. A. (1969). Kinetic disruption of optical texture: The perception of depth at an edge. *Attention, Perception and Psychophysics*, 6(4), 487–492.
- Kingdom, F. A. A. (2003). Color brings relief to human vision. *Nature Neuroscience*, 6, 641–644.
- Kromrey, S., Bart, E., & Hegd , J. (2011). What the 'moonwalk' illusion reveals about the perception of relative depth from motion. *PLoS ONE*, 6(6), e20951.
- Landy, M. S., & Kojima, H. (2001). Ideal cue combination for localizing texture-defined edges. *Journal of the Optical Society of America*, 18, 2307–2320.
- 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.
- Li, G., & Baker, C. L., Jr. (2012). Functional organization of envelope-responsive neurons in early visual cortex: Organization of carrier tuning properties. *Journal of Neuroscience*, 32(22), 7538–7549.
- Mareschal, I., & Baker, C. L., Jr. (1999). Cortical processing of second order motion. *Visual Neuro Science*, 16(3), 527–540.
- Marr, D. (1982). *Vision*. New York: Freeman.
- Mutch, K. M., & Thompson, W. B. (1985). Analysis of accretion and deletion at boundaries in dynamic scenes. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 3, 133–138.
- Nakayama, K., & Loomis, J. (1974). Optical velocity patterns, velocity-sensitive neurons, and space perception: A hypothesis. *Perception*, 3(1), 63–80.
- Nakayama, K., & Shimojo, S. (1990). Da Vinci stereopsis: Depth and subjective occluding contours from unpaired image points. *Vision Research*, 30, 1811–1825.
- Niyogi, S. A. (1995a). Detecting kinetic occlusion. In *Proceedings of Fifth International Conference on Computer Vision (ICCV 1995)* (pp. 1044–1049), June 20–23, 1995, Cambridge, MA, USA.
- Niyogi, S. A. (1995b). Spatiotemporal junction analysis for motion boundary detection. *Image processing, proceedings, International Conference on (ICIP 95)* (pp. 368–471), October 23–26, Washington, DC, USA.
- 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., Gartland, A. J., & Born, R. T. (2004). Integration of contour and terminator signals in visual area MT of alert macaque. *Journal of Neuroscience*, *24*, 3268–3280.
- Peterhans, E., & Von der Heydt, R. (1991). Elements of form perception in monkey prestriate cortex. In A. Gorea, Y. Fregnac, Z. Kapoula, & J. Findlay (Eds.) *Representations of vision: Trends and tacit assumptions in vision research* (pp. 111–124). Cambridge: Cambridge University Press.
- 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.
- Rogers, B. J., & Graham, M. (1983). Anisotropies in the perception of three-dimensional surfaces. *Science*, *221*(4618), 1409–1411.
- Rosenberg, A., & Issa, N. P. (2011). The Y cell visual pathway implements a demodulating nonlinearity. *Neuron*, *71*(2), 348–361.
- Shimojo, S., & Nakayama, K. (1990). Real world occlusion constraints and binocular rivalry interaction. *Vision Research*, *30*(1), 69–80.
- Sun, D., Sudderth, E., & Black, M. J. (2010). Layered image motion with explicit occlusions, temporal consistency, and depth ordering. *Advances in Neural Information Processing Systems*, *23*, 2226–2234.
- Tanaka, H., & Ohzawa, I. (2006). Neural basis for stereopsis from second-order contrast cues. *Journal of Neuroscience*, *26*(16): 4370–4382.
- Thompson, W. B., Mutch, K. M., & Berzins, V. A. (1985). Dynamic occlusion analysis in optic flow fields. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, *7*(4), 374–383.
- Tsirlin, I., Wilcox, L. M., & Allison, R. S. (2012). Da Vinci decoded: Does da Vinci stereopsis rely on disparity? *Journal of Vision*, *12*(12):2, 1–17, <http://www.journalofvision.org/content/12/12/2>, doi:10.1167/12.2.2. [PubMed] [Article]
- Ujike, H., & Ono, H. (2001). Depth thresholds of motion parallax as a function of head movement velocity. *Vision Research*, *41*(22), 2835–2843.
- Vaillant, R., & Faugeras, O. D. (1992). Using extremal boundaries for 3-D object modeling. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *14*(2), 157–173.
- Von der Heydt, R., Krieger, F. T. Q., & He, Z. J. (2003). Neural mechanisms in border ownership assignment: Motion parallax and gestalt cues. *Journal of Vision*, *3*(9):666, <http://www.journalofvision.org/content/3/9/666>, doi:10.1167/3.9.666. [Abstract]
- 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:10.1167/11.1.62. [Abstract]
- Young, M. J., Landy, M. S., & Maloney, L. T. (1993). A perturbation analysis of depth perception from combinations of texture and motion cues. *Vision Research*, *33*(18), 2685–2696.
- Zhou, H., Friedman, H. S., & Von der Heydt, R. (2000). Coding of border ownership in monkey visual cortex. *Journal of Neuroscience*, *20*(17), 6594–6611.