Published in final edited form as:

Perception. 2016 September ; 45(9): 991–1007. doi:10.1177/0301006616643679.

A pursuit theory account for the perception of common motion in motion parallax

Michael Ratzlaff and Mark Nawrot

Center for Visual and Cognitive Neuroscience, Department of Psychology, North Dakota State University, Fargo ND 58104

Abstract

The visual system uses an extra-retinal pursuit eye movement signal to disambiguate the perception of depth from motion parallax. Visual motion in the same direction as the pursuit is perceived nearer in depth while visual motion in the opposite direction as pursuit is perceived farther in depth. This explanation of depth sign applies to either an allocentric frame of reference centered on the fixation point or to an egocentric frame of reference centered on the observer. A related problem is that of depth order when two stimuli have a common direction of motion. The first psychophysical study determined whether perception of egocentric depth order is adequately explained by a model employing an allocentric framework, especially when the motion parallax stimuli has common rather than divergent motion. A second study determined whether a reversal in perceived depth order, produced by a reduction in pursuit velocity, is also explained by this model employing this allocentric framework. The results show that an allocentric model can explain both the egocentric perception of depth order with common motion, and the perceptual depth order reversal created by a reduction in pursuit velocity. We conclude that an egocentric model is not the only explanation for perceived depth order in these common motion conditions.

Keywords

Motion parallax; depth perception; pursuit eye movements; common motion; visual perception

Introduction

Motion parallax is an important cue for the visual perception of relative depth. As an observer of a visual scene translates, the retinal images of objects in the scene also translate depending on the observer's point of gaze and each object's spatial relation to that point of gaze. However, knowing the direction or speed of retinal image motion of each object is insufficient to recover each object's spatial relation to the point of fixation. Additional information about the particular viewing conditions is required by the observer's visual system to determine which directions of retinal image motion correspond to objects nearer and farther than the point of gaze.

In one solution to this problem, Nakayama and Loomis (1974) described how knowing the speed and direction of observer translation, an observer could determine the absolute distance to any object in the scene (see related solutions offered by Longuet-Higgins & Prazdny, 1980; Gordon, 1965; Koenderink & vanDoorn, 1976; Koenderink & vanDoorn, 1987; Perronne & Stone, 1994; Fermüller & Aloimonos, 1997; Hanes et al., 2008). However, it is not clear how the visual system can accurately determine the speed and direction of head or body translation relative to the point of gaze.

A different solution to this problem is provided by the pursuit theory, which posits that a disambiguating signal is generated by the pursuit eye movement system (Nawrot, 2003; Naji & Freeman, 2004; Nawrot & Joyce, 2006). As an observer translates laterally (actively or passively) and maintains fixation on a stationary point in space, objects nearer and farther than the fixation point generate motion parallax upon the observer's retina. To maintain stable fixation the visual system must rotate the eyes to compensate for the observer's translation. While the visual system has a few different eye-movement mechanisms to maintain stable fixation during this translation (Miles & Busetini, 1992), only the pursuit system appears to provide the requisite signal (Nawrot & Joyce, 2006). It is the direction of the pursuit eye movement signal that disambiguates which direction of retinal image motion corresponds to objects lying nearer or farther than the fixation point.

Employing this pursuit signal, the motion/pursuit ratio (M/PR) (Nawrot & Stroyan, 2009) provides a general model for the perception of depth from motion parallax based on the dynamic geometry of the motion parallax conditions. The ratio of an object's retinal image motion (termed $d\theta$) to the observer's pursuit eye movement velocity (termed $d\alpha$) provides information about that object's distance (d_{mp}) from the observer's point of fixation at a viewing distance of f (equation 1)

$$\frac{d\theta}{d\alpha} \approx \frac{d_{mp}}{f} \quad (1)$$

In this M/PR model, similar to binocular stereopsis, depth sign is allocentric¹ being compared to the point of fixation. Opposite directions of $d\theta$ produce opposite signs of d_{mp} while a reversal in the direction of $d\alpha$ (keeping $d\theta$ constant) also produces a reversal in the depth sign of d_{mp} .

An alternative model of depth order, also based on a pursuit eye movement signal, was suggested by Mitsudo and Ono (2007). In their "additive model", visual image motion (termed v_i) is combined additively with the pursuit signal (termed $v_{pursuit}$) with the object motion producing the largest sum reported being be closer to the observer than any smaller sum (equation 2).

¹It is important to note that an allocentric – egocentric dichotomy is not simple or clearly defined (Klatzky, 1998). While the M/PR expresses depth relative to the point of fixation, the proximal visual cues of retinal image motion and pursuit velocity are egocentric. Stroyan and Nawrot (2012, section 3) present a detailed discussion of these issues related to the M/PR model. Here we are concerned primarily with the spatial reference for the perceived depth from motion parallax, so we refer to the M/PR as allocentric for simplicity and clarity.

$$\text{if } |v_1 + v_{\text{pursuit}}| \geq |v_2 + v_{\text{pursuit}}|, D_1 \leq D_2, \text{ otherwise, } D_1 > D_2 \quad (2)$$

That is, this additive model produces an egocentric representation of relative depth, with the largest v_j in the same direction as the pursuit eye movement being nearer the observer, $v_j = 0$ occurring at the fixation point, and with increasing v_j in the direction opposite pursuit receding away from the observer². It is important to note that the M/PR model and the additive model differ in their definition of “depth”; with the M/PR model depth is allocentric and relative to the point of fixation (and scales with viewing distance), whereas the additive model defines depth egocentrically, meaning that a greater depth is farther away from the observer. However, the additive model does provide an accurate description of the instantaneous, head-centric depth from motion parallax with both common and divergent motion. Nevertheless, the additive model has a few shortcomings and their study results do not indicate that the visual system must be generating a head-centric velocity field (from the addition of visual and pursuit velocities) and a concomitant egocentric representation of depth.

First, the additive model is unable to generate an explanation for “stable depth from motion parallax” (Mitsudo & Ono, 2007, p.133). Instead, the ordinal depth representation is accurate instantaneously, but the representations do not allow comparisons across changes in viewing distance and pursuit velocity. Such stability is important as an observer shifts the point of gaze and changes their pursuit to explore the scene. Imagine an observer performing a two-interval forced choice task regarding the egocentric ordinal depth in two motion parallax stimuli. In one interval $v_1 = 2 \text{ deg s}^{-1}$ and $v_{\text{pursuit}1} = 8 \text{ deg s}^{-1}$, while in the other interval $v_2 = 1 \text{ deg s}^{-1}$ and $v_{\text{pursuit}2} = 4 \text{ deg s}^{-1}$. The additive model (equation 2) suggests that $D1 < D2$, but the M/PR model (equation 1) indicates that $D1 = D2$ (use $f - d$ to get egocentric depth from the M/PR). If the first interval were $v_1 = 1 \text{ deg s}^{-1}$ and $v_{\text{pursuit}1} = 8 \text{ deg s}^{-1}$, the additive model would still suggest that $D1 < D2$, but the M/PR model (equation 1) would indicate that $D1 > D2$. Moreover, the additive model: 1) is unable to account for changes in the viewing distance (it would predict $D1 = D2$ if $v_1 = v_2$ and $v_{\text{pursuit}1} = v_{\text{pursuit}2}$ but the viewing distances changed, $f_1 \neq f_2$), 2) suggests that higher velocity pursuit targets (when $v_1 = v_2 = 0$) are perceived as closer than slower pursuit targets, and 3) cannot account for depth scaling (Ono, Rivest & Ono, 1986) or depth constancy, important properties of motion parallax accounted for by the M/PR model (Nawrot & Stroyan, 2009).

Second, Mitsudo and Ono (2007) conclude that depth from motion parallax must be an egocentric process, without reference to an allocentric point of visual fixation. It is not clear whether their results support only their additive model, nor is it clear whether these results support only an egocentric computational framework. We address this issue in the two studies presented here by determining whether accurate egocentric judgments, with both divergent and convergent motion parallax stimuli, are accurately predicted using the allocentric framework of the M/PR model (Nawrot & Stroyan, 2009).

²The direction of v_j is governed by the object’s allocentric relationship with the point of fixation.

Third, Mitsudo and Ono's (2007) novel common motion condition presents an interesting challenge for any model of motion parallax, especially one that uses an allocentric framework like the M/PR model. That is, depth from motion is inherently depth-sign ambiguous, but additional information present in the viewing conditions for motion parallax is required for the visual system to generate an unambiguous depth percept. To study the processes responsible for this disambiguation, typical motion parallax research stimuli employ equal speed, but opposite direction, motion signals (termed divergent motion here) (Rogers & Graham, 1979) with the aim of determining how additional information (e.g., lateral head translation) causes a particular direction of motion to be perceived nearer and the other direction of motion farther. This near-far disambiguation was the goal of the pursuit theory "heuristic" (Nawrot, 2003; Nawrot & Joyce, 2006). Indeed, an explanation of the depth order perceived with common motion requires a model for the combination of image motion with pursuit. Mitsudo and Ono's (2007) additive model accomplishes this task, but with a distinct lack of generalizability to other conditions. In the studies presented here we determine whether the M/PR model (Nawrot & Stroyan, 2009) can also explain the perception of depth order with conditions of common motion.

Experiment 1

The current study has two goals: first, to determine whether the M/PR model predicts perceived depth order with common motion stimuli, and second, to determine whether this model makes accurate depth order predictions in an egocentric framework, even though the model computations operate at an allocentric framework. If the results indicate that both are true, this calls into question whether the visual system must be generating a head-centric velocity field in order to recover depth order from motion parallax (Mitsudo & Ono, 2007).

Methods

Apparatus—Stimuli were presented by a Macintosh computer via a 21-inch NEC CRT monitor with a resolution of 1920 pixels $\times$ 1440 pixels $\times$ 75 Hz that, at 45cm viewing distance, covered 51 deg $\times$ 38 deg of visual angle. Stimuli were created using Matlab with the Psychophysics Toolbox-3 (Brainard, 1997; Pelli, 1997; Kleiner et al., 2007). A stabilizing chin and headrest was used to center the observer's right eye on the monitor, restrict the observer's head movements, and maintain a fixed viewing distance.

Stimuli—Random-dot stimuli were composed of 5000, 1.6 $\times$ 1.6 min white dots in an 8.4 $\times$ 8.4 deg square stimulus window on a black background. This window contained a small central 15 min $\times$ 15 min fixation point. This stimulus window appeared at the monitor center and then translated laterally ($d\alpha$), either rightwards (+) or leftwards (−), at 12 deg s^{−1} for a distance of 21.4 deg.

Motion parallax was generated within this stimulus window by laterally translating ($d\theta$) the stimulus dots rightward (+) or leftward (−) simultaneous with the stimulus window translation. Drawing from a commonly used sinusoidal motion parallax stimulus (Rogers & Graham, 1979; Nawrot & Joyce, 2006), here the velocity of the stimulus dot translations varied with two vertically oriented half-cycles of a sinusoidal function. These half-cycles were centered in the upper and lower halves of the stimulus window with zero dot velocities

at the top and bottom edge of the stimulus window, zero velocity at the vertical midline, and peak stimulus dot velocity ($d\theta$) midway between the upper and lower sections. In different conditions, the direction of stimulus dot motion within the two stimulus halves could be in the same direction (common motion) or in opposite directions (divergent motion). Moreover, the peak stimulus dot velocity ($d\theta$) in the two halves was varied so that one had a velocity ($d\theta_1$) of $\pm 2.0 \text{ deg s}^{-1}$ while the other had a velocity of half-magnitude ($d\theta_{0.5}$) $\pm 1.0 \text{ deg s}^{-1}$. These values for $d\alpha$ and $d\theta$ give motion/pursuit ratios of ± 0.16 and ± 0.08 . In all conditions the dot stimuli were confined to the square stimulus window, with dots entering and leaving as necessary to maintain the particular motion parameters. Finally, the direction of pursuit ($d\alpha$) could be in the same or opposite direction of either stimulus dot motion ($d\theta_1$ and $d\theta_{0.5}$). As will be detailed below, this gives sixteen different stimuli, eight with common motion and eight with divergent motion.

Common Motion: In these eight stimuli, both $d\theta_1$ and $d\theta_{0.5}$ were in the same direction thereby generating motion parallax stimuli with only common motion. To determine whether the motion/pursuit ratio can account for cases in which observers perform an egocentric depth judgment on stimuli with common motion, the two conditions presented here will outline the predicted egocentric depth judgments and compare them to observer performance.

In the first condition, Common Motion Near (CMN), the direction of $d\alpha$ was in the same direction as both $d\theta_1$ and $d\theta_{0.5}$. This arrangement should give the appearance of two convex sinusoidal surfaces of two different magnitudes (Figure 1, top panels). For simplicity, we define the stimulus with $d\theta_1$ in the upper portion, and $d\theta_{0.5}$ in the lower portion, as being the standard phase (Figure 1, left panel), and the reverse as being the alternate phase (Figure 1, right panel). In this condition, both the standard-phase and alternate-phase stimuli can be generated with both rightward (+) and leftward (–) pursuit eye movements ($d\alpha$), as long as both $d\theta_1$ and $d\theta_{0.5}$ motions are in the same direction as $d\alpha$. This condition is most similar to the stimulus configuration used by Mitsudo and Ono (2007), which had v_1 , v_2 , and $v_{pursuit}$ in the same direction. For the current study, the tables in Figure 1 illustrate the four different CMN stimulus combinations of the variables ($d\alpha$, $d\theta$, f). The table also presents the M/PR predictions for allocentric and egocentric depth. The egocentric depth column reflects the observer's task in this study, following the procedure used by Mitsudo and Ono (2007), which indicates which part of the stimulus is nearest to the observer. In the last column, the table gives the M/PR predicted response (upper or lower) that an observer will make in the egocentric depth discrimination task for each stimulus. The common feature of these conditions is that the stimulus segment with $d\theta_1$ should appear nearer than the segment with $d\theta_{0.5}$.

The second condition, Common Motion Far (CMF), is similar to the CMN conditions above in that both $d\theta_1$ and $d\theta_{0.5}$ moved in the same direction. However, in the CMF conditions $d\alpha$ was in the opposite direction to both $d\theta_1$ and $d\theta_{0.5}$. This should create the appearance of a two concave surface segments. Again, the tables in Figure 2 give the stimulus parameters. Similar to the CMN stimuli, the standard phase stimulus had $d\theta_1$ in the upper segment while the alternate phase had $d\theta_1$ in the lower segment. However, because the task was an egocentric depth judgment, the M/PR predicted responses with these stimuli are different

from those in the CMN condition. This means that observers cannot perform the task by discriminating $d\theta_1$ from $d\theta_{0.5}$. Instead, given these conditions, observers can only perform the discrimination task by judging the egocentric relative depth of the two stimulus segments.

Divergent Motion: In these eight stimulus conditions, $d\theta_1$ and $d\theta_{0.5}$ were in opposite directions. While common motion was still present within each of the stimulus segments above and below the fixation point, the observer's task was related to the divergent (relative) motion in the two segments. While Mitsudo and Ono (2007) did not suggest that the pursuit theory could not account for these conditions, they were included as a comparison for how well egocentric depth discriminations are made with divergent stimuli. That is, performance on these divergent conditions provides as a baseline for comparison with the common motion conditions. Indeed, with the exception of the peak velocity differences between the $d\theta_1$ and $d\theta_{0.5}$ segments, these divergent motion stimuli were similar to the sinusoidal motion parallax stimuli used in previous research (e.g., Rogers & Graham, 1979; Nawrot & Joyce, 2006).

In the Common Motion conditions above, the “Near” and “Far” designations referred to both the $d\theta_1$ and $d\theta_{0.5}$ stimulus segments, and denoted both segments as appearing convex (near) or concave (far). However, in Divergent Motion conditions the dots in these segments move in opposite directions. To differentiate the stimuli here, the appearance of the full-magnitude ($d\theta_1$) stimulus will be used as reference.

In the Divergent Motion Near (DMN) condition, the direction of $d\alpha$ movement was always the same as $d\theta_1$, and opposite $d\theta_{0.5}$. This produces a roughly sinusoidal stimulus that should appear to have a “full-magnitude” convex surface segment and a “half-magnitude” concave surface segment. In the standard phase (Figure 3, left panel) the $d\theta_1$ stimulus segment was always above the fixation point, while in the alternate phase (Figure 3, right panel) the $d\theta_1$ stimulus segment was below the fixation point. In all of these stimuli, the higher velocity $d\theta_1$ stimulus segment should be perceived as nearer to the observer.

In the Divergent Motion Far (DMF) condition, the direction of $d\alpha$ was always opposite the direction of $d\theta_1$, but was in the same direction as $d\theta_{0.5}$. In contrast to the DMN condition, these stimuli should produce a sinusoidal depth percept with a full-magnitude concave surface segment and a half-magnitude convex surface segment. With these stimuli, the lower velocity $d\theta_{0.5}$ stimulus segments should be perceived nearer to the observer when they are below the fixation point in the standard phase (Figure 4, left panel) and above the fixation point in the alternate phase (Figure 4, right panel). When combined with the DMN condition stimuli, these stimuli ensure that observers are reporting perceived depth, not just which stimulus segment has the higher velocity, because in this condition the higher velocity $d\theta_1$ stimulus segment should be perceived as farther away than the slower velocity $d\theta_{0.5}$ segment.

By combining the four half-cycle segment types ($d\theta_1$ convex and concave, $d\theta_{0.5}$ convex and concave), four stimulus types were created. In addition, these conditions were presented both in standard ($d\theta_1$ above midline) and alternate ($d\theta_1$ below midline) phases.

Procedure—Six participants with normal or corrected-to-normal vision completed the study. Participants volunteered by giving informed consent, receiving either pay or research credit for their participation. This study was conducted with oversight by the local Institutional Review Board and in accordance with national regulations and legislation, along with the World Medical Association Helsinki Declaration as revised in October, 2008.

Participants viewed the stimulus monocularly with the right eye. Binocular disparity cues were eliminated by an eye patch occluding the left eye. Participants were instructed to maintain fixation on the fixation point throughout each trial. Trials began with an initial fixation point centered on the monitor. After a key press, the stimulus window was drawn centered on fixation and immediately began translating across the screen in a single direction, leftward or rightward. Stimulus presentation duration was 1.6 s, followed by a blank black screen. Following each trial, observers indicated, via key press, which segment, upper or lower, appeared nearer to them. Following the response, the fixation point reappeared and a new trial was prepared.

All participants completed 10 blocks of trials, each consisting of 160 trials; 20 for each of the eight stimulus types. The order of trial presentation within a trial block, including phase and direction of lateral stimulus window translation, was randomized.

Results

D-prime (d') values were calculated by comparing response accuracy between standard and alternate phases of each condition. Values over 2.0 indicate both good discrimination of depth order and lack of vertical bias; i.e., observers could accurately determine which segment, upper or lower, was perceived closer to them in the stimulus window, regardless of which side of the midline it appeared. 'Hits' consisted of correct responses in the standard position trials and 'misses' as incorrect responses. Conversely, 'false alarms' were incorrect responses in the alternate position trials and correct responses represented 'correct rejections'.

The d' scores for each participant were averaged to measure depth order discrimination between standard and alternate phases. In the common motion conditions, observers easily determined depth order (CMN: $d' = 4.137$, $SE = .019$; CMF: $d' = 4.338$, $SE = .009$). Similarly, divergent motion conditions produced high discrimination (DMN₁: $d' = 4.508$, $SE = .022$; DMF₁: $d' = 4.712$, $SE = .0164$). These d' values are compared in Figure 5. In three conditions observers correctly responded to all 400 trials, and in nine conditions no false alarms occurred. In order to calculate a d' value, it was necessary to amend these 12 perfect scores to categorize either a single hit as a miss or correct rejection as a false alarm, respectively. Thus, the data points were made useable at the expense of more conservative average d' values.

In all conditions, d' values exceed 4, showing that observers were able to accurately and consistently discriminate depth order. This performance shows close agreement with the predictions of the M/PR model for both common motion and divergent motion.

Conclusion

The results of the current study demonstrate that the allocentric framework of the M/PR model can account for an observer's judgments of egocentric distances with both common and divergent motion parallax stimuli. Therefore, it does not appear necessary that "...the visual system calculates the object velocity relative to the head by adding retinal image and pursuit velocities, and then uses it to determine the depth order of moving patterns" (p. 130, Mitsuto & Ono, 2007). The current results indicate that the perception of depth order from motion parallax could also be the product of an allocentric system as outlined by the M/PR.

Unfortunately, the relevant neurophysiology, which has revealed how individual neurons respond to motion parallax stimuli, does not help in determining whether these visual processes for depth ordering are egocentric or allocentric (Nadler et al., 2008; 2009). While Nadler et al. assumed an allocentric, fixation-based framework wherein depth was simulated as near or far relative to the point of fixation, their results could also be interpreted as consistent with the egocentric, additive model of Mitsuto and Ono (2007). That is, in Nadler et al.'s (2009) "eye only (EO)" condition, the neuronal response is the combination of the response to retinal motion (RM) and an extra-retinal pursuit signal. In their example neuron, shown in their Figure 2, the EO firing rate is higher than the RM firing rate when the visual motion and pursuit are in the same direction (remember, visual motion and retinal motion are in opposite directions), but the EO firing rate is much less when visual motion and pursuit are in opposite directions. While Nadler et al. (2009) describe the decline in EO firing rate as "monotonic from near to far", consistent with the M/PR, the firing rate could also be described as the product of $v_i + v_{\text{pursuit}}$, thereby exhibiting a monotonic change across a range of visual motion velocities from $+v_i$ to $v_i = 0$ to $-v_i$. In this way the firing rate could be interpreted as signaling egocentric depth as described by Mitsuto and Ono (2007). This inability to differentiate the models is likely because the additive model, the M/PR model, and Nadler et al.'s neuronal responses all use the same components to describe the same phenomenon. Indeed, this is not unexpected as both models generate the same predictions for the experiment described above. Therefore, there is little to differentiate the models when only depth sign and depth order are considered.

Experiment 2

The first experiment demonstrated that the perception of depth order from common and divergent motion are treated similarly within the allocentric framework of the M/PR. Given that the M/PR general depth order predictions are not different from those made by Mitsuto and Ono's (2007) additive model, can the M/PR offer an alternative explanation for Mitsuto and Ono's demonstrated effect of changing pursuit velocity? They found that a change in pursuit velocity between 0 and 20 deg s⁻¹ causes a reversal in perceived depth order for their common motion stimulus. Their explanation relies on "common motion", but it is unlikely that this perceptual reversal is confined only to these limited conditions.

The phenomenology of these computer-generated, random-dot, motion parallax stimuli has not been well understood. That is, not all combinations of computer-generated pursuit ($d\alpha$) and retinal image motion ($d\theta$) stimulus velocities generate a solid (rigid) and unambiguous depth percept. Instead, some motion parallax stimuli appear ambiguous. In these ambiguous

motion parallax stimulus conditions, the percept can appear either normal, flat, or depthless (with dots appearing to streak or smear), or appear in the opposite depth phase with an unpredictable depth magnitude and with apparent rotation. Mitsudo and Ono (2007) attempt to explain these ambiguous stimulus conditions with their additive model and generated the following rationale: when depth from motion parallax is depth-sign ambiguous, or flat, this reflects a perceptual condition in which $D1 = D2$. Therefore, based on their additive model, this ambiguous percept must be created by the conditions when the average retinal image velocity is equal and opposite to the direction of pursuit: $(v_1 + v_2)/2 + v_{\text{pursuit}} = 0$ (p. 126). That is, Mitsudo and Ono suggest that perceived depth is ambiguous when the average velocity of two motion parallax stimuli is equal and opposite the pursuit eye movement velocity. However, they do not provide a rationale for this equality, nor do they make it clear why such spatial averaging of velocities might be expected, or required, for the perception of depth from motion parallax. Moreover, it is unclear how their particular pattern of results, showing a reversal in perceived depth with a reduction in pursuit velocity (e.g., Fig. 4, p. 130), is explained by this common-motion equality. Indeed, Mitsudo and Ono's equality holds for two points beyond the point of fixation, only when the average distance of $D1$ and $D2$ approaches infinity (with the attendant mathematical difficulties), and the average of v_1 and v_2 approaches v_{pursuit} . As explained below, the specific velocity parameters required for their equality appear to violate the typical motion parallax parameters that the visual system faces in a rigid environment.

In an alternative to the additive model, the M/PR provides a comprehensive framework to consider these particular motion parallax stimulus conditions. With the M/PR, relative depth for each point in the scene is determined independently from that point's retinal image velocity; spatial velocity averaging, or common motion, has no special role in the perception of depth from motion parallax. The M/PR provides a useful framework to understand the relationship between viewing distance (f), relative depth (d), pursuit ($d\alpha$), and retinal image motion ($d\theta$). Using the basic quantity of the M/PR ($d\theta/d\alpha$), the M/P Law (Equation 3) can tell us about the simulated relative depth in a computer generated motion parallax displays.³

$$\frac{d\theta}{d\alpha} \frac{1}{1 - \frac{d\theta}{d\alpha}} = \frac{d_{mp}}{f} \quad (3)$$

Using a nomenclature in which rightward $d\theta$ and $d\alpha$ velocities are positive, and leftward velocities are negative (remember that the direction of retinal image motion is opposite apparent motion in the visual scene), one important implication of this geometric model is that retinal image velocities ($d\theta$) greater than pursuit velocities ($d\alpha$), meaning $M/PR > 1$, are never generated by observer translation within in a rigid environment. That is, for objects beyond the fixation point, $d\theta/d\alpha > 0$, and as d goes towards ∞ , $d\theta/d\alpha$ approaches a maximum of 1. This means that $d\theta < d\alpha$ for points farther than fixation in rigid

³The M/PR (Equation 1) provides a useful approximation for the M/PLaw (Equation 3) when the ratio is small (e.g., objects are near fixation), as the magnitude of the error of the M/PR is the magnitude of the ratio (Nawrot & Stroyan, 2009). Therefore, the approximation provided by the M/PR is less useful as the $d\theta$ velocity approaches the $d\alpha$ velocity, and the ratio becomes larger.

environments. In contrast, objects nearer than fixation have $d\theta/d\alpha < 0$, and as d goes to a maximum of f , $d\theta/d\alpha$ approaches a large negative number ($-\infty$). Perhaps the visual system constrains its own processing of depth within these limits. If so, the visual system might have difficulty representing computer generated stimuli in which $M/PR > 1$, and instead attempts to represent these stimuli as “near” depth where larger M/PR may be represented. It is possible that this perceptual limitation, predicted by the M/PR , is responsible for the perceived depth reversal reported by Mitsudo and Ono (2007). For instance, all the data points above the horizontal line in Mitsudo and Ono’s Figure 4 come from stimulus conditions in which $M/PR \leq 1$ (the black square at 8 deg s^{-1} has an M/PR of 0.94) and all of these stimuli are more likely to be perceived in “reversed” depth.

In the following set of six conditions the common-motion (CMF) stimuli from the previous experiment were systematically devolved to determine whether common motion plays a role in the perceptual changes that follow changes in pursuit velocity with motion parallax. Additionally, stimulus M/PR was systematically varied to determine whether the perceptual transitions found with changes in pursuit velocity can be reasonably explained by the M/PR model. Specifically, our hypothesis is that stimuli with $M/PR < 1$ will be perceived in the correct depth sign as this stimulus is “reasonable” to the visual system. Stimuli with $M/PR > 1$ will be perceived in a reversed (geometrically incorrect) depth sign because such a stimulus is unusual or “unreasonable” for the limits of the visual mechanisms processing the perception of depth from MP.

Method

Apparatus—The apparatus described for the Experiment 1 was used in all six conditions here.

Procedure—Five participants with normal or corrected-to-normal vision completed the study. One of these participants was a participant in Experiment 1. Again, participants viewed the stimulus monocularly with the right eye, while the left eye was patched to eliminate binocular disparity cues to depth. Participants were instructed to maintain fixation on the fixation point throughout each trial. Each began with an initial fixation point centered on the monitor. After a key press, the stimulus window was drawn centered on fixation and immediately began translating across the screen in a single direction, leftward or rightward. In most conditions, the stimulus presentation duration was 1.6 s, but in Condition 2 the stimulus presentation duration was unlimited. The particular responses made by the observers varied with the condition, and is outlined below.

In each condition, participants completed 2 blocks of 100 trials, yielding 20 trials for each of the 5 different pursuit velocities. The order of trial presentation within a trial block, including the direction of lateral stimulus window translation, was randomized.

Stimuli—The random-dot stimulus construction was similar to the Common Motion Far (CMF) stimulus outlined in Figure 2. The critical feature of this stimulus is that the direction of $d\alpha$ and $d\theta$ are always in opposite directions and therefore the stimulus should appear farther away than the point of fixation. However, this stimulus was subjected to a few changes here. First, to determine the role of common motion, over the six conditions

described below the random-dot stimulus configuration was reduced from a sinusoidal velocity profile, with a range of velocities, to a simple square-wave with a single velocity. Second, to understand the role of M/PR, it was necessary to increase the M/PRs in the current study so that the velocity ratio in each condition spans $M/PR = 1$. To do this, a range of pursuit ($d\alpha$) velocities, between 2 deg s^{-1} and 10 deg s^{-1} , was used in different trials. Also, the average retinal image motion ($d\theta$) velocities were increased from Experiment 1, and manipulated in the six different conditions below.

Condition 1: Stimuli were two sinusoidal half-cycles as described in the CMF condition of Experiment 1. The peak retinal image velocity ($d\theta$) within the two sinusoidal velocity profiles of the stimulus were 6 deg s^{-1} and 8 deg s^{-1} , with an average $d\theta$ of 5.25 and 6.5 deg s^{-1} respectively, and an average “common motion” velocity for the entire stimulus of 5.9 deg s^{-1} . As outlined above, the pursuit ($d\alpha$) velocities ranged between 2 deg s^{-1} and 10 deg s^{-1} in different conditions. This produced a range of average M/PR for the faster $d\theta$ stimulus segment (what should be perceived as the more distant part) of between 0.65 and 3.25.

Similar to the procedure of Experiment 1, a key press initiated a 1.6 second presentation that consisted of a single rightward or leftward stimulus window translation. The observer’s task was to report whether the upper or lower portion of the stimulus appeared farther away in depth.

Condition 2: Unlike the brief, single direction stimulus presentation used in the current study, Mitsudo and Ono (2007) used a continuous oscillating stimulus presentation. To determine whether this difference in stimulus presentation has a significant effect on the reported percept, this condition used a continuous stimulus presentation in which the stimulus translated back and forth (with reversing $d\alpha$ and $d\theta$ directions) until a judgment and response was made by the observer. In all other aspect, the stimulus was identical to that used in Condition 1.

Condition 3: To eliminate the element of “common motion” between the two halves of the stimulus, here the stimulus half with the lower $d\theta$ velocity was eliminated, leaving only the “faster” half of the stimulus with the parameters described in Condition 1 above. This resulting “half-cycle” stimulus could appear above or below the fixation point, and could translate ($d\alpha$) leftward or rightward. However, since this stimulus change necessitated a new response, the observer’s task in this condition was to report whether the stimulus appeared nearer or farther than the fixation point.

Condition 4: The sinusoidal velocity profile of the stimuli used in previous conditions could be considered as retaining an element of common motion. That is, the sinusoidal velocity profile of the stimulus in Condition 3 contained $d\theta$ velocities ranging between 4 deg s^{-1} and 8 deg s^{-1} , with an average $d\theta$ velocity of 6.5 deg s^{-1} . This creates a range of M/PRs spanning 1.0. The lowest $d\theta$ velocity, both nearest to and farthest from the fixation point, generates an M/PR much lower (62%) than at the peak of the sinusoid. Perhaps these border regions, with lower $d\theta$ velocities, influence the perceived depth sign of the entire sinusoid. That is, these regions have a $M/PR < 1$ and are therefore perceived as representing points

farther away than the fixation point, which may subsequently influence the perceived depth sign of the entire stimulus.

To address this issue caused by having a range of $d\theta$ velocities, this and two subsequent conditions replace the multiple-velocity, random-dot stimuli with a horizontally-oriented, laterally-translating square-wave stimulus. This stimulus had a single $d\theta$ velocity, the same lateral extent of the random-dot stimuli, and was readily perceived as a singular object translating laterally. This square-wave stimulus had a cycle of 64 min, a height of 32 min, and appeared 16 min above or below the fixation point. Against the black background the white squares of the stimulus comprised roughly the same number of pixels as the previous random-dot stimulus. In this condition the $d\alpha$ (pursuit) velocities again ranged between 2 deg s^{-1} and 10 deg s^{-1} , and the $d\theta$ (retinal image) velocity was fixed at 6 deg s^{-1} . These velocity parameters produced a range of M/PR values between 3 and 0.6, which, if the M/PR hypothesis is correct, should be perceived nearer and farther than fixation respectively.

Conditions 5 and 6: These conditions used the same translating square-wave grating stimulus as the previous condition, but here the $d\theta$ (retinal image) velocity was fixed at 4 deg s^{-1} (Condition 5) and 2 deg s^{-1} (Condition 6). These changes in $d\theta$ give M/PR ranges of 2 to 0.4, and 1 to 0.2 respectively. This stimulus change should shift the results of both conditions toward more “correct” depth judgments as the M/PR values (for the same pursuit velocities) are now less than 1.

Results

For Conditions 1 and 2 the proportion of trials in which the faster $d\theta$ velocity was perceived as farther away (than the slower $d\theta$ velocity) was determined for each of the 5 stimuli (Figure 6). In both conditions the stimuli differed in $d\alpha$ (pursuit) velocity, which conferred a change in M/PR ratio, shown on the horizontal axis in Figure 6. The results show that as M/PR increases, the faster $d\theta$ velocity is less likely to be perceived as farther away than the slower velocity. This means that with the same retinal stimulus ($d\theta$) a reduction in pursuit ($d\alpha$) velocity (moving rightward in Figure 6) produces a change in the perceived relative depth of the stimulus. One unusual aspect of this result is that a faster pursuit velocity produces the geometrically “correct” percept while the slower pursuit velocity does not. Of course, we contend that this is not due to any independent role of pursuit, but instead due to the slow pursuit velocity being part of an unusual set of motion parallax parameters that the visual system has difficulty interpreting because they are not encountered in natural viewing conditions.

The current results are similar to the pattern of results reported by Mitsudo and Ono (2007). However, here only at the largest M/PR of 3.25 is the faster $d\theta$ more likely to be perceived as nearer. Moreover, continuous presentation of the stimulus did not make that percept more likely (upward shift of function for Condition 2 shown with grey squares). As noted above, Mitsudo and Ono (2007), using a stimulus with only two $d\theta$ velocities, found that the faster $d\theta$ was only perceived farther away when its M/PR value was < 1 . This difference from the current results might actually be due to differences in stimulus velocities. The current stimulus, with a sinusoidal velocity profile, included a range of $d\theta$ velocities. Perhaps slower

$d\theta$ velocities, nearer the fixation point and generating lower M/PRs, helped produce a percept in which the faster $d\theta$ velocity was still perceived as farther away. This possibility is explored in Conditions 4, 5 and 6.

The results of Conditions 3 through 6 are shown in Figure 7. In a change from the previous figure, here the proportion of trials in which each stimulus was perceived as farther than the point of fixation is shown along the vertical axis. For Condition 3 (grey triangles), the change from “common motion” stimulus with two half-cycles (Condition 1) to single “half-cycle” stimulus results in the same change in perceived depth sign with a change in M/PR. The two stimuli with $M/PR > 1$ (3.25, 1.63) are more likely to be perceived “incorrectly” nearer than fixation than the three stimuli with $M/PR \leq 1$. Indeed, a similar pattern of results is seen with Conditions 4, 5, and 6: stimuli with $M/PR \leq 1$ are more likely to be perceived “correctly” as farther than the fixation point, while stimuli with $M/PR > 1$ are more likely to be “incorrectly” perceived as nearer than the fixation point. The only difference between the stimuli in these three conditions is that changes in $d\theta$ velocities are used to produce changes in the M/PR, rather than the changes in $d\alpha$ velocities between trials within each of the conditions.

Conclusion

These results indicate that changes in perceived depth sign accompanying changes in either retinal image ($d\theta$) or pursuit ($d\alpha$) velocity are consistent with the notion of typical/interpretable and atypical/uninterpretable motion parallax stimuli as predicted by the M/PR model. An explanation relying on Mitsudo and Ono’s (2007) additive model is untenable due to the elimination of “common motion” from the stimuli, and the lack of rationale for why a stimulus is perceived with a reversed depth when $v_{\text{pursuit}} > v_i$.

In each of these six conditions the components of the retinal image motion ($d\theta$) remained the same between trials; only the pursuit ($d\theta$) velocity, and therefore the M/PR, changed, a resulted in a change in perceived depth sign. In the last three conditions, a change in M/PR was produced by changes in $d\theta$ velocities, further supporting the proposal that the visual system has difficulty generating the proper percept for stimuli having $M/PR > 1$, regardless of the particular $d\theta$ or $d\alpha$ velocities.

Discussion

The results of the first experiment demonstrate that the allocentric M/PR model of motion parallax provides good quantitative predictions for all different types of common and divergent motion stimuli. The second experiment employs the M/PR model to produce a testable hypothesis as to why some parameter combinations are misperceived, while others are perceived geometrically correct. While Mitsudo and Ono (2007) suggest “...Nawrot’s model cannot predict depth order...in which a retinal stimulus contains common motion...” (p. 126), we conclude that the allocentric M/PR model has much greater utility in aiding our understanding of the perception of depth from motion parallax than does the egocentric additive model.

Another important advantage of the M/PR model is that it goes beyond the depth ordinal scale used by the additive model to provide a depth metric (ratio scale) for the perception of depth from motion parallax (Nawrot & Stroyan, 2009; Stroyan & Nawrot, 2012). With such a metric model it is possible to demonstrate that the perception of depth magnitude from motion parallax is greatly foreshortened compared to the same depth geometry presented with binocular disparity (Nawrot et al, 2014). Specifically, the comparison with binocular disparity is particularly apt. The M/PR model operates with the very same framework, here based on the dynamic geometry rather than the static geometry of binocular disparity. Consider that binocular disparity information provides depth information relative to the point of fixation, which has disparity = 0, while the M/PR provides the same depth information relative to the point of fixation, which has retinal motion = 0. Similarly, while opposite disparity signs correspond to opposite depth signs, opposite directions of retinal motion correspond to opposite depth signs. Moreover, both binocular disparity and motion parallax scale depth with regard to egocentric viewing distance, but neither provide information regarding viewing distance. Given the importance of the internal representation of viewing distance for both binocular disparity and motion parallax, the visual processes for egocentric distance perception remain an important, and interrelated, topic of study (Gajewski et al., 2014a; 2014b; Li & Durgin, 2012; Ooi & He, 2007).

Acknowledgments

This work was supported by an NIH Centers of Biomedical Research Excellence (COBRE) Grant: NIH P20 RR020151 and P20 GM103505. Special thanks to Dan Gu and the Technical Services Core for assistance with programming.

References

- Brainard DH. The Psychophysics Toolbox. *Spatial Vision*. 1997; 10:433–436. [PubMed: 9176952]
- Fermüller C, Aloimonos Y. On the geometry of visual correspondence. *International Journal of Computer Vision*. 1997; 21(3):223–247.
- Gajewski DA, Philbeck JW, Wirtz PW, Chichka D. Angular declination and the dynamic perception of egocentric distance. *Journal of Experimental Psychology: Human Perception and Performance*. 2014a; 40(1):361. [PubMed: 24099588]
- Gajewski DA, Wallin CP, Philbeck JW. Gaze behavior and the perception of egocentric distance. *Journal of Vision*. 2014b; 14(1):20. [PubMed: 24453346]
- Gordon DA. Static and dynamic visual fields in human space perception. *JOSA*. 1965; 55(10):1296–1302.
- Hanes DA, Keller J, McCollum G. Motion parallax contribution to perception of self-motion and depth. *Biological Cybernetics*. 2008; 98:273–293. [PubMed: 18365242]
- Klatzky, RL. *Spatial Cognition*. Springer; Berlin Heidelberg: 1998 Jan. Allocentric and egocentric spatial representations: Definitions, distinctions, and interconnections; p. 1-17.
- Kleiner M, Brainard D, Pelli D, Ingling A, Murray R, Broussard C. What's new in Psychtoolbox-3. *Perception*. 2007; 36:1.
- Koenderink JJ, van Doorn AJ. Invariant properties of the motion parallax field due to the movement of rigid bodies relative to the observer. *Optic Acta*. 1976; 22:773–779.
- Koenderink JJ, van Doorn AJ. Facts on optic flow. *Biological Cybernetics*. 1987; 56:247–254. [PubMed: 3607100]
- Li Z, Durgin FH. A comparison of two theories of perceived distance on the ground plane: The angular expansion hypothesis and the intrinsic bias hypothesis. *I-Perception*. 2012; 3:368–383. [PubMed: 22792434]

- Longuet-Higgins HC, Prazdny K. The interpretation of a moving retinal image. *Proceedings of the Royal Society of London Series B*. 1980; 208:385–397.
- Mitsudo H, Ono H. Additivity of retinal and pursuit velocity in the perceptions of depth and rigidity from object-produced motion parallax. *Perception*. 2007; 36:125–134. [PubMed: 17357710]
- Miles FA, Busetini C. Ocular compensation for self-motion. *Visual mechanisms. Annals of the New York Academy of Sciences*. 1992; 656:220–232. [PubMed: 1599145]
- Nadler JW, Angelaki DE, DeAngelis GC. A neural representation of depth from motion parallax in macaque visual cortex. *Nature*. 2008; 452:642–645. [PubMed: 18344979]
- Nadler JW, Nawrot M, Angelaki DE, DeAngelis GC. MT neurons combine visual motion with a smooth eye movement signal to code depth sign from motion parallax. *Neuron*. 2009; 63:523–532. [PubMed: 19709633]
- Naji JJ, Freeman TCA. Perceiving depth order during pursuit eye movements. *Vision Research*. 2004; 44:3025–3034. [PubMed: 15474575]
- Nakayama K, Loomis JM. Optical velocity patterns, velocity-sensitive neurons, and space perception: A hypothesis. *Perception*. 1974; 3:63–80. [PubMed: 4444922]
- Nawrot M. Eye movements provide the extra-retinal signal required for the perception of depth from motion parallax. *Vision Research*. 2003; 43:1553–1562. [PubMed: 12782069]
- Nawrot M, Joyce L. The pursuit theory of motion parallax. *Vision Research*. 2006; 46:4709–4725. [PubMed: 17083957]
- Nawrot M, Ratzlaff M, Leonard Z, Stroyan K. Modeling depth from motion parallax with the motion/pursuit ratio. *Frontiers in Psychology*. 2014; 5
- Nawrot M, Stroyan K. The motion/pursuit law for visual depth perception from motion parallax. *Vision Research*. 2009; 49:1969–1978. [PubMed: 19463848]
- Ono H, Ujike H. Motion parallax driven by head movements: Conditions for visual stability, perceived depth, and perceived concomitant motion. *Perception*. 2005; 34:477–490. [PubMed: 15943054]
- Ono ME, Rivest J, Ono H. Depth perception as a function of motion parallax and absolute distance information. *Journal of Experimental Psychology: Human Perception and Performance*. 1986; 12:331–337. [PubMed: 2943861]
- Ooi TL, He ZJ. A distance judgment function based on space perception mechanisms: revisiting Gilinsky's (1951) equation. *Psychological Review*. 2007; 114(2):441. [PubMed: 17500634]
- Pelli DG. The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*. 1997; 10:437–442. [PubMed: 9176953]
- Perrone JA, Stone LS. A model of self-motion estimation within primate extrastriate visual cortex. *Vision Research*. 1994; 34:2917–2938. [PubMed: 7975326]
- Rogers B, Graham M. Motion parallax as an independent cue for depth perception. *Perception*. 1979; 8:125–134. [PubMed: 471676]
- Stroyan K, Nawrot M. Visual depth from motion parallax and eye pursuit. *Journal of Mathematical Biology*. 2012; 64(7):1157–1188. [PubMed: 21695531]

Figure 1. Common Motion Near (CMN) conditions and the corresponding values according to the M/PR. The figures show three-dimensional representations of the common motion conditions in both standard and alternate phases, along with the M/PR values for both the upper and lower stimulus segments.

Figure 2. Common Motion Far (CMF) conditions and the corresponding values according to the M/PR.

Figure 3. Divergent Motion Near (DMN₁) conditions and the corresponding values according to the M/PR.

Figure 4. Divergent Motion Far (DMF₁) conditions and the corresponding values according to the M/PR.

Figure 5.

Average d' values for each condition, including standard error bars. The horizontal axis displays the four stimulus conditions while the vertical axis represents the average d' value in each condition. All conditions exhibited high d' values, showing accuracy in performing the depth discrimination task.

Figure 6.

Shown is the proportion of relative depth judgments for Conditions 1 and 2. The horizontal axis displays the M/PR of the 5 stimuli on a log scale. Error bars show ± 1 SE.

Figure 7.

Shown is the proportion of relative depth judgments for Conditions 3 through 6. The horizontal axis displays the M/PR of the 5 stimuli in each condition on a log scale. Error bars show ± 1 SE.