

Object-centered reference frames in depth as revealed by induced motion

Jasmin Léveillé

Center for Computational Neuroscience and Neural Technology, Boston University, Boston, MA, USA



Emma Myers

Center for Computational Neuroscience and Neural Technology, Boston University, Boston, MA, USA



Arash Yazdanbakhsh

Center for Computational Neuroscience and Neural Technology, Boston University, Boston, MA, USA



An object-centric reference frame is a spatial representation in which objects or their parts are coded relative to others. The existence of object-centric representations is supported by the phenomenon of induced motion, in which the motion of an inducer frame in a particular direction induces motion in the opposite direction in a target dot. We report on an experiment made with an induced motion display where a degree of slant is imparted to the inducer frame using either perspective or binocular disparity depth cues. Critically, the inducer frame oscillates perpendicularly to the line of sight, rather than moving in depth. Participants matched the perceived induced motion of the target dot in depth using a 3D rotatable rod. Although the frame did not move in depth, we found that subjects perceived the dot as moving in depth, either along the slanted frame or against it, when depth was given by perspective and disparity, respectively. The presence of induced motion is thus not only due to the competition among populations of planar motion filters, but rather incorporates 3D scene constraints. We also discuss this finding in the context of the uncertainty related to various depth cues, and to the locality of representation of reference frames.

Introduction

One of the notions having received considerable attention in the study of perception is that of the reference frame encoding the spatial structure of objects (Rock, 1990; Lappin & Craft, 2000). Research conducted in the past few decades suggests that a multiplicity of such reference frames are simultaneously at work in the visual system, indicating that which

frame is active at a given time depends in part on the information present in the surrounding environment (Indow, 2004). Many experiments have shown that the features of a given stimulus component may be coded relative to another, a case referred to as an object-centric reference frame (Wade & Swanston, 1987). In theory, object-centric reference frames could provide support for affine invariant object recognition (Marr & Nishihara, 1978; Hinton, 1981) and for structure-from-motion (Koenderink, 1994).

The phenomenon of induced motion has been widely used to probe the nature of object-centric reference frames (Duncker, 1929/1938). In induced motion, the motion of a frame in a particular direction induces a motion component in the opposite direction in a static dot. For example, in a display consisting of a vertically oscillating dot surrounded by a horizontally oscillating frame, the dot appears to be moving diagonally (Wallach, Bacon, & Schulman, 1978), as would be predicted by subtracting the motion component of the frame from that of the dot. Although a relatively new subject in laboratory experiments, induced motion was actually noted several times throughout history, the earliest instance of which may be found in the works of Ptolemy (Smith, 1996).

A considerable body of research has focused on characterizing the “low-level” underpinnings of induced motion, such as how it is affected by stimulus aspects like contrast (Murakami, 1999), speed and texture (Cohen, 1964). It is clear that it is important to have a full characterization of how variations in these stimulus dimensions affect induced motion in order to fully understand the phenomenon. The investigations we report in this paper concern instead the role of 3D geometry on the induced motion illusion. That is,

Citation: Léveillé, J., Myers, E., & Yazdanbakhsh, A. (2014). Object-centered reference frames in depth as revealed by induced motion. *Journal of Vision*, 14(3):15, 1–11, <http://www.journalofvision.org/content/14/3/15>, doi: 10.1167/14.3.15.

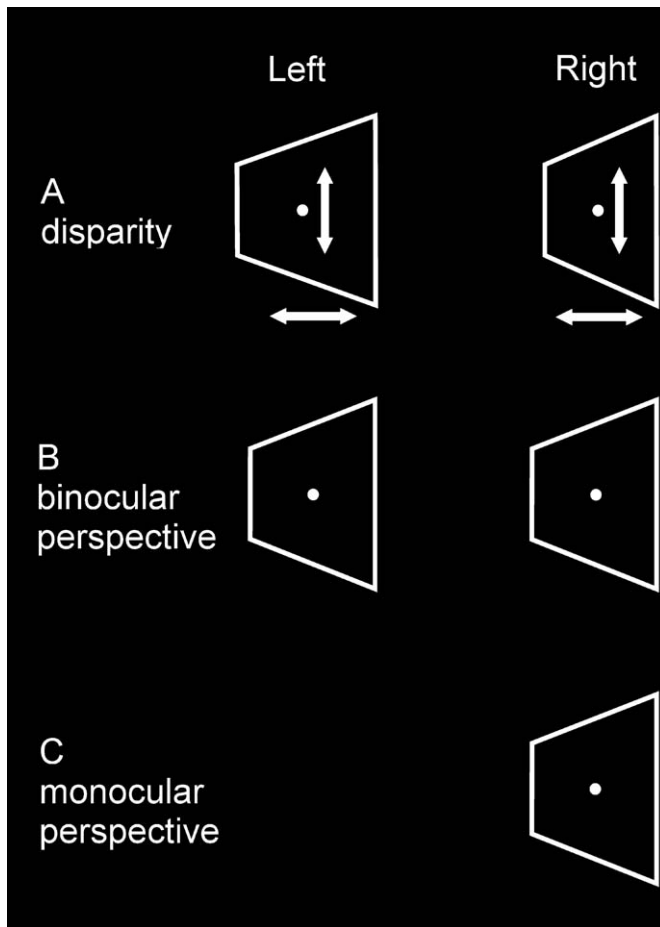


Figure 1. Experimental displays presented to the left and right eye. (A) Stimulus 1 had both disparity and perspective cues (DP). (B) Stimulus 2 had perspective cues but no disparity (BP), and (C) Stimulus 3 consisted of a perspective cue presented to only one eye (MP). Arrows in (A) denote the frontoparallel directions of motion, and are omitted from (B) and (C) for clarity. Actual stimulus boundaries were white on a black background, as shown.

rather than studying the effect of variations in “surface” properties (e.g., texture density, contrast), we assume those to be fixed and study how varying the 3D layout of the stimulus affects the percept of induced motion.

Induced motion and its relationship to depth have been studied in a number of experiments (e.g., Gogel & Tietz, 1976; Gogel & MacCracken, 1979; Di Vita & Rock, 1997; Léveillé & Yazdanbakhsh, 2010). One general finding is that coplanarity of the dot and frame is not at all necessary for motion induction. Other studies have shown that induced motion can occur along the depth dimension provided that the inducer can be seen as moving in depth (Farnè, 1972; Farnè, 1977; Gogel & Griffin, 1982; Harris & German, 2008; Nefs & Harris, 2008). Farnè (1972, 1977) showed the possibility of induced motion in depth using displays

that consisted of static target lines and a background surface oscillating in depth. One crucial difference between this setup and the one we propose here is that here the reference frame is not seen as moving in depth, despite that it triggers induced motion in depth. Similarly, Gogel and Griffin (1982) used a display consisting of target and reference dots and showed that induced motion could be perceived either in the frontoparallel plane or in depth, depending on whether the reference dots travelled along the corresponding dimension. Harris and German (2008) found that, for matching extents of retinal motion, observers perceive equal amounts of induced motion on the frontal plane and in depth, suggesting that the illusion does not depend on 3D scaling. Nefs and Harris (2008) compared the effects of size (looming) and binocular disparity as well as fixation condition in displays in which the display elements were perceived as moving either in depth or in the frontal plane. In the depth condition, whereas changes in size did not lead to induced motion in depth, binocular disparity did, but primarily when fixation was on the inducer. The finding that binocular disparity can lead to induced motion relates to our present experiments, although we do not specifically test for the effect of fixation, and the display elements here do not actually move along the depth dimension. Moreover, we chose a point-like size for the target dot (13 arcmin of visual angle) to further weaken the size-induced depth effect.

Here we show that it is possible to induce a motion component in depth in a target dot when using an induced motion display in which no stimulus element actually moves in depth. In order to do this, we propose a new stimulus display in which a component of slant is added to the inducer using either perspective or disparity as depth cue (Figure 1). Crucially, in our stimuli, motion of the inducer remains along the frontoparallel plane (i.e., left-right oscillation), and can only be seen as having a three-dimensional motion component when surface slant is taken into consideration. Unlike in the displays of previous 3D-induced motion experiments, the frame does not actually move in depth. The lack of an actual component of motion in depth of the frame, coupled with its slant, allows us to study the formation of object-centric coordinates in depth.

In particular, we can study whether the presence of slant in the oscillating frame will make the target dot appear as if it were moving in depth. Based on the *equidistance tendency* (Gogel, 1965), if the relative positions of the dot and frame are sufficiently ambiguous, the dot could be assimilated to the frame, in which case it would be seen as moving along the frame. On the other hand, if the relative depth of the dot and frame is unambiguous and computed locally, induced motion in depth could be perceived. The term

“locally” here refers to a local neighborhood that encompasses the intersection of the slanted inducer frame and the line of sight. Assuming a slanted frame whose leftmost vertical edge is farther in depth than its rightmost edge and a target dot with no discernible horizontal motion component, the distance between the dot and its projection along the line-of-sight on the frame will increase/decrease as the frame moves in the left direction. If the computations that lead to perceived motion are influenced by spatial displacement in depth, then such a change in distance could be perceived as the dot moving away from the observer.

Finally, if induced motion results only from the competition among opposite-directed motion filters coding for motion in the 2D frontal plane, neither induced motion nor motion assimilation to the surface of the frame should be observed in our stimulus. Our experiment specifically addresses whether the different depth cues used here lead to either of these perceptual outcomes.

Material and methods

Participants

Nine naive (to the purpose of the experiments) and three non-naive subjects participated in the experiment, all with normal or corrected-to-normal vision. Our experiment requires that subjects form stable percepts of slant of the inducer while also estimating the motion trajectory of the target. In various control pre-experiments, we have found that forming stable slant percepts tends to be rather difficult for nonexperienced subjects who may not be accustomed to viewing displays defined primarily by either binocular disparity or perspective. Data from six naive subjects were hence ultimately excluded due either to a bistable impression of some of the stimuli (two subjects) or to an unstable slant percept in the inducer frame across conditions (four subjects). See Results section for details. Most subjects took approximately 45 minutes to complete all experimental trials.

Stimuli

Stimuli were designed using Matlab’s Psychtoolbox, presented on a ColorEdge CGF241W 24-inch LCD monitor (507 mm × 317 mm; Eizo Nanao Corporation, Japan) with a resolution of 1920 × 1200 pixels, and viewed through a haploscope such that the total distance between the eye and the monitor was approximately $D = 63$ cm. Room lighting was adjusted according to each subject’s comfort level.

We designed three types of experimental trials in order to measure the effects of binocular disparity, binocular perspective, and monocular perspective. Each of the three stimuli made different use of depth cues. Stimulus 1 (binocular disparity cue, Figure 1A) and Stimulus 2 (binocular perspective cue, Figure 1B) were composed of two frame-and-dot images, which subjects binocularly fused while viewing through a haploscope. Perspective was added to Stimulus 1 because disparity alone did not lead to a strong impression of slant, as may be expected from using a rectangle as inducer (He & Ooi, 2000). Stimulus 3 (monocular perspective cue, Figure 1C) differed from Stimulus 2 in that the frame-and-dot display was shown only to one of the observer’s eyes. Although Stimuli 2 and 3 did not technically require the use of the haploscope, they were viewed through the haploscope to maintain consistency with the experimental setting of Stimulus 1. The monocular perspective condition was included in order to assess the possible influence of the zero-disparity cue in the binocular perspective condition. The three conditions are hereafter referred to as DP, BP, and MP (disparity and perspective as in Stimulus 1, binocular perspective as in Stimulus 2, and monocular perspective as in Stimulus 3).

Inducer frames were trapezoidal in shape and oscillated horizontally; target dots oscillated vertically. To avoid a cue conflict due to a possible size-induced depth effect, target dots were only 13 arcmin of visual angle. The vertical extent of the dot’s motion path and the horizontal extent of the frame’s motion path were 2.5 and 2.9 degrees of visual angle, respectively, and the time for one complete oscillation was roughly 3.4 s. This speed was determined in pilot studies as being comfortable for sustained fixation, as well as slow enough to avoid motion smear but rapid enough for appreciable change.

Depth cues were initially calculated so as to suggest a 7° (width) × 4° (height) rectangular wall-like surface tilted at 37° , with the right-hand side appearing nearer and the left-hand side farther than the screen. In pilot experiments we noticed that it was often more difficult for subjects to perceive the frame’s slant in depth when it was defined by perspective rather than by disparity. Since perception of *some* frame slant is critical to the present experiment, we added a calibration step whereby subjects were able to adjust the perspective cue of each condition so as to make the percept of slant consistent across conditions.

Disparity cue

Assuming fixation on the target dot at the monitor’s depth, the relative disparities for the near and far edges of the DP inducer frame were approximately -16 and 15 arcmins, respectively. This yielded horizontal extents

of 4.1 and 3.27 degrees of visual angle for the left and right half-images, respectively. The vertical lengths of the left and right edges were the same as for the BP and MP conditions (see below).

Perspective cue

The initial height of each vertical side of the inducer frame was computed using Thales' theorem in order to achieve a certain slant in depth. That is, the actual height (h_{act}) on the monitor of each vertical side of the inducer frame was given by $h_{\text{act}} = (D \times h_{\text{app}}) / (D + d)$, where D is the distance from eye to screen, d is the apparent depth of the vertical side relative to the monitor (negative when nearer, positive when farther), and h_{app} is the apparent height of the frame, that is, the height that is suggested by the slant configuration.

Procedure

A stationary stimulus was used at the beginning of the experiment to calibrate the haploscope and thereby ensure that sufficient frame slant be perceived during the binocular disparity trials. The calibration stimulus consisted of a white dot surrounded by a white square in the center of a black screen, a configuration not too different from the actual experimental stimuli. The disparity cue in the calibration trial was set so that the central dot appeared to be nearer than the surrounding square. Subjects confirmed verbally that they could comfortably fuse the images, and that the dot appeared a few centimeters nearer than the square.

Adjustment of perspective cue

In an earlier version of this experiment, the perceived angle in depth created by the DP inducer frame was consistently greater than that created by the other two inducer frames (BP and MP). We therefore attempted to create the same impression of slant in depth for the inducer frame across the three stimuli in conditions DP, BP, and MP. After calibrating the haploscope, participants completed a series of trials in which they compared a reference frame (a DP reference frame) to an adjustable frame (an inducer frame from any of the three stimuli). In each trial, these two frames were stationary and were viewed consecutively, with the ability to toggle between the reference and adjustable frames. In this section, the task (described below) required subjects to adjust the perspective cue of the adjustment frame until its perceived slant in depth matched that of the reference frame. Subjects had to ignore the resulting differences between the proportions of the adjustment frame and those of the reference frame. This is somewhat more difficult than other tasks

of the experiment, especially for subjects inexperienced with the haploscope; therefore frames in this section were stationary in order to ease that difficulty.

Subjects used the arrow keys to change the vertical edge heights of the adjustable frame such that the two frames appeared to have an equal degree of slant in depth, or as close to equal as the subject was able to make them. This session consisted of 15 randomized trials (five trials for each of the three possible comparisons: a DP frame compared to another DP frame, a BP frame, or an MP frame). In this way, the subject indicated the strength of perspective cue required to make the BP and MP frames appear to have an apparent slant in depth as close as possible to that of a DP frame. For each condition, the average vertical-edge heights chosen by the subject across the relevant five trials were used as edge heights for the inducer frames in BP and MP conditions for the rest of the experiments. The DP–DP comparisons served to confirm that subjects did not adjust the perspective cue when viewing two stimuli that were in fact the same.

We considered an alternative design in which the reference and adjustable frame would be viewed simultaneously, and were vertically aligned on the screen. The sequential viewing design was selected instead to avoid (a) possible effects of the two frames having different vertical angles, and (b) possible top-down effects of viewing a stimulus for a sustained period of time.

Estimate of the inducer frame slant

Following the initial static calibration and slant matching trials, observers were asked to adjust the angle of a 3D rotatable rod (defined with disparity cues) so as to match the perceived slant in depth of an inducer frame. This was to (a) determine whether the inducer frames in the three stimuli did have a similar apparent slant in depth after the adjustments of the previous session, and (b) enable a comparison between the perceived slant of the inducer frame and the dot's trajectory in the next session. In the current session, the frame moved horizontally as described in the Stimuli section. In order to facilitate matching of the rod to the slant of the inducer frame, subjects were able to freely switch between the rod display and the frame display. A rod orientation of 0° indicates that the frame appeared flat on the monitor. This section consisted of 15 randomized trials (five per condition).

We excluded subjects who had a bistable impression of the rod's orientation in depth by checking for inconsistency between the reported slant of the matched rod and that of the inducer frames. That is, if a subject sometimes left the rod with the right-hand end appearing more distant (despite the fact that the right hand edge of the inducing frame was set to appear

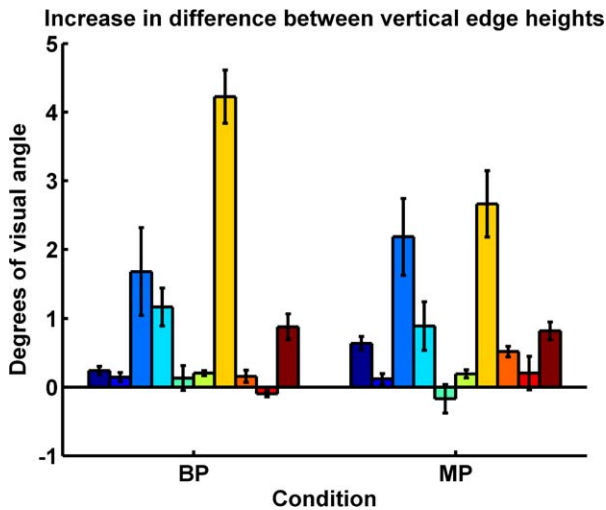


Figure 2. Subjects were asked to match the perceived slant in the MP and BP conditions to that in the DP condition by adjusting the lengths of the vertical edges of the slanted frame. Each bar shows the relative change in the difference between the right-hand and left-hand heights in either the MP or BP conditions compared to the DP condition. Each color corresponds to one subject in each respective condition. Error bars show the standard error computed across trials. As seen by the fact that most bars are above zero, most subjects needed to increase the difference between right-hand and left-hand edges in order to match disparity-defined slant.

closer), that subject was excluded from further analyses. For the final two subjects, a point at one end of the rod was colored red, because this appeared to help subjects avoid a bistable impression of the rod's orientation in depth. These subjects were also instructed to orient the rod such that the red end appeared farther away unless they were leaving the rod "flat" on the monitor. This gave us additional confirmation that these subjects were not getting a bistable percept of the rod.

Estimate of dot trajectory angle

This section resembled the frame-slant-estimation section, but in addition to the moving inducer frame subjects saw a dot moving vertically as described in the Stimuli section. Subjects were instructed to fixate on the dot and to rotate the rod so as to match the perceived direction of motion of the target dot in depth (cf. van Ee, van Dam, & Erkelens, 2002, for a similar method). A rod orientation of 0° indicates that the target dot was perceived as moving purely on the frontoparallel plane, whereas an orientation of 90° indicates motion parallel to line-of-sight. Stimulus presentations were randomized and each stimulus condition (DP, BP, or MP) was repeated five times, yielding a total of 15 trials.

Estimate of dot trajectory without inducer frame

Finally, six subjects completed an additional set of six control trials in which the oscillating dot on the frontal plane was presented alone and the subject's task was to match the rotatable rod to the dot trajectory in depth. It is possible that some amount of slant in the dot trajectory as reported in the experimental trials could simply reflect an inherent bias in the subject's reports of the slant. The purpose of these control trials was thus to measure the presence of that possible bias.

Results

Disparity – Perspective matching

Of the 12 subjects originally recruited, two had a bistable impression of the rod's orientation. Data from these subjects is discarded from further analysis since the goal of the present experiment is specifically to test for the perception of slant in the dot's trajectory *given* perceived frame slant and assuming a stable percept of slanted rod.

The 10 remaining subjects made a variety of adjustments to the perspective cue in the BP and MP conditions (no subject made any adjustment to the DP condition perspective cue). Each vertical bar in Figure 2 shows the change in the BP or MP frame's vertical edge length that was performed by individual subjects in order to match the perceived DP frame slant.

A bar above zero indicates that, in order to match the slant of a perspective-defined frame to that of the disparity-defined frame, the subject increased the length of the near (rightmost) vertical edge while reducing the length of the far (leftmost) edge. In most (but not all) cases, subjects increased the perspective cue in the perspective-only conditions, indicating that the DP condition created a more pronounced slant. In other words, the perspective cue had to be more pronounced in either BP or MP frames in order to match the overall slant in the DP frames.

Subject estimates of inducer frame slant

When asked to report their percepts of slant of the frame following the initial Disparity-Perspective matching adjustments, six subjects perceived the adjusted inducer frames from each condition to have similar slants in depth (Figure 3; subjects are displayed in the same order as in Figure 2). Of the last four subjects shown in Figure 3, three saw the DP frame as much more slanted than the other frames despite the prior adjustment of the perspective cue, and two of

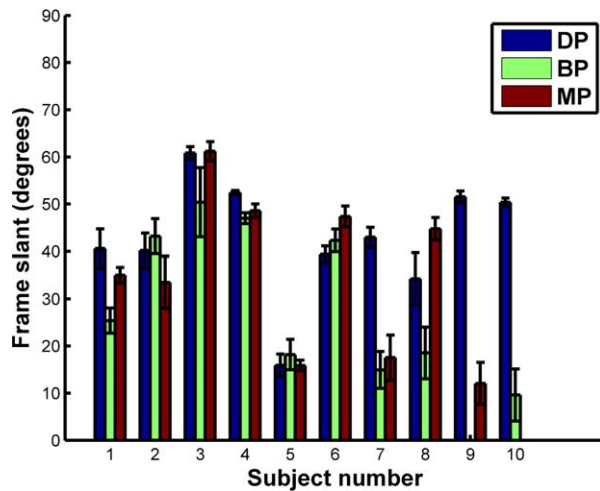


Figure 3. Subject estimates of frame slant under differing depth cue conditions, following initial matching with the DP frame. Perceived slants in the DP, MP, and BP conditions are shown in blue, green, and red, respectively. The four subjects who perceived different frame angles under different conditions are grouped at the end (for convenience, those subjects were not tested consecutively). Only subjects that were able to perceive similar amounts of slants across conditions were kept in the analysis of the induced motion trials. Error bars show the standard error computed across trials.

those in fact saw no slant in either the BP frame (Subject 9) or the MP frame (Subject 10). Subject 8 perceived the BP frame as much less slanted than the DP or MP frames. None of the four subjects had indicated difficulty in matching either BP or MP frames to DP frames (although all subjects were invited to do so).

Given that it was not possible to ascertain that the last four subjects would perceive a reasonably consistent frame slant during experimental trials—a necessary prerequisite to investigate slant-based induced motion in depth—their data was excluded from further analyses.

Induced motion in depth

Induced motion in depth can be determined by the perceived slant of the target dot’s trajectory relative to the frontal plane. Each vertical bar in Figure 4a shows the perceived target dot trajectory slant for the subjects not discarded during control trials and for the three experimental conditions—DP, BP, and MP—respectively. A negative slant indicates that the dot appeared to be moving farther in depth as the frame was moving to the left, and nearer in depth as the frame was moving to the right, as could be expected from induced motion (see also Figure 6b). One way to describe the associated

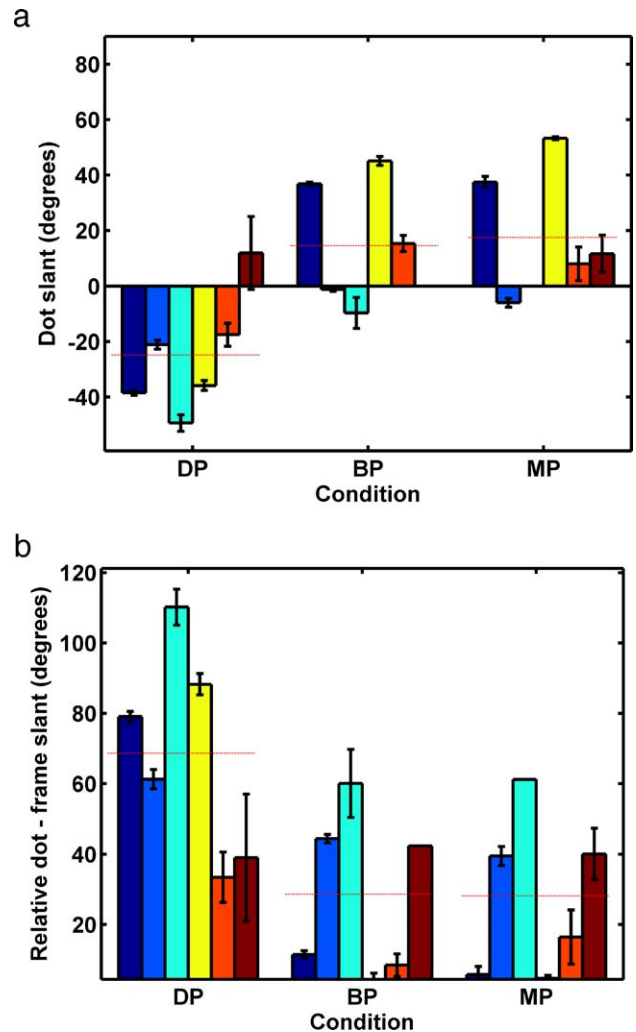


Figure 4. (a) Perceived angle in depth of the dot trajectory. A negative sign indicates that the dot was perceived as moving farther in depth as the frame was moving to the left, and nearer in depth as it was moving to the right, consistent with our definition of induced motion. A positive sign indicates that the dot was moving in a direction similar to the slant of the frame, which could result from assimilation. The DP condition thus resulted in induced motion whereas either the MP or BP conditions led to assimilation. (b) Absolute value of the perceived dot trajectory angle in depth (average across trials) minus perceived frame angle in depth (average frame estimates for a given subject within a given condition, cf. Figure 3). Error bars show the standard error computed across trials. Each color corresponds to an observer’s data. Pale green, orange, and red bars indicate data from naive subjects.

percept is to imagine that the dot moves “through” the frame.

According to this definition, five of the subjects clearly showed a pattern of induced motion in the DP trials. Hence, induced motion can be triggered by a slanted inducer frame that does not itself move in depth

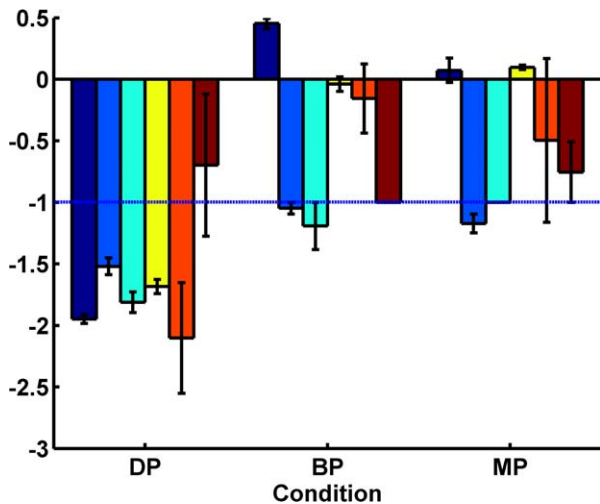


Figure 5. Dot trajectory angle minus frame angle, normalized by frame angle. Values below the line at -1 indicate induced motion in depth in the dot; values above it indicate assimilation of the dot's motion to the frame. Each color corresponds to an observer's data. Pale green, orange, and red bars indicate data from naive subjects.

but contains some amount of slant defined by a binocular disparity cue.

Motion assimilation, on the other hand, may be observed here in the trajectory of the target dot under perspective conditions, by noticing that the perceived trajectory reported seems pulled away from the frontal plane and toward the slanted frame (Figure 6a). According to the convention used in Figure 4a, this would be shown as a positive dot slant. Figure 4a shows that this was the case for most of the subjects, in either the BP or the MP condition. Hence, when the slant in the stimulus is defined primarily by perspective, motion assimilation occurs. A one-way, repeated-measures ANOVA confirms that the effect of condition on dot slant estimate is significant ($p = 0.01$, $F = 8.03$).

Figure 4b further shows that the dot's perceived motion in depth relative to the frame was greater in the DP condition than in either perspective-only condition for five of the six subjects (this effect is also significant; $p = 0.006$, $F = 8.72$).

Figure 5 shows the differences from Figure 4b normalized by perceived frame slant; i.e., (trajectory angle - frame angle) / frame angle. Here, if the inducer frame had no effect on the dot's trajectory in terms of depth, the resulting value would be -1 . If the dot's trajectory was assimilated to the slant of the frame, the value would be greater than -1 . If the frame induced motion in depth in the dot's trajectory, the value would be less than -1 .

Figure 5 clearly shows that, in five of the six subjects, the frame with both depth cues successfully induced motion in depth in the dot. In all six subjects, the frames with only the perspective cues resulted either in

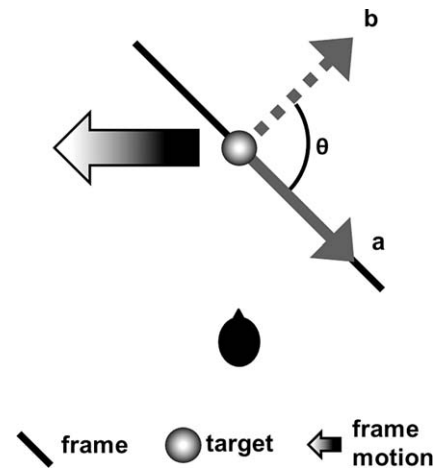


Figure 6. Different frames of reference lead to different dot motion percepts as the slanted frame moves to the left. (a) Perceived motion that occurs when perspective is dominant and frame slant is perceived. (b) Perceived motion of the dot that occurs in the disparity case.

assimilation to the frame's depth component or in little to no effect. A one-way repeated-measures ANOVA calculated on the normalized values indicates a significant effect of condition ($p = 0.01$, $F = 8.05$).

Observer bias in target dot slant estimate

It is in theory possible that part of the dot trajectory slant reported in Figure 4 originates in a form of observer bias that reveals itself in impoverished displays such as the one in which only the target probe is presented. To ensure that such a bias—if it exists—does not greatly affect our findings, five of the six remaining subjects performed an additional set of control trials in which their only task was to align on the frontal plane defined by the monitor, the trajectory of a sole moving dot in an otherwise completely empty background. Bias is then defined as the deviation from 0° of that adjusted dot trajectory. A negative bias indicates that subjects would have a natural tendency to align the dot trajectory in the direction that we identify with induced motion. Conversely, a positive bias would indicate a natural tendency to perceive the dot in the direction that we consider as motion assimilation. The group-averaged bias that we measured was -1.36° (average standard error 2.87), suggesting mainly that the deviations reported in Figure 4 might be slightly overestimating the effect of induced motion and underestimating the effect of motion assimilation. Nevertheless, since the bias is of such small magnitude compared to the reported dot trajectory slants in the induced motion stimulus, the presence of this bias does not change our conclusions.

Discussion

The experiments we conducted primarily aim to determine whether it is possible to trigger induced motion along the depth dimension with an inducing stimulus whose motion is within the frontoparallel plane. Our results show that this depends on the cue used to define depth in the inducer.

When slant is defined by disparity combined with perspective (Figure 1A, DP condition), the dot tends to be seen as moving in depth relative to the frame (Figure 6b). When slant is defined instead by perspective alone, the relative angle difference between the dot's trajectory and the frame is drastically reduced (Figure 6a), whether disparity information conflicts with the suggested perspective slant—as in the binocularly viewed perspective condition—or not. In some subjects, frame slant was simply not perceived, whereas in others, frame slant was still perceived (Figure 1B and C, BP and MP conditions). In either case, the dot's motion in depth relative to the frame was smaller than in the disparity case. The dot and frame configuration was thus perceived as either completely flat on the monitor or the dot's motion was assimilated to the slant of the frame.

In order to try to make motion in depth comparable across the various cue conditions, we asked subjects to match the slant of the frame in the BP and MP conditions to that of the DP condition (Figure 3), prior to performing the main experimental trials. Most subjects were able to reliably produce similar amounts of slant across frames (six subjects out of the initial 10) and the remaining four were discarded from further analysis. It is possible that further improvements to the slanted frame stimulus—such as adding a texture gradient, or varying the thickness of the various frame edges—would have produced more reliable percepts of slant in those subjects. On the other hand, such stimulus enhancements may have affected the relative motion and further complicated its analysis. Hence, we prefer instead to keep the stimulus as simple as possible and to interpret our results as demonstrating the possibility of induction/assimilation in depth given an assumed initial percept of slant. Due to the difficulty of perceiving slant, especially for frames where perspective is the cue that defines slant, it is not clear whether, in its present form, the stimulus we propose would easily lead to the 3D motion phenomenon in a large population of nonexperienced observers. The study of the possibility of motion induction/assimilation in such a population may require developing a more natural stimulus, perhaps using actual 3D objects rather than computer renderings. The fact that classic (planar) induced motion can be observed in rich, natural settings suggests that the 3D variant we propose may also be perceivable in similar settings.

On a related note, the fact that the diameter of the target dot was kept constant could be expected to have led to a cue conflict situation in which the dot's motion in depth would have been reduced. Although the results of Nefs and Harris (2008) suggest that looming does not lead to induced motion in depth, it remains possible that the perceived motion in depth in the current experiment would have been facilitated by changing the target dot's diameter in a manner consistent with the direction of motion in depth. Moreover, the target dot size in our experiments is 13 arcmin, a point-like size that could further weaken the looming cue. Another factor that might have contributed to reducing the magnitude of motion induction/assimilation is the luminance difference between the monitor's surface and the room, which could be perceived as a static, enclosing reference frame. Using a better calibrated monitor (or actual 3D stimuli) may thus yield stronger effects, although results of Di Vita and Rock (1997) for 2D stimuli suggest a limited influence on induced motion for such enclosing frames.

Note that the speed of the vertical frame boundaries was not adjusted so as to be consistent with the prediction of motion parallax for a slanted frame moving in depth. Rather, motion was generated by simple translation of rigid 2D shapes on the monitor. It is possible that this would again result in a cue conflict situation that would reduce the percept of frame slant and thereby reduce the perceived motion in depth.

It should be noted that, although the motion direction of the frame is in the frontoparallel plane, the fact that the frame is perceived as slanted in depth means that motion is also imparted to a region of the frame that spans its extent in depth. That is, despite that there is no motion other than in the frontal plane where the trapezoids are shown, motion is actually present in depth in the perceived slanted frame. This seemingly obvious fact could in turn mean that opposite-directed motion—along the frontal plane—in the target dot could be expected while its depth relative to the frame varies. According to this scenario, the target dot should be seen to move on the frontal plane in exactly the same way as in traditional 2D induced motion stimuli. This hypothesis is not corroborated by the present set of results. Rather, an additional illusory component of motion in depth is incorporated to the perceived trajectory of the target dot. Hence, this purely “two-dimensional” view of our display misses the fact that the dot's motion trajectory is not only the result of actual 2D motion opponency—something that may be considered as early in the visual processing hierarchy—but also of the relative displacement in depth.

Figures 4 and 5 suggest that the depth component is more robustly perceived in the moving target dot in the DP condition than in either perspective condition. This

could reflect a somewhat reduced impression of frame slant in perspective trials despite the fact that subjects matched the slant across conditions in initial control trials. Indeed, it is not possible to guarantee beyond any doubt that subjects perceived the same amount of frame slant during the experimental trials as in the control trials. Perhaps by attending more closely to the trajectory of the target dot, some of the frame's slant defined by perspective was lost. This possibility would be corroborated if, for example, it could be shown that the amount of slant perceived in a perspective stimulus increases/decreases as attention is shifted to/away from the slanted stimulus. We are not aware of such evidence at present. It is also possible that the less robust motion in depth seen in perspective stimuli is due to an inability of the visual system to resolve the fundamentally ambiguous position of the target dot relative to the frame as provided by a perspective cue (see below). The disparity cue, on the other hand, signals unequivocally the relative position of the various stimulus elements (up to a scaling), and in that case it may then be easier for the visual system to assign a particular 3D motion trajectory to dot.

We suggest that relative, rather than absolute, motion in depth is perceived due to general ambiguity in the depth of a visual scene. Figure 6 shows the different possible perceptual outcomes for the dot motion in the different cue conditions. As our results show, the exact nature of the perceived motion depends on which cues are present. In the case of perspective, we found that the target dot's depth was assimilated to the frame (Figure 6a). This is consistent with the equidistance tendency (Gogel, 1965) which in fact predicts that the dot's motion should be assimilated to the surface of the slanted frame when their exact depth ordering is sufficiently ambiguous. However, our results also show that in the case of the disparity cue, instead of assimilation of the dot to the frame, there is a clear component of induced motion in depth (Figure 6b). Hence, despite that the equidistance tendency does not apply in this case, there still remains sufficient ambiguity in the scene that the motion of a target dot will still be perceived relative to the frame. In both cases, the frame, rather than the dot, acts as the reference by virtue of its larger size.

Our results may also be interpreted from the viewpoint of the locality of the representation of relative depth. When slant is defined by disparity, as the frame moves to the left, the distance between the dot and a local neighborhood on the frame increases. Hence, if the dot is indeed represented relative to a *local* neighborhood of the frame, it will be seen as moving further away from the observer, consistent with the direction of *induced* motion and with the observations for the DP condition in Figure 4. A similar situation holds in the case of slant defined by perspective: if the

dot is represented locally relative to the frame, it will also be moving in depth but *along* the frame according to an equidistance tendency, as shown by the results for the BP and MP conditions in Figure 4.

The motion phenomenon we study here can be related to other perceptual phenomena found for static reference frame cues that have been investigated in a number of studies. He and Ooi (1999) showed in a Ternus display that static factors of perceptual organization affect apparent motion. Our results are consistent with a previous finding according to which the relative depth of two vertical lines is determined by their distance to a common slanted rectangle (He & Ooi, 2000). Specifically, the depth difference between two vertical lines is increased by separating them along the depth dimension relative to a slanted reference. The depth difference is smaller, however, if the lines are shown directly on top of a frontoparallel reference. The finding that the relative depth difference between the lines is smaller when they lie on top of a slanted, rather than a flat, reference frame can be seen as a static version of the phenomenon studied whereby the target dot is seen as moving in depth as its distance relative to the frame is varied.

In a set of studies on motion integration, Watamaniuk and McKee (1995) showed that when randomly moving dots were separated in depth from another moving target dot, it was no easier for observers to detect the motion of the target dot, suggesting that motion integration does not depend on depth assignment. The induced motion we study here is consistent with this finding in the sense that the motion of the reference frame—which is seen at a different depth than the target dot—nevertheless interacts with the motion of the target dot. The motion assimilation that we report for the perspective conditions is akin to the motion integration of Watamaniuk and McKee (1995), in that in both cases the target motion is mixed in with that of the reference frame, perhaps due to uncertainty in both kinds of displays. In a related study, He and Nakayama (1994) found that trajectory motion is guided by perceived surface belongingness in otherwise bistable apparent motion displays. The apparent motion trajectory perceived tends to be the one that preserves the relative depth relationship with the reference surface. This phenomenon is perhaps closest to our demonstration that motion assimilation occurs in the perspective conditions: In the case of uncertainty in the motion trajectory, the reference surface assimilates trajectory motion.

The experiments reported here did not explicitly address the roles of fixation condition or display element speed. It is possible that by fixating on the target, observers would have perceived frame slant differently and vice versa, and that the different slant would have affected the perceived motion in depth.

This hypothesis follows from recent theoretical and experimental results that the ratio of retinal motion and smooth pursuit determines relative depth (Stroyan & Nawrot, 2012). Although the current set of results can only represent the case of target fixation, more accurate predictions of the resulting motion in depth could be derived from combining the intuition in Figure 6 with this ratio, and these predictions tested by explicitly varying fixation condition and stimulus speed.

Conclusions

We have introduced a new induced motion stimulus in which a target dot can be seen as moving in depth despite that actual motion in depth is absent in both the target dot or the inducer frame. Rather, the combination of slant in depth and frontoparallel motion in the inducer is sufficient to trigger an illusory percept of motion in depth in the target dot. When slant is defined by disparity, the target dot appears to be moving away from the frame, in a way that may be described as induced motion in depth. When slant is defined instead by perspective and that slant in depth of the inducer is actually perceived, then the dot's motion is closer to the slant of the frame, in a way that may be called motion assimilation. One way to look at this discrepancy is to consider that perspective is more ambiguous than disparity in terms of the relative position of the target and frame. Our results then suggest that less ambiguity is conducive to motion induction in depth whereas greater ambiguity results in motion assimilation.

Keywords: induced motion, motion in depth, reference frames, depth cues, disparity, perspective, monocular depth cue, surface slant

Acknowledgments

The paper is supported in part by CELEST, an NSF Science of Learning Center (SBE-0354378 and OMA-0835976), Office of Naval Research (ONR N00014-11-1-0535), and AFOSR FA9550-12-1-0436, and by the SyNAPSE program of DARPA (HR0011-09-03-0001).

Commercial relationships: none.

Corresponding author: Arash Yazdanbakhsh.

Email: yazdan@bu.edu.

Address: Center for Computational Neuroscience and Neural Technology, Boston University, Boston, MA, USA.

References

- Cohen, R. L. (1964). *Problems in motion perception*. Uppsala: Appelbergs Boktryckeri.
- Di Vita, J. C., & Rock, I. (1997). A belongingness principle of motion perception. *Journal of Experimental Psychology: Human Perception and Performance*, *23*, 1343–1352.
- Duncker, K. (1938). Induced motion. In W. D. Ellis (Ed.), *A sourcebook of Gestalt psychology* (pp. 161–172). London: Routledge and Paul Kegan, Original work published 1929).
- Farnè, M. (1972). Studies on induced motion in the third dimension. *Perception*, *1*, 351–357.
- Farnè, M. (1977). Motion in depth induced by brightness changes in the background. *Perception*, *6*, 295–297.
- Gogel, W. C. (1965). Equidistance tendency and its consequences. *Psychological Bulletin*, *64*, 153–163.
- Gogel, W. C., & Griffin, B. W. (1982). Spatial induction of illusory motion. *Perception*, *11*, 187–199.
- Gogel, W. C., & MacCracken, P.J. (1979). Depth adjacency and induced motion. *Perceptual and Motor Skills*, *48*, 343–350.
- Gogel, W. C., & Tietz, J.D. (1976). Adjacency and attention as determiners of perceived motion. *Vision Research*, *16*, 839–845.
- Harris, J. M., & German, K. J. (2008). Comparing motion induction in lateral motion and motion in depth. *Vision Research*, *48*, 695–702.
- He, Z. J., & Nakayama, K. (1994). Apparent motion determined by surface layout not by disparity of three-dimensional surface. *Nature*, *367*, 173–176.
- He, Z. J., & Ooi, T.L. (1999). Perceptual organization of apparent motion in the Ternus display. *Perception*, *1999*, *28*(7), 877–892.
- He, Z. J., & Ooi, T.L. (2000). Perceiving binocular depth with reference to a common surface. *Perception*, *29*, 1313–1334.
- Hinton, G. E. (1981). *A parallel computation that assigns canonical object-based frames of reference*. Proceedings of the Seventh International Joint Conference on Artificial Intelligence, 2, Vancouver, BC.
- Indow, T. (2004). *The global structure of visual space*. Singapore: World Scientific Publishing Co.
- Koenderink, J. J. (1994). Vector analysis. In G. Jansson, S. S. Bergström & W. Epstein (Eds.), *Perceiving events and objects* (pp. 337–346). Hillsdale, NJ: Lawrence Erlbaum Associates.

- Lappin, J. S., & Craft, W.D. (2000). Foundations of spatial vision: From retinal images to perceived shapes. *Psychological Review*, *107*, 6–38.
- Léveillé, J., & Yazdanbakhsh, A. (2010). Speed, more than depth, determines the strength of induced motion. *Journal of Vision*, *10*(6):10, 1–9, <http://www.journalofvision.org/content/10/6/10>, doi:10.1167/10.6.10. [PubMed] [Article]
- Marr, D., & Nishihara, H.K. (1978). Representation and recognition of the spatial organization of three-dimensional shapes. *Proceedings of the Royal Society London B Bio*, *200*, 269–294.
- Murakami, I. (1999). Motion-transparent inducers have different effects on induced motion and motion capture. *Vision Research*, *39*, 1671–1681.
- Nefs, H. T., & Harris, J.M. (2008). Induced motion in depth and the effects of vergence eye movements. *Journal of Vision*, *8*(3):8, 1–16, <http://www.journalofvision.org/content/8/3/8>, doi:10.1167/8.3.8. [PubMed] [Article]
- Rock, I. (1990). The frame of reference. In I. Rock (Ed.), *The legacy of Solomon Asch: Essays in cognition and social psychology* (pp. 243–270). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Smith, M. A. (1996). Ptolemy's theory of visual perception: An English translation of the "optics" with introduction and commentary. *Transactions of the American Philosophical Society, New Series*, *86*(2).
- Stroyan, K., & Nawrot, M. (2012). Visual depth from motion parallax and eye pursuit. *Journal of Mathematical Biology*, *64*, 1157–1188.
- van Ee, R., van Dam, L.C.J., & Erkelens, C.J. (2002). Bi-stability in perceived slant when binocular disparity and monocular perspective specify different slants. *Journal of Vision*, *2*(9):2, 597–607, <http://www.journalofvision.org/content/2/9/2>, doi:10.1167/2.9.2. [PubMed] [Article]
- Wade, N. J., & Swanston, M.T. (1987). The representation of nonuniform motion: Induced movement. *Perception*, *16*, 555–571.
- Wallach, H., Bacon, J., & Schulman, P. (1978). Adaptation in motion perception: Alteration of induced motion. *Perception & Psychophysics*, *24*, 509–514.
- Watamaniuk, S. N. J., & McKee, S.P. (1995). Seeing motion behind occluders. *Nature*, *377*, 729–730.