Interactions between cues to visual motion in depth

Ian P. Howard

Yoshitaka Fujii

Robert S. Allison

Information about the motion in depth of an object along the midline of a stationary observer is provided by changes in image size (looming), changes in vergence produced by changes in binocular disparity of the images of the object, and changes in relative disparity between the moving object and a stationary object. Each of these cues was independently varied in the dichoptiscope, which is described in Howard, Fukuda, and Allison (2013). The stimuli were a small central dot and a textured surface moving to and fro in depth along the midline. Observers tracked the motion with the unseen hand. Image looming was normal or absent. The change in vergence was absent, normal, more than normal, or reversed relative to normal. Changing relative disparity between the moving stimulus and a stationary surface was present or absent. Changing vergence alone produced no motion in depth for the textured surface but it produced some motion of the dot. Looming alone produced strong motion in depth for the texture but not for the dot. When the direction of motion indicated by looming was opposite that indicated by changing relative disparity observers could use either cue. The cues dissociated rather than combined.

Introduction

We investigated the perception of motion in depth produced by three cues acting in various combinations. The three cues were:

1. Looming: The image of an object changes in size as the object approaches or recedes relative to an observer. We refer to changing image size as looming. Image size provides no information about the distance of a stationary unfamiliar object. However, changes in image size induce a compelling impression of motion in depth. Only a few objects, such as balloons, dilating pupils, or puffer fish, produce images that change in size when the objects are not moving in depth. In moving from distance \( D_1 \) to distance \( D_2 \) the angular subtense of the image changes from \( h_1 \) to \( h_2 \) according to the function

\[
\frac{\tan \theta_2}{\tan \theta_1} = \frac{D_1}{D_2}.
\]

2. Changing absolute disparity: When the eyes are converged on a stationary point, the disparity between the images of an object moving in depth changes. For an interocular distance, \( a \), and with the eyes converged symmetrically on a point at infinity, the angular disparity, \( \theta \), between the images of an object at distance \( D \) is given by \( \theta = 2\arctan(a/2D) \). Changes in absolute disparity provide the main signal for changing vergence. When the convergence of the eyes changes so as to keep the images of the moving object fused, the change in the angle of vergence is defined by this same function.

3. Changing relative disparity: The relative disparity between two objects at different distances equals the difference between the angles of convergence when the eyes fixate one and then the other object.

A paper by Erkelens and Collewijn (1985a) triggered a general interest in the perception of motion in depth. They varied the overall horizontal disparity of a large random-dot stereogram leaving the relative disparity within the stereogram constant. The fused image did not appear to move in depth even though changes in overall disparity must have been detected to drive changes in vergence required to keep the images fused. The display appeared to move in depth as soon as
changing relative disparity was introduced by adding a stationary stimulus. Regan, Erkelens, and Collewijn (1986) also showed that changes in vergence do not produce a sensation of motion in depth in a textured surface that does not loom.

The conclusion from both these studies was that neither changes in absolute disparity nor changes in vergence contribute to sensations of motion in depth. But a large textured stimulus that does not loom provides strong information that it is not moving in depth. Howard (2008) obtained some motion in depth from changing vergence when the stimulus was a radial pattern viewed through an aperture. The image of such a pattern does not provide any kind of looming signal because it does not change in size when the pattern moves in depth relative to the stationary aperture. Some motion in depth was also produced by a small dot, which does not provide a strong impression of looming. González, Allison, Ono, and Vinnikov (2010) reported similar results. One aim of our experiment was to investigate how the perception of real motion in depth is affected when the change in vergence is decreased, increased, or reversed with respect to normal.

Brenner, van den Berg, and van Damme (1996) investigated the effects of looming, vergence, and relative disparity on the perception of motion in depth. They presented a 1-s display of a computer-simulated polyhedral object moving at 21.6 cm/s towards the subject. The object moved from either 60 cm to 38 cm or from 82 cm to 60 cm with reference to a stationary textured background. Subjects adjusted the velocity of a laterally moving comparison object to match the approach velocity of the test object. In one condition, looming, vergence, and relative disparity all changed. In other conditions, only one or only two cues changed. Keeping any one cue constant, especially looming, reduced the perceived velocity of approach below the value when all cues changed. When only one cue changed, with the others held constant, looming produced the highest perceived velocity, changing only relative disparity produced less perceived velocity, and changing only vergence produced no motion in depth. Subjects differed in the weights assigned to the different cues. Subjects also set the comparison object to the perceived final position of the moving target. The final position of the moving stimulus was underestimated relative to the full-cue condition when one cue was kept constant and more so when two cues were kept constant.

One would expect perceived velocity to be proportional to perceived amplitude of motion in a given time interval. For example, if an object that moved 20 cm in a given interval of time is perceived to have moved only 10 cm, one would expect its perceived velocity to be half its actual velocity. But Brenner et al. (1996) found that perceived velocity was not closely related to perceived magnitude of motion. There are several possible reasons for this dissociation between velocity and motion amplitude. Judging amplitude of motion in terms of the judged final position of the moving object takes no account of the perceived starting position. On each trial the object moved only once in one direction at a velocity of 21.6 cm/s. It is well known that people misjudge the location of a moving object relative to a stationary object, especially when the object moves at a high velocity (Linares, Holcombe, & White, 2009). The movement required a change of vergence of 3.5°. At this velocity and with this vergence amplitude, vergence would certainly lag the change in stimulus disparity and the moving object would become diplopic at the end of its motion (Erkelens & Collewijn, 1985b).

In our experiment we overcame the problem of matching the locations of a moving object and a stationary object by having subjects track, with an unseen hand, a stimulus moving to and fro in depth. The to-and-fro motion allowed the observer to keep the motion of the hand synchronized with the target motion. This procedure provides concurrent estimates of perceived velocity and perceived magnitude of motion. We reduced the lag of vergence with respect to changing disparity by keeping the speed of motion in depth at the low value of 2 cm/s.

Brenner et al. (1996) did not examine the effects of having two cues indicate opposite directions of motion. This question was investigated by Regan and Beverley (1979). Subjects fixated on a stationary spot in a small square displayed in a stereoscope. The size of the square was sinusoidally modulated by ±2.5 arcmin about a mean of 1°. The disparity of the square relative to the fixation spot was modulated sinusoidally by up to 10 arcmin. When the cues indicated opposite motions in depth they tended to cancel. Thus a sensation of motion in depth produced by changing size was cancelled by an appropriate opposed change in disparity. Changing disparity became more effective relative to looming as the velocity of motion in depth increased or as exposure time increased. The relative effectiveness of the two cues varied widely between subjects. Regan and Beverley proposed that looming and changing disparity combine according to a simple weighted-sum model, even when they indicate opposite direction of motion.

Heuer (1987) reported that the cues of looming and changing vergence produced approximately additive effects on the perceived amplitude of motion in depth when the cues signaled motion in the same direction. However, when the cues indicated opposite motions in depth they rivaled rather than combined, with now one and then the other cue dominating. Heuer suggested that Regan and Beverley (1979) failed to observe this cue rivalry because they used very small modulations of
size and of disparity. However, the two experiments involved different cues to motion in depth. In Regan and Beverley’s experiment, both absolute and relative disparity changed but there was no changing vergence. In Heuer’s experiment, vergence changed but there was no change in absolute or relative disparity. One aim of the present experiment is to investigate the effects of conflicting cues using real motion in depth with all combinations of looming, relative disparity, and vergence.

In all previous experiments on motion in depth, stimuli were presented on flat screens in a stereoscope. Motion in depth was simulated by modulating disparity and/or looming. With simulated motion in depth, accommodation, perspective, disparity between different parts of the object, and parallax do not change as they do when a real object moves in depth. In our experiment we used a new type of instrument, which allowed us to independently vary cues to the motion in depth of a real object actually moving in depth. Since the object is really moving in depth, all features of the object change appropriately with changes in depth. The new instrument, which we call the dichoptiscope, is described in some detail in Howard, Fukuda, and Allison (2013).

The aims of the present experiments are (a) to use this new instrument to investigate how the perception of real motion in depth is affected when vergence demand is decreased, increased, or reversed with respect to normal, and (b) to investigate how looming, relative disparity, and vergence cues to motion in depth are integrated under conditions of cue conflict, relying on the capabilities of the dichoptiscope to vary high-fidelity depth cues independently.

### Methods

The stimulus that moved in depth, which we will call the moving stimulus, was either the frontal textured surface shown in Figure 1 or a small central spot of light. The textured surface consisted of rows of texture elements printed on an acetate sheet and mounted on an electroluminescent panel. The surface was 20 cm and 30 cm high. The luminance of the white texture elements was approximately 19 cd/cm² and that of the dark background was approximately 0.16 cd/cm². The spot of light was produced by an LED and conveyed to the center of the textured surface by a fiber-optic cable 2 mm in diameter.

One textured surface or one light spot was seen by one eye and an identical surface or spot was seen by the other eye. These stimuli, which we call monocular stimuli were mounted on tracks on either side of two semisilvered mirrors, as shown in Figure 2. The monocular stimuli were combined by the mirrors to form the impression of a single stimulus in front of the observer, which we call the dichoptic stimulus. The center of the textured surface and the spot were in the
midline at eye level. A motor moved the monocular stimuli to and fro in opposite directions along the tracks at a velocity of 2 cm/s. This caused the fused image of the textured surface or the point of light to move to and fro along the observer’s midline between a distance of 60 cm and 40 cm from the eyes.

The stimulus that did not move in depth, which we call the fixed stimulus, was a surface with rows of letters, as shown in Figure 1. This surface was also printed on an acetate sheet and mounted on an electroluminescent panel at a distance of 50 cm in front of the observer. It was seen by both eyes through the semisilvered mirrors. As the moving stimulus moved between distances of 60 cm and 40 cm it appeared to pass through the fixed stimulus. The rows of texture elements on the moving stimulus alternated with the rows of letters on the fixed stimulus so that the elements of two surfaces did not overlap.

We used four values of changing absolute disparity. When the eyes remain converged on the moving stimulus the changes in disparity correspond to four values of changing vergence. In keeping with convention, we use the term vergence demand to refer to the target disparity and vergence response to refer to the disjunctive eye movement made by the observer when attempting to fixate the target. We manipulated vergence demand of the stimulus but assume that the eyes approximately tracked the slowly moving stimulus. As the stimulus moved from 60 cm to 40 cm the changes in the angle of vergence demand were: normal (+3.1°), no vergence, increased with respect to normal (+5.6°), and reversed with respect to normal (−3.2°). For an interocular distance of 6.5 cm, the angles of vergence demand at three distances are shown in Table 1.

The value of changing vergence was determined by rotating each track about a vertical axis 50 cm from the eye. The change in vergence demand was normal when the racks were aligned with the apex of the two mirrors. Vergence demand remained at a constant angle of 7.4° when each track was aligned with an eye. Vergence demand was reversed when each track was set at an angle of −7° (in the temporal direction) and was greater than normal when each track was set at +3° (in the nasal direction) as shown in Figure 3.

The fixed stimulus was always at a distance of 50 cm. With the eyes converged on the fixed stimulus, the changes in relative disparity between the moving and

<table>
<thead>
<tr>
<th>Distance of monocular stimuli</th>
<th>Normal</th>
<th>Fixed</th>
<th>Increased</th>
<th>Reversed</th>
</tr>
</thead>
<tbody>
<tr>
<td>40 cm</td>
<td>9.3°</td>
<td>7.4°</td>
<td>10.8°</td>
<td>5.6°</td>
</tr>
<tr>
<td>50 cm</td>
<td>7.4°</td>
<td>7.4°</td>
<td>7.4°</td>
<td>7.4°</td>
</tr>
<tr>
<td>60 cm</td>
<td>6.2°</td>
<td>7.4°</td>
<td>5.2°</td>
<td>8.4°</td>
</tr>
<tr>
<td>Total change</td>
<td>3.1°</td>
<td>None</td>
<td>5.6°</td>
<td>−3.2°</td>
</tr>
</tbody>
</table>

Table 1. Target vergence angles of the moving stimulus at three distances of the monocular stimuli. The change in vergence demand was normal, fixed, increased, or reversed. The bottom row shows the total change in vergence for each condition.
fixed stimuli depended on the angular settings of the tracks.

For each changing vergence, the fixed stimulus was either present at a distance of 50 cm (relative disparity present) or absent (relative disparity absent). In these eight conditions, the image of the moving stimulus changed in size (looming present). As the stimulus moved from 60 cm to 40 cm the angle subtended by the width of the textured surface increased from 18.4° to 27.4°, a 50% increase. In two further conditions, vergence demand was normal but the image did not change in size (looming absent). The width of the surface subtended a constant angle of 22°. Relative disparity was present in one of these conditions and absent in the other. These 10 conditions are set out in Table 2. For each condition, the moving stimulus was either the textured surface or the central spot.

In the two conditions in which looming was absent each monocular object was fixed at 50 cm from the eye.

![Figure 3](image.jpg)

Figure 3. The green and blue solid lines indicate the reflected images of the tracks along which the monocular stimuli move in depth. Each track can be rotated about an axis 50 cm from the eye. The dashed green and blue lines are the visual axes. The angle at which the tracks are set determines the way in which vergence changes as the stimulus moves in depth. The solid black lines indicate the positions of the stimulus at 40 cm and 60 cm from the eyes. The 20-cm movement of the stimulus determined the change in the subtended angle of the stimulus (looming). Each dashed black line indicates the distance of the stimulus specified by the angle of vergence demand, for an interocular distance of 6.5 cm.

<table>
<thead>
<tr>
<th>Looming</th>
<th>Vergence demand</th>
<th>Relative disparity</th>
<th>Condition, reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>No</td>
<td>Normal</td>
<td>No</td>
<td>Vergence only, alone</td>
</tr>
<tr>
<td>Yes</td>
<td>Normal</td>
<td>Yes</td>
<td>Vergence only, relative</td>
</tr>
<tr>
<td>None</td>
<td>No</td>
<td>Normal, alone</td>
<td>Normal, relative</td>
</tr>
<tr>
<td>Reversed</td>
<td>No</td>
<td>Yes</td>
<td>Looming only, relative</td>
</tr>
<tr>
<td>Increased</td>
<td>No</td>
<td>Yes</td>
<td>Vergence reversed, alone</td>
</tr>
<tr>
<td></td>
<td>Yes</td>
<td>Yes</td>
<td>Vergence reversed, relative</td>
</tr>
</tbody>
</table>

Table 2. Ten combinations of cues to motion in depth used in the experiment.
The change in the disparity of the fused images was produced by moving the monocular stimuli in opposite directions along tracks at right angles to the lines of sight. This movement produced the same 3.1° change in the angle of vergence that was produced when the monocular stimuli moved in depth between 40 cm and 60 cm, with the normal change in vergence. The change in relative disparity was also the same as when the monocular stimuli moved in depth.

In all conditions, each monocular stimulus made 10 complete back and forth movements in depth between 40 cm and 60 cm. The observer grasped a rod and moved it along a track that extended in the midline under the moving stimulus. The hand and rod were hidden from view. The observer moved the rod 10 times back and forth in depth at the same perceived velocity and through the same perceived amplitude as the moving stimulus. These movements were recorded by an encoder and stored in a computer. The first two cycles of movement were discarded. The mean amplitude of the final 16 hand movements was taken as a comparative measure of the perceived amplitude of motion of the stimulus under the various experimental conditions.

Results

Overall analysis

Repeated-measures analyses with the subject treated as a random effect in a mixed-effects model (linear mixed-effects model using package nlme in R, http://www.r-project.org/) were run on the motion-in-depth estimation data. Independent variables were stimulus (texture vs. dot), reference (fixed stimulus pattern for relative disparity and looming signals present or not: alone vs. relative), and condition. Condition referred to the vergence-looming combination (looming only, normal, vergence only, vergence reversed, or vergence increased). For some analyses the condition was instead decomposed into separate independent variables for specified vergence (nominally 0, 20, 20, –20, and 40 cm for the five conditions, respectively) and the specified looming (20, 20, 0, 20, and 20 cm for the five conditions, respectively). As the conditions did not represent balanced or factorial combinations of looming and disparity, analysis was performed on subsets of conditions to investigate the experimental hypotheses. One outlier data point was identified in the reversed-vergence dot-relative condition and removed. In this case, the subject reported large motion in depth in the direction of the looming cue (compared to the typical percept in the direction of target vergence) that was more than two standard deviations from the mean of the rest of the data.

Analysis of the full dataset (except the effects of attention on the reference stimulus in the vergence-reversed condition, see below) demonstrated a significant interaction between stimulus and condition. $F(4, 95) = 19.58, p < 0.001$, as well as main effects of reference, stimulus, and condition, $F(1, 95) = 15.53, p < 0.001$; $F(1, 95) = 27.64, p < 0.001$; and $F(4, 95) = 28.10, p < 0.001$, respectively. There was also a marginal interaction between reference and condition, $F(4, 95) = 2.25, p < 0.07$. We will address this possible interaction below when looking at the effects of disparity and looming on perceived motion in depth. In the conditions with reference, there was no ambiguity about the motion and subjects reported that the moving dichoptic stimulus appeared to move while the fixed surface was stationary.

Due to the qualitative differences in various conditions it is difficult to interpret these interactions and it is more informative to look at planned contrasts in subsets of the data.

Effects of vergence and looming with different stimuli

To compare the effects of looming and vergence we looked at the subset of conditions where motion in depth was specified by vergence alone (vergence only), looming alone (looming only), or both congruent (normal). The reported motion in depth for the vergence-only, looming-only conditions is shown in Figure 4. For the looming-only case, strong motion in depth was reported with the textured stimulus but little motion in depth was perceived for the dot stimulus. Conversely, for the vergence-only conditions robust motion in depth was produced for the dot-alone stimulus but not for the texture alone. No subject reported seeing motion in depth from vergence only in the texture-alone condition. Addition of a reference surface enhanced depth from target vergence and more depth was perceived in the dot-relative and texture-relative than in the dot-alone and texture-alone conditions, respectively. Inspection of the figure shows that the strongest motion-in-depth reports were obtained for consistent cues (normal condition); depth estimates tended to be larger with congruent cues than when only a single cue specified motion in depth.

Linear mixed-effects model analysis was used to analyze the relationship between reported and specified motion in depth. There were significant interactions between condition and stimulus type, $F(2, 55) = 29.55, p < 0.001$, and reference and condition, $F(2, 55) = 6.44, p = 0.003$, as well as main effects of reference, stimulus, and condition.
The effect of stimulus type (dot vs. texture) differed for the looming and vergence cues as reflected in the significant interaction between stimulus and condition. For motion in depth specified by both looming and vergence (the normal condition) the estimated motion in depth was similar for both the texture and the dot. For the isolated stimuli in the normal conditions the mean estimate was near the specified 20 cm and somewhat, but not significantly, overestimated for the conditions with a stationary reference surface ($M = 26.23$ and $27.94$, not significantly different from 20.0, $p = 0.34$ and $p = 0.12$, respectively).

For the looming-only conditions, the dot produced little motion in depth while motion in depth near the value specified by looming was obtained with the textured stimulus. A reference surface providing relative texture cues produced no increase in settings for the texture stimulus and nonsignificant increases for the dot stimulus (mean difference 4.48, $p = 0.69$, adjusted for multiple comparisons).

In contrast, the vergence-only condition resulted in little depth except in the presence of a reference surface. The dot stimulus produced more motion in depth than the textured stimulus. In the absence of a reference, no subjects saw depth from disparity alone with the texture and depth with the dot was relatively small.

With the reference surface, motion in depth was larger for both stimulus conditions, particularly for the textured stimulus and estimates for both stimuli were approximately the specified 20 cm. The increase in motion in depth due to addition of the reference was particularly significant in the case of the texture stimulus (mean increase 16.28, $p < 0.001$) since no motion in depth was apparent without the reference. For the dot, motion in depth was increased by the reference surface as well but by a lesser amount (mean increase 9.03, $p = 0.04$).

In summary, motion in depth from looming only was similar to the specified value only when a textured stimulus was used, and only from vergence only in the presence of a reference surface. The normal condition containing both cues produced robust motion in depth for all stimuli. Movie 1 provides a demonstration of these phenomena with a simplified stimulus.

### Looming-vergence conflict

Figure 5 shows how apparent motion in depth varied as a function of vergence-demand-specified motion in depth when looming specified a motion in depth of 20 cm. This set of stimuli, with constant looming-specified motion in depth and varied vergence-specified motion...
in depth, allow us to assess the integration of the two cues with various types of cue conflict.

For the normal (congruent) case when both vergence-specified and looming-specified motion in depth was 20 cm, the average reported motion-in-depth was near the specified value for all conditions. Average motion-in-depth settings were slightly larger for the conditions with a reference (texture relative, dot relative) than for isolated stimuli (texture alone, dot alone).

Interestingly, very similar data to the normal condition were obtained when the vergence-specified motion in depth was doubled to 40 cm (vergence increased). This suggests that the disparity might have been simply ignored in favor of the looming cue when the cues were discrepant. However, data at negative and zero vergence-specified depth suggest an influence of the vergence demand of the stimulus under cue conflict.

When vergence-specified motion in depth was negative and thus opposite the looming signal (vergence reversed) the apparent depth depended on the stimulus. For the dot-alone and dot-relative stimuli that provided weak looming cues, motion in depth was consistently seen in the direction of vergence demand (negative values in Figure 5). Conversely for texture stimuli that provided robust looming signals average
apparent motion in depth was consistently in the direction specified by looming. Thus the conflicting vergence demand seemed to have little influence for the texture stimuli although average settings were smaller than in the normal condition.

Similarly, when vergence specified no motion in depth (looming only), the apparent motion in depth depended on the stimulus. For the texture-relative and texture-alone stimuli, apparent motion in depth was in the direction specified by the looming signal, while for the dot-relative and dot-alone stimuli the reported motion in depth was small and not significantly different from the zero value specified by vergence demand. Thus, as in the reversed condition, the motion in depth for the texture stimuli appeared to be mainly determined by the looming signal while motion in depth of the dot stimuli was mainly determined by vergence demand.

To analyze the effects of the degree of vergence-looming conflict we ran a repeated-measures linear mixed effects model with stimulus, reference, and conditions of increasing levels of vergence-looming conflict (vergence reversed, looming-only, normal, and vergence-increased conditions, respectively) as dependent variables and subject as a random effect. In this series looming always specified a 20-cm motion and the vergence specified motion varied from approximately −20 (in the opposite direction to looming) to +40 cm.

There was a significant interaction between stimulus and condition, \( F(3, 75) = 14.22, p < 0.001 \), and significant main effects of all three dependent variables, reference: \( F(1, 75) = 4.87, p < 0.030 \); stimulus: \( F(1, 75) = 46.41, p < 0.001 \); condition: \( F(1, 75) = 32.10, p < 0.001 \). No other interactions were significant (all \( p \)s > 0.40). Reported motion in depth was larger with increased specified depth from vergence, textured compared to dot stimuli, and with a reference surface compared to in isolation.

Profile analysis demonstrated that the interaction between stimulus and condition (and the effect of condition) was reflected in the linear component. Therefore to analyze the interaction we ran another repeated-measures model using the specified vergence angle for each condition (from −20 cm to +40 cm) as a dependent variable instead of the condition. This analysis indicated that the slope for the vergence term when the stimulus was a dot was 0.55 cm/cm, which was significantly different than zero (\( p < 0.001 \)) and predicted a 33 cm change in vergence over the 60 cm range in specified vergence. For the texture stimulus, the sensitivity of the motion in depth estimates to

Figure 5. Effects of vergence setting with looming held constant. Each curve represents the mean motion in depth setting (\( N = 6 \)) for each visual stimulus type by reference combination. The looming was always constant, specifying 20 cm motion in depth, with vergence specifying one of −20, 0, 20, or 40 cm of depth (vergence-reversed, looming-only, normal, and vergence-increased conditions, respectively). Negative values indicate responses in the opposite direction (180° phase) from the looming-specified motion in depth.
changes in vergence was weaker and nonsignificant (slope of 0.11, \( p = 0.07 \)) and only predicted a 6.71-cm change in motion in depth due to vergence over the entire range between the reversed and increased vergence conditions.

Post-hoc analysis confirmed that the effects of vergence with looming held constant (specifying 20 cm motion in depth) were more significant for the dot than the texture surface. There were no significant differences between any of the texture settings as a function of disparity when looming specified 20 cm motion in depth. For the dot there were significant differences between the normal and reversed, normal and looming only, increased and reversed, and the increased and looming only conditions. Movie 2 shows examples of these effects.

**Ambiguous response**

With a reference and the textured stimulus, perceived depth in the vergence-reversed conditions was ambiguous (Figure 6). In all cases the subjects fixated the moving stimulus. When subjects were instructed to attend to the motion of the moving dichoptic stimulus (moving texture pattern), while trying to ignore the fixed stimulus, then motion in depth was in the direction specified by looming for all participants and on average near the amplitude predicted from perspective (\( M = \)
19.74 cm, $SE = 4.28$ cm). However, when the subjects were instead asked to pay attention to the relative motion between the moving and fixed stimuli then the depth from changing relative disparity was apparent and five of six participants reported that the dichoptic stimulus surface moved in the direction specified by vergence, opposite to its looming-specified direction ($M = -19.77$ cm, $SE = 5.14$ cm). Note that the former data is shown in Figure 5. The difference between the two responses had a mean difference of 40.76 cm which is close to the 40-cm difference between the motion in depth specified by the two cues and significantly different than zero, paired $t(4) = 5.02$, $p = 0.007$. One subject found that judging the motion in depth while attending to the reference surface was not possible and so could not provide an estimate. Her data were excluded from this particular comparison and in Figure 6.

**Discussion**

**Absolute and relative looming and vergence signals**

We found that change in target vergence (vergence-only condition) gave robust motion in depth in the direction specified by the changing vergence when the target was an isolated dot but no observer perceived any motion in depth when the stimulus was an isolated, large, textured surface. Regan et al. (1986) reported this finding of motion in depth from target vergence change alone in isolated dots but not large textured surfaces. While initially attributed to lateral inhibitory interactions, recent papers have suggested that the difference in percept between the two conditions reflects the effect of the absence of the looming cue, which suppresses the response to the textured stimulus but only minimally suppresses the response to the dot (González et al., 2010; Howard, 2008).

The sensitivity to cue conflict for motion in depth from changing target vergence in an isolated stimulus is in line with general insensitivity of the stereoscopic system to absolute disparity or absolute vergence signals (Wallach & Lindauer, 1962). In contrast, the system is exquisitely sensitive to relative disparity signals (Westheimer, 1979). In the present experiments this sensitivity was apparent in the robust motion in depth responses to changing vergence signals for both the dot and the dichoptic texture in the presence of the fixed stimulus, which acted as a disparity reference. Thus vergence change can produce robust percepts of motion in depth even in the face of cue conflict with
looming, if changing relative disparity information is available.

A complementary set of results was obtained with the looming-only conditions. With isolated stimuli, robust percepts of motion in depth were obtained for the dichoptic texture but not for the dot. In both cases unchanging vergence indicated a stationary target and thus presented a cue conflict. The pattern of results was similar with the fixed surface present. This surface provides a reference for both relative looming and relative disparity. The fact that the looming of the dot produced slightly more motion in depth than the isolated dot might be due to having a reference for the looming signal, but the enhancement was small. As target vergence was unchanging, the reference for relative disparity would be expected to increase the cue conflict with the stereoscopic cues. However, the pattern of results did not depend on whether the reference was present or absent, which suggests that cue conflict was not the reason for the difference in the percepts for the dot and dichoptic surface stimuli in the looming-only conditions.

A more likely explanation is based on the effectiveness of the looming cue in the two stimuli. As the stimulus moved from 60 cm to 40 cm its retinal image expanded 50%. Such a change is readily apparent in the textured stimulus but corresponds to a few minutes of arc increase in the diameter of the spot. In the limit of an infinitesimal point source, the change in size would not be at all informative about motion in depth. In the present experiment with a finite spot, while the change in spot size was not negligible and looming of the spot did produce some weak motion in depth, the small stimulus size would have made looming a much weaker cue to motion in depth than it was with the texture. Similar conclusions have been reached for time-to-contact judgments from binocular stimuli. Heuer (1993) had subjects estimate time to contact with a computer-generated circle that loomed and/or changed disparity. When the cues were in conflict, the changing-size cue was dominant for objects larger than about 1°, but changing disparity (and/or vergence) was the dominant cue for objects smaller than 0.5°. Similarly Gray and Regan (1998) found that subjects could not make time-to-contact estimates with a looming small target. Looming is not an effective cue for small objects.

While the stimuli for motion in depth in these experiments was change in target vergence and change in optical distance, these may not be the effective binocular and monocular stimuli used by the visual system. The monocular and dichoptic signals to motion in depth can be further decomposed and there are several possible informative signals. For the change in target vergence, motion in depth could be signaled by the change in vergence (if the stimulus was tracked), change in disparity over time, and/or interocular velocity differences (Regan et al., 1995). Both interocular velocity differences and changing disparity seem to be sufficient to produce percepts of motion in depth (Allison, Howard, & Howard, 1998; Cumming & Parker, 1994; Regan, 1993; Shioiri, Saisho, & Yaguchi, 2000) and have been the subject of considerable recent attention (see Allison & Howard, 2011; Harris, Nefs, & Grafton, 2008, for recent reviews). Similarly monocular information to motion in depth could be obtained from cues such as optic flow, change in size (or texture density), looming, or time to contact (Regan & Hamstra, 1993; Regan et al., 1995). Assessing the relative contribution of these monocular and dichoptic components requires stimulus patterns designed to isolate or accentuate the subsidiary cues. While this goes beyond the scope of the current study it would desirable to investigate these issues in future research.

### Combinations of looming and vergence signals

The looming-only and vergence-only conditions showed the effect of motion in depth signals when the vergence or looming cue, respectively, was unchanging, not absent, and thus the response represents a response to the combination of a changing and unchanging depth cue. In everyday life the visual system needs to interpret these cues to judge the motion of objects through space and there are situations where the cues provide ambiguous or conflicting signals to motion in depth. For instance, the looming cue is complicated by the approach of non-spherical or tumbling targets, which produce varying and/or anisotropic looming (Gray & Regan, 2000). Similarly both cues need to be calibrated for distance to provide metric depth or motion in depth. In cases such as these and many others, the cue discrepancy is in the magnitude of the motion in depth rather than in its direction. Finally, the three cues considered here—looming, change of target vergence, and change of relative disparity—differ in their specificity for motion in depth. Except in rare cases, isotropic looming in the natural world always indicates motion in depth of the looming object relative to the observer. In contrast, target vergence (and absolute retinal disparity) changes every time we change vergence and therefore is common even when we gaze about a stationary environment. Changing disparity of one object relative to another always unambiguously indicates motion in depth relative to the observer although which object moves is not unambiguously specified.

We explored the effect of varying the magnitude and direction of vergence-specified motion in depth while the looming-specified motion in depth was the same across the conditions. Increasing the vergence-specified depth to a value consistent in direction but twice as
large as the looming-specified depth resulted in robust percepts of motion in depth in the expected direction. However, the magnitude was similar in the normal and vergence-increased cases. Thus, we found no strong evidence for additivity of directionally consistent cues either in an isolated dichoptic stimulus or in the presence of a fixed reference. In fact when the motion in depth was specified by a single cue (looming only and vergence only) the mean motion in depth responses for the most effective stimuli were near to the predicted value from the changing cue and similar to the normal condition. With the textured stimulus the looming cue tended to dominate. This dominance probably reflects the fact that (a) the texture was an effective stimulus for looming, and (b) the looming was produced by physical motion of the target so that residual cues such as accommodation and internal disparity were consistent with the looming cue.

For the dot stimulus—where the texture cue was weak—we found an influence of disparity/vergence and the stimulus appeared to move in the direction specified by vergence for the reversed-vergence condition. In contrast, the looming signal seemed dominant for the textured dichoptic stimulus even when the vergence signal was opposite to the looming signal. In the reversed vergence conditions subjects perceived depth in the direction of looming and near the predicted value from that cue. This might be expected from robust cue integration considerations that discount large discrepancies in cues (Landy, Maloney, Johnston, & Young, 1995). We have previously found evidence for such behavior in subjects judging static and changing surface slant under conditions of strong conflict (Allison & Howard, 2000; see also Hillis, Watt, Landy, & Banks, 2004). Consistent with the idea of cue dissociation in the reversed-vergence condition, we found that in the presence of the fixed reference surface, subjects could either perceive the motion in depth due to looming or due to changing relative disparity.

Thus, in agreement with Heuer (1987), we found that the cues of looming and changing vergence dissociated when they signaled opposite motion in depth. This result confirms cue dissociation rather than simple weighted averaging (Regan & Beverley, 1979) can occur when using real motion in depth with all combinations of looming, relative disparity, and vergence. Normally the cues are concordant and redundant and thus there is no reason to combine them in most natural scenarios. The relative disparity and the looming cues are very different, which makes combination less likely and beneficial. The looming cue is a powerful and primitive monocular cue to a dangerous approaching object. Human infants as young as four to eight weeks exhibit defensive blink responses to looming stimuli (Nánez & Yonas, 1994) and infant monkeys make avoidance responses to such stimuli (Schiff, Caviness, & Gibson, 1962). The response to looming stimuli arises early in evolution and seems to be preserved across species: Flies respond to looming stimuli (Borst & Bahde, 1988); locusts (Gabbiani, Krapp, Koch & Laurent, 2002) and pigeons (Sun & Frost, 1998) appear to have specialized neurons to process looming stimuli, and response to looming stimuli is common if not ubiquitous in visual animals. In contrast stereopsis and processing of relative disparity requires binocular combination, evolved later, and is much less common, being largely restricted to animals with frontal eyes and ecological drivers for enhanced depth perception. Given the different evolutionary history and information provided from the vergence and looming system it thus is not surprising that they do not fuse, particularly when they do not agree.

**Estimating motion in depth by manual tracking**

In these experiments subjects estimated motion in depth by tracking the apparent motion of the target with their unseen hand. This is a natural task that is easy for subjects to understand and execute. There are range and precision limitations due to reaching distance and arm kinematics but for the modest speed and amplitude of motion used in this experiment subjects found the task easy and were reliable.

As far as we know haptic tracking has not been previously used as a measure of motion in depth. However, haptic measures of depth and distance have often been used. In the normal condition haptic matches were not significantly different from the predicted value. Since the full range of depth cues was available and consistent with the specified motion, this finding suggests that the subjects were responding accurately with the tracking task. Pointing estimates of distance have been found to be about half as variable as verbal estimates (Foley, 1977). Moving the unseen hand between positions should be more reliable than pointing since kinesthetic feedback about position is more direct and angular errors are not amplified by distance as in pointing. Brown, Knauf, and Rosebaum (1948) asked subjects to move an unseen slider to match the distance between two previously seen visual markers. For target intervals between 0.6 and 40 cm, mean error in matching movements made in a direction extending toward or away from the body, such as we used here, were less than 1.5 cm. Small intervals were slightly overestimated and longer intervals slightly underestimated. Variability in the responses increased with size of the movement with a standard deviation of the settings of 2–3 cm for the largest 40-cm movement. These data suggest haptic matching of intervals in depth is a reasonably accurate and precise technique.
Consistent cues

We considered conditions of either cue congruence or strong cue conflict: The looming specified a given motion in depth while vergence was either held at zero, reversed, or significantly increased, or looming was held at zero while disparity specified motion in depth. In all cases the blur/accommodation cues were consistent with the looming cues since the optical distance of the stimulus determined these cues. In such cue conflict we found that cue dominance often occurred with looming cues generally determining the percepts. Strong cue conflict (particularly with a cue controlled by setting it to zero) is a very common experimental paradigm (e.g., Brenner et al., 1996; Harris & Rushton, 2003; Khuu, Lee & Hayes, 2010; Nefs, O’Hare & Harris, 2010; Schiller, Slocum, Jao, & Weiner, 2011). However, cue integration is theoretically more likely to occur with weak cue conflict when cues to motion in depth are unbiased and generally consistent with each other (Landy et al., 1995). Such optimal cue integration has been reported for stereoscopic depth perception (e.g., Hillis et al., 2004); although Rosas, Wichmann, and Wagemans (2007) found a failure of reliability-based weighting for stereoscopic slant discrimination and Scarfe and Hibbard (2011) found that weighted cue combination of motion and disparity occurred even when the estimates were biased. A perturbation analysis investigation of motion-in-depth cue integration has not been performed to our knowledge. The dichoptiscope is ideally suited to investigations of these questions for motion in depth since the discrepancy between motion-in-depth cues can be precisely controlled and conflict with extraneous cues not under study eliminated.

When both change in target vergence and looming specified the same motion in depth (normal condition) all subjects perceived strong motion in depth for both the dot and dichoptic texture stimuli, regardless of whether the fixed reference stimulus was present or not. This is not surprising since all information indicated the same motion in depth. The dichoptiscope has significant advantages to other techniques when attempting to provide consistent or near consistent stimuli. Since it depends on real physical stimuli and motion it presents high-fidelity cues to the actual motion. In the normal condition all cues were consistent with the motion in depth and the stimulus actually underwent the specified motion. There is no conflict between vergence and accommodation; furthermore, motion parallax and changing perspective are normal. Such a situation could also be obtained by physically moving a target through space on the desired profile. However, simple physical motion does not allow for dissociation of visual cues while the dichoptiscope instrument allows the experimenter to control independently and precisely each of several cues to motion in depth. This provides the ability to assess the effects of perturbing the sensory cues while the cues as a whole are generally consistent. Under such conditions cue integration becomes more likely than cue dissociation or veto (Howard, 1997). We believe the dichoptiscope has great promise for studying the effect of perturbations of sensory information under conditions representing ecologically valid combinations of visual cues and favorable to sensory fusion.

Conclusions

The dichoptiscope enabled the study of cue combination under cue conflict, while providing high-fidelity independent control of relevant motion-in-depth signals. Haptic tracking was found to be an accurate and precise measure of apparent motion in depth in this instrument.

Results for conditions when the cues opposed suggested cue dissociation rather than simple weighted averaging. Normally the cues are concordant and redundant and there is no reason to combine them. Dominance of the looming cue (when present) over changing absolute and relative disparity is likely due to the unique specificity of isotropic looming as a signal for motion in depth.

Keywords: motion in depth, vergence, stereopsis, depth perception

Acknowledgments

This research was supported by Discovery grants from the National Sciences and Engineering Research Council of Canada and IPH and RSA. YF was supported by the NSERC CREATE training program in Vision Science and Applications.

Commercial relationships: none.
Corresponding author: Robert S. Allison.
Email: allison@cse.yorku.ca.
Address: Centre for Vision Research, York University, Toronto, Canada.

References


