Short Communication

Disambiguating Velocity Estimates Across Image Space

MERCEDES BARCHILON BEN-AV,* MAGGIE SHIFFRAR†‡

A translating homogeneous edge viewed through an aperture is an ambiguous stimulus, while a translating edge discontinuity is unambiguous. Under what conditions does the visual system use unambiguous velocity estimates to interpret ambiguous velocity estimates? We considered a translating rectangle visible through a set of stationary apertures. One aperture displayed a rectangle edge while the other apertures displayed corners. Observers reported the direction in which the edge appeared to translate. The results suggest that collinearity and terminator proximity determine whether the unambiguous corner velocity was used to interpret the ambiguous edge velocity. These results suggest some of the ways in which the visual system controls the integration of velocity estimates across image space.

Motion   Aperture problem   Grouping   Collinearity   Proximity

Local motion measurements are often ambiguous. For example, a moving edge viewed through a small aperture is physically ambiguous because the component of translation parallel to the edge's orientation cannot be measured. This so-called "aperture problem" has been widely investigated because all known visual systems contain spatially limited receptive fields (Wallach, 1935; Adelson & Movshon, 1982; Hildreth, 1984). How does the visual system determine object motion when initial velocity estimates may be ambiguous? Wallach (1935, 1976) first suggested that unambiguous velocity estimates obtained from edge discontinuities are used by the visual system to overcome ambiguous velocity estimates.§ Because such feature or discontinuity based approaches are vulnerable to noise, subsequent linear models emphasized the information available from combinations of ambiguous velocity estimates (Adelson & Movshon, 1982; Mingolla, Todd & Norman, 1992). However, other studies have re-emphasized the importance of unambiguous velocity estimates since observers have difficulty accurately combining ambiguous velocity estimates across spatially disconnected contours (Adelson & Movshon, 1983; Ramachandran, 1990) even when those contours define a single, rigid object (Shiffrar & Pavel,

1991; Lorenceau & Shiffrar, 1992). Unambiguous velocity estimates influence the perceived direction of a translating line when discontinuities are high contrast, shown at long duration, part of a relatively short line (Lorenceau, Shiffrar, Wells & Castet, 1993; Castet, Lorenceau, Shiffrar & Bonnet, 1993), and not a result of occlusion (Shimojo et al., 1989; Lorenceau & Shiffrar, 1992). Discontinuities also influence the perception of plaid patterns (Derrington, Badcock & Henning, 1993).

Do unambiguous velocity estimates from contour discontinuities influence the interpretation of ambiguous velocity estimates across disconnected spatial locations? When non overlapping, unambiguous and ambiguous motion signals are present in the same aperture, the unambiguous signals control the interpretation of the entire display (Rubin & Hochstein, 1993; Shiffrar, Li & Lorenceau, 1995). Over what spatial extent can the visual system use unambiguous velocity estimates to interpret ambiguous velocity estimates? Nakayama and Silverman (1988) addressed this question by examining when a dash influenced the interpretation of a translating curve. When the dash was positioned off of the curve, it had only a minor effect on the perceived curve motion. Since accurate motion interpretations require both integration and segmentation, there may have been insufficient support for the coherent integration of the non-overlapping dash and curve stimuli (Braddick, 1993; Stoner & Albright, 1993). The purpose of the current experiments was to examine how the visual system interprets object motion under those conditions most conducive to motion integration across space. Since motion integration across contours is thought to be enhanced when contours define a rigid object (Ullman, 1979) and appear behind apertures

*Department of Neurobiology, Brain Research, The Weizmann Institute of Science, Rehovot 76100, Israel.
†Department of Psychology and Center for Neuroscience, Rutgers University, 197 University Avenue, Newark, NJ 07102, U.S.A.
‡To whom all correspondence should be addressed.
§This paper limits itself to motion in the fronto-parallel plane. Thus, the terms 'ambiguous' and 'unambiguous' refer only to two-dimensional motion measurements. See Rubin, Solomon and Hochstein (1995) for discussion of the three-dimensional case.

(Shimojo, Silverman & Nakayama, 1989; Lorenceau & Shiffrar, 1992; Trueswell & Hayhoe, 1993) we examined a rigidly translating polygon visible behind a set of stationary apertures. This stimulus, illustrated in Fig. 1, contained a spatially separated combination of ambiguous edge velocities and unambiguous corner velocities. We investigated the conditions under which the visual system used the corner motion to disambiguate the edge motion.

EXPERIMENT 1: DEPTH AND PROXIMITY

Research regarding the influence of depth on motion integration across edges has focused primarily on the perception of ambiguously translating line segments. Do depth cues also facilitate the integration of unambiguous velocity estimates across space? To test this idea we examined how different depth cues altered the influence of unambiguously translating corners on the interpretation of an ambiguously translating edge. Disparity (Shimojo et al., 1989) and edge length variability (Lorenceau & Shiffrar, 1992; Kooi, 1993) were manipulated through aperture shape and the use of three-dimensional glasses. Since proximity is thought to facilitate perceptual grouping (Wertheimer, 1923; Koffka, 1935; Ben-Av & Sagi, 1995), manipulations of spatial proximity were used to measure the disambiguating power of translating contour discontinuities.

Methods

Subjects. Three subjects (one author and two naive) with normal or corrected-to-normal visual acuity participated in this experiment. These same observers served as subjects in all of the subsequent experiments.

Stimuli. The stimuli were displayed on a 19 in. RGB Hitachi monitor with a 60 Hz refresh rate. The monitor was controlled by a Silicon Graphics Personal Iris model 4D/TG 30. This apparatus was used in all experiments.

The display consisted of an obliquely oriented, rigidly translating rectangular outline viewed through five stationary apertures. The rectangle was oriented 45 deg clockwise or counterclockwise from the horizontal and translated vertically with a constant speed of 0.7 deg/sec. Four apertures were positioned so that each displayed a rectangle corner. The remaining aperture displayed one of the rectangle's straight edges. The center of this critical aperture was positioned over the midpoint of the rectangle's longest edge. A fixation point was placed at the center of the aperture display, as shown in Fig. 1. The distance between the fixation point and the center of the critical aperture was 0.5 degrees of visual angle (dva). The edge appeared randomly with equal probability at one of the four cardinal positions relative to the fixation point: above left, above right, below left or below right. The polygon translated vertically upward when the test edge was above the fixation point and translated vertically downward when the edge was below the fixation point. This manipulation fixed the direction of the visible edge's translation so that edge eccentricity always increased. Rectangle width remained constant and equal to 1.6 dva. There were seven possible rectangle lengths: 1.6, 2.0, 2.4, 2.8, 3.2, 3.6 and 4.0 dva. During each trial, aperture location was linked to rectangle size so that the corners and edges were always visible through and centered in the apertures. The rectangle outline was white with a luminance of $12.5 \ cd/m^2$. The background was black while the apertures were cyan, 0.8 dva in diameter and had a luminance of $0.96 \ cd/m^2$. This color and luminance difference between the apertures and rectangle created monocular depth cues in the form of T-junctions.

Aperture shape and depth varied according to a within subjects design. Apertures were circular or square. The square apertures were oriented so that the visible edge's length remained constant. Visible edge length varied smoothly when viewed through a circular aperture. The second variable was relative aperture depth. Either the apertures were presented at the same zero disparity depth plane as the rectangle and were viewed monocularly or they were presented at 11 min arc uncrossed disparity and
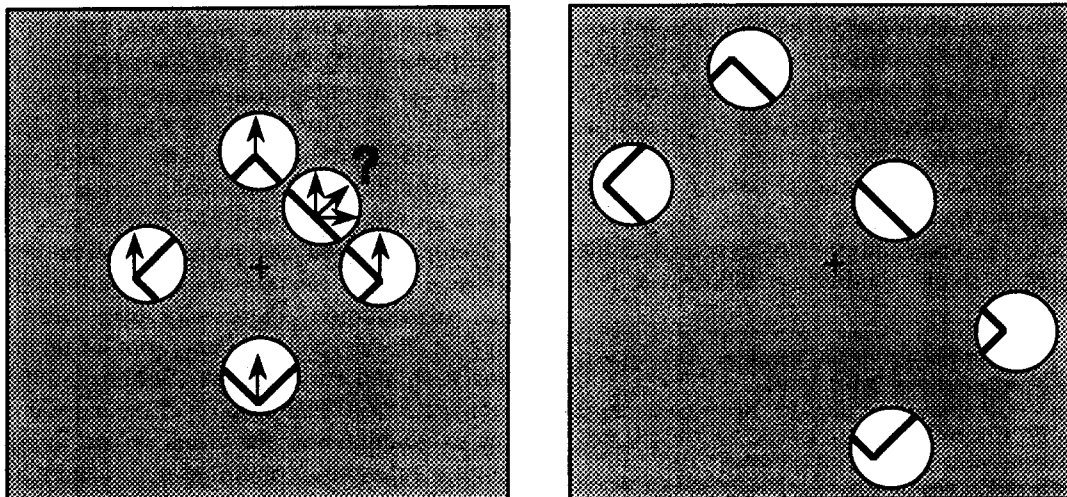

FIGURE 1. Stimuli used in Expt 1. Rectangular outline figures translated behind circular or square (not shown) apertures so that four corners and one edge were visible. The distance between the edge and corners was varied.

were viewed binocularly with three-dimensional glasses. Thus there were four conditions: square apertures with zero disparity; square apertures with binocular disparity; circular apertures with zero disparity; circular apertures with binocular disparity. Each aperture shape/disparity condition was run as a separate block of trials.

Procedure. The experiments were performed in a dark environment. Viewing distance was fixed at 95 cm from the monitor with a chin rest. Subjects wore three-dimensional glasses (CrystalEyes Eyewear by Stereo Graphics Corp.) in the disparity conditions. Subjects were informed that sometimes the edge would appear to translate in the same direction as the corners while other times it would appear to translate independently. Subjects were asked to report the visible edge's direction of translation while maintaining fixation. At the beginning of each trial a fixation point appeared at the center of the screen followed immediately by the translating rectangle. After 0.6 sec the stimulus disappeared and an arrow appeared. One end of the arrow was fixed to the fixation point and the other end could be moved with a mouse device. Observers rotated the arrow to select the perceived direction of edge translation and recorded their response with a button press. Each subject completed five blocks of 56 trials per condition. Block order was randomized across subjects. Each subject completed a few practice trials before beginning the experiment. No feedback was provided. These general procedures and stimuli were used in all subsequent experiments.

Results and discussion

Performance was identical across figure orientation (45 deg clockwise or counterclockwise from horizontal) and visible edge location (i.e. above, below, right, or left of fixation). Therefore, the results of this and the later experiments are collapsed across figure orientation and reported as if the figure was always oriented 45 deg clockwise from horizontal and the visible edge was always located above and to the right of fixation (as shown in Fig. 1). We report rightward translation along the horizontal as 0 deg and upward translation along the vertical as 90 deg. When the test edge was interpreted in isolation, it would appear to translate 45 deg from horizontal. The corners always appeared to translate vertically along the 90 deg axis. Therefore, whenever the edge was interpreted with the corners, it would appear to translate 90 deg from the horizontal. Intermediate directions of translation fall between 45 and 90 deg. The measures of spatial separation refer to the distance between the vertices of the nearest corners and the center of the test edge.

Because performance did not significantly differ across subjects, the results are shown in Fig. 2 collapsed across subjects. Across all depth conditions, the perceived direction of edge translation depended strongly on the proximity of the corners to the edge ($P < 0.01$). As the distance between the corners and edge increased, the perceived direction of the edge changed from vertical (90 deg) to orthogonal to its own orientation (45 deg). A separation distance of 1.4 deg between the center of

Experiment 1
Perceived Direction of Edge Translation
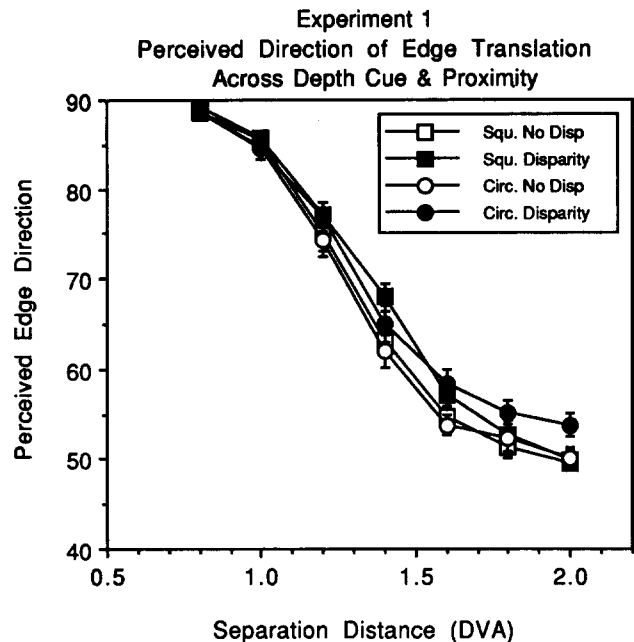**Across Depth Cue & Proximity**

FIGURE 2. Results of Expt 1. The data are collapsed across observers and the error bars represent SEs. The solid symbols show data from the disparity conditions with square (Squ.) or circular (Circ.) apertures. The open symbols show data from the monocular depth conditions. The perceived direction of edge translation changed with the distance separating the edge and corners.

the test edge and nearest corner marked the point of transition between the independent and the corner-dominated interpretation of the edge. These results suggest that the visual system can use unambiguous motion signals to overcome the aperture problem, but only within a limited spatial range. Indeed, proximity (Koffka, 1935) appears to be a more powerful constraint on motion interpretation than object rigidity. While the visual system is thought to select motion interpretations consistent with rigid objects (Ullman, 1979), subjects in this experiment often interpreted the rectangle non-rigidly even though a rigid interpretation was always available.

While contour length variability (Lorenceau & Shiffrar, 1992) and binocular disparity (Shimojo *et al.*, 1989) facilitate the integration of ambiguous motion signals across contours, no such facilitation appears for unambiguously translating corners. Although the disparity magnitude used in this experiment falls within the range used by previous researchers, our results do not rule out the possibility that greater disparities are needed to enhance the assimilation of ambiguous and unambiguous velocity estimates across space. Nonetheless, these results do suggest that facilitation of depth cues on motion integration across space may not extend to images containing unambiguous velocity estimates.

Another purpose of this study was to determine the spatial extent of the disambiguating effect of contour discontinuities under optimal conditions. Previous research using disconnected curves and dashes as stimuli suggested that proximity had little influence on the interpretation of ambiguously translating curves

(Nakayama & Silverman, 1988). Other "motion cooperativity" studies have suggested spatial extents of 50 (Nawrot & Sekuler, 1990) and 15 min arc (Chang & Julesz, 1984) within which velocity estimates are integrated. These variations in the spatial estimates of velocity integration suggest that distance alone may not determine motion integration. Some models use a single distance parameter to define what velocity estimates should be integrated (Grzywacz & Yuille, 1991). However, if figural or collinearity cues suggest that velocity estimates define the same object then the distance over which those velocities are integrated may shift (Braddick, 1993). The following experiment assesses the relative strength of two potential object cues.

EXPERIMENT 2: COLLINEARITY AND CLOSURE

Closure facilitates object detection (Kovacs & Julesz, 1993) and thus may also facilitate velocity integration. The influence of closure on motion integration was tested by removing two apertures, and hence two visible corners, from the translating rectangle stimulus. Is collinearity of the edge with the corners important? Previous work suggests that there is increased cooperative processing between aligned line segments or corners (Peterhans & von der Heydt, 1989; Grossberg & Mingolla, 1993). Thus, the corners may have a greater influence on perceived edge translation when the corners and edge are collinear.

Methods

Observers viewed translating rectangles through three circular apertures positioned so that one aperture displayed an edge and two apertures displayed corners, as shown in Fig. 3(A). There were two conditions: either the corners were collinear with the edge or they were noncollinear and belonged to the opposite rectangle side. The same corner–edge separation distances from Expt 1 were used. The white rectangle outline was positioned at 11 min arc uncrossed disparity while the circular apertures were at zero disparity. Subjects wore three-dimensional glasses and completed five blocks of 56 trials per condition. Again, subjects indicated the perceived direction of edge translation.

Results and discussion

The results, shown in Fig. 4, suggest that collinearity is a critical factor in the integration of unambiguous and ambiguous velocity estimates. When the corners were not collinear with the ambiguously translating edge, the corner velocities were not used to interpret the edge. However, when the corners were collinear with the edge, the corners influenced perceived edge translation for spatial separations < 1.4 dva. It is still possible that non-collinear corners influence perceived edge translation for spatial separations < 0.5 dva. However, the point that we wish to emphasize is that proximity alone does not determine velocity integration. Rather, the relative organization of features also influences velocity integration. Thus, our results are not consistent with models

of motion perception in which all velocity estimates falling within some fixed spatial region are integrated. These finding also support the enhancement of cooperative processing between aligned contour discontinuities (Wertheimer, 1923; Peterhans & von der Heydt, 1989; Grossberg & Mingolla, 1993). Finally, this collinearity effect might explain why previous studies found minimal proximity differences in the motion integration of non-collinear, non-overlapping curve and dash stimuli (Nakayama & Silverman, 1988).

The second factor that this experiment addressed was closure. Our results suggest that two visible corners are sufficient to disambiguate the edge motion when the corners are near and collinear to the edge. The spatial extent of corner influence for the two collinear corners did not significantly differ from the results obtained in Expt 1 for the same condition with four visible corners ($P > 0.10$). Thus, figural organization at the level of complete objects may have little effect on motion integration.

EXPERIMENT 3: GAP SIZE

In the previous experiments, the distance between the corner vertices and the center of the visible edge varied systematically with the distance between corner and edge terminators (gap size). To determine which of these distances is key, we conducted a final experiment in which gap size varied for a fixed rectangle size. If the distance between the corner vertex and the edge is the critical distance, then observers should always perceive the same edge translation across variations in gap size. However, if terminator separation is the critical distance in motion integration across contours, then perceived edge translation should vary with the distance between the edge and corner terminators.

Methods

A translating rectangle at 11 min arc uncrossed disparity was visible through three circular apertures at zero disparity, as shown in Fig. 3(B). Unlike previous experiments only one rectangle length (3.6 dva), or vertex–edge separation (1.8 dva), was used. Previously, observers did not perceive the edge to translate with the corners at this separation. We manipulated gap size by varying aperture diameter. There were five possible aperture diameters: 1.40, 1.25, 1.10, 0.95 and 0.80 dva; which resulted in five possible gap sizes: 0.40, 0.55, 0.70, 0.85 and 1.0 dva respectively. Each subject wore three-dimensional glasses and completed five blocks of 40 trials in which they indicated the perceived direction of edge translation.

Results and discussion

The results, shown in Fig. 5, indicate that perceived edge translation depended significantly on the distance separating the corner and edge terminators ($P < 0.05$), even though the distance between the corner vertex and the visible edge remained constant. The unambiguously
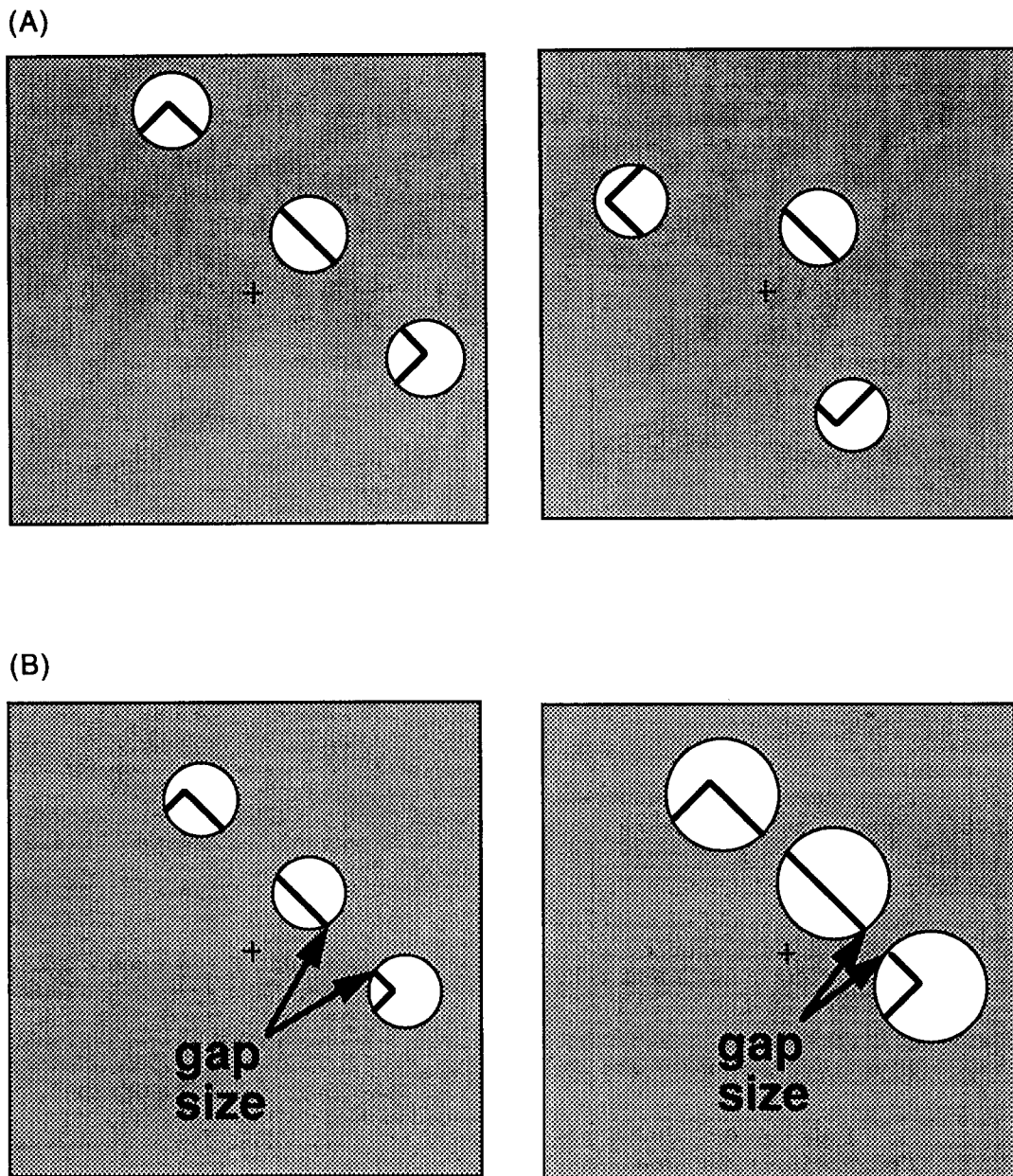
(A)

(B)

FIGURE 3. (A) Stimuli for Expt 2. The translating rectangle was viewed through three apertures. Two corners and one edge were visible. The corners were either collinear or non-collinear with the test edge. The distance between the corners' vertices and the center of the test edge was varied. (B) Stimuli for Expt 3. A single fixed size rectangle was viewed through three apertures. The size of the apertures varied to manipulate the distance between the terminators of the edge and corners (gap size).

translating corners influenced perceived edge translation when the gap size was <0.7 dva. It appears that the distance between the nearest points of any two contours may influence whether their velocity estimates are integrated. This finding suggests that motion is first integrated along contours (Hildreth, 1984) and only after this local integration occurs can integration across contours take place.

GENERAL DISCUSSION

The accurate interpretation of object motion requires the integration of velocity estimates from the same object as well as the segmentation of velocity estimates from different objects. Feature based models of motion analysis

suggest that velocity estimates from contour discontinuities dominate the interpretation of object motion (Shiffrar & Pavel, 1991; Lorenceau & Shiffrar, 1992; Rubin & Hochstein, 1993). An unrestricted reliance on such velocity estimates would result in image segmentation errors. The purpose of the current experiments was to identify how such feature based models could be adapted to accurately segment moving images.

Our studies suggest that unambiguously translating corners can influence the perceived direction of ambiguously translating edges within limited spatial separations. Separation distances appear to be defined by the minimum distance between two contours. Collinearity, but not object closure, also significantly influences the spatial range within which the visual system relies on

Experiment 2
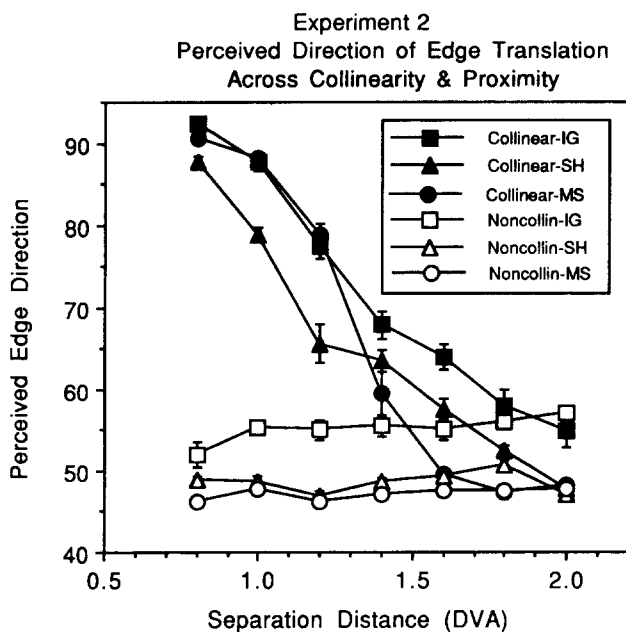### Perceived Direction of Edge Translation Across Collinearity & Proximity

FIGURE 4. Results of Expt 2. The data for each subject are shown. The solid symbols show the results of the collinear condition and the open symbols show the results of the non-collinear condition. The edge only appeared to translate with the corners when the corners were collinear to and near the visible edge.

unambiguous velocity estimates. Thus, distance alone does not determine motion integration. Lateral connections working within a limited spatial range may be involved in the interpretation of edge motion (Peterhans & von der Heydt, 1989; Castet *et al.*, 1993; Grossberg & Mingolla, 1993; Lorenceau *et al.*, 1993).

Another way to consider these results is that within a limited spatial range, the corner velocity "captures" the edge velocity. Motion capture refers to a biased motion analysis in which the perceived direction of one image

Experiment 3
Perceived Direction of Edge Translation Across Terminator Separation
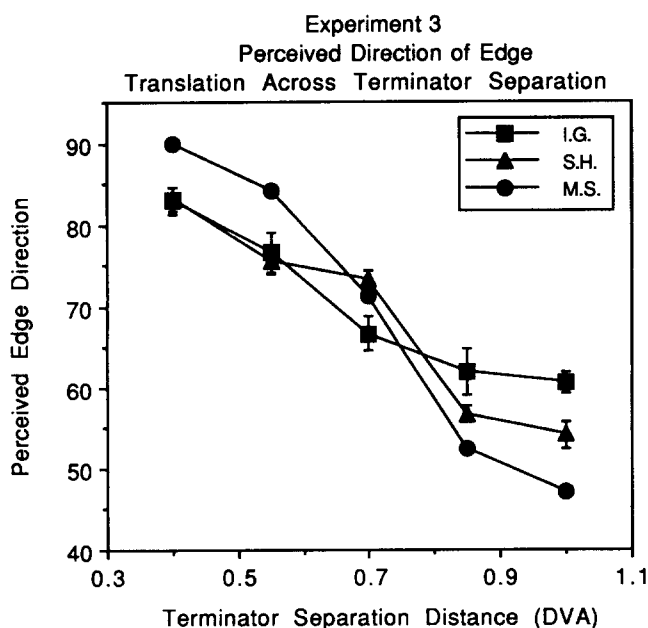

FIGURE 5. Results of Expt 3. The results for three observers are shown individually. The perceived direction on edge translation was directly related to the separation distance between the terminators.

feature is controlled or captured by the velocity of another feature (Ramachandran, 1985; Ramachandran & Cavanagh, 1987; Yo & Wilson, 1992). The analogy of a leopard running through a forest is often used to describe this phenomenon. Motion capture is thought to enable the visual system to more efficiently group the leopard's spots with the outline of the leopard's body (Ramachandran & Cavanagh, 1987). The limitations on motion capture described in this paper suggest how the visual system might avoid capturing the leaves on the forest's trees with the leopard's body.

REFERENCES

Adelson, E. H. & Movshon, J. A. (1982). Phenomenal coherence of moving visual patterns. *Nature, 300*, 523–525.

Adelson, E. H. & Movshon, J. A. (1983). The perception of coherent motion in two-dimensional patterns. *Proceedings of the ACM Siggraph Interdisciplinary Workshop on Motion: Representation and Perception*, Toronto (pp. 11–16).

Ben-Av, M. B. & Sagi, D. (1995). Perceptual grouping by similarity and proximity: experimental results can be predicted by intensity autocorrelations. *Vision Research, 35*, 853–866.

Braddick, O. (1993). Segmentation versus integration in visual motion processing. *Trends in Neuroscience, 16*, 263–268.

Castet, E., Lorenceau, J., Shiffrar, M. & Bonnet, C. (1993). Perceived speed of moving lines depends on orientation, length, speed and luminance. *Vision Research, 33*, 1921–1936.

Chang, J. & Julesz, B. (1984). Cooperative phenomena in apparent movement perception of random-dot cinematograms. *Vision Research, 24*, 1781–1788.

Derrington, A., Badcock, D. & Henning, G. (1993). Discriminating the direction of second-order motion at short stimulus durations. *Vision Research, 33*, 1785–1794.

Grossberg, S. & Mingolla, E. (1993). Neural dynamics of motion perception: Direction fields, apertures, and resonant groupings. *Perception & Psychophysics, 53*, 243–278.

Grzywacz, N. & Yuille, A. (1991). Theories for the visual perception of local velocity and coherent motion. In Landy, M. S. & Movshon, J. A. (Eds), *Computational models of visual processing* (pp. 231–252). Cambridge, Mass.: MIT Press.

Hildreth, E. C. (1984). *The measurement of visual motion*. Cambridge, Mass.: MIT Press.

Koffka, K. (1935). *Principles of Gestalt psychology*. New York: Harcourt Brace.

Kooi, F. (1993). Local direction of edge motion causes and abolishes the barberpole illusion. *Vision Research, 33*, 2347–2351.

Kovacs, I. & Julesz, B. (1993). A closed curve is much more than an incomplete one: Effect of closure on figure–ground segmentation. *Proceedings of the National Academy of Science, 90*, 7495–7497.

Lorenceau, J. & Shiffrar, M. (1992). The role of terminators in motion integration across contours. *Vision Research, 32*, 263–273.

Lorenceau, J., Shiffrar, M., Wells, N. & Castet, E. (1993). Different motion sensitive units are involved in recovering the direction of moving lines. *Vision Research, 33*, 1207–1217.

Mingolla, E., Todd, J. T. & Norman, J. F. (1992). The perception of globally coherent motion. *Vision Research, 32*, 1015–1031.

Nakayama, K. & Silverman, G. H. (1988). The aperture problem—II. Spatial integration of velocity information along contours. *Vision Research, 28*, 747–753.

Nawrot, M. & Sekuler, R. (1990). Assimilation and contrast in motion perception: Explorations in cooperativity. *Vision Research, 30*, 1439–1451.

Peterhans, E. & von der Heydt, R. (1989). Mechanisms of contour perception in monkey visual cortex. II. Contours bridging gaps. *Journal of Neuroscience, 9*, 1749–1763.

Ramachandran, V. S. (1985). Apparent motion of subjective surfaces. *Perception, 14*, 127–134.

Ramachandran, V. S. (1990). Visual perception in people and machines. In Blake, A. & Troscianko, T. (Eds), *AI and the eye* (pp. 21–77). New York: Wiley.

Ramachandran, V. & Cavanagh, P. (1987). Motion capture anisotropy. *Vision Research, 27,* 97–106.

Rubin, N. & Hochstein, S. (1993). Isolating the effect of one-dimensional motion signals on the perceived direction of two-dimensional objects. *Vision Research, 33,* 1385–1396.

Rubin, N., Solomon, S. & Hochstein, S. (1995). Restricted ability to recover 3D global motion from 1D local signals: Theoretical observations. *Vision Research, 35,* 569–578.

Shiffrar, M. & Pavel, M. (1991). Percepts of rigid motion within and across apertures. *Journal of Experimental Psychology: Human Perception and Performance, 17,* 749–761.

Shiffrar, M., Li, X. & Lorenceau, J. (1995). Motion integration across differing image features. *Vision Research, 35,* 2137–2146.

Shimojo, J., Silverman, G. H. & Nakayama, K. (1989). Occlusion and the solution to the aperture problem. *Vision Research, 29,* 619–626.

Stoner, G. & Albright, T. (1993). Image segmentation cues in motion processing: Implications for modularity in vision. *Journal of Cognitive Neuroscience, 5,* 129–149.

Trueswell, J. & Hayhoe, M. (1993). Surface segmentation mechanisms and motion perception. *Vision Research, 33,* 313–328.

Ullman, S. (1979). *The interpretation of visual motion.* Cambridge, Mass.: MIT Press.

Wallach, H. (1935). Uber visuell wahrgenommene Bewegungsrichtung. *Psychologische Forschung, 20,* 325–380.

Wallach, H. (1976). *On perception* (pp. 201–216). New York: The New York Times Book Co.

Wertheimer, M. (1923). Untersuchungen zur Lehre von der Gestalt. *Psychologische Forschung, 4,* 301–350. Available in abridged English translation (1958). Principles of perceptual organization. In Beardsler, D. C. & Wertheimer, M. (Eds), *Reading in perception* (Section 8). New York: Van Nostrand.

Yo, C. & Wilson, H. (1992). Moving two-dimensional patterns can capture the perceived directions of lower or higher spatial frequency gratings. *Vision Research, 32,* 1263–1269.