Pergamon

0042-6989(94)E0058-S

Vision Res. Vol. 34, No. 20, pp. 2701–2706, 1994
Copyright © 1994 Elsevier Science Ltd
Printed in Great Britain. All rights reserved
0042-6989/94 $7.00 + 0.00

Motion Capture of Luminance Stimuli by Equiluminous Color Gratings and by Attentive Tracking

JODY C. CULHAM,*† PATRICK CAVANAGH*

Received 2 September 1993; in revised form 14 March 1994

Two experiments demonstrated motion capture of luminance-defined dots by gratings with no net luminance-based motion. In a series of two-frame experimental trials, we superimposed bright dots and a color grating rotating in opposite directions. Capture was observed at equiluminance and was facilitated by the presence of color in gratings over a range of luminance contrasts. In a second experiment, observers noted that when a counterphase grating was tracked in either direction with attention, the superimposed dots were captured in that direction. These results suggest motion capture is supported not only by luminance-based motion, but also by color- and attention-based motion. Indeed, we suggest that the most parsimonious explanation is that all capture is mediated by attention.

Motion capture Color Luminance Equiluminance Attentive tracking

Motion capture occurs when features with no net motion of their own appear to move in synchrony with other salient moving features. In the first demonstration of capture, MacKay (1961) showed that as a wire loop was dragged across a the dynamic white noise of a detuned television screen, the noise elements appeared to move along with the loop. Ramachandran (1985) recognized the effect as purely perceptual, and not dependent on eye movements as MacKay had suggested. More recent investigations have shown that sinusoidal luminance gratings are also effective at capturing randomly-moving luminance dots (Ramachandran & Inada, 1985; Ramachandran & Cavanagh, 1987). Such capture is most effective with low spatial frequencies, with phase shifts of 1/4 cycle, and when grating shifts are synchronized with dot displacements (Ramachandran & Inada, 1985). In fact, low spatial frequency gratings can even generate capture strong enough to overcome coherent dot motion in the opposite direction (Ramachandran & Cavanagh, 1987). Capture is not limited to gratings but has also been generated by the motion of illusory contours such as those produced by Kanizsa squares (e.g. Ramachandran, 1987).

Motion capture has been observed with a variety of stimuli, but until now all successful demonstrations have used luminance contours as the capturing stimuli. However, motion capture may also provide clues about the contributions of other types of visual information to motion perception. Color-based motion and attention-based motion are potential sources of capture which we investigate here. Although Ramachandran (1987) was able to show that moving luminous dots easily captured an equiluminous color patch, he suggested that the converse, capture of luminance by color, did not occur. In his displays, illusory contours derived from luminance-defined corners of a Kanizsa square were able to capture a central color square, but those derived from equiluminous color corners were not. Capture by illusory contours is also suggestive of higher-level motion perception, though the loss at equiluminance implies dependence on luminance-based motion. Alternatively, the failure to obtain capture at equiluminance may have been due to the weakening of illusory contours at equiluminance (Ejima & Takahashi, 1988) rather than a lack of color-based or higher-level motion capture.

Our preliminary observations suggested that the disappearance of capture at equiluminance in Ramachandran's (1987) experiment is indeed specific to illusory contours. In these observations, 1000 bright dots alternated between two positions within an annulus to produce a rocking motion around the center when viewed alone. The dots were superimposed on a radial red–green grating. In synchrony with the dot displacements, the grating rotated by $\frac{1}{4}$ cycle phase steps. Across a range of luminance contrasts between the red and green of the grating, including equiluminance, we noted motion capture of the dots. Furthermore, this capture appeared facilitated when the grating was actively tracked with attention (Cavanagh, 1992).

---

*Department of Psychology, Harvard University, 33 Kirkland Street, Cambridge, MA 02138, U.S.A. [*Email* culham@isr.harvard.edu].
†To whom all correspondence should be addressed.

To separate the effects of capture by color and by attention suggested by the preliminary observations, two new experiments were conducted. In the first experiment, we used brief two-frame presentations to prevent observers from scrutinizing the display. Between the two frames, the dots and the grating rotated in opposite directions and we quantified the strength of capture using the motion coherence of the dots (Newsome & Paré, 1988) needed for observers to overcome the effect of capture and see dot motion veridically. In the second experiment, we examined the capture of oscillating dots during attentive tracking of ambiguous grating motion.

EXPERIMENT 1

Method

Stimuli were generated on a Macintosh IIcx computer with a Macintosh 8.24 display card and displayed on an AppleColor 13 in. High-Resolution RGB monitor. The 640 × 480 pixel video signal had 8 bits of intensity resolution for each of the red, green and blue signals. The image was refreshed at 66 Hz and the luminance output on the screen was calibrated for linearity.

One of the authors, JC, and two naive observers, SS and JI, all with normal color vision and corrected-to-normal acuity, binocularly viewed a series of experimental trials. Observers fixated on a central black-white bull's-eye (1.1 deg) surrounded by an annular display (inner radius 3.5 deg, outer-radius 8 deg) on a black background as illustrated in Fig. 1. For each trial, two 250 msec frames were presented with no interstimulus interval. In the first frame, an eight-cycle radial sinusoidal grating and 1000 superimposed bright yellow dots appeared within the annulus. The phase of the grating was randomly chosen to prevent observers from anticipating its position. In the second frame, the grating was displaced by a phase shift of 2.81 deg of rotation (22.5 deg of phase) in one direction, clockwise or counterclockwise, while the dots were displaced in the opposite direction by the same rotation (2.81 deg, equivalent to dot displacements of 10–24 arc min, depending on eccentricity).

We tested capture of the dots by gratings with different compositions of color and luminance in separate blocks of trials. The average luminance (45 cd/m^2) and chromaticity (yellow) were the same for all gratings and for the intertrial intervals. Chromatic contrast was generated by sinusoidally modulating red and green phosphors in opposite phase and was defined as the percentage of the maximum modulation between the red and green phosphors, 50% in the first experiment. Luminance contrast was generated by modulating the red and green phosphors in phase and was calculated as the Michelson contrast: the difference between maximum and minimum luminance divided by their sum and expressed as a percentage. For gratings with both chromatic and luminance contrast, luminance contrast was arbitrarily labeled as positive when red was more luminous than green and negative when green was more luminous than red. The CIE x,y coordinates of the phosphors were (0.641, 0.327) for the red and (0.284, 0.599) for the green, with a mean chromaticity of yellow (0.434, 0.483).

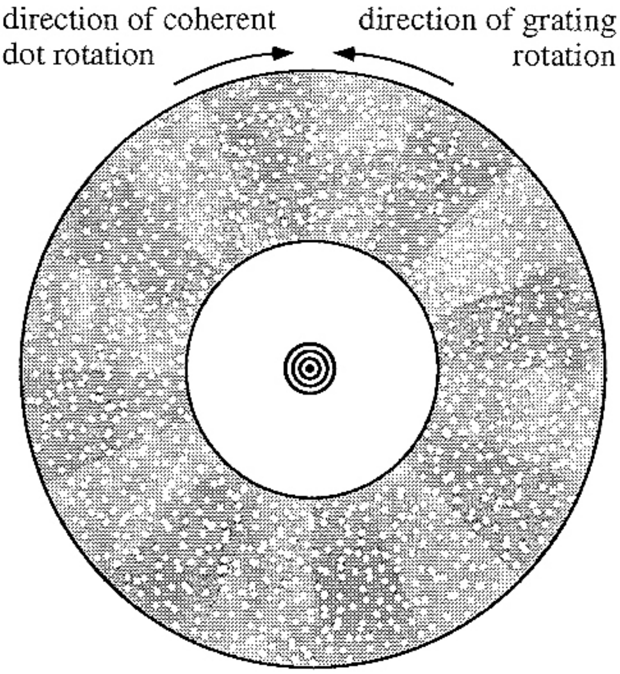

FIGURE 1. Schematic illustration of a sample trial in Expt 1. In the first frame of the trial, an array of bright dots appeared superimposed randomly upon a radial color grating of variable luminance contrast (illustrated as a square-wave grating, but actually sinusoidal). In the second frame, the grating was rotated by a small displacement in one direction (counterclockwise here) while a subset of the dots were rotated by the same amount in the opposite direction (clockwise here) and a remainder of the dots were randomly replaced. Subjects then indicated the apparent direction of the dots which could either appear to move veridically (clockwise) or to be captured by the grating (counterclockwise).

One thousand bright yellow dots were randomly placed within the annulus. Each dot was 4.2 × 4.2 arc min and the overall dot density was 3% of the area of the annulus. The dots had the same mean chromaticity as the grating but their luminance was doubled to 90 cd/m^2 (luminance contrast = 33%). The strength of the dot motion signal could be varied by altering the percentage of dots which were displaced coherently. Eight levels of dot signal were used, ranging in equal steps between complete noise (0% coherence) and the maximum signal (100% coherence). For example, with 50% coherence, half the dots would provide a motion signal and half would not.

Observers judged the apparent direction of the dot motion while fixating on the central bull's-eye. They were instructed to press one of two keys to indicate whether the dots appeared to move clockwise or counterclockwise and to ignore the motion of the grating. The naive observers were not told about the opposite motion of the dots and gratings.

Observers' accuracy in detecting the true motion of the dots was evaluated for combinations of chromatic and luminance contrast. Chromatic gratings were tested across a range of luminance contrasts centered at the nominal equiluminance point. Each observer's nominal equiluminance point for the chromatic gratings was determined by adjusting the red/green luminance balance until adjacent colors were minimally distinct and the perceived motion when viewed at 0.3 Hz appeared slowest. For comparison, accuracy was tested for achromatic gratings with similar luminance contrasts and for dots alone with no grating present. Observers performed at least 16 (SS) or 32 (JC, JI) trials at each of the eight dot coherence levels for each of 8 (SS), 11 (JI) or 14 (JC) conditions.

Results

Psychometric functions were plotted for each observer in each condition and used to interpolate measures of capture strength. Figure 2 shows one observer's functions relating the frequency of veridical dot motion judgments to the proportion of coherently moving dots. Data are shown for dots which were presented alone (open circles), superimposed on an equiluminous grating (solid circles), or superimposed on a grating with 10.5% luminance contrast (solid squares). With the dots alone, accuracy is at chance when the dots are completely incoherent and improves as the dot signal increases. When luminance or color gratings are present, accuracy again improves with increasing dot signal but is actually *below* chance when dot coherence is weak, indicating capture as the dots appeared to move with the grating rather than in their true direction.

Capture strength was quantified for each condition by interpolating the point at which each curve reached 50% accuracy. This threshold indicates the matching point at which the observer was equally likely to see the dots moving with the grating as to see them moving veridically. Thus, for the results in Fig. 2, an equiluminous grating generated capture until dot coherence reached 37%; whereas, a 10.5% luminance contrast grating generated capture until dot coherence reached 91%.

The capture thresholds for each observer are shown as a function of grating luminance contrast in Fig. 3. Upper curves (solid circles) show data for color gratings over a range of values of luminance contrast between the red and green bars which includes the nominal equiluminance point (luminance contrast = 0). Lower curves show data for a range of contrast values of the luminance grating, including zero contrast when the dots are presented alone. Note that data points are duplicated at negative values of contrast to allow comparison with color gratings. In these graphs, higher values indicate stronger capture; namely, a greater coherent motion signal strength was required to overcome capture. For both color and luminance gratings, the degree of capture increased as luminance contrast was added. In comparing the upper and lower curves, however, capture was almost always stronger when chromatic contrast was present, suggesting facilitation of capture by color across a range of luminance contrasts, not just at equiluminance.

Although observers varied in the overall strength of capture, all observers experienced capture at equiluminance and demonstrated facilitation from color across a narrow (JI), intermediate (JC), or broad (SS) range of luminance contrasts.

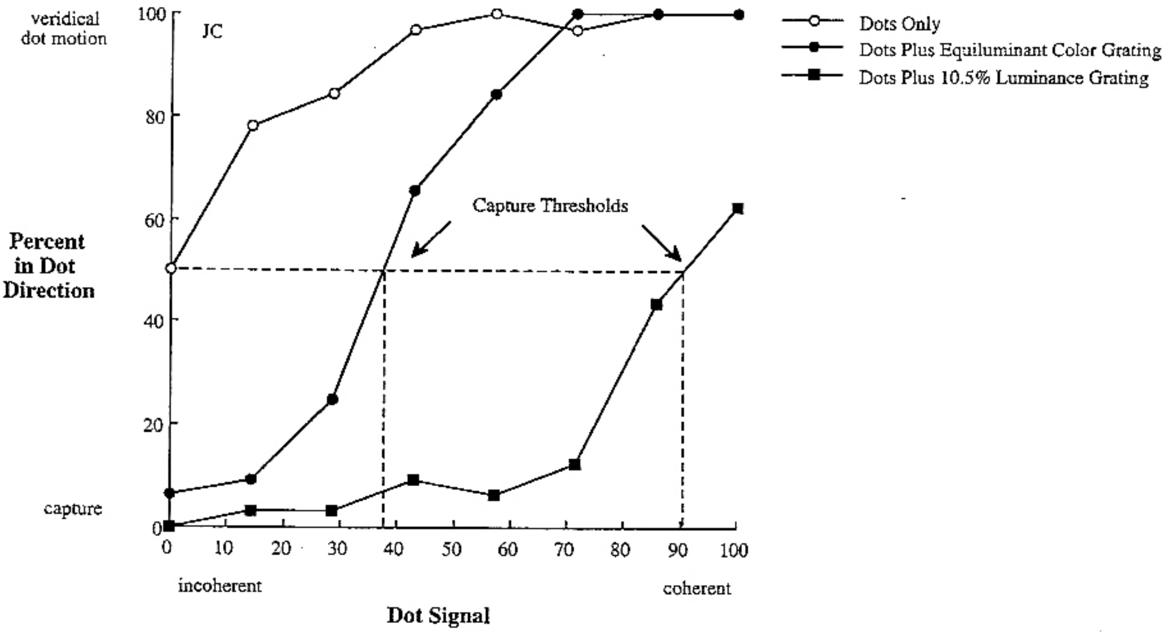

FIGURE 2. Percentage of trials for which veridical dot motion was perceived as a function of the dot signal (percent coherence) for one observer, JC. The 50% capture thresholds have been interpolated for the equiluminous grating (solid circles) and the luminance grating with 10.5% contrast (solid squares). Points below the threshold indicate dominance of motion capture; points above threshold indicate dominance of veridical dot motion perception. Performance is also shown for the case with dots alone (open circles).

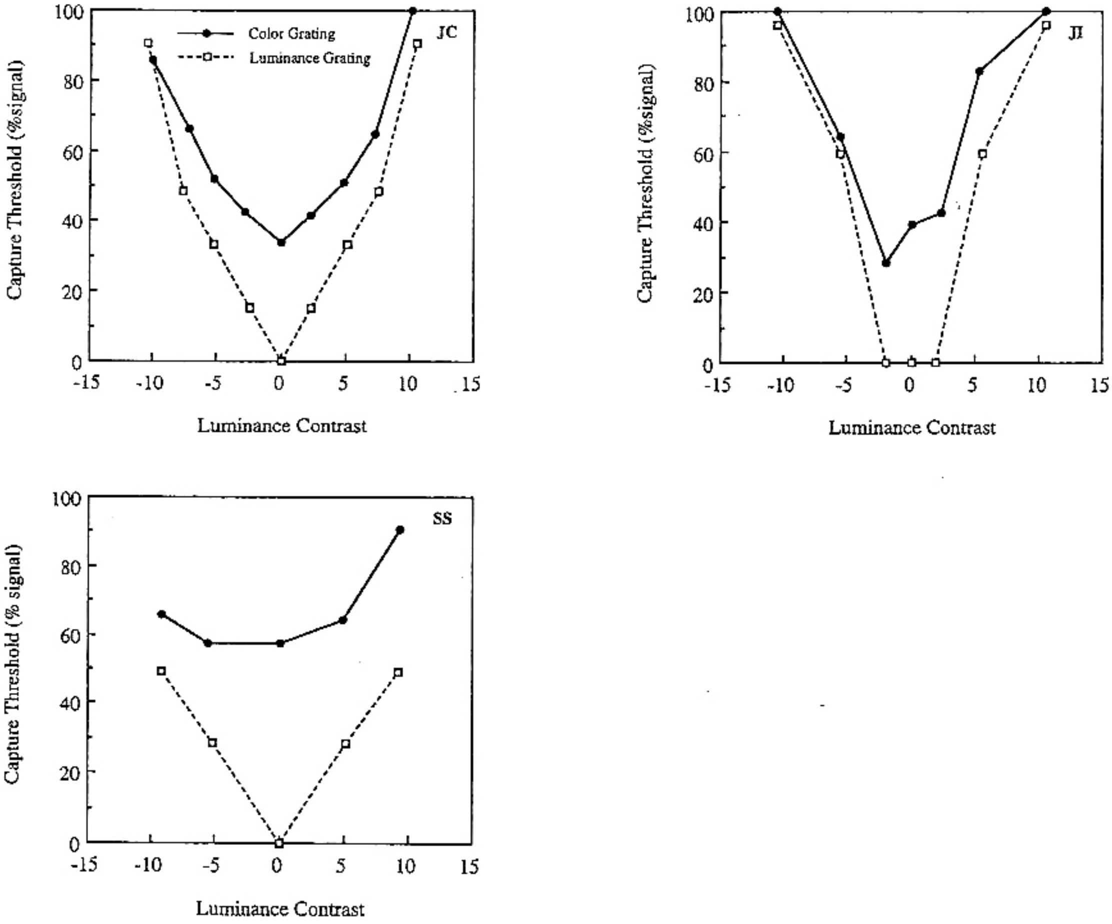

FIGURE 3. Capture thresholds as a function of luminance contrast for color and luminance gratings shown for each observer. Upper curves (solid circles) show data for color gratings over a range of luminance contrast levels including the nominal equiluminance point (0), positive contrasts where the red component was brighter than green, and negative contrasts where green was brighter than red. Lower curves (open squares) show data for a range of contrast levels for the luminance grating including zero contrast (dots only) and positive contrast values which have been duplicated at negative contrast values to allow comparison with color gratings.

EXPERIMENT 2

We noted in preliminary tests that capture was facilitated when the bars of the moving grating were actively tracked with attention (Cavanagh, 1992). In the second experiment, the capacity of attentive tracking to generate capture was examined by having observers track an ambiguously-moving stimulus.

Method

Radial equiluminous or achromatic sine-wave gratings, as in the first experiment, were now set in counterphase flicker (making discrete phase shifts of $\frac{1}{2}$ cycle three times per second for a temporal frequency of 1.5 Hz) to produce a directionally ambiguous signal. One-thousand bright dots were then added and these alternated between two positions in synchrony with grating jumps. Dot positions were separated by 2.81 deg of rotation (equivalent to 10–24 arc min linear displacement, depending on eccentricity).

Two observers, JC and SS, first viewed the equiluminous or achromatic gratings at high color contrast (50%, as defined in Expt 1) and high luminance contrast (20%), respectively. Observers were able to track the grating while maintaining fixation on the central bull's-eye by using attention to select a pair of bars on opposite sides of the grating and actively following their changing positions in one direction, clockwise or counterclockwise, rather than the other. While alternately tracking the bars in either direction they noted the apparent direction of dot motion. Next, they adjusted the luminance or color contrast to the minimum setting which still permitted attentive tracking of the grating. They then noted the perceived direction of dot motion at this threshold contrast level for attentive tracking. This adjustment was repeated eight times.

Results

Observers saw capture of the dots whenever the color or luminance contrast of the grating was sufficient for attentive tracking. Further, the perceived direction of capture always corresponded to the direction of tracking. At low grating contrast, only the dots superimposed on the tracked bars appeared captured; whereas, at higher levels of contrast, the entire annulus appeared to move in the direction of tracking.

DISCUSSION

These experiments extend the set of salient motion cues that can capture finer details to include those with no net luminance-based motion. Namely, both equiluminous color contours and attentive tracking can generate motion capture.

Color is able to generate capture, though the mechanisms of its contribution are debatable. We will consider three possibilities. First, color capture could arise from several sources of residual luminance contrast in *nominally* equiluminous gratings (see Cavanagh & Anstis, 1991 for a review): chromatic aberration, poor accommodation, and variability in equiluminance points among individual neurons and with retinal eccentricity. However, our use of very low spatial frequencies and a high-contrast achromatic fixation point reduces the effects of chromatic aberration and poor accommodation to negligible levels. Further, variability in equiluminance points across individual units did not seem to be a factor in our results. While variable equiluminance points can produce facilitation in a narrow range around equiluminance, two of our three observers showed facilitation across a fairly wide range of luminance contrasts (Cavanagh & Anstis, 1991).

A second possibility is that motion capture at equiluminance arises from a true low-level contribution of opponent-color signals to motion. Although some studies report little contribution of color contours to motion (Anstis, 1970; Ramachandran & Gregory, 1978; Livingstone & Hubel, 1987), other results suggest that color provides a weak input to motion (Cavanagh, Tyler & Favreau, 1984; Cavanagh & Favreau, 1985). In fact, motion signals may even be several times stronger for chromatic motion than achromatic motion with the same absolute cone contrast (Stromeyer, Eskew & Kronauer, 1990).

A third alternative is that both luminance-based and color-based motion capture may have arisen from higher-level attentive mechanisms. Indeed, motion capture was initially interpreted in terms of a cognitive long-range motion process (Braddick, 1980; Anstis, 1980). For example, Ramachandran and Cavanagh (1987) proposed that capture is a long-range function to preserve the identity of a moving object in spite of spurious motion signals from finer features. Further, Cleary and Braddick (1990) suggested motion capture may result from a different mechanism than the short-range process involved in perceiving kinematograms. In motion capture, low spatial frequencies most effectively mask high frequencies (Ramachandran & Cavanagh, 1987); whereas, in the short-range perception of kinematograms, Cleary and Braddick demonstrated that high spatial frequencies masked low frequencies. They suggested that this difference would not be contradictory if motion capture involves a different high-level or "long-range" process rather than the low-level or "short-range" process thought to be involved in perceiving kinematograms.

Given many problems with the short-range vs long-range dichotomy (Cavanagh & Mather, 1989), motion capture may be better understood in terms of an active attention-based motion process (Cavanagh, 1991, 1992). Active motion perception explains the interpretive properties attributed to long-range processes as effects of attentive tracking strategies. Moreover, attention-based motion can be demonstrated independently, and in an opposite direction, from low-level motion processes (Cavanagh, 1992). Certainly the results of our second experiment imply that attention can modulate the perception of capture. Additional evidence for capture as a high-level process includes its dependence on figure–ground relationships (Ramachandran & Anstis, 1986; Eagle & Rogers, 1991) and object completion (Meyer & Howard, 1993).

Attentive tracking shares some similarities with other attention-based motion phenomena. In particular, tracking is believed to be a limited capacity mechanism (e.g. Pylyshyn & Storm, 1988). We would therefore expect tracking-mediated effects such as motion capture to be hampered when attention is distracted and experiments to test this conjecture are underway. On the other hand, the aspects of attention which support tracking are quite different from those underlying recently reported attention-based phenomena such as the line motion effect of Hikosaka, Miyauchi and Shimojo (1993) or the illusory temporal order effects of Stelmach and Herdman (1991). Both of the other effects have been explained by speeded processing due to the *distribution* of attention drawn by sustained effort or transient events; in contrast, attention-based motion appears to arise from the *displacement* of attention itself.

Attentional displacements alone, however, appear to be insufficient for motion capture without an appropriate surface representation (Ramachandran & Anstis, 1986; Eagle & Rogers, 1991; Meyer & Howard, 1993). In Ramachandran's (1987) experiment, a stationary colored square was placed in the center of a larger illusory square formed by four "pac men". When the illusory square was moved, the inner square was captured and appeared to move alone with it. However, when the brightness of the pac men was adjusted to be equiluminous with the background, the central square failed to be captured. In this case, we claim that it is the illusory contours which failed at equiluminance, not motion capture. Without the illusory contours, the pac men are interpreted as four independent items. We duplicated Ramachandran's stimulus and found that even though these pac men could be easily tracked, the central square was not captured because it was not seen as part of a moving surface. In comparison, the dots in our study were captured by regions which were perceived as moving surfaces—the entire grating at high contrasts or individual bars at low contrasts.

Overall, the most parsimonious explanation for our results would be that all capture results from a higher-level process. Such a process would not necessarily rely on luminance contrast but may also be supported by the motion of other attributes such as color, or even by the movement of attention alone. Indeed, if capture is an active motion process for maintaining object identity in the face of conflicting motion signals, multiple visual attributes could provide more reliable input than luminance information alone. In Ramachandran's (1986) example, an "arboreal ancestor" seeing a leaping leopard extracts the motion of its luminance-based outline from low spatial frequencies and attributes that motion to the higher frequency spots which have no net signal for large displacements. However, as the leopard leaps through the forest, the shadows of the tree trunks and leaves falling on it produce spurious luminance-based motion cues. The chromatic border between the leopard and the trees in the background does not change as radically and may provide a more reliable clue to the leopard's position and motion. According to our results, the hypothetical arboreal ancestor trying to escape the leopard would do well determine its motion using the most reliable moving contours, regardless of whether they arise from luminance or color, and in cases of ambiguity, to emphasize attended features.

REFERENCES

Anstis, S. M. (1970). Phi movement as a subtraction process. *Vision Research, 10,* 1411–1430.

Anstis, S. M. (1980). The perception of apparent movement. *Philosophical Transactions of the Royal Society of London B, 290,* 153–168.

Braddick, O. J. (1980). Low-level and high-level processes in apparent motion. *Philosophical Transactions of the Royal Society of London B, 290,* 137–151.

Cavanagh, P. (1991). Short-range vs. long-range motion: Not a valid distinction. *Spatial Vision, 5,* 303–309.

Cavanagh, P. (1992). Attention-based motion perception. *Science, 257,* 1563–1565.

Cavanagh, P. & Anstis, S. (1991). The contribution of color to motion in normal and color-deficient observers. *Vision Research, 31,* 2109–2148.

Cavanagh, P. & Favreau, O. E. (1985). Color and luminance share a common motion pathway. *Vision Research, 25,* 1595–1601.

Cavanagh, P. & Mather, G. (1989). Motion: The long and short of it. *Spatial Vision, 4,* 103–129.

Cavanagh, P., Tyler, C. W. & Favreau, O. E. (1984). Perceived velocity of moving chromatic gratings. *Journal of the Optical Society of America, 1,* 893–899.

Cleary, R. & Braddick, O. J. (1990). Masking of low frequency information in short-range apparent motion. *Vision Research, 30,* 317–327.

Eagle, R. A. & Rogers, B. J. (1991). Does figure–ground organisation influence motion capture? *Investigative Ophthalmology and Visual Science (Suppl.), 31,* 520.

Ejima, Y. & Takahashi, S. (1988). Illusory contours induced by isoluminant chromatic patterns. *Vision Research, 28,* 1367–1377.

Hikosaka, O., Miyauchi, S. & Shimojo, S. (1993). Focal visual attention produces illusory temporal order and motion sensation. *Vision Research, 9,* 1219–1240.

Livingstone, M. L. & Hubel, D. H. (1987). Psychophysical evidence for separate channels for the perception of form, color, movement and depth. *Journal of Neuroscience, 7,* 3416–3468.

MacKay, D. M. (1961). Visual effects of non-redundant stimulation. *Nature, 192,* 739–740.

Meyer, G. E. & Howard, C. (1993). Texture capture in stationary and moving Necker cubes. *Investigative Ophthalmology and Visual Science (Suppl.), 34,* 1238.

Newsome, W. T. & Paré, E. B. (1988). A selective impairment of motion perception following lesions of the middle temporal area (MT). *Journal of Neuroscience, 8,* 2201–2211.

Pylyshyn, Z. W. & Storm, R. W. (1988). Tracking multiple independent targets: Evidence for a parallel tracking mechanism. *Spatial Vision, 3,* 179–197.

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

Ramachandran, V. S. (1986). Capture of stereopsis and apparent motion by illusory contours. *Perception & Psychophysics, 39,* 361–373.

Ramachandran, V. S. (1987). Interaction between color and motion in human vision. *Nature, 328,* 645–647.

Ramachandran, V. S. & Anstis, S. M. (1986). Figure–ground segregation modulates apparent motion. *Vision Research, 26,* 1969–1975.

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

Ramachandran, V. S. & Gregory, R. (1978). Interaction between colour and motion in human vision. *Nature, 275,* 55–56.

Ramachandran, V. S. & Inada, V. (1985). Spatial phase and frequency in motion capture of random-dot patterns. *Spatial Vision, 1,* 57–67.

Stelmach, L. B. & Herdman, C. M. (1991). Directed attention and perception of temporal order. *Journal of Experimental Psychology: Human Perception and Performance, 17,* 539–550.

Stromeyer, C. F. III, Eskew, R. T. & Kronauer, R. E. (1990). The most sensitive motion detectors in humans are spectrally opponent. *Investigative Ophthalmology and Visual Science (Suppl.), 31,* 241.

Acknowledgements—This research was supported by grant NEI EY09258 to PC. The results were presented in part at the annual meeting of the Association for Research in Vision and Ophthalmology in Sarasota, Fla, 1993. We are grateful to James Intriligator, Ron Rensink, Satoru Suzuki, Peter Tse, and Qin Qin Wang for helpful comments on the manuscript.