11

Is There Low-Level Motion Processing for Non-Luminance-Based Stimuli?

Patrick Cavanagh

How do we see the motion of "equiluminous" features? By equiluminous, I mean regions that have the same luminance as the background but differ in some other property such as color, texture, depth, or relative motion. The physiological literature has given ample evidence of directionally selective units at early levels of visual processing but these have been overwhelmingly tested with luminance-defined stimuli. More recent studies have revealed directionally selective responses to equiluminous color stimuli (Saito et al., 1989; Dobkins and Albright, 1993) and texture-defined stimuli (Albright, 1992) but it has been argued that these may be weak residual responses mediated by nonlinearities in the luminance pathway. In this chapter, I will examine whether there are specialized low-level motion detectors for non-luminance-based stimuli. The alternative is that the motion of non-luminance-based stimuli is detected only by a second motion system, which relies on the attentive tracking of visible features (Cavanagh, 1992). In the first series of experiments, I hunted for various types of low-level detectors by testing whether there was a common, low-level motion pathway that responded to both luminance and non-luminance-defined stimuli. In the second series of experiments, I diverted attention from moving, non-luminance-defined stimuli to see whether motion processing remained viable in the absence of attention. I am reporting only preliminary observations in the case of drifting stereo- or motion-defined structure as these patterns were created by local modulations of a static dot field. This static component may have biased the low-level motion responses. Further studies with dynamic random-dot stereograms and motion fields with short dot life times are underway to extend these preliminary observations.

Two Motion Streams

The notion of two separate motion systems was initially suggested by Julesz (1971). He claimed that the low-level movement detectors found by Hubel and Wiesel (1968) were different from higher-level movement analyzers that operate following pattern matching. Julesz described experiments on the motion of figures defined in random-dot stereograms and concluded that the motion perceived for these stimuli was mediated by a higher-order process. This early description of two motion streams was followed by similar claims by Anstis (1980) and Braddick (1974, 1980). Recently, I described these two motion streams as passive (low-level) and active (high-level) processes and considered how each of them might analyze the motion of different stimulus types (Cavanagh, 1991). Figure 11.1 shows the possible combinations of passive and active motion processes with five types of stimuli. Of these five, I have classified color and luminance as first-order stimuli and the three others as second-order stimuli (Cavanagh and Mather, 1989). These classifications are based only on stimulus structure. First-order statistics specify the frequency with which individual points in an image have specified intensity or color values. Two areas in an image differ in their first-order statistics if they have different mean luminances or spectral compositions. Motion detectors specialized for first-order patterns (or at least for luminance) therefore correspond to the extensively studied directionally selective units of the striate cortex. Recent psychophysical (Cavanagh and Anstis, 1991) and physiological studies (Dobkins and Albright, 1993) have also argued for low-level motion detectors for equiluminous color.

Two areas may have the same mean luminance and color, but differ in their spatial, temporal, or ocular distributions of luminance and color. The two areas are then differentiated by second-order properties such as texture, motion, or binocular disparity. The difference between first-order and second-order structure lies in the stimulus. One goal of this chapter is to determine if there are passive, low-level motion detectors for second-order stimuli defined by texture, relative motion, or binocular disparity.

Motion Process	Stimulus Factors				
	First Order		Second Order		
	Lum.	Color	Texture	Stereo	Motion
Passive					
Active					

Figure 11.1
Stimulus and process factors in motion perception. A stimulus factor divides stimulus types into first order or second order. First-order stimuli (luminance or color) can be defined at a single point. Second-order stimuli require two points, separated in space for texture, separated by eye for binocular disparity, and separated by space and time for motion. Two types of motion processes—active and passive—can respond to both of these stimulus types. Passive motion processes involve arrays of localized motion detectors that monitor all areas of the retina. Active processes involve tracking individual targets with attention as they move about the visual field. Short-range motion, as originally described by Braddick (1980) and Anstis (1980), corresponds only to the responses of passive motion processes to luminance stimuli. Long-range motion corresponds to the remaining combinations.

In addition to the low-level, passive motion detectors there is also a second stream of motion processing that I have called active motion perception because it involves the use of attention to track moving stimuli. This attention-based motion process can obviously respond to the same stimuli (first- or second-order) that activate passive detectors; that is, as long as the stimuli are visible they can be tracked. The active/passive distinction is therefore independent of the stimuli present in the display. Compared to low-level motion processes (Braddick, 1980; Anstis, 1980), this attention-based process appears to have a limited capacity (Pylyshyn and Storm, 1988) and very different thresholds but supports more accurate velocity judgements (Cavanagh, 1992). These differences have suggested that the tracking process itself does not rely on, or at least does not require, low-level motion signals to maintain tracking of the target. For example, observers can accurately track equiluminous color targets even when their apparent velocity judged by low-level mechanisms is grossly underestimated (Cavanagh, 1992). Any given stimulus may engage either, or both, passive and active motion mechanisms and the observed performance can be interpreted meaningfully only if it is known which are involved (Cavanagh, 1991).

Julesz (1971) suggested that when cues for both motion streams are present, the low-level stream usually dominates and the operation of the higher-order mechanism is concealed. He claimed that using cyclopean stimuli avoided engaging the low-level process, allowing the higher-level one to be isolated. His conjecture will be explicitly tested in this chapter using a more general test to isolate the two motion streams. The test was developed to examine the role of low- and high-level motion in the perception of motion for color stimuli (Cavanagh and Anstis, 1991; Cavanagh, 1992). In the test, the two motion streams see opposite directions of motion and the observer can report either at will. Luminance and color gratings were superimposed and set in motion in opposite directions. Because of masking from the color grating, the bars of the luminance grating were not visible and they could not be tracked; nevertheless, their motion was visible and it determined the observed direction of rotation. This "disembodied" motion was striking because no features could be seen actually moving in the direction of the overall motion.

On the other hand, the bars of the color grating were visible and could be tracked at will by the high-level, attention-based process. The tracking of the color bars could not have been based on the motion signals from low-level detectors because the low-level motion response was dominated by the luminance grating. If the low-level signal alone were sufficient for tracking, then the luminance bars should have been tracked at least as easily as the color bars; but the luminance bars could not be tracked at all (they were not visible to the pattern pathway that supposedly mediates tracking).

In this composite stimulus, the color bars could be seen to move (in the opposite direction to the overall stimulus rotation) only when they were being tracked with attention. Evidently, the motion impressions during tracking of the color bars in this stimulus must represent the output of a distinct motion process, one that cannot be based on the motion signals from low-level detectors (these are signaling motion in the opposite direction). The results therefore indicate that there must be two independent sources for the impressions of motion in this stimulus: one for the judgment of overall motion (passive) and a separate one for tracking (active).

The attention-based motion process might derive motion signals from position information that is read out from the focus of attention. Motion might be derived from monitoring or noticing the change in this attention-selected position signal (figure 11.2b). Alternatively, motion impressions could be based on information about the focus of attention itself, either its position or its displacement (figure 11.2c). Specifically, to perform a tracking task, some servomechanism must be comparing the position of the attentional focus to that of the tracked target and repositioning attention to follow the target. The signals from this control process, either the current position of attention or perhaps its tracking rate, would be sufficient to generate an appropriate motion impression.

Attentive tracking may be required to perceive the motion of non-luminance-based stimuli, or, to put it another way, there may be no low-level motion detectors for non-

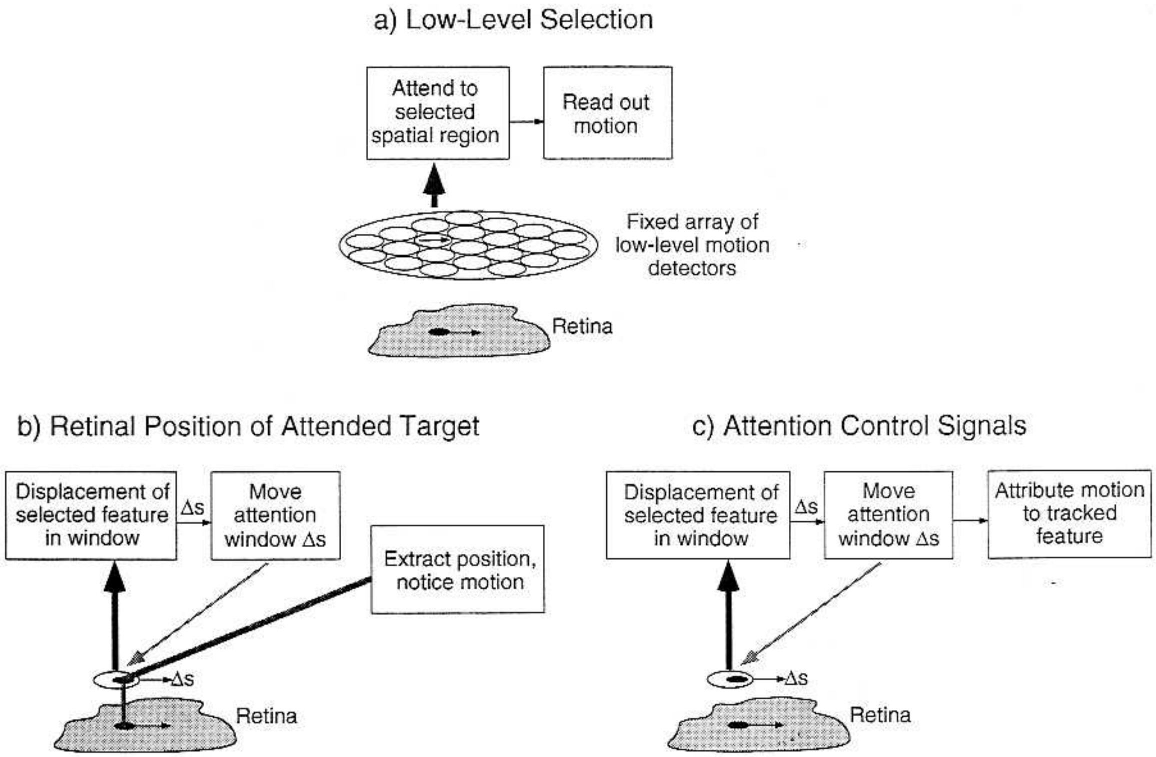

Figure 11.2
The motion sensations for an object tracked with attention could arise in a number of ways. Although they could be based on a selection from the low-level signals available in the attended region as depicted in (*a*), specific stimuli where this option is not available (Cavanagh, 1991) still support motion impressions independently of low-level signals. These high-level motion impressions may be derived from the position of features individuated and tracked within the focus of attention (*b*), or from the signals that keep the focus of attention centered over the tracked object (*c*).

luminance-based stimuli. If this is the case, the three top rightmost boxes of figure 11.1 would not exist. I tested for the presence of these low-level detectors by trying to get them to interact with luminance-based low-level detectors in a motion nulling task and in a plaid motion task. I also tried to divert attention from first- and second-order stimuli to see whether motion impressions and motion aftereffects could survive the removal of attention. If motion of, say, stereo-defined gratings could be perceived in the absence of attention, I would conclude that there are low-level passive motion detectors available for this attribute.

Motion Nulls between Luminance and Non-Luminance-Based Motion

If low-level detectors do operate on non-luminance-based stimuli, the different types of low-level detectors may all contribute to a common motion pathway. If this is the case, the motion of a luminance grating drifting in one direction should be able to cancel or null the motion of an oppositely moving, non-luminance-based stimulus. Nothing requires a common motion pathway for the different types of low-level detectors, so a negative outcome (e.g., if two attributes do not null) is inconclusive. On the other hand, a positive outcome (e.g., two attributes do null) does demonstrate the existence of low-level detectors for both attributes. As an example, we have been able to demonstrate motion nulling between color and luminance (Cavanagh and Anstis, 1991). When color is moving, say, to the left and luminance is moving to the right, the overall motion will appear to be in the direction of the color grating when the luminance grating has low contrast. At much higher contrasts the overall motion is in the direction of the luminance grating. At some intermediate value (between 10 and 20% contrast in our experiments), the two motions cancel and a flickering stimulus is seen. The fact that the two stimuli cancel suggests that they contribute to a common mechanism. If they did not contribute to a common mechanism, we might expect a perception of transparency with the two gratings appearing to slide through each other.

This sliding or transparency can be seen with two opposing luminance gratings if they are sufficiently separated in spatial frequency (about a factor of 4). Similarly, color and luminance gratings slide over each other if they have dissimilar spatial and temporal frequencies.

It is easy to show that the nulling of the overall motion impression between the two gratings is independent of the operation of high-level tracking processes and so requires the participation of low-level processes for both the attributes being opposed. The evidence for this is that even though the overall motion of the superimposed gratings shows this nulling behavior, attentive tracking is never "nulled." The observer can switch between reporting an impression of overall motion (which shows low-level nulling) or the direction of tracking individual features. At moderate luminance contrasts (10–20%), the color bars can be tracked even though the overall motion is in the opposite direction. At higher luminance contrasts, either the color or the luminance grating can be tracked at will. These results suggest that motion nulling between two gratings of different types is a signature of a low-level detectors for both types that contribute to a common pathway. I therefore extended our first test of motion nulling for color and luminance to tests of motion nulling between luminance gratings and gratings defined by color, texture, binocular disparity, or relative motion.

Motion nulling was clear between color and luminance. It was possible to get a motion null between texture and luminance at slower speeds, although at higher speeds, they appeared to slide through each other. It was never possible to get a motion null between stereo-defined gratings and luminance or between motion-defined gratings and luminance. In both cases, the gratings appeared completely transparent and moved independently in opposite directions. These results imply that there may be a low-level motion detector responding to texture-defined gratings (see also Chubb and Sperling, 1991). The negative results for stereo-defined and motion-defined gratings may indicate that there are no low-level motion detectors for these two attributes. Contrary evidence has been reported by Patterson et al. (1993) who claim to find a motion aftereffect following adaptation to drifting stereo-defined gratings. Nishida and Sato (1993) also reported motion aftereffects following adaptation to drifting second-order stimuli.

Motion Plaids between Luminance and Non-Luminance-Based Motion

A motion plaid (Adelson and Movshon, 1982) is produced by the superposition of two gratings at different orientations. When both gratings are defined by luminance there is a range of different contrasts, speeds, and spatial frequencies of the two gratings for which they cohere and a compound pattern is seen drifting in an intermediate direction. Outside this coherence range, the two gratings appear to slide through each other. Stoner and Albright (1992) have tested gratings composed of different attributes with the assumption that coherence between different attributes shows that they activate a common motion pathway or analysis. They did find coherence between texture gratings and luminance gratings implying, as I argued above, that texture is analyzed by low-level motion detectors. I tested plaids made up of luminance gratings and either stereo-defined gratings or motion-defined gratings. It was not possible in either case to find any setting of speed or contrast that produced a coherent, compound pattern. The gratings appeared to drift over each other as if completely transparent.

Although these results are consistent with the nulling results above, other results from Krauskopf and Farell (1990) are not. These authors reported that plaids made up of color and luminance did not cohere even though, as mentioned above, we found that color and luminance would null each other (Cavanagh and Anstis, 1991) when superimposed and drifting in opposite direction. Why would the two attributes interfere when superimposed and drifting in opposite directions (let us call this a 180° plaid) but not when they were superimposed at 90°? There is no obvious answer to this discrepancy but it is possible that the results for plaid stimuli may be mediated by other processes in addition to motion (for more details, see Farell, this volume). The ability to see the two components as separate or not may also be an extension of the monocular rivalry phenomenon (Georgeson, 1984). In this stimulus, two orthogonal, superimposed gratings can be seen separately, alternating back and forth between one and the other, even though neither is moving. If this is the source of the lack of coherence for the gratings in Krauskopf and Farell's study, then these same stimuli should be easily separated when stationary as well. In this case, positive results from cross-cue plaid experiments would indicate common motion processes (e.g., texture and luminance gratings cohere, Stoner and Albright, 1992), whereas negative results (e.g., color and luminance do not cohere, Krauskopf and Farell, 1990) would not be informative.

Summary for Common Pathway Studies

Overall, we have strong evidence for low-level detectors that contribute to a common pathway for luminance-, color-, and texture-defined gratings. The lack of interaction between luminance and either stereo-defined or motion-defined gratings is clear in both the motion null and motion plaid experiments. This result might indicate that there are no low-level motion detectors for these two stimulus types but strictly speaking it may indicate only that if these detectors are present, they do not contribute to a final pathway shared by luminance. Moreover, the lack of coherence in the moving plaids made up of luminance and either stereo- or motion-defined gratings may be a result of strong monocular rivalry. It remains to be seen whether there are low-level detectors for stereo-defined and motion-defined stimuli. The next tests considered the alternative hypothesis that the motion of these two stimulus types is mediated only by attentional processes (all five stimulus types were tested). Attention was diverted from the moving stimuli to see whether any impressions of motion would persist in the absence of attention.

Motion Perception without Attention for Non-Luminance-Based Stimuli

Gratings defined by color, texture, stereo, or relative motion can be seen to move, but all produce a degraded or slowed perception of motion. It could be argued that there are no low-level detectors for these types of stimuli and that the motion is perceived only because of the attentive tracking of the stimuli. This experiment examined the attention hypothesis by engaging attention in a secondary task (attentive tracking in an inner, counterphasing ring) and measuring the effect on motion impressions for the different stimulus types moving smoothly in an outer ring.

In the outer annulus, a grating defined by luminance, color, stereo, motion, or texture was presented rotating either CW or CCW (figure 11.3). In the inner annulus, a luminance grating was presented in counterphase flicker.

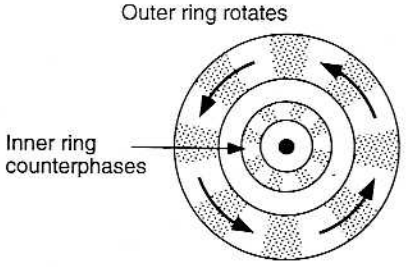

Figure 11.3
A tracking task is run in the center of the display. The luminance grating is in counterphase flicker and the observer can track it either clockwise or counterclockwise. During a given trial, tracking is in one direction only. In the outer ring, a grating defined by luminance, color, texture, stereo, or motion is presented. If attention is necessary to perceive the motion of any of these stimuli, the perception of motion in the outer ring should be severely compromised, especially when its motion is in the direction opposite to the tracking direction.

The temporal frequencies of the moving and flickering gratings were the same (2 Hz). Observers were trained to track the central grating in either direction and to note the impressions of motion in the outer grating during tracking.

Observers reported clear impressions of motion in the outer grating when defined by luminance, color, or texture. It was not important whether the outer grating was moving in the same direction as the tracking for the inner grating.

Observers reported a large loss in the impressions of motion in the outer ring when the stimulus was a stereo- or motion-defined grating. The loss was not so large when the outer grating was moving in step with the tracking of the inner grating, but it was almost total when the outer grating moved in opposition to the tracking.

These results suggest that the perception of motion for stereo- and motion-defined stimuli is mediated solely by attention and that there are no low-level detectors signaling motion for these stimuli.

Motion Aftereffects for Non-Luminance-Based Stimuli

I tested for simple motion aftereffect with drifting gratings of the five stimulus types. In each case the aftereffect was verified on a static grating of the same type as the adapting grating. Aftereffects were found only for luminance and color gratings, replicating many earlier studies. Other studies have reported aftereffects for texture-defined (Nishida and Sato, 1993) and stereo-defined gratings (Patterson et al., 1993). Nishida and Sato (1993) claim that motion aftereffects can be revealed for second-order stimuli like texture if a counterphase test is used. Patterson et al. (1993) used a static luminance grating as a test so their result is quite striking. An additional difference between his studies and ours is that his stereo-defined gratings use dynamic random dots whereas ours use static random dots (modulated by disparity).

Motion Aftereffects for Luminance-Based Stimuli in the Absence of Attention

As a final test, I examined whether attention was required to produce a motion aftereffect even in the case of luminance stimuli (Chaudhuri, 1990). The double annulus stimulus shown in figure 11.3 was used with a luminance grating moving continously in the outer ring during adaptation. Observers tracked the inner ring in the direction opposite to the motion of the outer ring for about 30 sec. The motion in the outer ring was then stopped. All observers reported a motion aftereffect in the outer ring, supporting Wohlgemuth's (1911) original observations.

Conclusions

How do we see the motion of "equiluminous" features? For at least two types of non-luminance-based stimuli, color-defined and texture-defined, the evidence presented here argues that there are low-level motion detectors that signal their motion. These detectors, along with those for luminance-defined stimuli, contribute to a common motion pathway, accounting for their mutual interference in the motion nulling task. Attentive tracking would always be available to mediate the perception of motion for color and texture-defined stimuli, but motion could be seen for both even in the absence of attention.

It could be argued that the low-level responses to color and texture are due to some residual nonlinearities in luminance-based motion detectors. This possibility has been discounted for the case of color (Cavanagh and Anstis, 1991). In fact, Stromeyer et al. (1990) have shown that the threshold for drifting equiluminous gratings is as much as four times lower (in cone contrast units) than it is for a luminance grating. This finding rules out the notion that the motion response to color could be mediated by a residual distortion in the response of the luminance pathway to color. Their result has been confirmed recently by Metha et al. (1993). Chubb and Sperling (1991) have considered the nature of low-level detectors that would respond to their texture gratings and they suggest detectors with half- or full-wave rectification of image contrast. Neither of their suggested operators is present in standard motion detectors for luminance-based stimuli as contrast polarity plays a significant role in low-level motion response (Anstis and Rogers, 1975).

In each experiment, motion-defined and stereo-defined stimuli acted as if there were no low-level motion detectors specialized for these stimuli. They did not null the motion of a luminance grating moving in the opposite direction; they did not cohere with a luminance grating in a plaid; their motion was no longer seen when attention was diverted away from them.

These results would appear to validate Julesz's (1971) conjecture that the motion of cyclopean stimuli was analyzed solely by high-level mechanisms. I would add a caution, however, based on the nature of the stereo- and motion-defined stimuli used here. As mentioned in the introduction, the drifting stereo- or motion-defined structure used in these experiments was created by local modulations of a static dot field. Further studies with dynamic random-dot stereograms and motion fields with short dot life times are underway to extend these preliminary observations reported here.

Acknowledgments

This research was supported by NEI grant EY09258.

References

Adelson, E. H., and Movshon, J. A. (1982). Phenomenal coherence of moving gratings. *Nature (London) 300*, 523–525.

Albright, T. D. (1992). Form-cue invariant processing in primate visual cortex. *Science 255*, 1141–1143.

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

Anstis, S. M. (1980). The perception of apparent movement. *Phil. Transact. R. Soc. London B290*, 153–168.

Anstis, S. M., and Rogers, B. J. (1975). Illusory reversal of visual depth and movement during change of contrast. *Vision Res. 15*, 957–962.

Braddick, O. (1974). A short-range process in apparent motion. *Vision Res. 25*, 839–847.

Braddick, O. (1980). Low-level and high-level processes in apparent motion. *Phil. Transact. R. Soc. London B290*, 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., and Anstis, S. M. (1991). The contribution of color to motion in normal and color-deficient observers. *Vision Res. 31*, 2109–2148.

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

Chaudhuri, A. (1990). Modulation of the motion aftereffect by selective attention. *Nature (London) 344*, 60–62.

Chubb, C., and Sperling, G. (1991). Texture Quilts: Basic tools for studying Motion-from-Texture. *J. Math. Psychol. 35*(4), 411–442.

Dobkins, K. R., and Albright, T. D. (1993). What happens if it changes color when it moves?: The nature of chromatic input to Macaque visual area MT. *J. Neurosci.*

Georgeson, M. A. (1984). Eye movements, afterimages and monocular rivalry. *Vision Res. 24*, 1311–1319.

Hubel, D. H., and Wiesel, T. N. (1968). Receptive fields and functional architecture of monkey striate cortex. *J. Physiol. 195*, 215–243.

Julesz, B. (1971). *Foundations of Cyclopean Perception*. Chicago: University of Chicago Press.

Krauskopf, J., and Farell, B. (1990). Influence of colour on the perception of coherent motion. *Nature (London) 348*, 328–331.

Metha, A. B., Vingrys, A. J., and Badcock, D. R. (1993). Cone contrast detection and direction thresholds. *Invest. Ophthalmol. Visual Sci. 34*, 1032.

Nishida, S., and Sato, T. (1993).Two kinds of motion aftereffect reveal different types of motion processing. *Invest. Ophthalmol. Visual Sci. 34*, 1363.

Patterson, R., Bowd, C., Phinney, R., Pohndorf, R., Barton-Howard, W. J., and Angilletta, M. (1993). Properties of the stereoscopic (cyclopean) motion aftereffect. Unpublished manuscript.

Pylyshyn, Z. W., and Storm, R. W. (1988). Tracking multiple independent targets: evidence for a parallel tracking mechanism. *Spatial Vision 3*, 151–224.

Saito, H., Tanaka, K., Isono, H., Yasuda, M., and Mikami, A. (1989). Directionally selective response of cells in the middle temporal area (MT) of the macaque monkey to the movement of equiluminous opponent color stimuli. *Exp. Brain Res. 75*, 1–14.

Stoner, G. R., and Albright, T. D. (1992). Motion coherency rules are form-cue invariant. *Vision Res. 32*, 465–475.

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

Wohlgemuth, J. M. (1911). On the after effect of seen motion. *Br. J. Psychol. Monogr. Suppl. 1*, 1–117.