

Commentary

More than one way to see it move?

Thomas D. Albright

Howard Hughes Medical Institute, Salk Institute for Biological Studies, La Jolla, CA 92037

A dominant 19th-century view on the nature of visual perception, known as “elementism,” held that any percept is no more and no less than the sum of the internal states caused by the individual sensory parts (or elements), such as brightness, color, and distance. Elementism is easily falsifiable, but the idea of linearly independent internal states is so compelling to a reductionist neuroscience that it has been dragged from the dustbin repeatedly over the last 100 years—in conjunction with the more familiar 19th-century doctrine on brain organization known as “localization of function.” A sensational case in point is a hypothesis put forward in the 1980s that maintains that there are multiple independent channels in the primate visual system, extending from retina through several stages in the cerebral cortex, each of which is specialized for the processing of distinct elements of the visual image (1). Although this hypothesis has rather general implications for the neuronal bases of perceptual experience, nowhere have these implications been made more explicit, nor inspired greater controversy, than with regard to the neuronal representations of color and motion. In a report in this issue of the *Proceedings*, Sperling and colleagues (2) adopt an intriguing approach to this topic. Although they are sure to fan the flames of controversy, the results of Lu *et al.* (2) do offer a stimulating perspective and emphasize a broad theoretical framework for visual motion processing in which many related phenomena can be understood.

One trail of debate surrounding the relationship between color and motion processing can be traced to the fact that motion perception is necessarily secondary to the detection of some type of contrast in the visual image; quite simply: if you can't see it, you can't see it move. Visible moving objects commonly contrast with their backgrounds in multiple ways, such as by differences in the intensity (luminance), wavelength composition (chrominance), and pattern (texture) of reflected light. Inasmuch as these differences are manifested in the retinal image, they are all potential cues on which a motion detector may operate. Indeed, because chromatic differences are commonly reliable indicators of the edge of an object (3), one might expect chromatic contrast in an image to be among the most robust cues for motion detection.

Teleology notwithstanding, a number of observations, interpreted through the lens of a 20th-century neurobiological elementism, have led to a rather different view, in which the neuronal representations of color and motion are thought to be very limited in their interactions. This view began to crystallize approximately 25 years ago with discovery of the multiplicity of cortical visual areas in primates. In particular, based on a survey of the visual response properties of cortical neurons, Semir Zeki (4) reported the existence of separate visual areas specialized for the analysis of motion (area MT or V5) and color (area V4). This apparent independence led to the novel prediction that, if a moving object were distinguished from background solely on the basis of a difference in color—an infrequent occurrence in the natural world—its motion would not be perceived. Early tests of this hypothesis (made possible, in part, by the advent of sophisticated and inexpensive computer graphics) appeared to support it in a limited form: Perceived motion of chromatically defined stimuli was said to

be possible, but the quality of the percept was often poor (5, 6). In 1987, Livingstone and Hubel (1) upped the ante by noting that not only was this perceptual degradation marked under selected and well controlled conditions, but it could be explained by evidence that neuronal signals giving rise to color and motion percepts are channeled separately through several stages of processing, from retina to higher cortical visual areas.

These conflicting portraits of the relationship between color and motion—one drawn from simple functional expectations, the other from modern neuroscience—have inspired a large number of anatomical, physiological, and behavioral studies in recent years (e.g., refs. 1 and 7–20; see ref. 21 for review), which are perhaps most notable for their failure to achieve a broad consensus of opinion. Nonetheless, the one thing most generally agreed on is that the sensitivity of both human observers and cortical neurons to the motion of chromatically defined stimuli is poorer than sensitivity to motion of stimuli defined by other contrast cues. In their new report, Lu *et al.* (2) refute this conclusion, alleging instead that motion of chromatically defined stimuli can be perceived quite clearly, provided the chromatic contrast in the moving stimulus is sufficiently “salient” to draw the attention of the observer.

This provocative claim, of course, begs the question of what is meant by “salience” and “attention.” The answer can be understood, at least in part, by reference to a larger theoretical framework for motion processing previously championed by Lu and Sperling (22). According to this framework, there are three basic forms of motion detection, which are defined, to a degree, by the types of moving stimuli they detect. The “first-” and “second-order” motion systems, which have been studied for many years and for which both neuronal substrates (23–26) and computational mechanisms (27–31) have been identified (see ref. 24 for review), are sensitive to motions of luminance-defined and texture-defined patterns, respectively.

The hypothesized “third-order” motion system, as described by Lu and Sperling (22), “tracks” the motions of salient image features, regardless of how they are physically defined in the visual image. The concept of a third form of motion detection is not novel (see ref. 32), but Sperling and colleagues have made important advances in defining its properties and underlying mechanisms (2). According to Lu *et al.* (2), the input to this system is from a “salience map,” in which image features are assigned weights based on a combination of their inherent physical contrasts and the observer's attentional focus. All else being equal, image features that are interpreted perceptually as “foreground,” rather than “background,” are assigned the greatest weights.

Under normal environmental conditions, visual image motions that activate the proposed third-order system also will activate the first- and/or second-order system. By careful elimination of image cues that drive first- and second-order systems, however, Lu and Sperling (22) found it possible to selectively activate the third-order motion system in human observers. In addition, they catalogued a set of perceptual criteria by which the operations of this system could be distinguished reliably from those of first- and second-order systems. These criteria include a variety of filtering

properties: By contrast with the first- and second-order systems, the third-order system is more sluggish in its responsiveness, is entirely insensitive to the eye of stimulation, and tracks the motions of features defined by the intersections of moving and nonmoving patterns.

In their new study, Lu *et al.* (2) applied these criteria to identify the motion system(s) activated by chromatically defined stimuli. Heretofore, it was believed that such stimuli—inasmuch that they were seen to move—could drive both first- (12, 33) and third-order systems (32). Lu *et al.* (2) now report that human sensitivity to chromatically defined stimuli satisfies the aforementioned criteria for third-order motion processing. Such evidence is correlative and circumstantial by nature, but it suggests that chromatically defined stimuli may be processed to a great extent by a system that is distinct from the classically defined first- and second-order motion systems.

If that suggestion alone is not sufficient food for thought (or rebellion), real intrigue lies in the predictions that can be made from the mechanistic principles thought to underlie the third-order system. As defined by Lu and Sperling (22), these principles predict that manipulations of the salience of chromatically defined stimuli should markedly affect perceived motion. To test this hypothesis, Lu *et al.* (2) were required to alter stimulus salience without altering its detectability by first- and second-order systems. They sought that goal by two related means, the second of which illustrates the point. In this experiment, Lu *et al.* (2) compared perceptual sensitivity for motion of a red/green striped pattern (Fig. 1A), viewed on a gray background, to sensitivity for motion of red/gray stripes (Fig. 1B) on the same gray background. In both cases, the pattern was defined solely by a chromatic difference (red/green or red/gray). The two patterns differed considerably, however, with respect to the relative salience of their stripes. Although the former stimulus was composed of red and green stripes of approximately equal salience, the latter was composed of red and gray stripes of markedly different salience. Specifically, because the gray stripes were physically coextensive with the larger background, they were seen as such whereas the red stripes appeared as moving foreground features.

The effects of this manipulation were striking: Human observers reported little or no perceived motion in the first instance (low salience) whereas the motion percept was robust and compelling in the second (high salience), despite the fact that the only cue available for motion detection in both instances was chromatic contrast. Compelling though this demonstration and the underlying logic may seem, there are some loose ends that deserve greater attention. In particular, the intended stimulus differences also render quite different degrees of chromatic contrast within the moving pattern. These contrast differences, in turn, lead to (i) differences in the activations levels of long-wavelength-sensitive (L) and mid-wavelength-sensitive (M) cone photoreceptors and (ii) different chromatic adaptation levels at the boundary of the moving stripes. The possibility certainly exists that one or both of these by-products could yield differential activation of the first-order motion system.

Lu *et al.* (2) conclude their report with a demonstration of the way in which chromatically defined features can subserve motion detection, provided they are segmented from background. The stimulus used for this purpose is one in which two types of patterns are presented on alternate frames of a movie (Fig. 2). Odd frames contain chromatically defined stripes (e.g., red and green); even frames contain stripes defined by different textures (e.g., coarse and fine). If only odd or even frames of the movie are viewed, no physical motion occurs—only flicker is perceived. If odd and even frames are viewed in the full sequence, however, motion is clearly seen. Moreover, the direction of perceived motion depends on which of the stripes in each pattern type (red vs. green and coarse vs. fine) is seen as foreground. The unambiguous message here is that perceived motion is only possible by matching salient features, which is, of course, the *modus operandi*

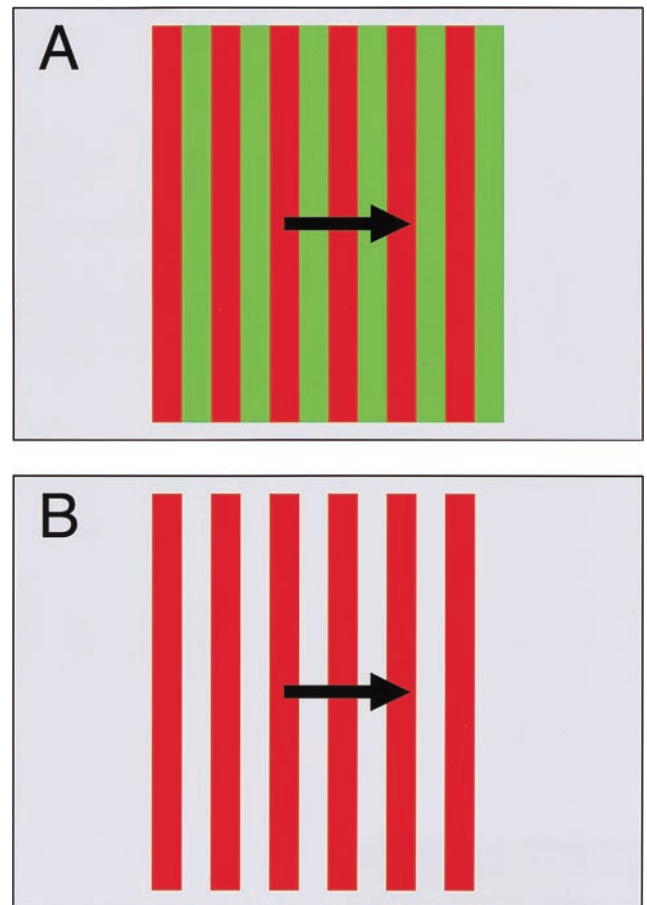


FIG. 1. Schematic illustration of one method used by Lu *et al.* (2) to manipulate salience within a chromatically defined pattern. (A) The standard stimulus configuration, which has been used in many previous experiments, consists of a pattern of red and green stripes on a chromatically intermediate (gray) background. The luminances of the stripes and background are equivalent (i.e., isoluminant). When this stimulus is moved, its motion generally appears slower and more irregular than it truly is. Lu *et al.* (2) attribute this nonveridical percept to the fact that the red and green stripes are of nearly equal salience. (B) If the green stripes are replaced with a gray color that is identical to background, the red stripes are more likely to be perceived as foreground and hence more salient than the gray. The result is a robust percept of motion, despite the fact that the red/gray pattern—like the red/green pattern—is isoluminant and defined only by a chromatic difference. Lu *et al.* (2) argue that this percept of “high-quality isoluminant motion” is a product of the proposed third-order motion system.

of the proposed third-order system. What’s more, the matching occurs between salient features that are defined by two completely different forms of physical contrast. These observations refine claims regarding the mechanistic underpinnings of the third-order system: (i) the system operates on feature “tokens”—a generic representational currency for salient features, regardless of how they are physically defined—and (ii) chromatic contrast is sufficient to establish such tokens.

The tableau of evidence delivered by Lu *et al.* (2) makes a plausible case for third-order motion processing and its role in the perception of motion of chromatically defined patterns. It will not go unnoticed, however, that these authors additionally claim such motion to be “computed by the third-order motion system, and only by the third-order motion system” (ref. 2, p. 8289). What then are we to make of the many recent behavioral and neurophysiological experiments (see ref. 21 for review) that imply sensitivity by a first-order mechanism? Lu *et al.* (2) dismiss these evidences en masse, citing the possibility of unbalanced salience, as well as potential calibration errors that could lead to unintended lumi-

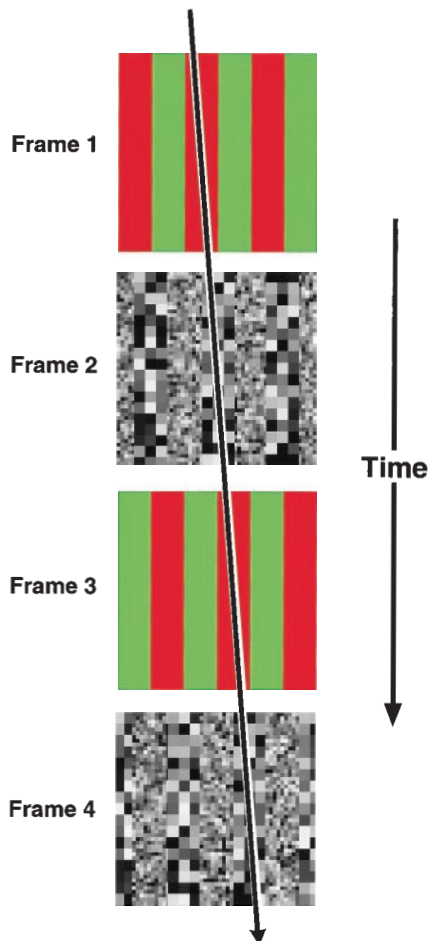


FIG. 2. Schematic illustration of method used by Lu *et al.* (2) to demonstrate contribution of chromatic contrast to third-order motion detection. A movie consisted of four sequentially presented frames. Odd frames contained a red/green striped pattern; even frames contained a striped pattern formed of coarse and fine textures. If colored or textured frames are viewed alone, only flicker is perceived. If the complete sequence is viewed, however, coherent motion is seen. The direction of perceived motion is determined by the stripes that are interpreted as foreground within each frame type. For example, if red stripes are seen as foreground in chromatic frames, and coarse-texture stripes are seen as foreground in texture frames, motion will be perceived to the right: i.e., in the direction that these foreground stripes move (indicated by arrow). This phenomenon argues for a type of motion detection that operates on perceptually defined foreground features, regardless of how they are physically defined.

nance contrast in the visual stimuli used. True, it is rather difficult to create chromatic stimuli that lack such luminance artifacts, but it is disingenuous, at best, to suppose that all others have not been equal to the task. A more parsimonious explanation is that both first- and third-order mechanisms operate under these conditions, although the latter may well play the more important role.

The proposal of Lu *et al.* (2) also inevitably leads to questions about neuronal substrates: At what level in the visual processing hierarchy is third-order motion represented? What neuronal structures and events underlie the salience map, thought to provide input to the third-order system? At this point in time, relevant physiological data are scarce, and there exists plenty of room for healthy speculation. One possibility is that established motion detection substrates, such as that known to exist in cortical visual area MT, are multifunctional and thus capable of operating on low-order signals (luminance, chrominance, texture) as well as on inputs representing feature salience. Recent experiments (34) demonstrating that motion sensitivity of neurons in cortical visual area MT can be markedly modulated by attention to a moving

target are consistent with this view. The salience map is a much larger and thornier problem, both because the mechanistic principles [as defined by Lu *et al.* (2)] are underconstrained and because physiological experiments are only beginning to approach the issue effectively (35). The issue is an important one, however, that has implications that extend far beyond the realm of motion processing (e.g., ref. 36) and is likely to be a major focus of future research in this area.

The persistent controversy over the relationship between color and motion processing is not likely to diminish with the new report of Lu *et al.* (2). On the contrary, it seems apt to increase the volume of the debate in certain areas, which is not at all a bad thing. Indeed, perhaps the greatest merit of the report is that it brings a stimulating perspective to a contentious field in which battle lines have become noticeably stagnant and self-serving. More generally, it further reinforces the belief that sensory elements interact in many complex ways to yield perceptual experience.

1. Livingstone, M. S. & Hubel, D. H. (1987) *J. Neurosci.* **7**, 3416–3468.
2. Lu, Z.-L., Lesmes, L. A. & Sperling, G. (1999) *Proc. Natl. Acad. Sci. USA* **96**, 8289–8294.
3. Mollon, J. D. (1989) *J. Exp. Biol.* **146**, 21–38.
4. Zeki, S. M. (1978) *Nature (London)* **274**, 423–428.
5. Ramachandran, V. S. & Gregory, R. L. (1978) *Nature (London)* **275**, 55–56.
6. Cavanagh, P., Tyler, C. W. & Favareau, O. E. (1984) *J. Opt. Soc. Am. A* **1**, 893–899.
7. Livingstone, M. & Hubel, D. (1988) *Science* **240**, 740–749.
8. Saito, H., Tanaka, K., Isono, H., Yasuda, M. & Mikami, A. (1989) *Exp. Brain Res.* **75**, 1–14.
9. Cavanagh, P. & Anstis, S. (1991) *Vision Res.* **31**, 2109–2148.
10. Papathomas, T. V., Gorea, A. & Julesz, B. (1991) *Vision Res.* **31**, 1883–1892.
11. Dobkins, K. R. & Albright, T. D. (1993) *Vision Res.* **33**, 1019–1036.
12. Dobkins, K. R. & Albright, T. D. (1994) *J. Neurosci.* **14**, 4854–4870.
13. Dobkins, K. R. & Albright, T. D. (1995) *Vision Neurosci.* **12**, 321–332.
14. Palmer, J., Mobley, L. A. & Teller, D. Y. (1993) *J. Opt. Soc. Am. A* **10**, 1353–1362.
15. Gegenfurtner, K. R., Kiper, D. C., Beusmans, J. M., Carandini, M., Zaidi, Q. & Movshon, J. A. (1994) *Vision Neurosci.* **11**, 455–466.
16. Hawken, M. J., Gegenfurtner, K. R. & Tang, C. (1994) *Nature (London)* **367**, 268–270.
17. Stromeyer, C. F., III, Kronauer, R. E., Ryu, A., Chaparro, A. & Eskew, R. T., Jr. (1995) *J. Physiol. (London)* **485**, 221–243.
18. Sawatari, A. & Callaway, E. M. (1996) *Nature (London)* **380**, 442–446.
19. Cropper, S. J. & Derrington, A. M. (1996) *Nature (London)* **379**, 72–74.
20. Thiele, A., Dobkins, K. R. & Albright, T. D. (1999) *J. Neurosci.*, in press.
21. Dobkins, K. R., Stoner, G. R. & Albright, T. D. (1998) *Perception* **27**, 681–709.
22. Lu, J. L. & Sperling, G. (1995) *Vision Res.* **35**, 2697–2722.
23. Albright, T. D. (1992) *Science* **255**, 1141–1143.
24. Albright, T. D. (1993) *Rev. Oculomot Res* **5**, 177–201.
25. Zhou, Y. X. & Baker, C. L., Jr. (1993) *Science* **261**, 98–101.
26. Mareschal, I. & Baker, C. L., Jr. (1998) *Nat. Neurosci.* **1**, 150–154.
27. Reichardt, W. (1961) in *Sensory Communication*, ed. Rosenblith, W. A. (MIT Press, Cambridge, MA), pp. 303–317.
28. Adelson, E. H. & Bergen, J. R. (1985) *J. Opt. Soc. Am. A* **2**, 284–299.
29. van Santen, J. P. & Sperling, G. (1985) *J. Opt. Soc. Am. A* **2**, 300–321.
30. Heeger, D. J. (1987) *J. Opt. Soc. Am. A* **4**, 1455–1471.
31. Chubb, C. & Sperling, G. (1988) *J. Opt. Soc. Am. A* **5**, 1986–2007.
32. Cavanagh, P. (1992) *Science* **257**, 1563–1565.
33. Cavanagh, P. & Mather, G. (1989) *Spatial Vision* **4**, 103–129.
34. Treue, S. & Maunsell, J. H. (1996) *Nature (London)* **382**, 539–541.
35. Gottlieb, J. P., Kusunoki, M. & Goldberg, M. E. (1998) *Nature (London)* **391**, 481–484.
36. Koch, C. & Ullman, S. (1985) *Hum. Neurobiol.* **4**, 219–227.