

## The influence of figural interpretation on the selective integration of visual motion signals

Thomas D. Albright and Gene R. Stoner

The Salk Institute for Biological Studies  
La Jolla, CA 92037

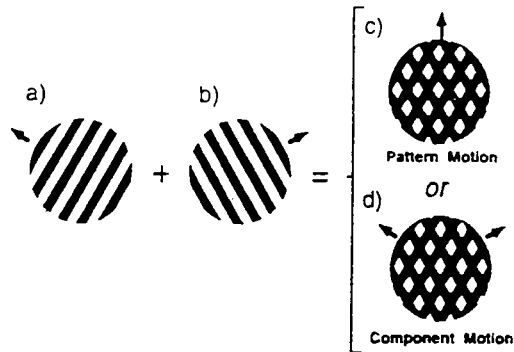
### ABSTRACT

The solution to the computational problem of reconstructing object motion from retinal image motion is underconstrained. In an effort to converge on a solution to this problem, the primate visual system appears to rely upon image cues that lead to an interpretation of the spatial relationships between objects in a visual scene. Psychophysical experiments illustrate this phenomenon through the apparent dependence of motion signal integration on luminance-based cues for occlusion and perceptual transparency. Neurophysiological studies of the cell populations thought to underlie motion signal integration reveal a change in directional selectivity that precisely parallels the perceptual phenomenon. Among obstacles faced in attempts to understand the neural bases of primate vision, the integration of motion signals holds a unique position: The computational problem is well-defined, a specific neural substrate has been identified, and the solution to the integration problem is absolutely critical for visually-guided behavior. As such, it stands as a model system for exploring the relationships between neuronal phenomena, perception, and behavior.

### 1. MOTION SIGNAL INTEGRATION

The motions of objects in the world often give rise to a complex pattern of moving and overlapping features in the retinal image. From such intangibles it is clearly possible for the primate visual system to construct a veridical representation of moving objects. Because the solution is otherwise grossly underconstrained, we have proposed that this process relies upon tacit knowledge of the "rules" by which two-dimensional (2D) retinal image features are formed from their real-world 3D counterparts<sup>1,2</sup>. Such information is essential for perceptual interpretation of the spatial relationships between moving image features, which in turn allows moving features to be integrated according to object of origin.

This hypothesis regarding the integration of visual motion signals can be readily tested in psychophysical and neurophysiological experiments using stimuli that have been termed "moving plaid patterns"<sup>3,4</sup>. These 2D patterns are formed, as illustrated in Figure 1, by superimposition of two overlapping and drifting 1D gratings. Plaids provide a simple laboratory counterpart to real-world situations that give rise to overlapping contours in the retinal image. Their value in this context comes from the fact that under some conditions the two grating components are seen to move independently or "non-coherently", while under other conditions the two components are seen to form part of a single 2D pattern that moves "coherently". By manipulating various image parameters it becomes possible to identify the conditions that lead to these two different percepts. Our hypothesis predicts that these conditions will correspond to those that influence perceptual parsing of the image into two surfaces vs. one.



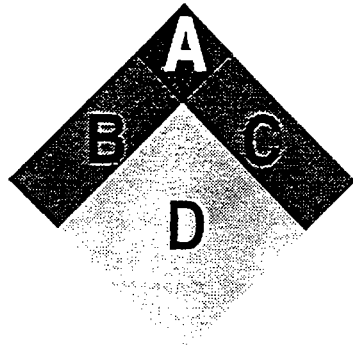
**Figure 1:** Moving plaid patterns are produced by superimposition of two drifting periodic grating. The resultant percept is either that of a coherently moving two-dimensional plaid pattern or two one-dimensional gratings sliding past one another, depending on a variety of stimulus parameters.

## 2. PSYCHOPHYSICAL STUDIES OF MOTION SIGNAL INTEGRATION

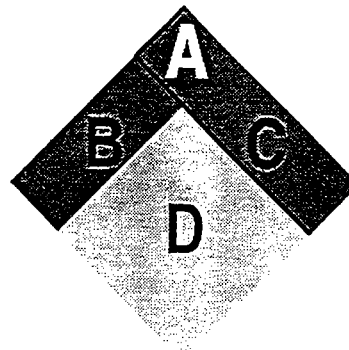
Using drifting plaid patterns as stimuli, it was originally shown that the likelihood of perceptual motion coherence decreases when the component gratings differ significantly along the dimensions of spatial frequency or luminance contrast<sup>3</sup>. Subsequent psychophysical experiments demonstrated that components having different binocular disparities (thus appearing to lie in different depth planes) are also less likely to cohere<sup>5</sup>, as are those created by modulation along different color-opponent axes<sup>6</sup>. These observations have typically been explained by invoking relatively simple channel-based mechanisms for selective integration<sup>7</sup>. The observations are nonetheless consistent with our functional proposal, whereby motion signal integration hinges upon determination of the figural origins of moving image features.

To explore this possibility more extensively, we chose to manipulate luminance cues that directly influence perceived depth ordering of surfaces<sup>1</sup>. The appropriateness of these cues (and their ubiquity in natural images) can be evaluated by considering how retinal images are formed. It is often the case that when one moving object passes in front of another, the nearer object occludes the distant object. At the point of overlap the luminance may be exclusively that of the nearer surface. In other instances, a *transparent* foreground object may attenuate, but not occlude, light from the distant surface. A special case of such attenuation is that characteristic of shadows. Bearing in mind these properties of image formation, there are some elementary luminance relationships that dictate whether simple patterns, such as those shown in Figure 2, are physically consistent with two overlapping surfaces or four distinct surfaces. These luminance relationships are captured by the "rules" of perceptual transparency<sup>8,9</sup>. Accordingly, luminance relationships falling between the extremes of occlusion and shadow-like transparency can arise from independent but spatially overlapping surfaces. Luminance variations occurring at the region where two independent surfaces intersect (region A in Figure 2) are considered "extrinsic"<sup>10</sup>, i.e., a consequence of the figural relationships between surfaces rather than a property of the surfaces themselves. Luminance variations that lie outside of this occlusion-transparency range cannot be attributed to object interrelationships and must, therefore, result from "intrinsic" surface properties, such as differences in surface reflectance. In accordance with our hypothesis, we predicted that the figural relationships implied by these luminance variations would gate the selective integration of motion signals.

"PURE" TRANSPARENCY

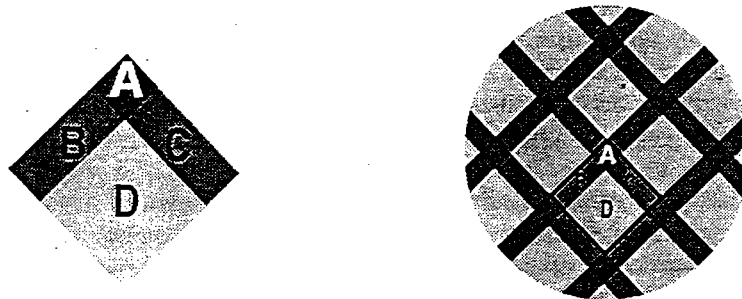


"PURE" OCCLUSION

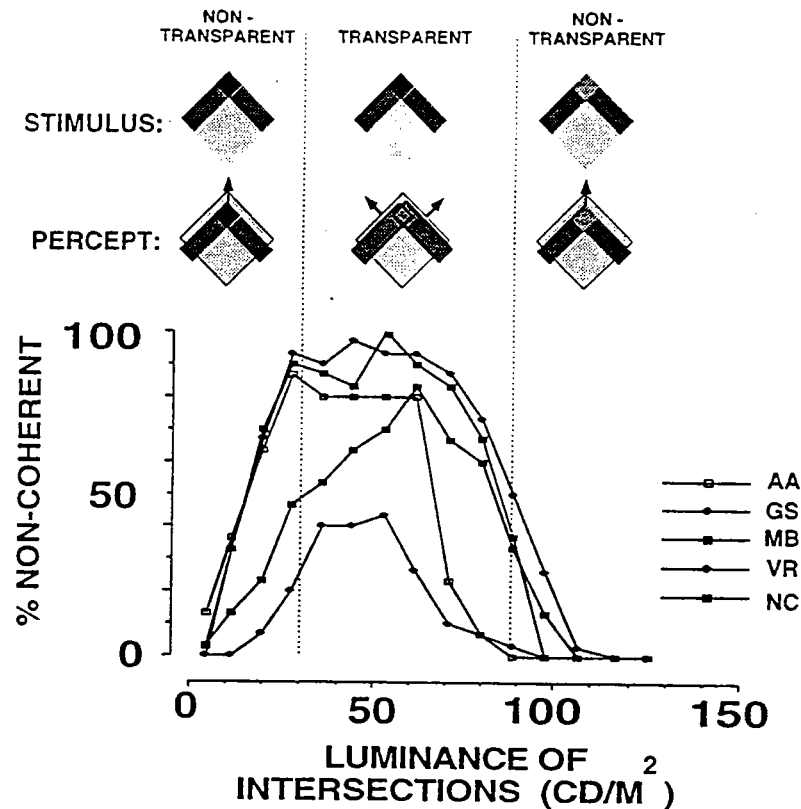


**Figure 2:** The procedures for creating perceptually transparent plaid patterns are derived from the physics of transparency<sup>8,9</sup>. Simply put, luminance ratios within the pattern must be physically consistent with the transmittance of light from a far surface *through* a near surface. Appropriate luminance ratios convey a sense of depth ordering (and hence image segmentation) in a pattern devoid of other depth cues. The zone of perceptual transparency is bounded by two extremes: "pure" transparency and "pure" occlusion. Pure transparency (**left**) occurs when the near transparent surface AC reflects no light but transmits light from the surface behind it. Pure occlusion (**right**) occurs when the near surface AC reflects light but transmits no light from the surface behind it. From Stoner and Albright<sup>2</sup>.

This prediction was tested using plaid patterns constructed with reference to the laws of perceptual transparency (Figure 3). As predicted, when the luminance relationships were configured such that the component gratings appeared occlusive or transparent, human subjects generally reported a percept of non-coherent motion (i.e., the two gratings appeared to slide across one another) (Figure 4). Alternatively, when the luminance configuration was incompatible with transparency or occlusion, subjects generally reported a percept of coherent motion<sup>1</sup>.



**Figure 3:** Each plaid can be viewed as a tessellated image composed of four distinct repeating subregions, identified as A, B, C, and D. Region D is normally seen as background. Regions B and C are seen as narrow overlapping surfaces, and the remaining region A is seen as their intersection. Perceptual transparency was manipulated in both psychophysical<sup>1</sup> and neurophysiological<sup>17</sup> experiments by adjusting the luminance of region A, while the luminances of regions B, C, and D were held constant.



**Figure 4:** Results from psychophysical experiments examining the effects of perceptual transparency on motion coherency. Probability of the component motion percept is plotted as a function of the intersection luminance for appropriately configured plaid patterns (see Figures 2 and 3). Both gratings were of the same spatial frequency (1.75 cyc/°). On each trial the individual gratings were moved at an angle of 135° relative to one another at a speed of 3°/s, resulting in a pattern speed of 8°/s. Pattern direction was either up or down, and varied on a random schedule. Intersection luminance was varied in equal steps, such that it was either compatible or incompatible with transparency. The "transparency zone" extends from pure (multiplicative) transparency (35 cd/m<sup>2</sup>) up to the point of occlusion (90 cd/m<sup>2</sup>). A percept of non-coherent component motion is most likely within a region roughly centered on the transparency zone. Each data point represents the mean of 30 trials for each intersection luminance value. Data are shown for five subjects. Adapted from Stoner et al.<sup>1</sup>.

These results support our general hypothesis regarding the contribution of image segmentation cues to motion signal integration. They tell us little about the mechanism involved, however. One common proposal<sup>2,11,12</sup> is based upon the fact that the luminance manipulations involved in simulating transparency and occlusion also vary the strength of fourier components that move in the coherent direction. Indeed, if one allows for an early logarithmic signal compression<sup>13</sup>, the resultant strength of such phantom fourier components roughly accounts for the results of Stoner et al.<sup>1</sup>. This low-level explanation is called into question, however, by the results of other recent experiments, which show that the perception of transparency -- and, in turn, motion coherence -- is also dependent upon image cues that do not affect the strength of phantom fourier components. For example, perceptual transparency is inherently dependent upon figural cues that influence the visual system's ability to interpret the relationship between foreground and background in a visual scene. The percept of transparency in the left panel of Figure 2 relies upon the fact that the observer interprets surface *AC* as foreground and surface *BD* as the unattenuated background. Clearly, the converse figural interpretation (which can be willed with some effort) does not lead to a percept of transparency. The explanation for this phenomenon is rooted in the fact that transparent surfaces typically do not

enhance the contrast of surfaces seen through them. Hence, the luminance variations in surface *BD* cannot arise solely by virtue of it being transparent and overlying surface *AC*.

To further explore this phenomenon and examine its influence over motion signal integration, Stoner and Albright<sup>14</sup> used pictorial cues to bias foreground interpretation. Two methods were used (Figure 5). Human subjects viewed plaid patterns for which foreground/background interpretation was manipulated and, as in the original transparency experiments, they reported whether they perceived coherent or non-coherent motion. Only the luminance of one region of the pattern was varied (region *A* in Figure 5). Using the means indicated, foreground/background interpretation was manipulated such that region *A* was likely to be perceived as either (1) the intersection of the two component gratings, or (2) the unobstructed background. As expected, both the percept of transparency and motion coherence were heavily dependent upon foreground/background interpretation, as manipulated by either technique. This result is not compatible with explanations based upon simple detection of the motions of phantom Fourier components. Although the details of an alternative mechanism have yet to be worked out, the result adds further weight to our claim that motion integration has access to image segmentation processes built upon the rules governing retinal image formation from natural scenes.

### 3. NEUROPHYSIOLOGICAL STUDIES OF MOTION SIGNAL INTEGRATION

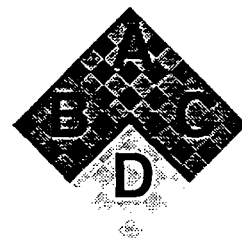
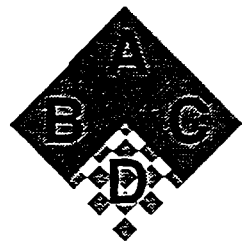
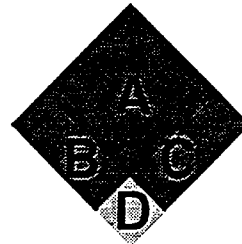
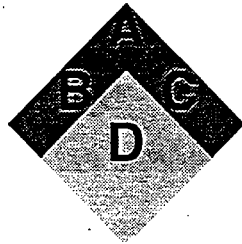
The conceptual framework described above implies the existence of at least two stages of motion processing in the primate brain, which provide the roles of motion detection and integration, respectively. Neurophysiological studies employing plaid patterns as visual stimuli have allowed a tentative identification of the neuronal populations corresponding to these two stages. Movshon and colleagues<sup>7</sup> examined the directional selectivity of V1 neurons to perceptually coherent plaid patterns. Consistent with their orientation tuning, V1 neurons were found to signal only the motion of the 1D components. Such neurons have been referred to as "component type" and are believed to represent the first motion processing stage. The integration process appears to take place in the middle temporal visual area (area MT), an area that receives direct input from V1 and is thought to play a crucial role in motion processing<sup>15</sup>. While many MT neurons (40%) also appear to be component type when tested under these conditions, a small population (25%) respond in a way that reflects sensitivity to pattern motion<sup>7,16</sup>. Neurons of this latter type have been called "pattern type" and are presumed to constitute the second stage of motion processing, at which motion signal integration takes place (Figure 6).

Although the factors affecting motion signal integration have been studied in some detail psychophysically, until recently virtually nothing was known of the neural interactions underlying these effects. Previous studies that attempted to classify directionally selective neurons on the basis of responses to component or pattern motion used plaid patterns that were *always perceptually coherent*. Since these neurons are believed to play some significant role in the integration process, we hypothesized that their behavior would be altered by stimulus attributes known to influence perceptual integration of motion signals. As in our earlier psychophysical experiments described above<sup>1</sup>, perceptual motion coherence was manipulated by altering the stimulus conditions such that plaid patterns were either consistent or inconsistent with transparency. Directional selectivity of single MT neurons was assessed using each of three different plaid configurations<sup>17</sup>. Two of these stimuli elicited a percept of coherent pattern motion; the third elicited a percept of independently moving components. Data obtained from a typical neuron are illustrated in Figure 7. When stimulated using either of the perceptually coherent plaids, this cell responded more strongly when the 2D pattern moved in the neurons's preferred direction than when either of the 1D components moved in that same direction. As can be seen in Figure 6, this type of tuning is characteristic of pattern type neurons<sup>7,16</sup>. When stimulated using the transparent and perceptually non-coherent plaid, however, this cell's behavior underwent a marked transformation: the pattern response dropped while component responses became elevated. The resultant bilobed directional tuning curve is characteristic of component type neurons (Figure 6). As was the case for the majority of neurons in our sample, this cell's sensitivity to component motion increased (and sensitivity to pattern motion decreased) when the visual stimulus was configured such that a percept of component motion became more likely.

LUMINANCE CONTRAST VIEWED THROUGH  
 TRANSPARENT FOREGROUND SURFACE [AC] SHOULD BE  
 LESS THAN THAT OF BACKGROUND [BD]

PERCEPTUALLY  
 TRANSPARENT

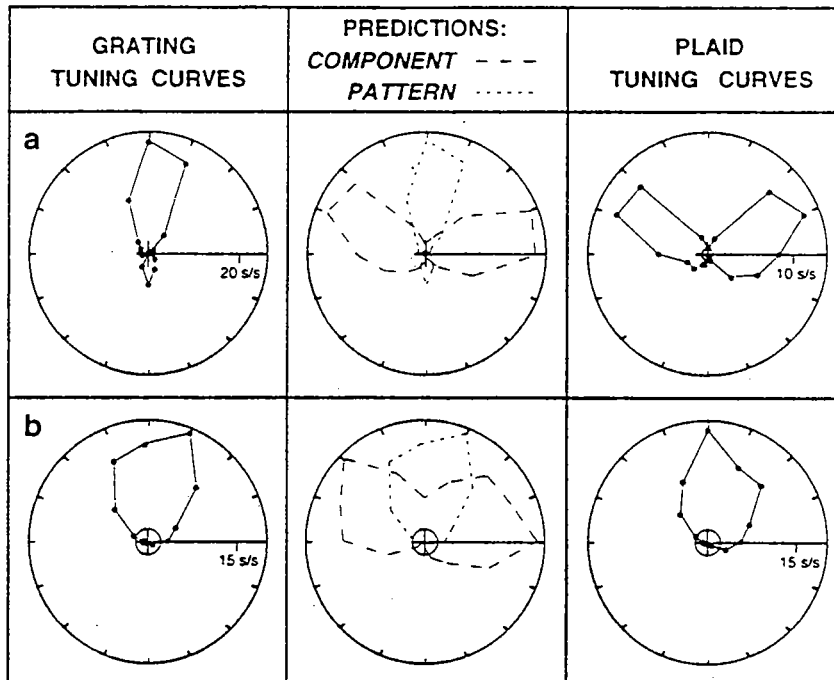
PERCEPTUALLY  
 NON-TRANSPARENT



FOREGROUND CONTRAST (A/C)  
 IS LESS THAN  
 BACKGROUND CONTRAST (B/D)

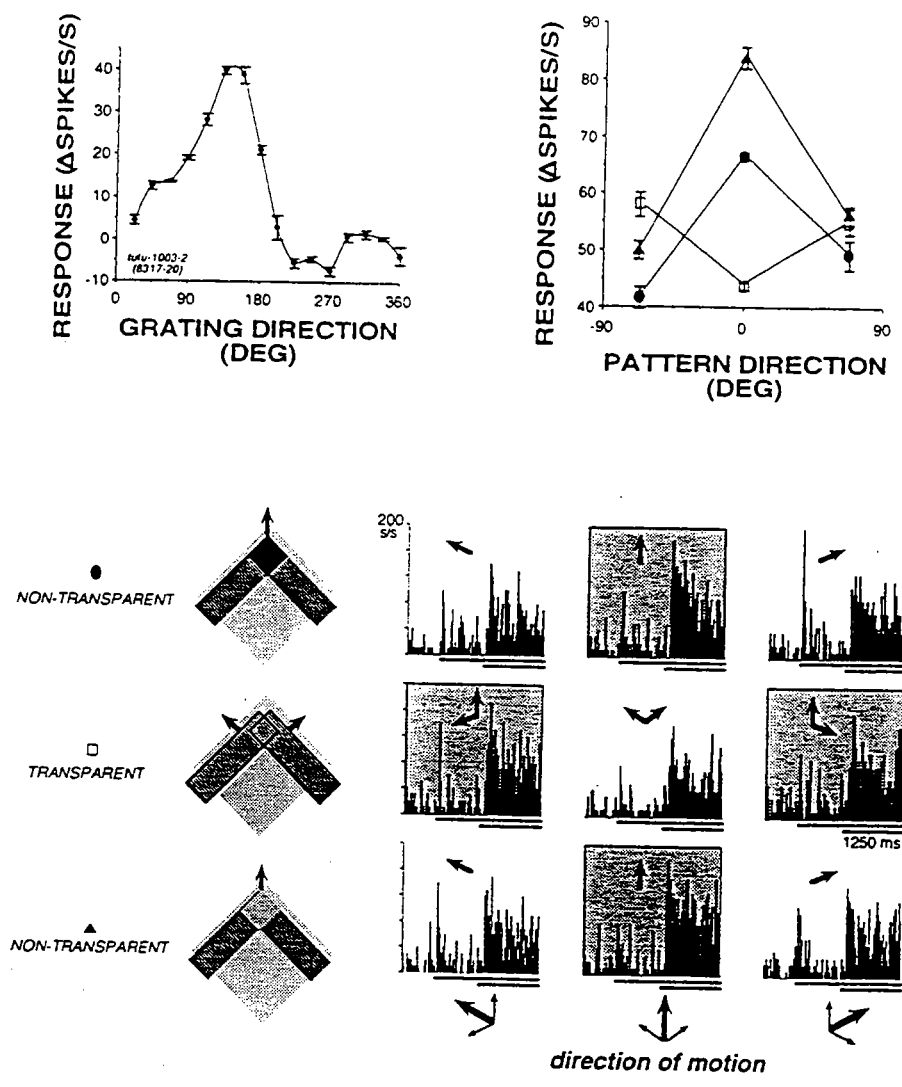
FOREGROUND CONTRAST (B/D)  
 IS GREATER THAN  
 BACKGROUND CONTRAST (A/C)

**Figure 5:** Schematic illustration of two methods used by Stoner and Albright<sup>14</sup> to manipulate the perception of foreground and background in plaid patterns. The plaids were tessellated versions of these basic patterns (see Figure 3). One method, depicted in the **top row**, involved manipulating the relative sizes of the plaid sub-regions. The larger regions (region *D* on the **left** and region *A* on the **right**) are usually seen as background<sup>18,19</sup>. A second method, shown in the **bottom row**, was to place a static checkerboard in the putative background region. The plaid motion progressively occluded/disoccluded this pattern, causing the textured region (region *D* on the **left** and region *A* on the **right**) to be seen as background (note that the "occlusion" of this checkerboard by the putatively transparent gratings is physically consistent with the contrast reduction associated with transparent surfaces or the blurring common with translucent surfaces such as smoked glass). Both methods reliably influence perceptual assignment of foreground/background, while leaving the space-averaged luminance of the four regions constant. The reversal of foreground/background assignment, in turn, had a profound effect on both perceptual transparency and motion coherency judgements by human observers. From Stoner and Albright<sup>14</sup>.



**Figure 6:** Data from two MT neurons representing "component" (*a* row) and "pattern" (*b* row) stages of motion processing in cortical visual area MT of the rhesus monkey. Direction tuning curves were acquired using a drifting sine-wave grating (**first column**) or a perceptually coherent plaid pattern (**third column**). Responses elicited by each stimulus type, moving in each of 16 different directions, are plotted in a polar format. The radial axis represents response amplitude (s/s = mean spike rate during presentation of the stimulus within the receptive field), the polar angle corresponds to the direction of motion, and the small circle in the center of each polar plot represents the level of spontaneous activity. Both cells exhibit a single peak in the grating tuning curve. From these curves, responses to the moving plaid pattern were predicted in accordance with either component or pattern assumptions (**second column**). The component predictions reflect sensitivity to both oriented components in the plaid pattern. The pattern predictions reflect sensitivity to the composite appearance of the plaid. By definition, the behavior of the component motion neuron conforms to the component prediction while that of the pattern motion neuron conforms to the pattern prediction. Adapted from Rodman and Albright<sup>16</sup>.

These neurophysiological data demonstrate that visual stimulus conditions that influence image segmentation and perceptual motion coherence, also lead to systematic differences in the directional tuning of the neurons that are believed to underlie motion signal integration. Although the neuronal circuits responsible for this perceptually-dependent gating of directional selectivity are entirely unknown, these new results suggest that image segmentation signals might interact with motion signals at the integration stage.



**Figure 7:** Neural correlates of perceptual motion signal integration. **A:** Differential responses of an MT pattern-type neuron to coherent vs. non-coherent plaid patterns. Directional tuning for a single drifting grating is plotted at left. Responses to coherent and non-coherent plaids are plotted at center. When stimulated with coherent plaid patterns, response was maximal when the pattern moved in the neuron's preferred direction (0° in graph at center; highlighted histograms in top and bottom rows at right). However, when stimulated with non-coherent plaid patterns, responses were maximal when either component moved in the preferred direction ( $\pm 67.5^\circ$  in graph at center; highlighted histograms in center row at right). Adapted from Stoner and Albright<sup>17</sup>.

#### 4. ACKNOWLEDGEMENTS

The research reported herein was supported in part by a grant from the National Institute of Mental Health and a Development Award from the McKnight Endowment fund for Neuroscience. Gene Stoner was partially supported by a Research Fellowship from the McDonnell-Pew Center for Cognitive Neuroscience at San Diego. We thank J. Costanza for superb technical assistance.



## 5. REFERENCES

1. Stoner, G.R., Albright, T. D., and Ramachandran, V. S. Transparency and coherence in human motion perception. *Nature*, **344**: 153-155, 1990.
2. Stoner, G.R. and Albright, T.D. Image segmentation cues in motion processing: Implications for modularity in vision. *J. Cog. Neurosci.*, **5**(2):129-149, 1993.
3. Adelson, E.H. and Movshon, J.A. Phenomenal coherence of moving visual patterns. *Nature*. **300**:523-525, 1982.
4. De Valois, K.K., De Valois, R.L., and Yund, E.W. Responses of striate cortex cells to grating and checkerboard patterns. *J. Physiol. (Lond.)*, **291**:483-505, 1979.
5. Adelson, E.H. Binocular disparity and the computation of two-dimensional motion. *J. Opt. Soc. Am. A*, **1**:1266, 1984.
6. Krauskopf, J. and Farell, B. Influence of colour on the perception of coherent motion. *Nature*. **348**:328-331, 1990.
7. Movshon, J.A., Adelson, E.A., Gizzi, M. and Newsome, W.T. The analysis of moving visual patterns. In (Eds.) C. Chagas, R. Gattass and C. G. Gross. *Study Group on Pattern Recognition Mechanisms*, Pontificia Academia Scientiarum: Vatican City pp. 117-151, 1985.
8. Metelli, F. The perception of transparency. *Scientific American*, **230**: 91-95, 1974.
9. Beck, J., Prazdny, K. and Ivry, R. The perception of transparency with achromatic colors. *Percept. Psychophysics*. **35**:407-422, 1984.
10. Shimojo, S., Silverman, G.H. and Nakayama, K. Occlusion and the solution to the aperture problem for motion. *Vision Res.* **29**:619-626, 1989.
11. Mulligan, J.B. Non-linear combination rules and the perception of visual motion transparency. *Vision Res.* **33**:2021-2030, 1992.
12. Trueswell, J.C., and Hayhoe, M.M. Surface segmentation mechanisms and motion perception. *Vision Res.* **33**:313-328, 1993.
13. MacLeod, D.I.A. Visual sensitivity. *Annual Review of Psychology*. **29**:613-645, 1978.
14. Stoner G.R. and Albright T.D. The influence of foreground/background assignment on transparency and motion coherence in plaid patterns. In preparation, 1994.
15. Albright, T.D. Cortical processing of visual motion. In F.A., Miles and J., Wallman (Eds.), *Visual Motion and its Role in the Stabilization of Gaze*. Amsterdam: Elsevier, pp.177-201, 1993.
16. Rodman, H.R. and Albright, T.D. Single-unit analysis of pattern-motion selective properties in the middle temporal visual area (MT). *Exp. Brain Res.* **75**:53-64, 1989.
17. Stoner, G.R. and Albright, T.D. Neural correlates of perceptual motion coherence. *Nature*. **358**:412-414, 1992.
18. Petter, G. Nuove ricerche sperimentali sulla totalizzazione percettiva. *Rivista di psicologia*, **50**: 213-227, 1956.
19. Vallortigara, G. and Bressan, P. Occlusion and the perception of coherent motion. *Vision Res.* **31**:1967-1978, 1991.