

Neural correlates of perceptual motion coherence

Gene R. Stoner & Thomas D. Albright

Vision Center Laboratory, The Salk Institute for Biological Studies, La Jolla, California 92037, USA

THE motions of overlapping contours in a visual scene may arise from the physical motion(s) of either a single or multiple surface(s). A central problem facing the visual motion system is that of assigning the most likely interpretation. The rules underlying this perceptual decision can be explored using a visual stimulus formed by superimposing two moving gratings. The resultant percept is either that of a single coherently moving 'plaid pattern' (coherent motion) or of the two component gratings sliding noncoherently across one another (noncoherent motion)^{1,2}. When plaid patterns are configured to mimic one transparent grating overlying another, the percept of noncoherent motion dominates³. We now report that neurons in the visual cortex of rhesus monkeys exhibit changes in direction tuning that parallel this perceptual phenomenon: sensitivity to the motions of the component gratings is enhanced under conditions that favour the perception of noncoherent motion. These results challenge models of cortical visual processing that fail to take into account the contribution of figural image segmentation cues to the analysis of visual motion.

The perceptual phenomenon of motion coherence may be mediated by a subset of neurons in the middle temporal visual area (MT) of the primate cerebral cortex^{2,4}. In rhesus monkeys, these directionally selective 'pattern neurons' are distinguished by their ability to signal correctly the motion of a perceptually coherent plaid pattern^{2,4}. Directionally selective neurons in primary visual cortex (V1) and the majority of MT neurons ('component neurons'), by contrast, signal only the motions of the component gratings in a perceptually coherent plaid pattern^{2,4}. To investigate the contribution of area MT to visual motion coherence, we examined whether the responses of MT neurons can be modified by the same factors known to influence this perceptual decision. We used the phenomenon of perceptual transparency as a means to generate both coherent and noncoherent plaid patterns³ (Fig. 1). We studied 105 MT neurons in two rhesus monkeys. Directional tuning was first assessed using a drifting grating. The peak in this tuning curve defined the preferred direction for each cell. Directional selectivity was then examined using each of three different plaid configurations. Two of these stimuli were nontransparent and judged to move coherently by human observers³. The third plaid stimulus was transparent and judged to move noncoherently by human observers³.

Data obtained from a typical neuron are shown in Fig. 2a. When presented with either of the perceptually coherent plaids, this cell responded more strongly when the pattern moved in

the preferred direction (the 'pattern response') than when either of the components moved in this same direction (the 'component responses'). This type of directional selectivity (pattern response larger than component responses) is characteristic of pattern neurons^{2,4} and reflects integration of locally derived motion signals. When presented with the perceptually noncoherent plaid, however, this neuron's behaviour changed: the pattern response decreased by 42% and component responses were elevated slightly (by an average of 15%). These changes resulted in a type of direction tuning (component responses larger than pattern response) that is normally characteristic of component neurons in areas MT and V1 (refs. 2, 4). Hence the cell's responses became more component-like when the stimulus was configured to render noncoherent motion as the dominant percept.

Changes in directional selectivity that accompany perceptual changes were not restricted to neurons classified as pattern type. This can be illustrated effectively by subtracting responses obtained using coherent plaids from those obtained using noncoherent plaids. Difference curves of this sort are plotted in Fig. 2b for three additional MT neurons, which fell into component, pattern, and 'unclassifiable' categories using established criteria^{2,4}. Despite different classifications, all three neurons exhibited enhanced component responses and diminished pattern responses when stimulated with noncoherent plaids.

If these coherence-related changes in MT are to account for the perceptual phenomenon, then the responses of the neural population as a whole should become more component-like and less pattern-like in the presence of visual stimuli that are more apt to yield a percept of noncoherent motion. To test this prediction, we quantified coherence-related changes in directional tuning using an extension of a procedure introduced by Movshon *et al.*² (see legend to Figure 3). Briefly, predictions of both component- and pattern-type direction tuning to plaids were formed for each cell on the basis of responses to drifting gratings. The degree of similarity between observed plaid tuning curves and each of the component and pattern predictions was quantified by computing a measure of the variance accounted for by each prediction. Finally, to acquire a single index of how 'component-like' a neuron's tuning was for each stimulus condition, we computed the difference between the variance accounted for by the two predictions: the larger the value of this 'component index', the more component-like the response. The range and extent of coherence-related changes across our sample can be inspected by plotting the component index computed for each neuron under noncoherent conditions as a function of that obtained under each of the coherent conditions (Fig. 3a, b). Points falling above the diagonals in these plots reflect a shift toward component-like behaviour under stimulus conditions that usually yield a percept of component motion. Although there were varying degrees of coherence-related shifts, pairwise comparisons of the distributions revealed the component index to be systematically larger for the noncoherent condition than for either of the two coherent conditions (one-tailed paired *t*-test, $P < 0.001$).

The selective integration of locally derived motion signals is a perceptual phenomenon of profound behavioural significance. Retinal images with locally identical motions can render very different motion percepts, depending upon how the visual scene is parsed. Consider, for example, the dynamic retinal image elicited by the movements of a primate through a forest canopy. Moving image contours arise from a variety of real object boundaries as well as a combination of variations in surface reflectance, shadows, and occlusion. Accurate assessment of the motion of any of the objects in this scene (the primate, the swaying foliage, or perhaps the shadows cast by either) depends not only on integrating component motion signals but also on integrating only those signals common to each moving object. Our results suggest that the visual motion system accomplishes

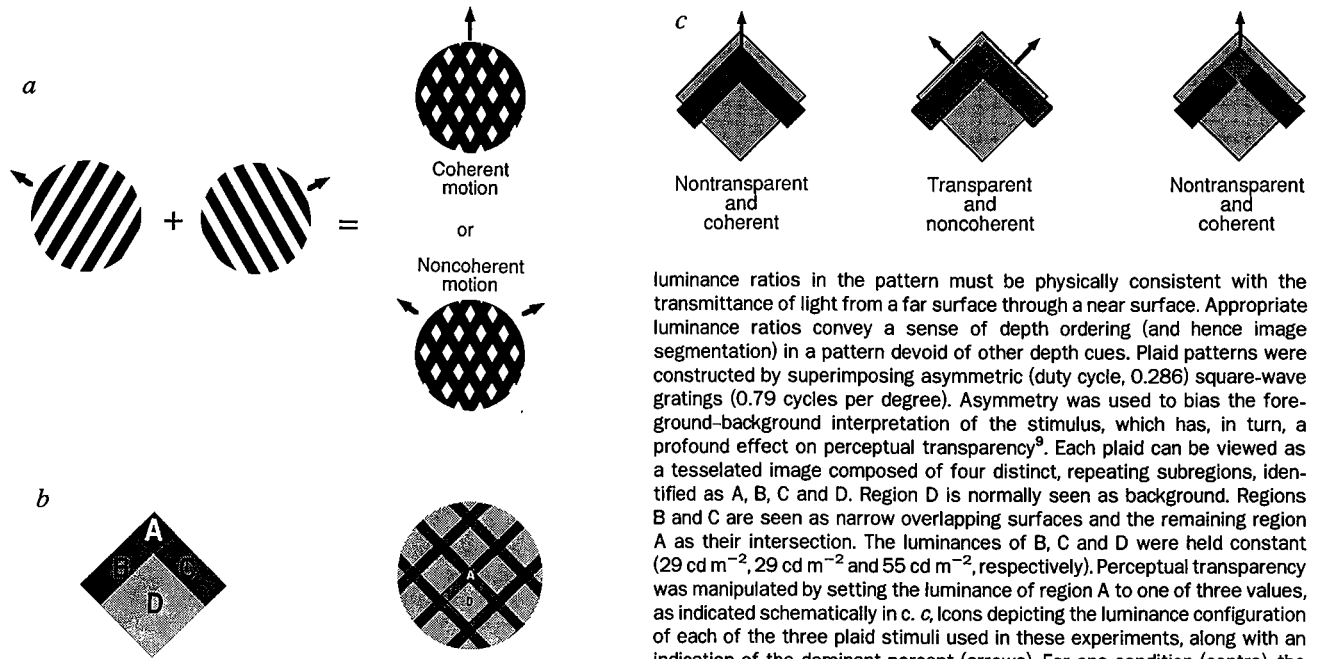


FIG. 1 *a*, Moving plaid patterns are produced by superposition of two drifting periodic gratings. The resultant pattern can be perceived to move either coherently or noncoherently depending on a variety of stimulus parameters^{1-3,5-9}. *b*, The procedures for creating perceptually transparent plaid patterns are derived from the physics of transparency^{10,11}. Simply put,

luminance ratios in 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. Plaid patterns were constructed by superimposing asymmetric (duty cycle, 0.286) square-wave gratings (0.79 cycles per degree). Asymmetry was used to bias the foreground-background interpretation of the stimulus, which has, in turn, a profound effect on perceptual transparency⁹. 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 as their intersection. The luminances of B, C and D were held constant (29 cd m^{-2} , 29 cd m^{-2} and 55 cd m^{-2} , respectively). Perceptual transparency was manipulated by setting the luminance of region A to one of three values, as indicated schematically in *c*. *c*, Icons depicting the luminance configuration of each of the three plaid stimuli used in these experiments, along with an indication of the dominant percept (arrows). For one condition (centre), the luminance of region A was chosen to be consistent with transparency (19 cd m^{-2}), yielding a percept of noncoherent motion in human observers³. For the remaining two conditions, the luminance of region A was either too dark (left, 3 cd m^{-2}) or too bright (right, 34 cd m^{-2}) to be compatible with transparency. These nontransparent plaids generally elicit a percept of coherent motion in human observers³. Other procedures for stimulus generation were as before³.

FIG. 2 Neural correlates of perceptual motion signal integration. *a*, Differential responses of an MT pattern-type neuron to coherent against noncoherent plaid patterns. Left, Directional tuning for a single drifting grating. Right, Responses to coherent and noncoherent plaids. Response amplitude ($\Delta \text{ spikes s}^{-1}$) was computed as the mean spike rate in the period of stimulus presentation minus the average baseline spike rate from all prestimulus periods. Error bars indicate s.e.m. response. When stimulated with either of the two types of perceptually coherent plaid patterns (filled circles, intersection region A too dark for transparency; open squares, intersection region A too bright for transparency; Fig. 1), response was maximal when the pattern moved in the neuron's preferred direction (0° in graph at right). But when stimulated with perceptually noncoherent plaid patterns (triangles), responses were maximal when either component moved in the preferred direction ($\pm 67.5^\circ$ in graph at right). *b*, Data from three additional MT neurons showing 'difference curves' (bottom row), computed by subtracting tuning curves obtained with coherent and noncoherent plaids (open squares and filled triangles, respectively, in top row). Pattern direction (P) is defined as the peak in the grating curve (filled squares, top row) and component directions (C) are displaced (67.5°) to each side. All three conventional neuronal types (component, pattern and unclassified) exhibited enhanced component responses and attenuated pattern responses when stimulated with noncoherent against coherent plaids. Visual stimuli were presented in a 10.5° circular aperture centred on the receptive field of each neuron. Gratings presented singly were moved at 10.26 degrees per s. Gratings composing plaid patterns were moved at 3.93 degrees per s and in directions differing by 135° , resulting in a pattern speed of 10.26 degrees per s. Stimuli were presented for five trials of 1.25 s exposure in each direction on a random schedule. Animals were alert and fixating during data acquisition. Eye position was monitored using a scleral search coil. Data from trials in which fixation deviated more than 0.5° were rejected.

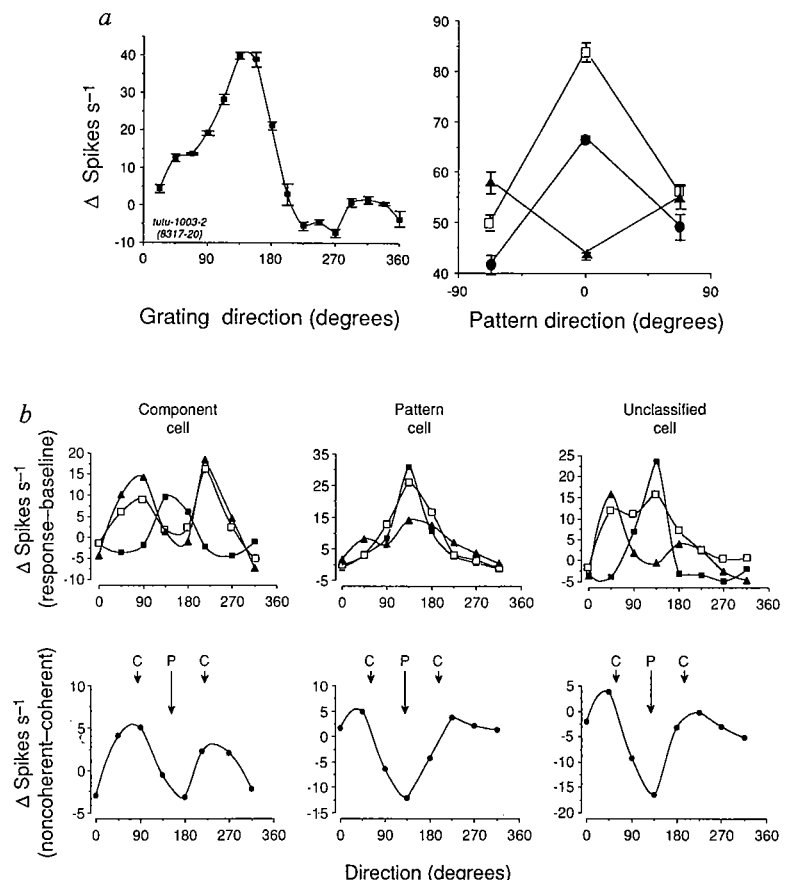


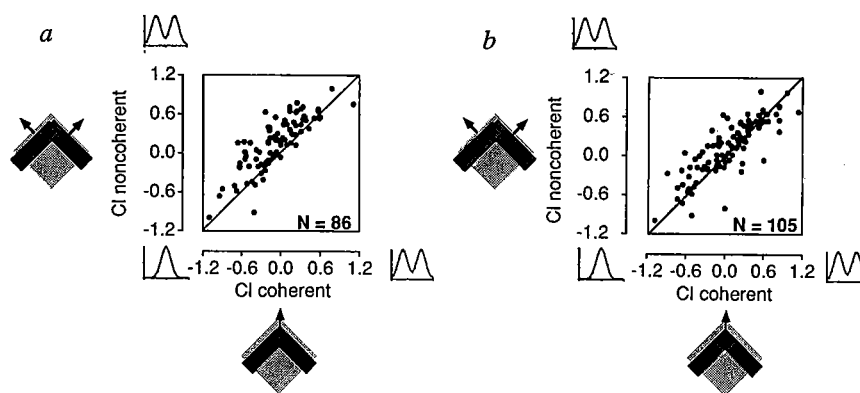
FIG. 3 Quantitative comparison of MT direction tuning for coherent against noncoherent plaid patterns. A 'component index' (CI) was computed by subtracting the variance accounted for by the pattern prediction from that accounted for by the component prediction. Each point in these graphs represents the value of the CI obtained for noncoherent (vertical axes) against one of the two coherent (horizontal axes) conditions. The larger the CI the more component-like (more bilobed, generally) the tuning curve. The graph at left compares CIs for the single noncoherent (transparent) case and one of the two coherent cases (nontransparent, region $A=34 \text{ cd m}^{-2}$; see Fig. 1). The graph at right compares CIs for the single noncoherent case and the other coherent case (region $A=3 \text{ cd m}^{-2}$; see Fig. 1). The tendency for points in both graphs to fall above the diagonal reflects a highly significant shift among our sample of MT neurons toward component-like responses in the presence of stimulus conditions that normally elicit a percept of noncoherent motion.

METHODS. The component index was computed using an extension of a procedure introduced by Movshon *et al.*² (see also ref. 4). Briefly, both a component and a pattern prediction were derived for each neuron from the observed responses to conventional drifting gratings. The actual direction tuning curves obtained using each plaid pattern were then compared with predicted component and pattern curves by computing the appropriate partial correlations as follows:

$$R_p = (r_p - r_c r_{pc}) / [(1 - r_c^2)(1 - r_{pc}^2)]^{0.5}$$

this selective integration by using segmentation cues, such as transparency, to classify motion signals according to physical origin. Furthermore, this selective integration is expressed in the behaviour of those neurons presumed to underlie perceptual integration of motion signals. Thus, an MT neuron that would have been classified as pattern-type using perceptually coherent plaids and established criteria^{2,4} will express component-type behaviour when stimulated with patterns that are more likely to result in a percept of noncoherent motion.

These results have two important implications. First, there is the implied correlation between neural and perceptual state, providing a first peek at the neural events underlying the selective integration of component motion signals. Although we have emphasized coherence-related shifts in the sample distribution, it remains to be seen whether behavioural reports of motion coherence are correlated on a trial-by-trial basis with the neuronal response from a single cell. Second, our data, together with mounting psychophysical evidence bearing on the factors that influence motion signal integration^{1-3,5-9}, demonstrate that figural aspects of a visual image unrelated to motion *per se* (such as perceptual transparency) have profound modulatory effects on the processing of motion signals in the primate visual system. We suggest that strict notions of modularity in visual processing must be modified to take these factors into account. □



where R_p = partial correlation for the pattern prediction, r_c = raw correlation of the data with the component prediction, r_p = raw correlation of the data with the pattern prediction, r_{pc} = correlation of the two predictions. The partial correlation coefficient for the component prediction (R_c) was produced by exchanging r_p and r_c . The variance in each plaid direction tuning curve that can be accounted for by component and pattern predictions is thus R_c^2 and R_p^2 , respectively. The component index was computed as the difference between component and pattern variances ($R_c^2 - R_p^2$). CI is therefore a composite measure reflecting the degree to which a cell expresses component-type selectivity more strongly than pattern-type selectivity, for a specified stimulus condition.

Received 11 February; accepted 15 June 1992

- Adelson, E. A. & Movshon, J. A. *Nature* **300**, 523-525 (1982).
- Movshon, J. A., Adelson, E. H., Gizzi, M. & Newsome, W. T. in *Study Group on Pattern Recognition Mechanisms* (eds Chagas, C., Gattass, R. & Gross, C. G.) (Pontificia Academia Scientiarum, Vatican City, 1985).
- Stoner, G. R., Albright, T. D. & Ramachandran, V. S. *Nature* **344**, 153-155 (1990).
- Rodman, H. R. & Albright, T. D. *Expl Brain Res.* **75**, 53-64 (1989).
- Krauskopf, J. & Farell, B. *Nature* **348**, 328-331 (1990).
- Kooi, F. L., DeValois, K. K. & Switkes, E. *Perception* (in the press).
- Vallortigara, G. & Bressan, P. *Vision Res.* **31**, 1967-1978 (1992).
- Adelson, E. A. *J. opt. Soc. Am.* **A1**, 1266 (1984).
- Stoner, G. R. & Albright, T. D. *Invest. Ophthalmol. Vis. Sci.* **33**, 1050 (1992).
- Metelli, F. *Scient. Am.* **230**, 91-95 (1974).
- Beck, J. in *Human and Machine Vision II* (ed. Rosenfeld, A.) 1-12 (Academic, New York, 1986).

ACKNOWLEDGEMENTS. We thank J. Costanza for technical assistance and V. S. Ramachandran for discussion. This work was supported by a grant from the NIH, a McKnight Foundation Development Award to T.D.A., and a Research Fellowship to G.R.S. from the McDonnell-Pew Center for Cognitive Neuroscience at San Diego.