

Motion Integration by Neurons in Macaque MT Is Local, Not Global

Najib J. Majaj, Matteo Carandini, and J. Anthony Movshon

Center for Neural Science, New York University, New York, New York 10003

Direction-selective neurons in primary visual cortex have small receptive fields that encode the motions of local features. These motions often differ from the motion of the object to which they belong and must therefore be integrated elsewhere. A candidate site for this integration is visual cortical area MT (V5), in which cells with large receptive fields compute the motion of patterns. Previous studies of motion integration in MT have used stimuli that fill the receptive field, and thus do not test whether motion information is really integrated across this whole area. For each MT neuron, we identified two regions (“patches”) within the receptive field that were approximately equally effective in driving responses. We then measured responses to plaids whose component gratings overlapped within a patch, and compared them with responses to the same component gratings presented in separate patches. Cells that were selective for the direction of motion of the whole pattern when the gratings overlapped lost this selectivity when the gratings were separated and became selective instead for the direction of motion of the individual components. If MT cells simply pooled all of the inputs that endow them with a receptive field, they would encode all of the motions in the receptive field as belonging to a single object. Our results indicate instead that critical elements of the computations underlying pattern-direction selectivity in MT are done locally, on a scale smaller than the whole receptive field.

Key words: vision; visual motion; extrastriate; MT; V5; pattern; component

Introduction

Neurons in the primary visual cortex (V1) analyze the retinal image through receptive fields that are tightly tuned for stimulus position and orientation and do not signal whether the contours that activate them are part of a coherent figure. For example, their responses to plaids made by adding two gratings of different orientations are well predicted from the sum of their responses to the individual gratings (Movshon et al., 1985). V1 cells thus do not use information contained in two-dimensional features such as the intersections of the gratings and are unable to integrate across spatial positions to obtain a unified estimate of object motion.

A candidate site for this spatial integration of motion information is macaque cortical area MT, which plays a central role in the analysis of visual motion (Dubner and Zeki, 1971; Albright, 1984; Newsome et al., 1985; Britten et al., 1992; Salzman et al., 1992). Cells in this area have receptive fields ~10 times larger than in V1 and are all selective for the direction of motion. Some

are component-direction selective and, like V1 cells, respond when any component of a complex pattern is moving in their preferred direction. Others, however, are pattern-direction selective and combine information from overlapping components to compute the true direction of movement of a coherent pattern (Movshon et al., 1985). The responses of such a pattern cell to a plaid stimulus are therefore typically quite different from the sum of its responses to the individual gratings alone.

The basic computations that allow MT to integrate the direction of motion of multiple features are beginning to be understood. A widely accepted model (Simoncelli and Heeger, 1998) proposes that selectivity for pattern direction is the result of two mechanisms. The first mechanism is pooling of inputs from appropriate component-direction-selective cells, which could lie in area MT (Movshon and Newsome, 1996), in V1, or elsewhere (e.g., visual areas V2 or V3); this pooling mechanism provides the basic integration of motion signals. The second mechanism is opponent inhibition, which reduces the responses to motion in the direction opposite to the preferred. A third factor contributing to pattern-direction selectivity has been proposed more recently (Rust et al., 2006) and consists of a contrast gain control mechanism akin to the one operating in V1, which effectively increases responses to plaids moving in the preferred direction of the cell. In combination, these three mechanisms are sufficient to accurately predict the responses of MT neurons (Rust et al., 2006).

Neither the Simoncelli–Heeger model nor its refinements, however, specify how features are spatially integrated: overlapping and nonoverlapping features are treated identically. We studied this integration by comparing the direction selectivity of

Received July 25, 2006; revised Nov. 21, 2006; accepted Nov. 22, 2006.

This work was supported in part by an investigatorship from the Howard Hughes Medical Institute and by National Eye Institute Grant EY02017 (J.A.M.). N.M. was supported in part by a grant from the Alfred P. Sloan Foundation. We are grateful to Matthew Smith for his assistance during some of the experiments and to him and Josh McDermott for their comments on a previous version of the manuscript. Suzanne Fenstemaker provided valuable assistance with histology.

Correspondence should be addressed to Najib J. Majaj, Center for Neural Science, New York University, 4 Washington Place, Room 809, New York, NY 10003. E-mail: najib@cns.nyu.edu.

M. Carandini's present address: Smith-Kettlewell Eye Research Institute, 2318 Fillmore Street, San Francisco, CA 94115.

DOI:10.1523/JNEUROSCI.3183-06.2007

Copyright © 2007 Society for Neuroscience 0270-6474/07/270366-05\$15.00/0

MT neurons to overlapping and nonoverlapping gratings presented within the receptive field. We reasoned that if the computation of pattern motion were local, the ability of the neuron to integrate the motions of the two gratings would be compromised when we separated them. If the pooling were global, and the neuron simply combined all of the inputs that contribute to its receptive field without regard to their receptive field location, then the neuron would signal a coherent direction of motion whether the gratings overlapped or not.

Materials and Methods

We studied the responses of 54 MT neurons in 12 anesthetized, paralyzed macaque monkeys. Our methods for animal preparation and data collection are standard in our laboratory and were described in detail previously (Cavanaugh et al., 2002). All experimental work with animals was conducted according to procedures approved by the New York University Animal Welfare Committee.

We initially determined the location and size of receptive fields on a tangent projection screen. All cells had receptive fields centered within 25° of the fovea; most were within 15° . After mapping the receptive fields, we presented luminance-modulated grating and plaid stimuli on a gray background to the preferred eye; the space- and time-averaged luminance of the stimuli was 33 cd/m^2 , which was also the luminance of the background. For each cell, we determined the direction, speed, spatial frequency, and size of the luminance-modulated sine wave that evoked the strongest response from the cell. We then identified two regions ("patches") within the receptive field that were approximately equally responsive. We chose patch sizes that gave reliable responses to gratings, typically half the diameter (range 25–50%) of the most effective stimulus, and arranged them so that the patches abutted but never overlapped. The center-to-center separation between the two patches was 50–75% of the diameter of the receptive field. We placed the patches along an axis parallel to the preferred direction of the cell.

We studied the direction selectivity of each cell using three patterns: gratings, plaids, and "pseudoplaids". For gratings and plaids, we tested each patch separately. Plaids were composed of two superimposed gratings whose direction of motion differed by 120° . For pseudoplaids, we presented one component of the plaid in each patch (this can be done in two configurations, $+120^\circ$ or -120°). All three patterns were presented drifting in 12 directions of motion, in three to five pseudorandomly ordered blocks. Each stimulus was presented for 2 s followed by a brief interstimulus interval of ~ 300 ms. Response was measured by counting spikes over the entire stimulus interval.

Our analysis of directional selectivity was conventional (Movshon et al., 1985). Using the directional tuning of each cell for gratings, we constructed predictions of responses to plaids for idealized pattern-direction-selective and component-direction-selective cells. For pseudoplaids, the predictions were based on the separate grating tuning curves measured for each patch (Fig. 1*a*). We computed partial correlations of the actual responses with the predicted tuning curves (Albright, 1984; Movshon et al., 1985) and transformed them into normal deviates using Fisher's *r*-to-*Z* transformation [see equations 13.13.3–5 in Hays (1981)] (Smith et al., 2005).

To allow histological confirmation of the recording sites, we made small electrolytic lesions at the end of each electrode track by passing DC current ($2 \mu\text{A}$ for 5 s) through the recording electrode. At the end of each experiment, the monkey was killed with an overdose of Nembutal and perfused through the heart with 0.1 M phosphate buffer solution and 4% paraformaldehyde. Sections ($40 \mu\text{m}$) were stained for Nissl substance with cresyl violet or for myelin using the method of Gallyas (1979). Most recording locations were confirmed directly; in the few remaining cases, we relied on the proximity to histologically confirmed recording sites and the high proportion of directional cells with receptive field sizes characteristic of MT as evidence that the recordings were from MT (Desimone and Ungerleider, 1986).

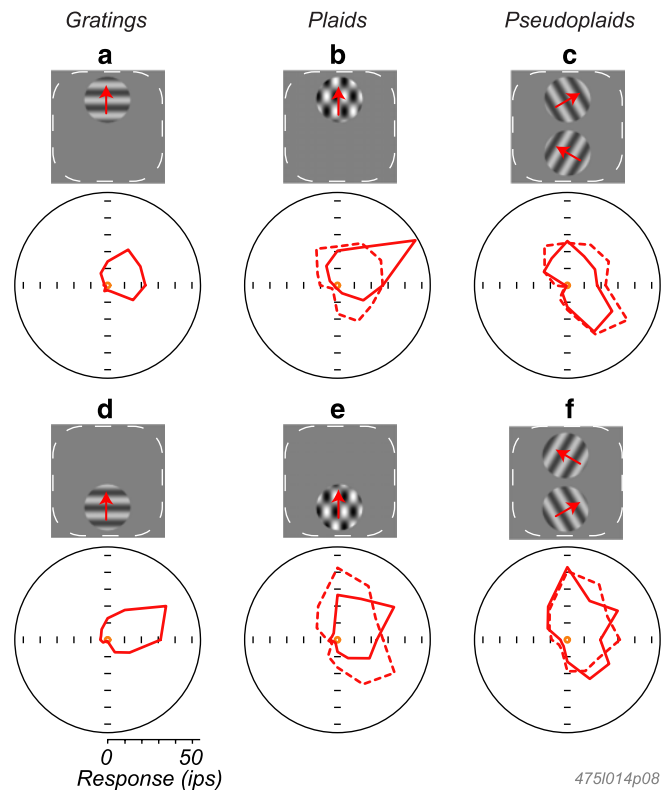


Figure 1. Responses of an MT cell to gratings, plaids, and pseudoplaids. Polar plots express cell response in spikes per second as distance from the origin, with the angle indicating the stimulus direction of motion. The small orange circles indicate spontaneous firing rate. **a, d**, Responses to a drifting grating covering one or the other of two patches within the receptive field, as indicated by the stimulus icons. **b, e**, Responses to the plaids created by summing two of the gratings tested in **a**, with orientations differing by 120° . Dashed red curves indicate the predicted response of a component-direction-selective cell. **c, f**, Responses to pseudoplaids obtained with gratings in the two patches arranged with a direction difference of $+120^\circ$ (**c**) or -120° (**f**).

Results

To determine whether a neuron was component- or pattern-direction selective, we measured the direction selectivity within each patch using gratings and plaids composed of superimposed gratings whose direction of motion differed by 120° (Fig. 1*a, b, d, e*). We used the grating response to make two predictions for the plaid response (Movshon et al., 1985). A component-direction neuron would have a bi-lobed tuning curve that peaks when either of the component gratings matches the preferred direction of motion of the cell (Fig. 1*b, d*, dashed lines). Conversely, a pattern-direction neuron would only respond when the pattern moves in the preferred direction, in which case the tuning curve for plaids (Fig. 1*b, d*) would be similar to that for gratings (Fig. 1*a, d*). The latter was the case for the example cell in Figure 1, which we thus classified as pattern-direction selective.

To test the spatial integration properties of the neuron, we measured direction selectivity with the components of the plaids delivered separately to each patch (pseudoplaids). In response to pseudoplaids, the example cell was not selective for the direction of the pattern but rather responded when either of the components was moving in the preferred direction (Fig. 1*c, f*). Indeed, the tuning curves measured with pseudoplaids were substantially broader than those obtained with plaids (Fig. 1*b, d*) and agreed well with the expectation for a component-direction cell (Fig. 1*c, f*, dashed lines).

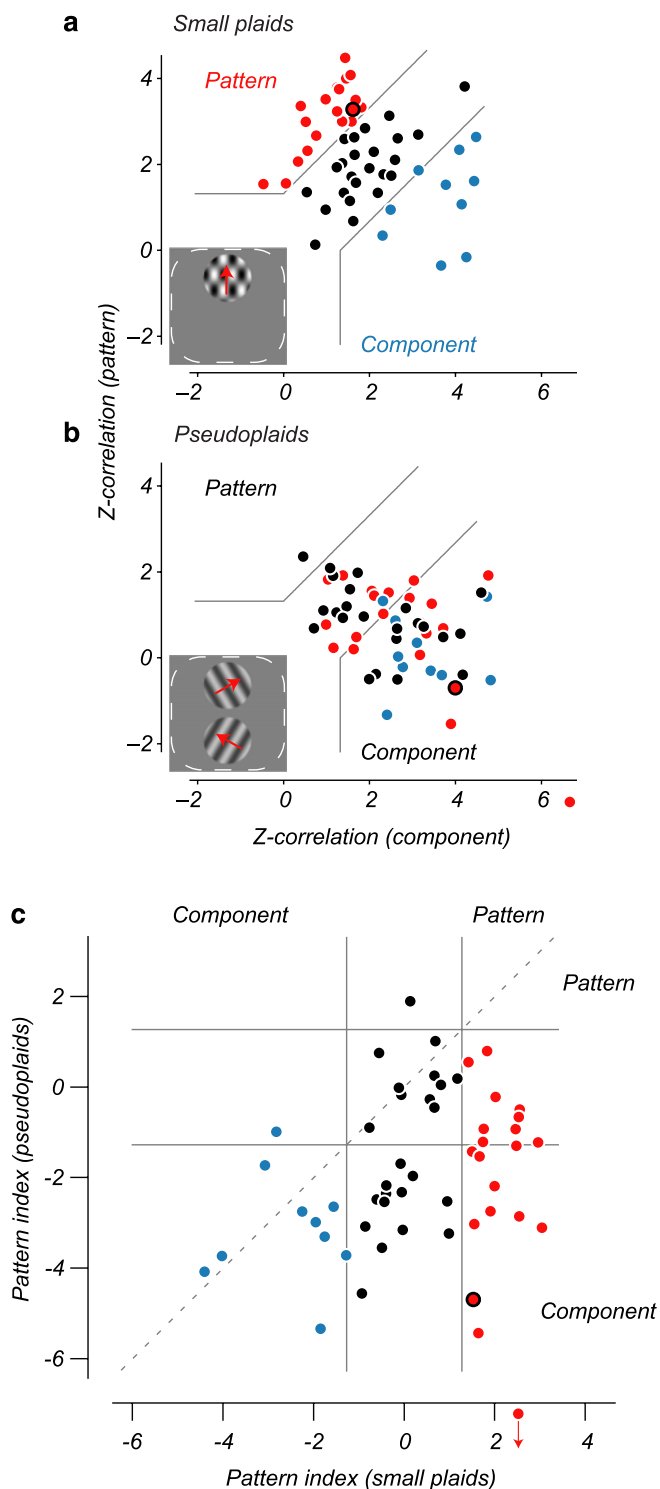


Figure 2. Summary of the results across the cell population. *a*, The degree to which cells are selective for the direction of a whole pattern or of the individual components as determined with small plaid stimuli. The Z-transformed partial correlations between the data and the “component” and “pattern” models are plotted against one another. Gray lines separate regions within which cells are classified as pattern-direction selective or component-direction selective, according to a conservative statistical criterion (Movshon et al., 1985). Cells classified as pattern selective are indicated in red, as component selective in blue–green, and as unclassified by this method in black. *b*, A similar plot made from data taken from the same cells using pseudoplaids. The colors indicating the classification of the cells using small plaids are retained from *a*. *c*, The distribution of pattern indices for the two conditions, pseudoplaids and plaids. The pattern index is the difference between Z-transformed pattern and component partial correlations ($Z_p - Z_c$). The gray lines bound regions of significance. We used a criterion difference of +1.28 or –1.28, equivalent to $p = 0.90$, for this purpose. Data points falling in the leftmost column

To quantify the degree to which cells were selective for the direction of motion of a pattern rather than of its individual components, we computed the partial correlation coefficients between the pattern responses and the predictions of two models (Movshon et al., 1985). The first model defines selectivity for component direction: its response to a pattern is the sum of its responses to the components. The second model defines selectivity for pattern direction: its response to a pattern is the same as that to a single component moving in the pattern direction. The partial correlation between the actual response and the predictions of these two models determines the classification of the cell. Because the confidence intervals around correlation coefficients are strongly dependent on the correlation, we used Fisher’s variance-stabilizing r -to- Z transform to convert the correlations into Z -values that can be treated as normal deviates (Hays, 1981). Transformation into Z space has the particular virtue for this analysis that distances between values have the same meaning everywhere in the space.

Figure 2*a* shows the resulting distribution of correlation values, as measured with small plaids; the plotted values are the average of those measured for the two patches of the receptive field. The gray lines divide each plot into three zones. Points falling in the region marked “Component” indicate cells whose component correlation coefficients significantly exceeds either zero or the pattern correlation coefficient, whichever is larger. Similarly, cells falling in the “Pattern” region are significantly pattern-direction selective. Cells falling in the intermediate region are unclassified by this method, usually because their response is well predicted by both models. For small plaids (Fig. 2*a*), the distribution of correlations is very similar to that previously reported, showing approximately equal numbers of cells as pattern-direction selective (red, 20 of 54), component-direction selective (blue–green, 10 of 54), and unclassified (black, 24 of 54) (Movshon et al., 1985; Rodman and Albright, 1989; Smith et al., 2005).

This distribution changed when pseudoplaids were used (Fig. 2*b*). The data point for each cell retains the color given on the basis of the classification in Figure 2*a*. All but one of the 24 pattern-direction cells (red) migrated from the pattern region into the other two regions.

This migration can be visualized by computing and comparing a pattern index, the difference between the Z -transformed pattern and component correlation coefficients ($Z_p - Z_c$), for the two conditions (plaids and pseudoplaids) for each cell. The distribution of pattern indices for pseudoplaids and plaids is shown in Figure 2*c* (colors retained from Fig. 2*a*). The gray solid lines mark the boundaries of classification regions derived from those in Figures 2*a,b*. Data points falling on or close to the identity line (the dashed line) mark cells that had the same pattern index for pseudoplaids as that for plaids. It is clear from the graph that cells that were component selective in response to plaids remained component in response to pseudoplaids (all but one of the blue–green data points are in the bottom left quadrant). However, all of the cells that were pattern selective in the plaid condition changed their behavior in the pseudoplaid condition. Approximately half of them become component selective, and the other half became

← or the lowest row represent partial correlations that are significantly component. Data points in the rightmost column and the top row represent partial correlations that have a significant pattern. The colors indicating the classification of the cells using small plaids are retained from *a*. The outlying point with an arrow at the bottom right had a pseudoplaid pattern index of –9.8. In each plot, the datum circled in black corresponds to the cell whose responses are shown in Figure 1.

unclassified. Among the cells that were unclassified when tested with plaids, half remained unclassified, and the other half behaved like pattern cells and became component. Thus, separating the components of a plaid into separate regions of the receptive field abolishes pattern motion selectivity in MT cells.

Discussion

These results do not support a simple model in which motion signals are pooled across an entire MT receptive field (Movshon et al., 1985; Snowden et al., 1991; Simoncelli and Heeger, 1998; Britten and Heuer, 1999), because such a model would treat all of the motions in the receptive field as if they belong to a single object. Instead, our results show that the computation underlying pattern-direction selectivity in MT occurs on a scale finer than the whole receptive field and suggest that spatial coincidence of the components of a moving object is required. How can we incorporate these local computations into thinking about the organization of MT receptive fields and, in particular, the creation of pattern-direction selectivity? These computations could begin in earlier visual areas such as V2 and V3, where cells have smaller receptive fields and where small numbers of pattern cells are found (Levitt et al., 1994; Gegenfurtner et al., 1997). But it does not seem reasonable to suggest that the small numbers of pattern cells in earlier areas directly create the much higher prevalence of pattern selectivity of MT. Alternatively, the local computation might be done within MT, perhaps through local dendritic interactions creating local pattern-direction selectivity by combining suitable inputs.

We prefer to think that the computation of pattern direction is begun in V1 but completed in MT. Rust et al. (2006) showed that they could account for MT response to gratings and additive plaids by linearly pooling the nonlinear outputs of directional neurons like complex cells in V1. They found that three factors were crucial for the creation of pattern-direction selectivity: a broad pooling of excitatory inputs from cells with varied preferred directions, a strong motion-opponent inhibition, and a tuned form of contrast gain control in the input V1 neurons. One of these mechanisms, broad pooling, is presumed to be a global feature of MT connectivity. But the V1 gain control would act locally, and there is also reason to think that opponent-motion suppression occurs in V1 (Rust et al., 2002) and is expressed locally within MT receptive fields (Qian et al., 1994; Rust et al., 2002). Two of the three key elements of pattern selectivity may therefore depend on mechanisms whose spatial scale is finer than the MT receptive field, accounting for the result reported here.

It has been suggested that MT signals underlie the perception of coherent motion (Movshon et al., 1985). The pseudoplaids used in our experiments do not cohere perceptually and are not combined by MT pattern neurons; thus, our results are not inconsistent with this idea. The key issue is how the visual system decides when, and whether, to combine signals over space.

Faced with multiple separated features, the visual system has to integrate over the ones that belong to a single object and segment the ones that belong to different objects. In psychophysical experiments, it has been shown that the arrangement of features influences whether observers group them into a single object or perceive them as separate objects moving in separate directions – it is harder to integrate features that are far apart (Wallach, 1935, 1976; Nakayama and Silverman, 1988; Lorenceau and Zago, 1999). But the separation of features is only one of several factors that influence perceived coherence. Several studies have shown that cues such as occlusion that influence coherence also change neural activity in MT (Stoner and Albright, 1992, 1994; Duncan

et al., 2000; Pack et al., 2004), and it might be that under different perceptual conditions, an experiment like ours would yield different results, reflecting the integration of information about scene organization with MT signals.

However, in the absence of contrast or depth cues, or special features like terminators, the visual system defaults to simple pooling over space to produce estimates of global velocity (Mingolla et al., 1992; Rubin and Hochstein, 1993; Lorenceau and Zago, 1999). Neurons in MT behave similarly, their response to spatially separate features presented within their receptive fields being the average of their responses to each of the features presented alone, whether the features differ in contrast (Britten and Heuer, 1999; Heuer and Britten, 2002) or direction of motion (Britten and Newsome, 1990; Ferrera and Lisberger, 1997; Recanzone et al., 1997). This pooling over the entire receptive field suggests that the segmentation computation is left for higher cortical areas, whose signals might feed back into MT under appropriate conditions.

In either case, the receptive field of MT cells is not the unit of motion integration for simple situations like the one we studied. Under our conditions, separation of the grating components abolished the computation of pattern motion, and the stimuli did not appear perceptually coherent. We therefore conjecture that signals from MT simply provide local motion measurements, which are integrated elsewhere with scene information to determine the final percept of coherent or incoherent motion.

References

- Albright TD (1984) Direction and orientation selectivity of neurons in visual area MT of the macaque. *J Neurophysiol* 52:1106–1130.
- Britten KH, Heuer HW (1999) Spatial summation in the receptive fields of MT neurons. *J Neurosci* 19:5074–5084.
- Britten KH, Newsome WT (1990) Responses of MT neurons to discontinuous motion stimuli. *Invest Ophthalmol Vis Sci [Suppl]* 31:238.
- Britten KH, Shadlen MN, Newsome WT, Movshon JA (1992) The analysis of visual motion: a comparison of neuronal and psychophysical performance. *J Neurosci* 12:4745–4765.
- Cavanaugh JR, Bair W, Movshon JA (2002) Nature and interaction of signals from the receptive field center and surround in macaque V1 neurons. *J Neurophysiol* 88:2530–2546.
- Desimone R, Ungerleider LG (1986) Multiple visual areas in the caudal superior temporal sulcus of the macaque. *J Comp Neurol* 248:164–189.
- Dubner R, Zeki SM (1971) Response properties and receptive fields of cells in an anatomically defined region of the superior temporal sulcus in the monkey. *Brain Res* 35:528–532.
- Duncan RO, Albright TD, Stoner GR (2000) Occlusion and the interpretation of visual motion: perceptual and neuronal effects of context. *J Neurosci* 20:5885–5897.
- Ferrera VP, Lisberger SG (1997) Neuronal responses in visual areas MT and MST during smooth pursuit target selection. *J Neurophysiol* 78:1433–1446.
- Gallyas F (1979) Silver staining of myelin by means of physical development. *Neurol Res* 1:203–209.
- Gegenfurtner KR, Kiper DC, Levitt JB (1997) Functional properties of neurons in macaque area V3. *J Neurophysiol* 77:1906–1923.
- Hays WL (1981) *Statistics*, Ed 3. New York: CBS College Publishing.
- Heuer HW, Britten KH (2002) Contrast dependence of response normalization in area MT of the rhesus macaque. *J Neurophysiol* 88:3398–3408.
- Levitt JB, Kiper DC, Movshon JA (1994) Receptive fields and functional architecture of macaque V2. *J Neurophysiol* 71:2517–2542.
- Lorenceau J, Zago L (1999) Cooperative and competitive spatial interactions in motion integration. *Vis Neurosci* 16:755–770.
- Mingolla E, Todd JT, Norman JF (1992) The perception of globally coherent motion. *Vision Res* 32:1015–1031.
- Movshon JA, Newsome WT (1996) Visual response properties of striate cortical neurons projecting to area MT in macaque monkeys. *J Neurosci* 16:7733–7741.
- Movshon JA, Adelson EH, Gizzi MS, Newsome WT (1985) The analysis of moving visual patterns. *Pontificia Academia Scripta Varia* 54:117–151.

- Nakayama K, Silverman GH (1988) The aperture problem—II. Spatial integration of velocity information along contours. *Vision Res* 28:747–753.
- Newsome WT, Wurtz RH, Dursteler MR, Mikami A (1985) Deficits in visual motion processing following ibotenic acid lesions of the middle temporal visual area of the macaque monkey. *J Neurosci* 5:825–840.
- Pack CC, Gattland AJ, Born RT (2004) Integration of contour and terminator signals in visual area MT of alert macaque. *J Neurosci* 24:3268–3280.
- Qian N, Andersen RA, Adelson EH (1994) Transparent motion perception as detection of unbalanced motion signals. III. Modeling. *J Neurosci* 14:7381–7392.
- Recanzone GH, Wurtz RH, Schwarz U (1997) Responses of MT and MST neurons to one and two moving objects in the receptive field. *J Neurophysiol* 78:2904–2915.
- Rodman HR, Albright TD (1989) Single-unit analysis of pattern-motion selective properties in the middle temporal visual area (MT). *Exp Brain Res* 75:53–64.
- Rubin N, Hochstein S (1993) Isolating the effect of one-dimensional motion signals on the perceived direction of moving two-dimensional objects. *Vision Res* 33:1385–1396.
- Rust NC, Majaj NJ, Simoncelli EP, Movshon JA (2002) Gain control in macaque area MT is directionally selective. *Soc Neurosci Abstr* 28:657.10.
- Rust NC, Mante V, Simoncelli EP, Movshon JA (2006) How MT cells analyze the motion of visual patterns. *Nat Neurosci* 9:1421–1431.
- Salzman CD, Murasugi CM, Britten KH, Newsome WT (1992) Microstimulation in visual area MT: effects on direction discrimination performance. *J Neurosci* 12:2331.
- Simoncelli EP, Heeger DJ (1998) A model of neuronal responses in visual area MT. *Vision Res* 38:743–761.
- Smith MA, Majaj NJ, Movshon JA (2005) Dynamics of motion signaling by neurons in macaque area MT. *Nat Neurosci* 8:220–228.
- Snowden RJ, Treue S, Erickson RG, Andersen RA (1991) The response of area MT and V1 neurons to transparent motion. *J Neurosci* 11:2768–2785.
- Stoner GR, Albright TD (1992) Motion coherency rules are form-cue invariant. *Vision Res* 32:465–475.
- Stoner GR, Albright TD (1994) Visual motion integration: a neurophysiological and psychophysical perspective. In: *Visual motion detection* (Smith AT, Snowden RJ, eds), pp 253–290. London: Academic.
- Wallach H (1935) Über visuell wahrgenommene Bewegungsrichtung. *Psychologische Forschung* 20:325–380.
- Wallach H (1976) On perceived identity: I. The direction of motion of straight lines. In: *On perception* (Wallach H, ed), pp 201–216. New York: Quadrangle, The New York Times Book Co.