

End-Stopping and the Aperture Problem: Two-Dimensional Motion Signals in Macaque V1

Christopher C. Pack,* Margaret S. Livingstone, Kevin R. Duffy, and Richard T. Born
Harvard Medical School
Department of Neurobiology
220 Longwood Avenue
Boston, Massachusetts 02115

Summary

Our perception of fine visual detail relies on small receptive fields at early stages of visual processing. However, small receptive fields tend to confound the orientation and velocity of moving edges, leading to ambiguous or inaccurate motion measurements (the aperture problem). Thus, it is often assumed that neurons in primary visual cortex (V1) carry only ambiguous motion information. Here we show that a subpopulation of V1 neurons is capable of signaling motion direction in a manner that is independent of contour orientation. Specifically, end-stopped V1 neurons obtain accurate motion measurements by responding only to the endpoints of long contours, a strategy which renders them largely immune to the aperture problem. Furthermore, the time course of end-stopping is similar to the time course of motion integration by MT neurons. These results suggest that cortical neurons might represent object motion by responding selectively to two-dimensional discontinuities in the visual scene.

Introduction

Humans and other primates are able to perceive the speed and direction of moving objects with great precision. This is remarkable in light of the fact that, at the earliest levels of the visual system, local measurements of velocity are often confounded with the object's shape. The problem is based on simple geometry: if the middle of a moving contour is viewed through a small aperture, only the component of velocity perpendicular to the contour can be recovered (Wallach, 1935; Wuerger et al., 1996). The component of velocity that is parallel to the orientation of the contour can be recovered only at the endpoints. Because neurons in the primary visual cortex (V1) have small receptive fields, this "aperture problem" (Marr, 1982) must be overcome if the visual system is to function properly.

In the example shown in Figure 1, a vertically oriented bar moving diagonally is shown at two consecutive time points. These successive "snap shots" make it clear that a receptive field located near the middle of the bar would be fooled by the aperture problem (Figure 1A) and see only the rightward component of motion. A receptive field located at the endpoints of the bar, however, would be able to recover the correct direction of motion (Figure 1B). This exercise suggests that a sensi-

ble first step toward measuring velocity would be to filter out long contours, and to respond selectively to endpoints. Such selectivity for endpoints was first observed by Hubel and Wiesel (1965), who identified it with end-stopped neurons in V1. End-stopped neurons respond well to short contours, but the responses are suppressed by long contours.

In recent years, it has become apparent that most V1 neurons are end-stopped to some degree (DeAngelis et al., 1994; Jones et al., 2001; Sceniak et al., 2001). It is not clear, however, how the property of end-stopping relates to direction selectivity and how this combination of properties bears on the aperture problem. We have therefore investigated these issues, using an efficient receptive field mapping technique (Livingstone et al., 2001). Our findings indicate that end-stopped V1 neurons are capable of encoding two-dimensional motion signals in a manner that is largely independent of the orientation of the stimulus—a property that is also found in neurons within the middle temporal visual area (MT or V5) (Movshon et al., 1986; Pack et al., 2001; Pack and Born, 2001; Rodman and Albright, 1989), which receive their most significant inputs from V1 neurons. Finally, the temporal dynamics of neural responses in V1 and MT suggest that the earliest responses in both regions encode information about entire contours, while the later responses selectively encode information about endpoints.

Results

V1 neurons were first screened with drifting bars, and if they were both direction selective (direction index > 0.5) and end-stopped, they were studied further. Our criterion for end-stopping was a 30% suppression of response to a long (9.5°) bar, as compared to a shorter bar of optimal length. We recorded a total of 38 neurons in V1, and 17 neurons in MT of two alert monkeys. Nine V1 cells failed to respond to the noise stimulus or were lost before sufficient data could be collected. The remaining 29 end-stopped V1 cells were analyzed in detail. These cells were generally quite end-stopped: the average suppression was 59%. For comparison, we also studied a few V1 cells that were direction selective, but not end-stopped. All of the V1 neurons were complex cells, as evidenced by the fact that one-bar maps (defined below) for light and dark bars were nearly identical. We chose not to include simple cells because they tend to be less direction selective than complex cells (Conway and Livingstone, 2003; Hubel and Wiesel, 1968) and apparently do not project to MT (Movshon and Newsome, 1996).

Figure 2 illustrates the stimulus used in the sparse noise experiments. In each frame, the stimulus consisted of two long bars, one black, one white, on a gray background, flashed at 75 Hz at random positions. For end-stopped cells, the bar length was chosen to be much longer than the excitatory length of the receptive field, as judged from the peak of the length-tuning

*Correspondence: cpack@hms.harvard.edu

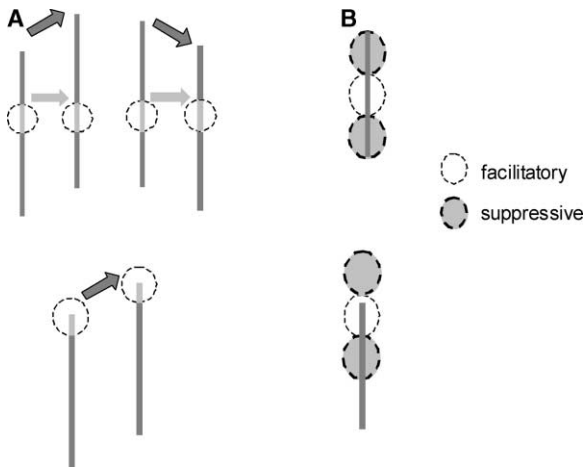


Figure 1. The Aperture Problem

(A) A small aperture (dotted circle) positioned at the center of a moving bar allows detection of only the rightward component of motion (light gray arrow), regardless of the actual motion direction (dark gray arrows). However, an aperture positioned at the endpoint of the bar permits measurement of the correct motion direction.

(B) An end-stopped cell responds poorly to a long bar centered on the receptive field, because the bar stimulates both the excitatory (dotted oval) and inhibitory (shaded oval) regions of the receptive field. However, the same end-stopped neuron responds better to the endpoint, because only part of the inhibitory region is stimulated.

curves. For 19/29 cells, the bar exceeded the receptive field length by a factor of four or more; for the remaining cells, it was at least twice as long. We generally tried to use the longest bar that evoked consistent responses during sparse noise stimulation. The centers of the bars were constrained to appear within a square stimulus range, which was centered on the receptive field. By presenting these stimulus sequences for approximately 30 min, we obtained a spike train from each cell, along with the corresponding record of several thousand stimulus presentations. The stimulus positions were corrected offline according to the monkey's eye position on each frame (Livingstone, 1998). We therefore measured stimulus position on the retina, not on the display monitor.

A distinct advantage of the sparse noise stimulus is that it provides estimates of the one-bar and two-bar responses from the same spike train. The one-bar response is simply the cell's likelihood of firing as a function of the position of either the black or white bar. The two-bar response is the likelihood of firing as a function of the displacement of bars on sequential stimulus frames (i.e., an apparent-motion stimulus). Because the bars could move in two dimensions, the two-bar maps allowed us to evaluate the extent to which neural responses were affected by the aperture problem.

Both the one-bar and two-bar maps were computed by reverse correlating the spike train with the stimulus sequence. For the one-bar maps, we reverse correlated the spikes with the position of the light or the dark bar on any given frame. The reverse correlation procedure thus measured the likelihood that a spike was preceded by a bar at a given position in space and a particular latency in time. The latency is called the correlation

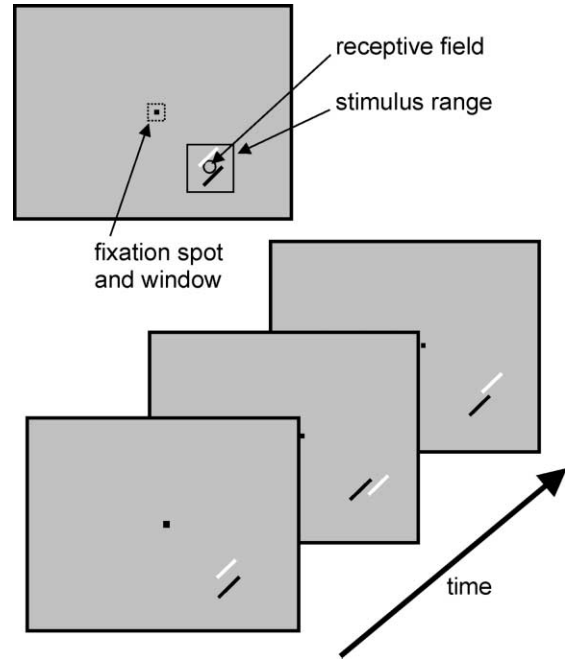


Figure 2. Sparse Noise Mapping Stimulus

Monkeys were required to maintain fixation within 1° (dotted square) of a small spot (black square). On each screen refresh (~ 13 ms), one black and one white bar appeared in random positions on a gray background. The bars were constrained to appear within a limited stimulus range (box with solid lines) that was centered on the receptive field (small circle) of each neuron. By reverse correlating the resulting spike train with the stimulus sequence, we obtained one-bar and two-bar maps for each neuron.

delay; except where explicitly stated, we used the time to peak response (usually between 40 and 60 ms). The spatial positions in two dimensions are the coordinates for the one-bar maps. For each spike train, this analysis yielded two one-bar maps, one for the black bars and one for the white bars. Because our cells were complex, the black and white bars produced similar maps, which were then averaged to generate the final one-bar maps.

For the two-bar maps, spikes were reverse correlated with the positions of two different stimuli on successive video frames, one of which we refer to as the reference stimulus and the other as the probe stimulus. The reference stimulus was defined for each spike as the stimulus that preceded the spike by the correlation delay. The other, probe stimulus, is one of the two (black or white) stimuli in the immediately preceding frame. Activity was mapped as a function of probe stimulus position minus the reference stimulus position, in 2D retinal coordinates, with the horizontal axis corresponding to the horizontal displacement and the vertical axis corresponding to the vertical displacement. Thus, position (0,0) represents occasions when the two successive stimuli fell in exactly the same retinal location, regardless of which absolute retinal location it was. For each spike train, this analysis yields four different two-bar maps, corresponding to the four possible bar sequences on successive frames (white-to-white, black-to-black, white-to-black, and black-to-white). Each one of these maps contains one-bar responses that depend on the position

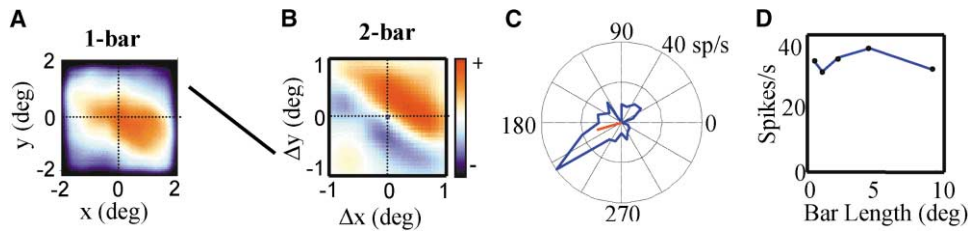


Figure 3. Receptive Field Maps for a Non-End-Stopped V1 Cell

(A) Response as a function of the position of the center of a 3° bar. The bar is drawn to scale to the right of the figure. The orientation was chosen to be optimal for this cell. The color of each pixel indicates the likelihood of a spike in response to a bar at the corresponding position. (B) Response as a function of relative positions of bars on sequential frames. The position of the dark red region indicates that the response was facilitated by downward and leftward sequences, while the position of the blue region indicates that the neuron was suppressed by upward and rightward sequences. (C) Polar direction-tuning curve for an optimally sized bar moving in a direction perpendicular to its orientation. The angle represents motion direction, and the radius represents time-averaged firing rate. The red line represents the mean vector. The mean vector angle is the preferred direction and its length is related to the sharpness of tuning. (D) Length tuning. The time-averaged response of this neuron to a bar of varying lengths moving at the preferred velocity through the receptive field.

and contrast of the individual bars, and two-bar responses that depend only on interactions between pairs of bars. To eliminate the contribution of the one-bar responses, we summed the same-contrast maps (white-to-white and black-to-black) and subtracted the opposite contrast maps (white-to-black and black-to-white), according to a second-order Wiener-like calculation (Emerson et al., 1987). This procedure preserves the directional responses, but eliminates the responses that depend only on one bar (Emerson et al., 1987; Livingstone et al., 2001).

Figure 3A shows the one-bar map for one of the V1 neurons that was *not* end-stopped. The position (0,0) indicates the center of the stimulus range, and each pixel in the map indicates the relative likelihood that a 3° long bar centered at that position elicited a spike from the neuron. Red regions indicate high probability, and blue regions indicate low probability. In this map, the red region is elongated along an axis parallel to the orientation of the bar stimulus, indicating that the cell responded well when any part of the bar fell on the receptive field.

Figure 3B shows the two-bar map for the same cell. Here the dark red region in the upper right portion of the map indicates that the cell's response to a reference stimulus was facilitated when it was preceded by a probe stimulus that appeared to its right and upward. In other words, the cell responded well to a down-left motion sequence. However, the two-bar response was not very selective for motion direction. The elongation of the red region indicates that the cell was responsive to almost any stimulus sequence that contained either a leftward or a downward component: a clear manifestation of the aperture problem. Similarly, any stimulus sequence that contained an upward or rightward component suppressed the cell's response, as indicated by the long blue streak in the two-bar map. Figure 3C shows a direction-tuning curve obtained by sweeping an optimally sized bar through the receptive field. In these conventional direction-tuning curves, the direction of bar motion was always perpendicular to its orientation. In this example, the response peaks for motion down and to the left, in agreement with the two-bar map. It is

important to appreciate, however, that the stimuli used to generate the two-bar map differ from those used to generate the conventional direction-tuning curve, in that the orientation of the bars for the former were always the same regardless of the direction of the displacement of its center. Thus, the conventional direction-tuning curves do not explicitly probe for the effects of the aperture problem, whereas the two-bar maps do.

Figure 3D shows the length-tuning curve for the same cell. The peak response occurred for bars 5° in length, but increasing the bar length to 9.5° only suppressed the response by 11.5%. When this neuron was tested with a sparse noise stimulus in which the bars were rotated 45° away from the preferred orientation, it was almost completely unresponsive (data not shown).

Figure 4 shows the results of a similar mapping experiment, but in this case the neuron was end-stopped, as indicated by the length-tuning curve in panel D. This cell responded well to short bars, but was completely unresponsive to long bars (96.0% suppression), even when the orientation and motion direction were optimal. Figure 4A shows the one-bar map calculated from the response to a sparse noise stimulus consisting of 2.4° long bars. The best response was obtained when the center of the bar was far from the center of the receptive field. The positions of the red regions show that the cell responded well when either one of the endpoints of the bar was in the receptive field. The lack of excitation at position (0,0) in this map indicates that the cell failed to respond when the bar was centered on the receptive field, as would be predicted from the length-tuning curve in Figure 4D. The slight asymmetry in the positions of the red regions with respect to the origins of the maps in Figure 4A is most likely due to small errors in our alignment of the stimulus range with the receptive field.

The selectivity of this cell for endpoints suggests that it might provide useful two-dimensional motion information (Lorençeau et al., 1993; Mingolla et al., 1992; Rubin et al., 1995) (Figure 1). This possibility is confirmed by the two-bar map shown in Figure 4B. Here the most facilitatory portion of the map (shown in dark red) is confined to a discrete region of the map. The position of this region suggests that the cell responded best to

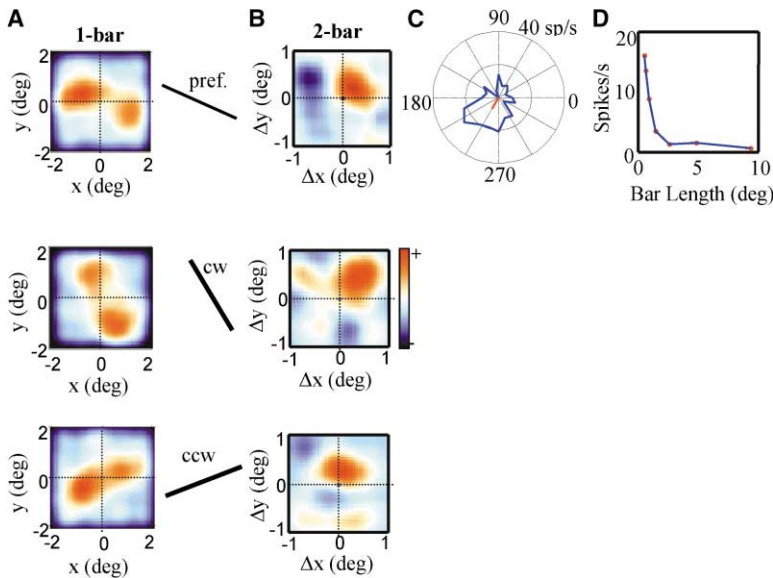


Figure 4. Receptive Field Maps for an End-Stopped V1 Cell

(A) Response as a function of the position of the center of a 2.4° bar. The bars are drawn to scale to the right of each figure. The color of each pixel indicates the likelihood of a spike in response to a bar at the corresponding position. The first row shows the response when the bar orientation was optimal. The second and third rows show the responses when the bar was rotated by 45° clockwise and counterclockwise from this orientation. This cell responded well when one of the endpoints of the bar was in the center of the receptive field, but not when the entire bar covered the receptive field.

(B) Responses as a function of relative positions of bars on sequential frames. The dark red regions indicate that the response was facilitated by down-left sequences. This was largely independent of the orientation of the bar, as shown by the similar maps obtained in all three rows.

(C) Polar direction-tuning curve for an opti-

mally sized bar moving in a direction perpendicular to its orientation. See Figure 3 caption for a description.

(D) Length tuning. This neuron responded to bars centered on the receptive field, but only if they were very short.

a down-left sequence of motion, in agreement with the conventional direction-tuning curve shown in Figure 4C. Furthermore, the dark blue regions on either side of the red region indicate that the neuron was suppressed by motion perpendicular to its preferred direction. In other words, the neuron was suppressed by motion parallel to the orientation of the bar. This type of selectivity allows the neuron to avoid the aperture problem by directly computing motion direction in two dimensions.

The two-bar maps in Figures 3 and 4 differ primarily in the degree of elongation of the facilitatory (red) regions. A very elongated region suggests that the cell responds well to any motion parallel to the orientation of the bar (as in Figure 3), while a more circular facilitatory region suggests limited responsiveness for motion parallel to the bar. In order to quantify this type of selectivity, we fit each of our two-bar maps with an oriented Gaussian function. The oriented Gaussian provides a measure of the elongation, or aspect ratio, of the subregions of the two-bar maps, which in turn provides a measure of the cells' selectivity for motion parallel to the orientation of the bars. We computed the aspect ratio of the facilitatory (red) subregions for all of the neurons in our sample that had a Gaussian fit with $R^2 > 0.9$. This criterion excluded only 4 of the 29 neurons.

The aspect ratio for the neuron in Figure 3 was 3.1, indicating that the neuron was approximately one-third as sensitive to motion parallel to the bar as it was to motion perpendicular to the bar. The aspect ratio for the neuron in Figure 4 was 1.4. Across the end-stopped population, the aspect ratios ranged from 1.2 to 3.0, with a geometric mean of 2.0 ± 1.08 . There was a weak, but statistically significant, negative correlation between the aspect ratio and the degree of end-stopping ($p < 0.05$; $R^2 = 0.22$). In other words, strongly end-stopped cells were less responsive than weakly end-stopped cells to motion parallel to the orientation of the bar. This relationship is shown in Figure 5. In this plot, the weakly end-stopped neuron from Figure 3 is marked by an X, and the neuron shown in Figure 4 is marked by an asterisk.

and the strongly end-stopped neuron from Figure 4 is marked by an asterisk.

Effects of Changing Orientation

Most end-stopped cells continue to respond even when the stimulus orientation is far from optimal (Hammond and Pomfrett, 1989; Henry et al., 1974; Schiller et al., 1976). Furthermore, in these cells, the property of end-stopping is largely independent of orientation (Orban et al., 1979). It therefore seems likely that cells whose receptive fields are both direction selective and end-stopped might respond to the endpoints of bars in a manner that is independent of bar orientation. We tested this possibility by repeating the sparse noise experiment on the cell shown in Figure 4, but using a bar orientation that was rotated 45° clockwise. The second row of Figure 4A shows the one-bar map for this experiment. Again, the cell responded well when the endpoints of the bar were in the receptive field, and poorly when the bar was centered on the receptive field.

The two-bar map in the second row of Figure 4B

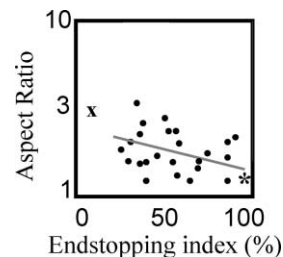


Figure 5. Relationship between End-Stopping and the Elongation of the Facilitatory Regions in the Two-Bar Maps

Each dot represents one neuron. The x axis indicates the degree of end-stopping, and the y axis indicates the degree of elongation of the facilitatory regions in the two-bar maps. The neuron shown in Figure 3 is marked by an X, and the neuron shown in Figure 4 is marked by an asterisk.

shows a similar direction preference to that shown in Figure 4A, and to the direction selectivity observed with conventional bar stimuli moving perpendicularly to their orientation (Figure 4C). The neuron responded well to down-left motion and poorly to other directions. The third row of Figure 4 shows the results of rotating the bars in the noise stimulus 45° counterclockwise. Again, the cell responded well to the endpoints and to down-left motion sequences. It is also interesting to note that the positions of the suppressive (blue) regions in the two-bar maps varied with bar orientation. In some cases, the location of the suppressive region seemed to result from end suppression, as in the top row of Figure 4B. Here the neuron was suppressed for motion parallel to the orientation of the bar, presumably because such displacements brought the center of the long bar into the receptive field. In contrast, for the counterclockwise rotation (bottom row of Figure 4B), the neuron was suppressed when the bar moved in the null direction. In the middle row of Figure 4B, the position of the suppressive region is not what would be expected based on surround suppression or null-direction inhibition, and may involve the interaction of multiple nonlinearities. The two-bar maps for this neuron were typical of the population in that the facilitatory regions were better predictors of the neuron's directional preference than the inhibitory regions.

The two-bar interactions shown in Figure 4B suggest that the neuron responded to a narrow range of motion directions, and that this range was similar across stimulus orientations. However, there were clearly some differences in the peak positions of the facilitatory regions for different stimulus orientations, as can be seen by comparing the top and bottom rows of Figure 4B. To understand how common these differences were across the population, we obtained two-bar maps at more than one orientation for 29 end-stopped V1 cells. To quantify the effect of orientation on the directional response, we used the peak position of the Gaussian fits to the two-bar maps as one measure of each cell's preferred direction (PD). For instance, the position of the red region in Figure 4B suggests that the preferred direction was down and to the left. For each cell, we also calculated the mean vector of the conventional direction-tuning curve, as a second measure of the PD. In Figure 4C, the mean vector is represented as the red line pointing downward and leftward. We then compared the PDs derived from two-bar maps with the mean vectors obtained from conventional direction-tuning curves. We have shown previously that these measures are well correlated for noise maps obtained with small spot stimuli (Pack et al., 2003).

Figure 6A plots the PDs obtained from conventional bar sweeps against those predicted from two-bar maps in which the bar orientation was perpendicular to each cell's preferred direction. The correlation is significant (angular-angular correlation [Mardia, 1975]; $p < 0.001$), with a median absolute angular difference of 24.4°. For the population, the circular standard deviation (cSD) (Mardia, 1975) of the angular difference was 42.8°. A histogram of the absolute angular differences is shown in the lower right corner of Figure 6A. Similar results were obtained for the instances in which the bar was rotated clockwise by 45° from the optimal orientation

(Figure 6B). Here the median angular difference between the PD derived from the noise map and that obtained from the conventional tuning curve was 21.7° (cSD 52.9°). This correlation was also highly significant (angular-angular correlation, $p < 0.0001$). When the bar was rotated 45° counterclockwise, similar results were obtained (Figure 6C; median difference 16.4°; cSD 35.6°; $p < 0.0001$). These results suggest that the directional responses of end-stopped V1 neurons are largely independent of stimulus orientation.

Temporal Dynamics of End-Stopping

Many surround effects in V1 take time to develop (Hupe et al., 2001; Knierim and van Essen, 1992; Lamme et al., 1998). That is, early responses generally reflect only the stimulus that is in the "classical receptive field," while later responses take into account the larger stimulus context. Because end-stopping involves an interaction between the receptive field center and surrounding regions, it seems likely that end-stopping might exhibit a similar time course. To investigate this issue, we computed one-bar maps like those in Figures 3 and 4 at different correlation delays. Figure 7A shows successive time slices of one-bar maps obtained from the cell in Figure 4. At early correlation delays, the cell responded well to a long bar, even when the bar was centered on the receptive field. The end inhibition was only apparent approximately 20 ms later, as evidenced by the gradual emergence of the "dumbbell"-shaped red regions. Below these maps are shown the averaged temporal response to every bar that was presented with its center over the receptive field (green line), or positioned so that only an endpoint was in the receptive field (black line). The baseline firing has been subtracted from this plot. Figure 7B shows a delay in end-stopping for the time course of the response averaged across the V1 population. Thus, similar to other surround effects, end-stopping requires some small amount of time (20–30 ms) to become fully manifest.

Figure 8A shows the temporal evolution of the two-bar map for the same neuron shown in Figures 4 and 7. An early response (at 50 ms) shows an elongated region of facilitation, with a prominent response near the topmost portion of the map, slightly to the left of the vertical meridian. This suggests a strong response for motion down and to the right. In other words, at short time intervals, the neuron was fooled by motion parallel to the orientation of the bar, as was the weakly end-stopped neuron in Figure 3. However, over time, this facilitatory region faded, and eventually was replaced by suppression, as indicated by the blue regions that flank the red regions. This later part of the response is consistent with the neuron's conventional direction-tuning curve (rightmost panel), which showed a preference for motion down and to the left.

MT Neurons

V1 neurons project to numerous cortical areas. One of these, the middle temporal area (MT or V5) is highly specialized for computing motion, and many MT cells respond to moving stimuli in a manner that is independent of the orientations of the edges that make up the stimuli (Movshon et al., 1986; Pack et al., 2001; Pack

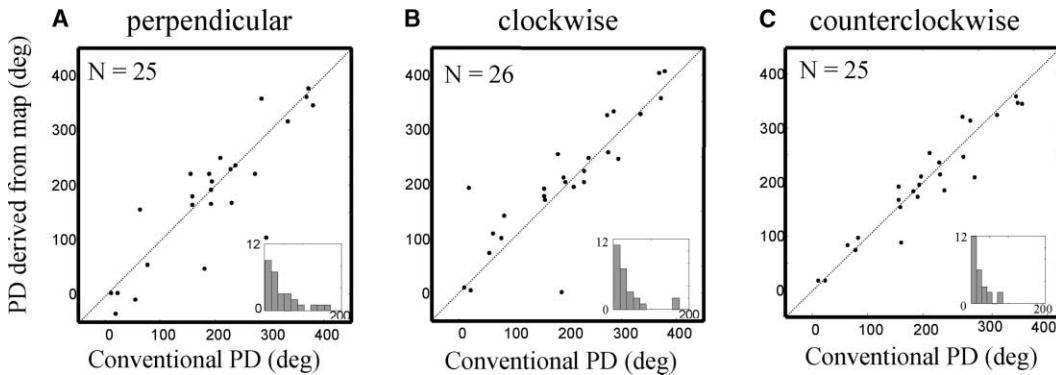


Figure 6. Consistency of V1 Directional Responses

Preferred directions were derived from conventional tuning curves and compared with those derived from Gaussian fits to the sparse noise maps. Each data point represents one neuron. The three columns show these comparisons for three different bar orientations. The small histograms at the bottom right of each plot indicate the absolute values of the angular differences between the two measures of preferred direction.

and Born, 2001; Rodman and Albright, 1989). Previous results have suggested that the integration of motion signals in MT is a gradual process. That is, when a stimulus moves through the receptive field of an MT neuron, the cell responds with a stereotyped time course: the early responses reflect predominantly the component of motion perpendicular to the orientation of edges, just as the cell in Figure 3 only encodes one dimension of stimulus motion. However, the later responses in MT are more like those in Figure 4, in that they respond to the correct direction of motion irrespective of the shape of the stimulus (Pack et al., 2001; Pack and Born, 2001). These previous experiments were conducted with stimuli consisting of plaids or fields of bars.

In light of our current findings, one possible explana-

tion for the temporal dynamics in MT is that end-stopping in V1 takes time to develop. Figure 8B shows an example of an MT cell that showed a similar temporal evolution of the two-bar map for the sparse noise stimulus. The cell at first responded to any motion sequence that contained a component in its preferred direction (downward). At longer time intervals it responded only to straight downward motion and was unresponsive to directions that contained a substantial horizontal component.

We quantified the directional interactions in 17 MT neurons by fitting the two-bar maps with an oriented Gaussian function. The geometric mean of the aspect ratios of the map subregions was 1.65 ± 0.54 , somewhat smaller than those in V1 end-stopped neurons. For each

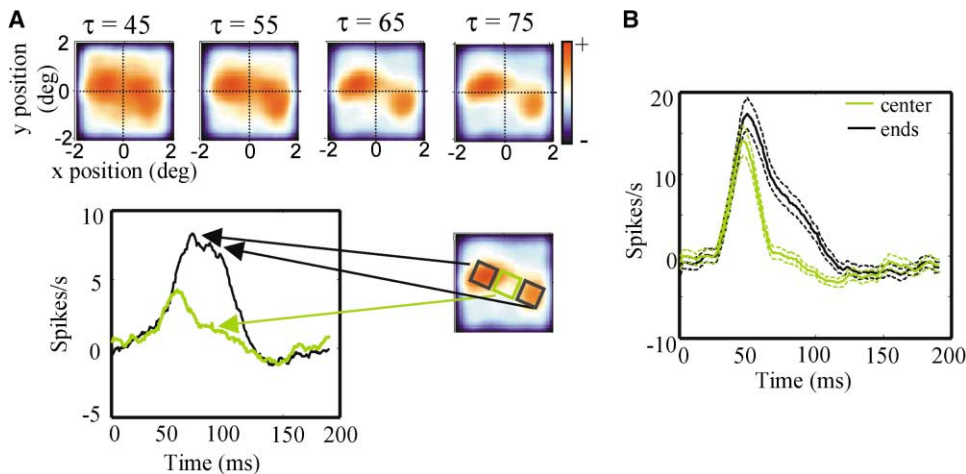


Figure 7. Timing of V1 End-Stopping

(A) The first row shows successive snapshots of the one-bar map for a strongly end-stopped V1 cell, for the preferred bar orientation. The cell initially responds well to a bar placed anywhere in its receptive field, but after 20–30 ms, it responds only when the endpoints are in the receptive field. The second row shows the time course of the response when the analysis was limited to specific regions of the stimulus range. When the analysis was limited to instances when the bar was centered on the receptive field (green square), the neuron responded transiently (green line). When the analysis was limited to instances when the endpoint of the bar appeared in the receptive field (black square), the cell exhibited a strong maintained discharge (black line).

(B) The time course of the response, averaged across the population of 29 neurons for the preferred bar orientation. Dotted lines indicate standard error of the mean.

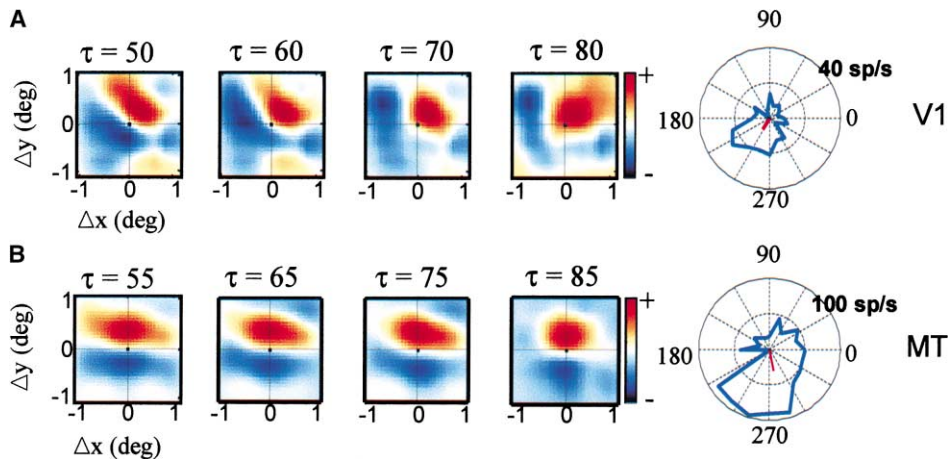


Figure 8. Timing of Two-Bar Responses
(A) Successive snapshots of the two-bar map for a V1 cell, with the bar orientation perpendicular to the preferred direction. The polar histogram on the right shows the neuron's conventional direction-tuning curve.
(B) Same as (A), but for an MT cell.

neuron, we measured conventional tuning curves with moving dot fields and compared the preferred direction from these tuning curves with those predicted from the peak positions of the Gaussians. For the maps obtained with bars that were perpendicular to each cell's preferred orientation, the median difference between the predicted and actual PD was 18.6° , with a circular standard deviation of 16.4° . For bars rotated 45° clockwise, the angular difference averaged 17.3° (cSD 14.0°). For counterclockwise rotations, the difference averaged 19.9° (cSD 19.7°). These median errors for bar stimuli in MT were similar to those obtained in a previous study (Pack et al., 2003) using small spot stimuli. On average, the magnitudes of the directional differences appear to be similar between V1 and MT, although there is much more variability in V1.

Discussion

We have shown that end-stopped V1 neurons (those whose responses are suppressed by extended contours) are capable of encoding two-dimensional motion information that is independent of the stimulus orientation. Our findings also demonstrate that end-stopping, like other contextual influences (Hupe et al., 2001; Knierim and van Essen, 1992; Lamme et al., 1998), takes time to develop. Consequently, end-stopped V1 neurons initially respond to a long bar placed anywhere in their receptive fields, but after 20–30 ms they only respond to the endpoints of the bar. Similarly, MT neurons initially respond to all of the motion signals generated by a moving bar, but later responses reflect predominantly the motion of the endpoints.

The measurement of local motion direction can be described as a problem of calculating the correct correspondence between points in successive views of a moving object (Marr, 1982). A similar statement can be made about the measurement of retinal disparity, in which the goal is to match corresponding points in the two eyes. Because features in the visual input exist in

two dimensions, a receptive field that is sensitive to two-dimensional luminance discontinuities is useful for solving both types of problems. End-stopped neurons have previously been shown to serve this purpose for retinal disparity (Maske et al., 1986), and our results suggest a similar competency for motion.

Our results suggest that end-stopped V1 neurons measure two-dimensional correspondence in motion by responding preferentially to line endings. An obvious question is whether these cells also measure the motion of other orientation discontinuities, such as corners, T junctions, and intersections. While it seems clear that many V1 neurons respond to static orientation discontinuities (Kapadia et al., 1999; Knierim and van Essen, 1992; Sillito et al., 1995), it is still not clear to what extent this selectivity is used for motion processing. Early studies of V1 responses to moving plaid stimuli (Movshon et al., 1986) found little support for the encoding of intersections, but the nature of the stimuli used to perform these studies—extended gratings and plaids—may have selected against the characterization of the most end-stopped neurons. More recent work suggests that the V1 responses to plaids depend on receptive field shape (Tinsley et al., 2003), but as far as we know, no study has related plaid responses to the property of end-stopping. Furthermore, both plaid responses (Pack et al., 2001) and surround interactions (Bolz and Gilbert, 1986; Lamme et al., 1998) can be affected by certain types of general anesthetics, so it would be interesting to determine how these properties are related in V1 of the alert animal.

In describing our cells as end-stopped, we do not mean to imply that end-stopped neurons form a separate class in V1. Rather, it appears that most V1 neurons exhibit some end-stopping, and more generally, surround suppression. Although there is a great deal of variability in V1, a large, high-contrast stimulus generally drives neurons about half as effectively as a stimulus that is confined to the receptive field center (Jones et al., 2001; Sceniak et al., 2001). The prevalence of end-

stopping varies across layers, with neurons in layer 4B being the most strongly end-stopped, and layer 6 neurons being rarely end-stopped (Sceniak et al., 2001).

The laminar dependence of surround suppression is of interest for motion processing because the projections from V1 to MT originate in layers 4B and 6 of V1, with 4B accounting for more than 90% of the V1 afferents (Maunsell and van Essen, 1983; Shipp and Zeki, 1989). Layer 4B also contains a high incidence of direction-selective neurons (Hawken et al., 1988). It therefore seems reasonable to infer that a substantial portion of the input to MT neurons is end-stopped. For stimuli that contain endpoints, our results suggest that these inputs would provide two-dimensional motion signals. MT neurons could therefore measure object motion by performing a vector average of these two-dimensional inputs (Groh et al., 1997; Recanzone and Wurtz, 2000; Rubin and Hochstein, 1993). Our observations on the temporal dynamics of end-stopping could then account for at least some of the MT temporal dynamics we observed in this work and elsewhere (Pack et al., 2001; Pack and Born, 2001). Of course, this interpretation is complicated by the fact that some of the surround suppression in V1 is dependent on feedback from MT (Angelucci et al., 2002; Hupe et al., 1998). The timing we observed in MT could also reflect top-down influences or interactions within MT, which are thought to be useful for resolving conflicting local motion signals (Grossberg et al., 2001; Liden and Pack, 1999; Nowlan and Sejnowski, 1995; Recanzone and Wurtz, 2000; Wilson et al., 1992). Our limited observations on weakly end-stopped V1 neurons suggest that conflicting local signals are not resolved within V1. That is, non-end-stopped neurons suffer from the aperture problem, even at long time intervals (C.C.P. and R.T.B., unpublished data).

To the extent that neurons in layer 6 of V1 are weakly end-stopped (Jones et al., 2001; Sceniak et al., 2001), our results suggest that the projection from layer 6 to MT would mostly carry one-dimensional motion information. To use this information, MT neurons could perform something like an intersection-of-constraints computation (Simoncelli and Heeger, 1998) to recover two-dimensional object motion. The intersection-of-constraints calculation could also be useful for integrating motion signals generated by stimuli of low contrast, for which surround suppression is reduced (Hupe et al., 1998; Kapadia et al., 1999; Levitt and Lund, 1997; Sceniak et al., 1999). The large receptive fields of neurons in layer 6 also have been implicated in the generation of end-stopping in layer 4 of V1 (Bolz and Gilbert, 1986).

The functional role of V1 neurons has often been interpreted in terms of edge detection, based on their pronounced orientation selectivity. In contrast, end-stopped neurons have been shown to signal curved (Dobbins et al., 1987; Hubel and Wiesel, 1965; Versavel et al., 1990) or discontinuous (Hubel and Wiesel, 1965; Versavel et al., 1990) stimulus regions. Theoretical considerations suggest that this type of selectivity is useful for shape recognition (Attneave, 1954; Pasupathy and Connor, 1999; Schwartz et al., 1983), depth perception (Heitger et al., 1992), and motion processing (Barth,

2000; Lorenceau et al., 1993; Rubin et al., 1995). Our results provide direct support for the latter hypothesis.

Experimental Procedures

Electrophysiology

Monkeys were prepared for chronic recording from V1 and MT, as described elsewhere (Born et al., 2000; Livingstone, 1998). All procedures were approved by the Harvard Medical Area Standing Committee on Animals.

We recorded from single units in V1 and MT of two alert rhesus macaque monkeys while they performed a fixation task. The monkeys were rewarded for maintaining fixation within 1° of a small fixation spot. Each single unit was isolated by spike height and waveform.

Visual Stimuli

Neurons were screened with a bar stimulus. The bar luminance was 39 cd/m² on a gray background (20 cd/m²). We first estimated the optimal bar length and then obtained a direction-tuning curve by sweeping the bar through the receptive field. Length-tuning curves were obtained by sweeping the bar through the receptive field at the optimal direction and speed. Neurons that were both end-stopped (suppressed by at least 30%) and direction selective (direction index > 0.5) were studied further.

The sparse noise stimuli were identical to those used previously (Livingstone et al., 2001). All cells were studied with two-dimensional white noise stimuli, consisting of pairs of bars flashed at 75 Hz. We used white and black stimuli that were 19 cd/m² above and below the mean background gray luminance of 20 cd/m². When black and white stimuli overlapped, the resultant stimulus was the same gray as the background.

Analysis

The reverse correlation analysis used here is identical to the analysis used in previous studies (Livingstone et al., 2001). Briefly, a computer recorded the evoked spike train (1 ms resolution), each stimulus position, and the monkey's eye position (4 ms resolution). For each map, between 5,000 and 30,000 spikes were collected over a 20–40 min period. Spikes were reverse correlated with the stimulus sequence to obtain the one-bar and two-bar maps. For the one-bar maps, we correlated the spike train with the position of the white or black bar on the frame that preceded each spike by a given delay. Because complex cell responses are not dependent on contrast sign, we averaged the white and black bar maps to obtain the one-bar maps. For the two-bar maps, the spike train was correlated with the relative positions of sequential stimuli at a delay corresponding to the peak of the response to the second stimulus (typically between 40 and 60 ms for V1 cells). The analysis thus related the spike train to the difference in retinal position between sequential pairs of bars. The two-bar maps were calculated by summing the same-contrast individual maps (white-to-white and black-to-black), and subtracting the different-contrast maps (white-to-black and black-to-white), as described in Livingstone et al. (2001). This is similar to a second-order Wiener-like calculation (Emerson et al., 1987). One- and two-bar maps were normalized to the maximal response in each map, in order to display the structure of the interactions independently of the amplitude of the response.

Gaussian fits to the two-bar maps were optimized using a least-squares criterion with the Levenberg-Marquardt algorithm in Matlab (Mathworks, Natick, MA). The aspect ratio was defined as the ratio of the standard deviations of the Gaussian along axes parallel and perpendicular to the orientation of the bar stimuli used to generate each map.

To compute the impulse responses for the one-bar maps (Figure 7), we generated a series of maps at 1 ms intervals. These maps were restricted to subregions of the stimulus range, each of which was a square with sides of length 0.7°. The orientation of the square was aligned with that of the stimulus bars. One subregion was intended to capture the instances when the bar was centered on the receptive field. To the extent that the stimulus range was perfectly centered on the receptive field, the center of this region corresponded to the origin of the one-bar map. In most cases, however,

the stimulus was not perfectly centered. This was corrected manually by centering a mouse-controlled cursor on the center of mass of the excitatory regions of the one-bar map. The other subregion was displaced from this point by half the length of the stimulus bar, in a direction parallel to the orientation of the bar. This corresponded to instances when one of the endpoints of the bar was in the receptive field. The responses from the two endpoints were averaged together.

Acknowledgments

David Freeman developed all of the computer programs. Tamara Chuprina and Phillip Hendrickson provided excellent technical assistance. We thank David Hubel and Aaron Seitz for helpful discussions. Supported by NIH NS07484 (C.C.P.), NIH EY13135 (M.S.L.), and NIH EY11379 and EY12196 (R.T.B.).

Received: March 31, 2003

Revised: May 14, 2003

Accepted: July 3, 2003

Published: August 13, 2003

References

- Angelucci, A., Levitt, J.B., Walton, E.J., Hupe, J.M., Bullier, J., and Lund, J.S. (2002). Circuits for local and global signal integration in primary visual cortex. *J. Neurosci.* 22, 8633–8646.
- Attneave, F. (1954). Some informational aspects of visual perception. *Psychol. Rev.* 61, 183–193.
- Barth, E. (2000). A geometric view on early and middle level visual coding. *Spat. Vis.* 13, 193–199.
- Bolz, J., and Gilbert, C.D. (1986). Generation of end-inhibition in the visual cortex via interlaminar connections. *Nature* 320, 362–365.
- Born, R.T., Groh, J.M., Zhao, R., and Lukasewycz, S.J. (2000). Segregation of object and background motion in visual area MT: effects of microstimulation on eye movements. *Neuron* 26, 725–734.
- Conway, B.R., and Livingstone, M.S. (2003). Space-time maps and two-bar interactions of different classes of direction-selective cells in macaque v-1. *J. Neurophysiol.* 89, 2726–2742.
- DeAngelis, G.C., Freeman, R.D., and Ohzawa, I. (1994). Length and width tuning of neurons in the cat's primary visual cortex. *J. Neurophysiol.* 71, 347–374.
- Dobbins, A., Zucker, S.W., and Cynader, M.S. (1987). Endstopped neurons in the visual cortex as a substrate for calculating curvature. *Nature* 329, 438–441.
- Emerson, R.C., Citron, M.C., Vaughn, W.J., and Klein, S.A. (1987). Nonlinear directionally selective subunits in complex cells of cat striate cortex. *J. Neurophysiol.* 58, 33–65.
- Groh, J.M., Born, R.T., and Newsome, W.T. (1997). How is a sensory map read Out? Effects of microstimulation in visual area MT on saccades and smooth pursuit eye movements. *J. Neurosci.* 17, 4312–4330.
- Grossberg, S., Mingolla, E., and Viswanathan, L. (2001). Neural dynamics of motion integration and segmentation within and across apertures. *Vision Res.* 41, 2521–2553.
- Hammond, P., and Pomfrett, C.J. (1989). Directional and orientational tuning of feline striate cortical neurones: correlation with neuronal class. *Vision Res.* 29, 653–662.
- Hawken, M.J., Parker, A.J., and Lund, J.S. (1988). Laminar organization and contrast sensitivity of direction-selective cells in the striate cortex of the Old World monkey. *J. Neurosci.* 8, 3541–3548.
- Heitger, F., Rosenthaler, L., von der Heydt, R., Peterhans, E., and Kubler, O. (1992). Simulation of neural contour mechanisms: from simple to end-stopped cells. *Vision Res.* 32, 963–981.
- Henry, G.H., Bishop, P.O., and Dreher, B. (1974). Orientation, axis and direction as stimulus parameters for striate cells. *Vision Res.* 14, 767–777.
- Hubel, D.H., and Wiesel, T.N. (1965). Receptive fields and functional architecture in two non-striate visual areas (18 and 19) of the cat. *J. Neurophysiol.* 28, 229–289.

- Hubel, D.H., and Wiesel, T.N. (1968). Receptive fields and functional architecture of monkey striate cortex. *J. Physiol.* 195, 215–243.
- Hupe, J.M., James, A.C., Payne, B.R., Lomber, S.G., Girard, P., and Bullier, J. (1998). Cortical feedback improves discrimination between figure and background by V1, V2 and V3 neurons. *Nature* 394, 784–787.
- Hupe, J.M., James, A.C., Girard, P., and Bullier, J. (2001). Response modulations by static texture surround in area V1 of the macaque monkey do not depend on feedback connections from V2. *J. Neurophysiol.* 85, 146–163.
- Jones, H.E., Grieve, K.L., Wang, W., and Sillito, A.M. (2001). Surround suppression in primate V1. *J. Neurophysiol.* 86, 2011–2028.
- Kapadia, M.K., Westheimer, G., and Gilbert, C.D. (1999). Dynamics of spatial summation in primary visual cortex of alert monkeys. *Proc. Natl. Acad. Sci. USA* 96, 12073–12078.
- Knierim, J.J., and van Essen, D.C. (1992). Neuronal responses to static texture patterns in area V1 of the alert macaque monkey. *J. Neurophysiol.* 67, 961–980.
- Lamme, V.A., Zipser, K., and Spekreijse, H. (1998). Figure-ground activity in primary visual cortex is suppressed by anesthesia. *Proc. Natl. Acad. Sci. USA* 95, 3263–3268.
- Levitt, J.B., and Lund, J.S. (1997). Contrast dependence of contextual effects in primate visual cortex. *Nature* 387, 73–76.
- Liden, L., and Pack, C. (1999). The role of terminators and occlusion cues in motion integration and segmentation: a neural network model. *Vision Res.* 39, 3301–3320.
- Livingstone, M.S. (1998). Mechanisms of direction selectivity in macaque V1. *Neuron* 20, 509–526.
- Livingstone, M.S., Pack, C.C., and Born, R.T. (2001). Two-dimensional substructure of MT receptive fields. *Neuron* 30, 781–793.
- Lorençeau, J., Shiffrar, M., Wells, N., and Castet, E. (1993). Different motion sensitive units are involved in recovering the direction of moving lines. *Vision Res.* 33, 1207–1217.
- Mardia, K.V. (1975). Statistics of directional data. *J. R. Stat. Soc. B* 34, 102–113.
- Marr, D. (1982). *Vision* (New York: W.H. Freeman & Co.).
- Maske, R., Yamane, S., and Bishop, P.O. (1986). End-stopped cells and binocular depth discrimination in the striate cortex of cats. *Proc. R. Soc. Lond. B. Biol. Sci.* 229, 257–276.
- Maunsell, J.H., and van Essen, D.C. (1983). The connections of the middle temporal visual area (MT) and their relationship to a cortical hierarchy in the macaque monkey. *J. Neurosci.* 3, 2563–2586.
- Mingolla, E., Todd, J.T., and Norman, J.F. (1992). The perception of globally coherent motion. *Vision Res.* 32, 1015–1031.
- Movshon, J.A., and Newsome, W.T. (1996). Visual response properties of striate cortical neurons projecting to area MT in macaque monkeys. *J. Neurosci.* 16, 7733–7741.
- Movshon, J.A., Adelson, E.H., Gizzi, M.S., and Newsome, W.T. (1986). The analysis of moving visual patterns. *Exp. Brain Res.* 11, 117–151.
- Nowlan, S.J., and Sejnowski, T.J. (1995). A selection model for motion processing in area MT of primates. *J. Neurosci.* 15, 1195–1214.
- Orban, G.A., Kato, H., and Bishop, P.O. (1979). End-zone region in receptive fields of hypercomplex and other striate neurons in the cat. *J. Neurophysiol.* 42, 818–832.
- Pack, C.C., and Born, R.T. (2001). Temporal dynamics of a neural solution to the aperture problem in visual area MT of macaque brain. *Nature* 409, 1040–1042.
- Pack, C.C., Berezovskii, V.K., and Born, R.T. (2001). Dynamic properties of neurons in cortical area MT in alert and anaesthetized macaque monkeys. *Nature* 414, 905–908.
- Pack, C.C., Born, R.T., and Livingstone, M.S. (2003). Two-dimensional substructure of motion and stereo interactions in primary visual cortex of alert macaque. *Neuron* 37, 525–535.
- Pasupathy, A., and Connor, C.E. (1999). Responses to contour features in macaque area V4. *J. Neurophysiol.* 82, 2490–2502.
- Recanzone, G.H., and Wurtz, R.H. (2000). Effects of attention on MT

- and MST neuronal activity during pursuit initiation. *J. Neurophysiol.* **83**, 777–790.
- Rodman, H.R., and Albright, T.D. (1989). Single-unit analysis of pattern-motion selective properties in the middle temporal visual area (MT). *Exp. Brain Res.* **75**, 53–64.
- Rubin, N., and 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.
- Rubin, N., Solomon, S., and Hochstein, S. (1995). Restricted ability to recover three-dimensional global motion from one-dimensional local signals: theoretical observations. *Vision Res.* **35**, 569–578.
- Sceniak, M.P., Ringach, D.L., Hawken, M.J., and Shapley, R. (1999). Contrast's effect on spatial summation by macaque V1 neurons. *Nat. Neurosci.* **2**, 733–739.
- Sceniak, M.P., Hawken, M.J., and Shapley, R. (2001). Visual spatial characterization of macaque V1 neurons. *J. Neurophysiol.* **85**, 1873–1887.
- Schiller, P.H., Finlay, B.L., and Volman, S.F. (1976). Quantitative studies of single-cell properties in monkey striate cortex. V. Multivariate statistical analyses and models. *J. Neurophysiol.* **39**, 1362–1374.
- Schwartz, E.L., Desimone, R., Albright, T.D., and Gross, C.G. (1983). Shape recognition and inferior temporal neurons. *Proc. Natl. Acad. Sci. USA* **80**, 5776–5778.
- Shipp, S., and Zeki, S. (1989). The organization of connections between areas V5 and V1 in macaque monkey visual cortex. *Eur. J. Neurosci.* **1**, 309–332.
- Sillito, A.M., Grieve, K.L., Jones, H.E., Cudeiro, J., and Davis, J. (1995). Visual cortical mechanisms detecting focal orientation discontinuities. *Nature* **378**, 492–496.
- Simoncelli, E.P., and Heeger, D.J. (1998). A model of neuronal responses in visual area MT. *Vision Res.* **38**, 743–761.
- Tinsley, C.J., Webb, B.S., Barraclough, N.E., Vincent, C.J., Parker, A., and Derrington, A.M. (2003). The nature of V1 neural responses to 2D moving patterns depends upon receptive field structure in the Marmoset Monkey. *J. Neurophysiol.*, in press.
- Versavel, M., Orban, G.A., and Lagae, L. (1990). Responses of visual cortical neurons to curved stimuli and chevrons. *Vision Res.* **30**, 235–248.
- Wallach, H. (1935). Über visuell wahrgenommene Bewegungsrichtung. *Psychol. Forsch.* **20**, 325–380.
- Wilson, H.R., Ferrera, V.P., and Yo, C. (1992). A psychophysically motivated model for two-dimensional motion perception. *Vis. Neurosci.* **9**, 79–97.
- Wuerger, S., Shapley, R., and Rubin, N. (1996). "On the visually perceived direction of motion" by Hans Wallach: 60 years later. *Perception* **25**, 1317–1367.