

Attentional Modulation of Motion Integration of Individual Neurons in the Middle Temporal Visual Area

Erik P. Cook and John H. R. Maunsell

Howard Hughes Medical Institute and Division of Neuroscience, Baylor College of Medicine, Houston, Texas 77030

We examined how spatially directed attention affected the integration of motion in neurons of the middle temporal (MT) area of visual cortex. We recorded from single MT neurons while monkeys performed a motion detection task under two attentional states. Using 0% coherent random dot motion, we estimated the optimal linear transfer function (or kernel) between the global motion and the neuronal response. This linear kernel filtered the random dot motion across direction, speed, and time. Slightly less than one-half of the neurons produced reasonably well defined kernels that also tended to account for both the directional selectivity and responses to coherent motion of different strengths. This subpopulation of cells had faster, more transient, and more robust responses to visual stimuli than neurons with kernels that did not contain well defined regions of integration. For those neurons that had large attentional modulation and produced well defined kernels, we found attention scaled the temporal profile of the transfer function with no appreciable shift in time or change in shape. Thus, for MT neurons described by a linear transfer function, attention produced a multiplicative scaling of the temporal integration window.

Key words: attention; vision; MT; motion; macaque monkey; reverse correlation

Introduction

Neurons throughout the visual cortex are modulated by spatial attention (Desimone and Duncan, 1995; Braun et al., 2001; Treue, 2003). The number of action potentials produced by a neuron typically increases when attention is directed to the region in space that coincides with the receptive field (RF) of a neuron. How this modulation affects the way neurons integrate visual stimuli is currently an active area of investigation (McAdams and Maunsell, 1999; Treue and Martinez-Trujillo, 1999; Reynolds et al., 2000; Fries et al., 2001; Martinez-Trujillo and Treue, 2002; Niebur et al., 2002). In this study, we used linear systems identification to understand how attention modulates the temporal integration of neurons in the middle temporal (MT) visual area.

Recent studies suggest that attention operates in a multiplicative manner and does not change stimulus selectivity (McAdams and Maunsell, 1999; Treue and Martinez-Trujillo, 1999; Recanzone and Wurtz, 2000). Thus, attentional modulation (AM) can be thought of as a gain change that scales the tuning curves of cortical neurons without altering their shape. This observation has important implications for the way AM of neuronal activity produces attentional effects on the behavior (Cook and Maunsell, 2002a). What is not known, however, is the extent to which the

multiplicative action of attention also applies to other aspects of neuronal processing, such as temporal integration.

One possibility is that attention increases the sensitivity of neurons without altering the time course of their response. This would result in the gain change that has been observed in tuning curves. Alternatively, because tuning curves are typically measured by counting spikes over hundreds of milliseconds, it is possible that attention alters both sensitivity and the time course of the neuronal response. Biophysical mechanisms that may underlie a change in the gain of neuronal responses, such as the opening or closing of membrane conductances, would also produce changes in the membrane time constant that could lead to changes in the temporal integration of synaptic inputs (Rall, 1959). Thus, attention might have measurable effects on the time course of neuronal responses. Knowing whether or not attention affects the temporal integration properties of cortical neurons is important for understanding the way attention exerts its influence on neuronal activity and has not been previously examined in detail.

We recorded the activity of single neurons in area MT in response to random dot motion from monkeys under two attentional states. We used the correlation between the motion stimulus and neuronal response to estimate the way the neurons integrated the motion in our random dot stimulus. Neurons that produced linear kernels with well defined regions of integration had distinctly different temporal response properties compared with cells that produced kernels with less-defined integration regions. For the subset of MT neurons that produced good linear kernels and experienced high AM, we found that attention scaled the temporal integration profile without changing its shape. This suggests that attention exerts a multiplicative effect on the temporal integration properties of MT neurons.

Received Nov. 18, 2003; revised July 28, 2004; accepted July 28, 2004.

This work was supported by the Howard Hughes Medical Institute (HHMI) and National Institutes of Health Grant R01 EY05911. J.H.R.M. is an Investigator with the HHMI. We thank W. Bosking, C. Boudreau, J. DiCarlo, G. Ghose, and T. Yang for helpful discussion on all aspects of this project. We also thank J. A. Movshon for comments on this manuscript, X. Pitkow for help with the data analysis, and D. Murray and T. Williford for technical assistance.

Correspondence should be addressed to Dr. Erik P. Cook, Department of Physiology, McGill University, 3655 Sir William Osler, Montreal, Quebec, H3G 1Y6 Canada. E-mail: erik.cook@mcgill.ca.

DOI:10.1523/JNEUROSCI.5102-03.2004

Copyright © 2004 Society for Neuroscience 0270-6474/04/247964-14\$15.00/0

Materials and Methods

Behavioral task. Two monkeys (*Macaca mulatta*) were trained to perform a motion detection task. While the animal fixated on a central point, two patches of random motion were presented (see Fig. 1A). Initially, there was no net motion (0% coherent) in the stimulus, and at a random time between 500 and 8000 msec, coherent motion began in one of the patches (motion onset times were exponentially distributed). Once coherent motion began, the monkeys had a reaction time window from 200 to 750 msec in which to release the lever to obtain a juice reward. The strength of the coherent motion was varied between three preset levels (low, medium, and high) that spanned the animal's detection threshold. The proportion of trials the animals correctly detected the coherent motion across all experiments was 50, 92, and 99% for the low, medium, and high coherences, respectively. The location and size of one patch was set to overlap the RF of the neuron under study, and the other was diametrically placed on the other side of the fixation point. The direction and speed of the coherent motion was matched to the preferred direction and speed of the neuron. Directional tuning curves were constructed using the motion detection task with high (50%) coherent motion in one of eight directions. Only trials in which the animals maintained fixation within 1° of the fixation point were analyzed. Trials in which the lever was released before the reaction time window were discarded.

At the beginning of each trial, a static cue of stationary dots was presented to indicate which patch was most likely to contain the coherent motion. The animals covertly directed their attention to the cued location during the trial. We refer to the condition in which the animal was cued to direct its attention to the patch of dots in the RF as the "attend-in" condition, whereas attention directed to the opposite patch is referred to as the "attend-out" condition. To verify that the animals were covertly directing their attention to the cued patch of dots, we used an invalid cue in 20% of the trials (Posner, 1980). Motion detection was better and reaction times were faster in the valid cueing conditions, when the animals were instructed to direct their attention to the patch of dots that contained the coherent motion stimulus (Cook and Maunsell, 2002a). For example, when the coherent motion occurred in the uncued patch, the animal's average detection performance dropped from 92 to 59%, indicating the animal used the spatial cue to direct its attention to the patch of dots that would most likely contain the coherent motion. Invalid cueing only occurred for trials containing the medium strength coherent motion. The static cue was presented in the same location for blocks of 15 completed trials.

The monkeys were also trained to perform a standard memory-delayed saccade task (White and Sparks, 1986). In this task, the monkey fixated on a central point while a peripheral target (0.25° in diameter) appeared for 500 msec. To obtain a reward, the monkey had to remember the target location for 500–2500 msec and then, after the central fixation point was extinguished, saccade to within 2.5° of its location within 300 msec.

Visual stimulus. The animal sat 62 cm from a computer monitor ($\pm 17 \times \pm 13^\circ$ of visual angle; 1600×1200 pixels; 75 Hz refresh). The stimuli consisted of two patches of white dots (each 0.25° diameter; 78 cd/m²) on a dark gray background (12 cd/m²) with a dot density of 2.1 dots/degree². Each patch of dots was updated every other frame (i.e., every 26.6 msec) using the following procedure. The dots in each patch were evenly divided into two groups. On each update, one group was replaced with new, randomly positioned dots, whereas dots in the other group were displaced by a fixed distance. The dots in this latter group determined the motion coherence. For 0% coherence, all the dots in this group moved a fixed distance in a random direction. For coherent motion greater than zero, a proportion of the dots moved with a fixed distance in the same direction. This proportion determined the strength of the coherent motion. On the next update (26.6 msec later), the groups were switched. This arrangement insured that all the dots had a lifetime of four video frames (two updates or 53.2 msec) before they were replaced and that there would be no changes in the apparent dot density associated with the onset of coherent motion. Because half the dots are always randomly replotted, regardless of the proportion of dots moving coherently, our motion had a maximum strength of 50% coherent. For

example, at 25% coherent motion, half the dots are randomly replotted, one-quarter are moving with the same fixed distance and direction, and one-quarter are moving with the same fixed distance in a random direction.

Data collection. Using standard extracellular recording techniques, we recorded from single neurons in area MT in both animals. When a neuron was isolated, the RF was mapped using a manually controlled bar while the animal fixated on a central spot. The diameter of the RFs ranged from 3.9 to 10.7° (median, 7.4). RF center eccentricities ranged from 3.9 to 11.1° (median, 7.9). The preferred speed was also judged using a bar moved by hand. The animals were trained to perform the task at slow or moderate motion speeds, so we usually selected neurons with a preferred speed between 4 and 12°/sec. For most cells, once the RF location, size, preferred direction, and speed were determined, the memory saccade task was run with the targets at the centers of where the random dot patches would be located. Five to 30 (median, 12) correctly completed trials were collected for this task. The motion detection task was then run, and we recorded from the neuron as long as possible. The number of completed trials per coherence level for the motion detection task ranged from 15 to 175 (median, 35). The monkey's performance varied with patch location, size, and motion speed, which were determined by the response properties of the neuron under study. Consequently, different neurons were tested with different coherence levels. The animal's eye position was measured every 5 msec using a scleral search coil (Robinson, 1963; Judge et al., 1980), and the occurrence of action potentials was recorded to the nearest millisecond.

Extracting the global motion from the stimulus. Given two frames of a random dot stimulus, what is the motion stimulus? To answer this, we used a method described by Barlow and Tripathy (1997) that used the correspondence between dots in two sequential frames. Figure 1B illustrates this approach using two consecutive frames that each contains four dots. Because it is unknown how each dot moved from frame 1 to frame 2, we calculated all possible motion vectors. The possible ways in which the center dot (labeled 1) could have moved are shown by the four motion vectors. Each motion vector has a direction and magnitude, and there are 16 motion vectors in this example (four possible vectors for each of four dots). Thus for N dots, we computed N^2 motion vectors between consecutive updates of the random dot motion stimulus. The set of N^2 motion vectors represents our description of the global motion from one update to the next. Because the size and location of the patch of dots was set to overlap the RF of the neuron under study, this model of the motion assumes spatial uniformity of motion vectors within the RF. Thus, our motion description has no spatial component.

Figure 1C shows six updates of a 0% coherent random dot stimulus used during an experiment. Each patch of dots was updated every 26.6 msec (every other video frame). N^2 motion vectors were computed for each time point using the current and proceeding update, and their distributions are shown in the polar histograms below each patch. In these histograms, the location of a bin corresponds to motion direction, and the distance a bin is from the origin corresponds to the magnitude of the motion vector (dot speed). Each plot has 96 bins (12 directions \times 8 speeds), and the number of vectors per bin is indicated by the grayscale of the bin. For example, the transition from 756 to 783 msec shown in Figure 1C had, by chance, a relatively strong rightward motion component at $\sim 15^\circ$ /sec as shown by the two nearby light gray bins (arrow).

One important issue was how to represent the sequence of motion vectors in time. The scanning process of the cathode-ray tube (CRT) display resulted in the dots located at the top of each patch being updated first. We used the simplifying assumption that the dots would be updated simultaneously at the time our CRT display scanned the vertical midpoint of the RF of the neuron. Thus, our motion stimulus was modeled as a set of motion impulses occurring every 26.6 msec (Fig. 1C).

Motion vectors with speeds $> 23.5^\circ$ /sec were not used in the analysis because most of the MT neurons we recorded from preferred slower speeds (median, 10.7° /sec; maximum, 16° /sec) and were likely not sensitive to high speeds (Van Essen et al., 1981; Lagae et al., 1993). The reason for the bias toward slower speed preferences was that the monkeys were trained with only low speed coherent motion and had difficulty detecting

the coherent motion at higher speeds. Thus, during recording sessions, we selected neurons that preferred the lower motion speeds.

Estimation of the linear transfer function. We assumed that the neuronal response of our MT neurons, R , was equal to the motion in our random dot stimulus, M , convolved with a linear kernel, K , or as follows:

$$R = K * M. \quad (1)$$

We did not use a spike-triggered average approach to estimate K because our motion stimulus, M , contained autocorrelations (i.e., was non-white). To estimate the linear kernel, we instead used the discrete formulation approach (DiCarlo et al., 1998; Theunissen et al., 2001; Blake and Merzenich, 2002). This approach expresses the convolution in Equation 1 as a linear regression problem that compensates for any autocorrelations in the motion stimulus.

Only the 0% coherent motion was used to estimate the linear kernel, K . The neuronal response is modeled as the sum of the output of 96 filters convolved in time for each motion direction and speed (Fig. 1D). The spike response of the neuron was represented as a sequence of zeros and ones (indicating spikes) at the 1 msec sampling interval.

Using digital signal processing routines in MATLAB, we filtered and resampled the sequence of motion impulses and the recorded spikes at a Nyquist sampling rate of 100 Hz. This corresponds to a sampling interval of 10 msec and limited our maximum frequency to 50 Hz. The filtered and resampled neuronal response was assumed to represent the underlying rate function that drives spike production. Other studies have usually binned the neuronal response (e.g., 10 msec bins) to estimate the rate function (DiCarlo et al., 1998). Although both methods are qualitatively similar, Nyquist filtering and resampling eliminates aliasing that could potentially add noise to the estimated kernels.

We defined each of our 96 motion filters to have nine weights (or taps) to cover a 90 msec range (Fig. 1D). We varied the latency of the kernel, t_D , to account for the neuronal latency and to insure our kernel overlapped the neuronal integration window. We computed a separate kernel for each value of t_D and selected the kernel that had the best signal-to-noise ratio (S/N; described below). The neuronal response r at time sample index t can be expressed as the weighted sum of the past motion vectors for each direction and speed, or as follows:

$$r_t = c + \sum_{d=1}^{12} \sum_{s=1}^8 \sum_{\tau=1}^9 k_{d,s,\tau} m_{d,s,\tau}(t - \tau - t_D), \quad (2)$$

where $m_{d,s,t}$ is the motion strength and $k_{d,s,t}$ is the filter (or kernel) weight corresponding to direction index d and speed index s at time sample index t . The value c is a constant to account for any static offsets in r , and t_D ranged from 20 to 100 msec in 10 msec intervals. Thus, for each of the 96 direction and speed bins in our motion stimulus, there is a linear filter with nine unknown weights corresponding to 10 msec sample intervals. Combining the output from all the linear filters results in $96 \times 9 = 864$ total unknown weights plus a constant (Fig. 1D). We refer to this collection of weights as the linear kernel of the neuron.

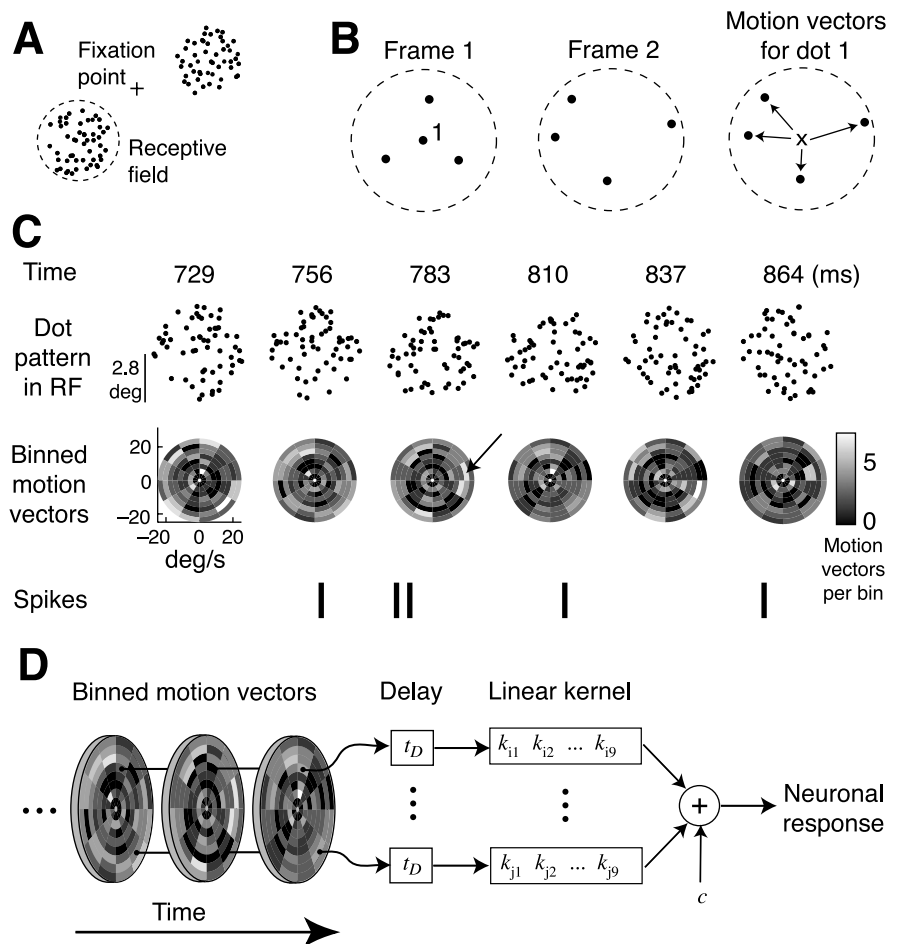


Figure 1. Description of the motion stimulus. *A*, The random dot stimuli consisted of two diametrically opposed patches of random dots, one of which overlapped the RF of the neuron under study. *B*, Example of calculating vectors between corresponding dots to describe the global motion in the stimulus. In the sequence from frame 1 to frame 2, dot 1 could have moved in one of four possible ways as illustrated by the motion vectors. There are a total of 16 motion vectors between frames 1 and 2. *C*, Sequence of random dots appearing in the RF of an MT neuron during an experiment. Each patch was updated every other frame on a 75 Hz monitor (every 26.6 msec). Below each patch is the polar histogram of all motion vectors between the previous and current patch. The bin location corresponds to the direction of the motion vector and the distance from the origin is the speed. The grayscale indicates the number of motion vectors per bin. Motion vectors corresponding to speeds $>23.5^\circ/\text{sec}$ were not included. Motion vectors were modeled as impulses occurring every 26.6 msec. The arrow indicates a strong rightward motion component occurring at 783 msec at $\sim 15^\circ/\text{sec}$. The corresponding recorded spike times are shown below the motion vector histograms. *D*, Structure of the linear kernels. Each bin of the motion vector histograms was convolved with a filter made up of nine weights (k) at 10 msec intervals. The response of the neuron was modeled as the sum of the individual filter outputs plus a constant. A variable time delay t_D was inserted to account for the latency in the neuronal response.

Equation 2 can be written as a single linear equation with 865 unknowns for every observation of $r(t)$. For example, a neuron that produced 5 min of data would result in $\sim 30,000$ observations (or equations). From this set of linear equations, the kernel weights were determined using built-in MATLAB routines for solving overdetermined systems of linear equations in a least squares manner. Mathematically, this is equivalent to linear regression with 865 unknowns and estimates the first-order Wiener kernel that provides the best approximation to the linear impulse response function of the neuron in direction, speed, and time.

Because the noise in the kernel estimates is uncorrelated among the coefficients (DiCarlo et al., 1998), kernels were smoothed by convolving with a three-dimensional low-pass filter (Hamming-windowed-based, linear phase, with a cutoff set at 0.7 of the normalized sampling frequency; from the Digital Signal Processing toolbox in MATLAB). The impulse response function of this smoothing filter is similar in width to a Gaussian filter with a SD of 1 bin but provides better filtering of the high-frequency noise in our kernels with less distortion of the low-

frequency kernel shape. Reducing the noise in our estimated kernels was important for analyzing how the kernels were altered by attention. Using other smoothing filters, such as a 1 bin Gaussian, produced kernels with more noise but did not qualitatively change the results of our analysis.

In displaying the kernel weights, we used grayscale plots that show the kernel coefficients as a function of direction and speed. The center of each of the 96 direction and speed bins corresponds to the value of the kernel weight, and linear interpolation was used between the bin centers.

Assessing the kernel quality. We assessed the quality of our kernels using a S/N. We assumed the total variance of our estimated kernel (σ_{total}^2) was equal to the variance of the underlying kernel (σ_{kernel}^2) plus the variance of the noise (σ_{noise}^2), or as follows:

$$\sigma_{\text{total}}^2 = \sigma_{\text{kernel}}^2 + \sigma_{\text{noise}}^2. \quad (3)$$

To estimate the variance of the noise, σ_{noise}^2 , we computed a noncausal kernel for each neuron. The noncausal kernel correlated the neuronal response with motion vectors forward in time from 0 to 80 msec. Because future motion vectors should be uncorrelated with the current neuronal response, the true noncausal kernel is zero for all coefficients, and any variations represents noise in our estimation process. To check that the variance of the noncausal kernel provided an accurate estimate of the noise variance, we also computed noise variance using the SD of each kernel coefficient that was returned by our regression package in MATLAB. Both methods produced nearly identical estimates of the noise variance in our kernels ($r^2 = 0.96$ across our population of neurons). However, we used the noncausal kernel to estimate σ_{noise}^2 because this allowed us to first smooth the noncausal kernels in the same manner as the causal kernels.

Using the variance of our smoothed noncausal kernel as our estimate of the variance of the noise (σ_{noise}^2), we computed the variance of the underlying kernel of the neuron as follows:

$$\sigma_{\text{kernel}}^2 = \sigma_{\text{total}}^2 - \sigma_{\text{noise}}^2. \quad (4)$$

We then defined the S/N of the estimated kernel as follows:

$$S/N = \frac{\sigma_{\text{kernel}}}{\sigma_{\text{noise}}}. \quad (5)$$

Thus, for neurons with causal kernels containing well defined regions of integration and noncausal kernels that were mostly flat, the S/N would be large. If no underlying kernel existed for a particular neuron, the causal and noncausal estimates would have equal variances ($\sigma_{\text{total}}^2 = \sigma_{\text{noise}}^2$), and the S/N would be zero.

Parametric fitting of the kernels. We fit each smoothed causal kernel with a product of Gaussian function that described the tuning along direction (d), speed (s), and time (t):

$$f(d, s, t) = \left(C + G \times e^{-\frac{(d - \mu_d)^2}{2\sigma_d^2}} \right) \times e^{-\frac{(s - \mu_s)^2}{2\sigma_s^2}} \times e^{-\frac{(t - \mu_t)^2}{2\sigma_t^2}}, \quad (6)$$

where G is the gain of the kernel, μ_d is the preferred direction, σ_d is the directional tuning width, μ_s is the preferred speed, σ_s is the speed tuning width, μ_t is the peak temporal integration, σ_t is the width of the temporal integration, and C is an offset term to allow for inhibition. Both attentional conditions were fit separately using optimization routines in MATLAB. We limited σ_t to a lower bound of 4 msec because this is approximately the smallest Gaussian that could be represented without distortion at our 10 msec sampling interval.

Multiplicative scaling of kernels. We scaled the kernels derived for the attend-out condition (K_{OUT}) to match the kernels derived for the attend-in condition (K_{IN}). Thus, we computed a scale factor, β , that best satisfied $K_{\text{IN}} = \beta K_{\text{OUT}}$. For this, we minimized the χ^2 merit function (Press et al., 1986):

$$\chi^2 = \sum_{i=1}^{864} \frac{(k_{\text{IN}i} - \beta k_{\text{OUT}i})^2}{(\sigma_{\text{noise}_{\text{IN}}}^2 + \beta^2 \sigma_{\text{noise}_{\text{OUT}}}^2)}, \quad (7)$$

where $k_{\text{IN}i}$ and $k_{\text{OUT}i}$ are the i th kernel weights for the attend-in and attend-out conditions and $\sigma_{\text{noise}_{\text{IN}}}^2$ and $\sigma_{\text{noise}_{\text{OUT}}}^2$ are the variances of the kernel noise estimated from the smoothed attend-in and attend-out noncausal kernels.

We also computed a second scaling factor using the estimated kernel variances $\sigma_{\text{kernel}_{\text{IN}}}^2$ and $\sigma_{\text{kernel}_{\text{OUT}}}^2$ (from Eq. 4) for the attend-in and attend-out conditions. In this case, a scaling factor, γ , that satisfied $K_{\text{IN}} = \gamma K_{\text{OUT}}$ was estimated as follows:

$$\gamma = \frac{\sigma_{\text{kernel}_{\text{IN}}}}{\sigma_{\text{kernel}_{\text{OUT}}}}. \quad (8)$$

To examine whether kernels changed their shape with attention, we examined the residuals as follows:

$$K_{\text{res}} = K_{\text{IN}} - \beta K_{\text{OUT}} \quad (9)$$

where K_{res} is the kernel residual and β is the multiplicative scale factor attributable to attention. If attention produced a true multiplicative scaling of the kernels, then K_{res} would be flat (or zero). However, if the attention produced a shape change in the kernels, then K_{res} would have some weights that are non-zero.

A difficulty in using the kernel residuals to determine the effects of attention is that any noise in the estimated kernels also contributes to non-zero elements in K_{res} . However, given our estimates of the noise in the attend-in and attend-out kernels ($\sigma_{\text{noise}_{\text{IN}}}^2$ and $\sigma_{\text{noise}_{\text{OUT}}}^2$), we can predict the variance of K_{res} that we would expect given a pure scaling with no change in shape. This predicted variance, σ_0^2 , of the residual kernel is as follows:

$$\sigma_0^2 = \sigma_{\text{noise}_{\text{IN}}}^2 + \beta^2 \sigma_{\text{noise}_{\text{OUT}}}^2, \quad (10)$$

where β is our estimated multiplicative scale factor. If attention produced a change in kernel shape, the variance of K_{res} (referred to as σ_{res}^2) would be larger than predicted by Equation 10. This hypothesis can be tested using the following statistic:

$$L = \frac{\nu \sigma_{\text{res}}^2}{\sigma_0^2}, \quad (11)$$

where L is a χ^2 value with $\nu = 864 - 1$ degrees of freedom (Zar, 1999).

Results

The goal of this study was to understand the effects of attention on the neuronal integration of motion in MT. To accomplish this, our analysis proceeded in three steps. First, we determined the linear transfer function (or kernel) that described the integration of the motion stimulus by MT neurons. Second, we selected neurons that produced kernels with clear regions of integration as indicated by their good S/N. Finally, we used this subpopulation of neurons to examine the effects of attention on the kernel. Thus, by examining how attention affected the linear transfer function, we could see how attention affects the temporal integration of MT neurons. The neurophysiological recordings analyzed here have been used in two other studies that examined the link between neuronal activity in MT and the perceptual capabilities of the subjects (Cook and Maunsell, 2002a,b). This new study focuses exclusively on how attention affected the way MT neurons processed the motion stimulus. Portions of our results have been previously published in abstract form (Cook and Maunsell, 2003).

Estimating the MT motion integration kernels

Using neuronal responses to 0% coherent motion and an optimal estimation procedure, we constructed linear models that described how MT neurons integrated the motion stimulus. This type of analysis has been used to examine the integrative properties of neurons in many sensory cortical areas (Jones and Palmer,

1987; DeAngelis et al., 1995; DiCarlo et al., 1998; Blake and Merzenich, 2002) and recently in MT (Bair et al., 1997; Livingstone et al., 2001; Borghuis et al., 2003). The kernel weights in these models describe how an MT neuron integrates the motion signal. We wanted to know how these weights were affected by attention to the stimulus.

The optimal set of kernel weights (also referred to as the first-order Wiener kernel) was determined for both attentional states in each neuron using standard numerical methods (see Materials and Methods). Figure 2*A* shows the computed kernel weights for one MT neuron. Each column shows the time evolution of the kernels for one attentional condition. The kernel weights are shown using the same polar format as the stimulus in Figure 1*C*, except that the plots have been smoothed by linear interpolating between bin centers. This interpolation makes it easier to distinguish excitatory (light shades of gray) and inhibitory (dark shades of gray) regions. This particular neuron began integrating the motion ~40 msec after the motion onset. At 40 msec, the kernel weights showed a clear preference for motion up and to the right at a speed of ~8.5°/sec (white). There was pronounced inhibition for motion in the opposite direction (black). By 60 msec after stimulus onset, the neuron no longer integrates the motion appreciably. The sequence of kernel weights over time provides a model of how the MT neuron integrated the motion.

Figure 3*A* shows the time evolution of the kernel weights for another MT neuron. For this cell, the integration window also starts at ~40 msec, and the preferred direction is toward the left. Note that the grayscale range used for displaying the kernels was optimized individually for each neuron. The structure of our kernels is very similar to that calculated for MT neurons by Livingstone et al. (2001) using a sparse stimulus of two random dots. For both example neurons, the kernels look similar between the two attentional conditions. To more closely address the effects of attention on the kernels, however, we needed to first evaluate the quality of our estimated kernels.

Assessing the estimated kernels

The example neurons in Figures 2 and 3 produced kernels with clear regions of excitatory and inhibitory integration. Not all of our MT neurons produced kernels with as well defined regions of integration as shown by these two example cells. To separate neurons that produced kernels with well defined regions of integration from those that did not, we computed the S/N in our kernels (see Materials and Methods).

To estimate the noise in our kernels, we first computed a noncausal kernel using the same methods to compute the causal kernels described above. A noncausal kernel is the optimal linear transfer function between future motion and current neuronal activity. Because there should be no correlations between the spike rate of a neuron and stimuli that will occur in the future, the expected noncausal kernel is flat with all coefficients equal to zero. Any deviations from zero represent chance correlations. These chance correlations are a measure of the noise in our kernels and can be expressed as the variance of the noncausal kernel or σ_{noise}^2 .

For each neuron, we assumed the total variance of our estimated kernel (σ_{total}^2) is attributable to the variance of the underlying kernel (σ_{kernel}^2) plus noise (σ_{noise}^2). Because we calculate σ_{noise}^2 from the noncausal kernel, we can estimate the variance of the underlying kernel by subtracting the noise variance from the total variance of our kernel (Eq. 4). We then computed a S/N for our kernel by dividing the SD of the true kernel by the SD of the noise ($\sigma_{\text{kernel}}/\sigma_{\text{noise}}$). An attractive property of this measure of

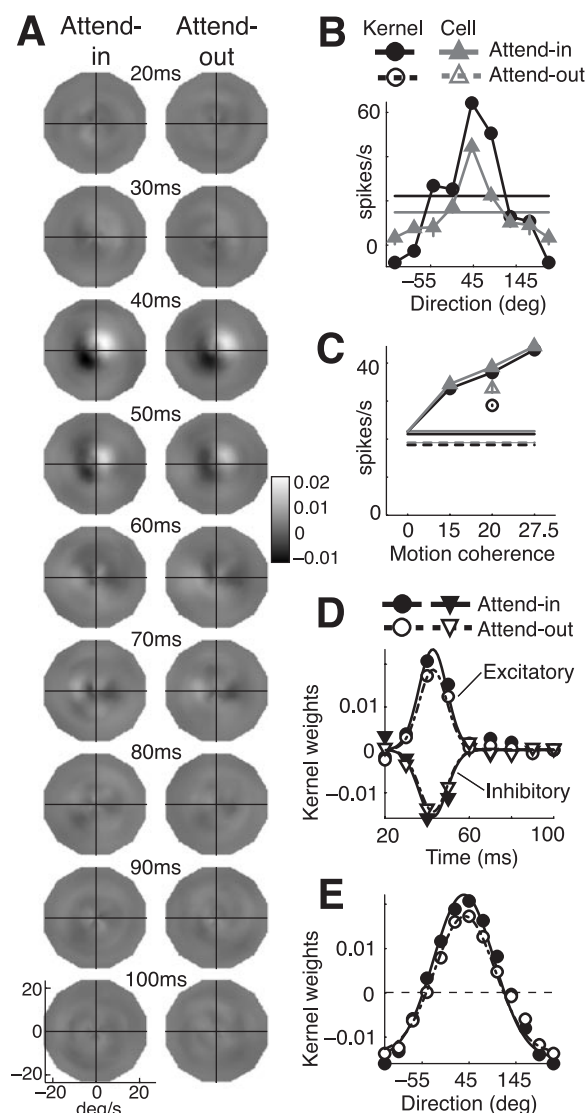


Figure 2. Linear kernel for an example MT neuron. *A*, Kernel weights for the attend-in (left) and attend-out (right) conditions. The weights of the kernels are shown in the same polar format as the motion vectors in Figure 1*C*. Each time point corresponds to 96 kernel weights with linear interpolation applied between bin centers. At 40 msec, the kernel has an excitatory region (shown in white) corresponding to a motion direction of 45° and speed of 8.5°/sec. An inhibitory region (shown in black) corresponds to motion in the null direction (225°). The kernels derived from the two attentional states are nearly identical for this cell. *B*, Directional tuning of the neuron and linear kernel. The average neuronal response of the cell (gray filled triangles) and the predicted response of the linear kernel (black filled circles) are shown for coherent motion in different directions. The stimuli were coherent motion (50%) in each of eight directions. For the kernel, the response was calculated by convolving the linear transfer function in *A* with the motion vectors calculated from the random dot stimulus. The horizontal lines are the responses of the model and cell to the 0% coherent motion in the directional selectivity trials. The error bars are SEM. *C*, Average firing rate of a cell (gray) and linear kernel (black) to different motion coherences in the preferred direction. The open symbols correspond to the trials where the animal's attention was directed away from the RF. The horizontal lines show the neuron and the response of the model to the 0% coherent motion for the attend-in (solid) and attend-out (dashed) conditions. The average neuronal responses in *B* and *C* were calculated using the 300 msec period just before and after the coherent motion began. *D*, Excitatory and inhibitory temporal profiles of the linear kernel for the two attentional conditions. The excitatory temporal profiles are shown for a constant direction of 45° and a constant speed of 8.5°/sec. The inhibitory temporal profiles are for a constant direction 225° and a constant speed of 8.5°/sec. The curves are from the Gaussian function (Eq. 6) fitted to the entire kernel. Both attend-in (solid symbols) and attend-out (open symbols) profiles are shown. *E*, Directional tuning profile of kernels for the two attentional states. The directional profiles correspond to 40 msec and 8.5°/sec. The lines are the fitted Gaussian function.

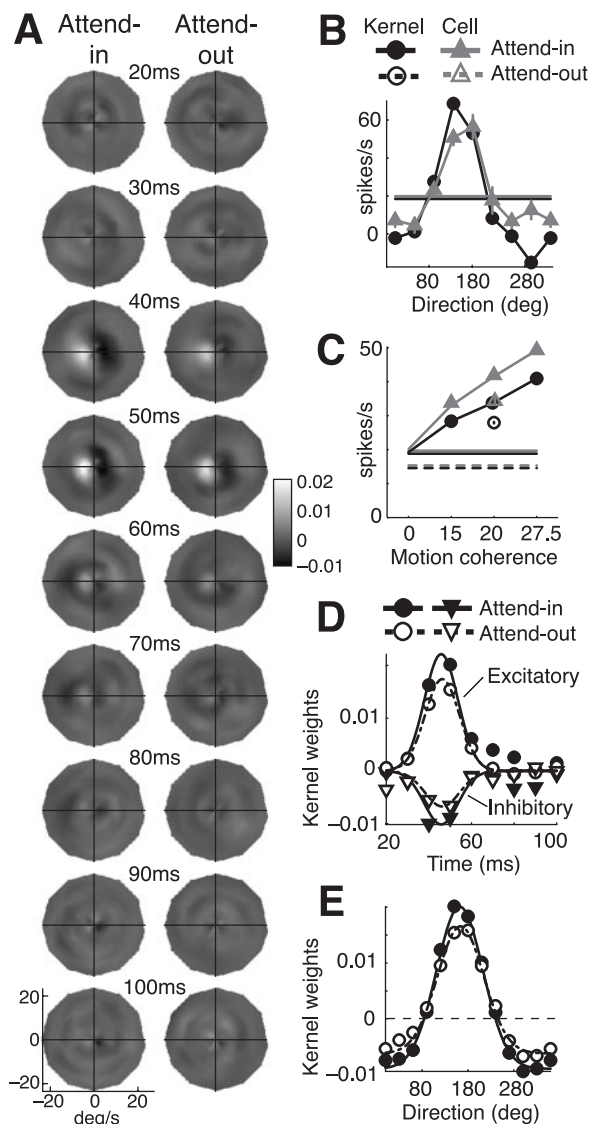


Figure 3. Linear kernel for a different MT neuron. The same format as in Figure 2 is shown, except as noted. *D*, Temporal profiles correspond to the direction of 180° (excitatory) and 0° (inhibitory) and 8.5°/sec. *E*, Directional tuning profiles computed at 40 msec and 8.5°/sec.

kernel quality is that no assumptions need to be made regarding the shape of the underlying kernel.

If an estimated kernel produced no well defined regions of integration, then the S/N would be expected to be small. Likewise, kernels with well defined regions of integration as in Figures 2 and 3 would be expected to produce relatively high values for the S/N. For our measure of how well the kernel contained well defined regions of integration, we averaged the S/N from both attentional conditions. The distribution of the average S/N is shown in Figure 4*A*. For reference, the kernels in Figures 2 and 3 produced an average S/N of 2.4 and 1.8, respectively. Figure 5 illustrates three other example cells with a progressively lower S/N. For kernels with a low S/N, there is less of a discernable region of motion integration. In Figure 5*C*, a S/N of 0.6 corresponds to a kernel that is dominated by noise and was not included in the analysis.

By inspection, we found that kernels with a S/N < 0.75 lacked clear structure. We therefore set an arbitrary S/N acceptance criterion of 0.75, resulting in 44 of 93 MT cells classified as “good S/N” (Fig. 4*A*, gray bars). The remaining cells were considered to

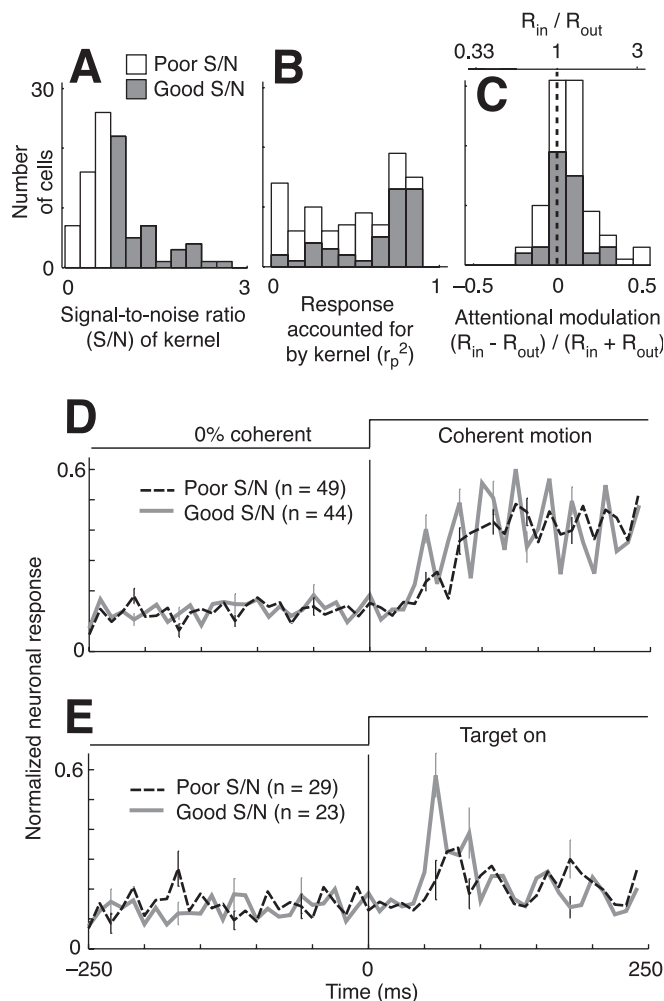


Figure 4. Comparison of kernel S/N estimates. *A*, Distribution of the kernel S/N estimates for all cells. The filled bars correspond to kernels with an S/N ≥ 0.75 (good S/N). The open bars are those kernels classified as poor S/N. *B*, Distribution (stacked histogram) of the proportion of the average response to the coherent motion that was accounted for by the linear transfer function (r_p^2). The filled bars correspond to kernels with a good S/N. *C*, Distribution (stacked histogram) of AM in a ratio metric form. Modulation was calculated from the average response to the 0% coherent motion using an attentional index of $(R_{in} - R_{out}) / (R_{in} + R_{out})$, where R_{in} and R_{out} are the average response for the attend-in and attend-out conditions, respectively. The top axis shows the equivalent ratio of responses. *D*, Average normalized response to the onset of coherent motion (at 50% strength). *E*, Average normalized response to the onset of a small target located in the center of the RF. Responses are shown for cells with a good S/N (thick gray line) and a poor S/N (thin dashed black line). The error bars are SEM and, for clarity, are not shown for all points.

have “poor S/N” and not useful for examining the effects of attention on temporal integration. This is because we could not assess the effects of attention on the structure of the kernels if the kernels lacked any discernable structure to begin with.

None of our results were sensitive to the particular criterion used. Using other measures of kernel quality, such as the goodness of a Gaussian fit or the ability of the kernel to account for the average stimulus selectivity of the neuron (shown below) produced very similar sets of acceptable neurons and did not affect our results or conclusions.

To further assess the kernels, we compared the predicted response of the kernels to that of the actual response of the neurons. We wanted to know whether the kernels captured the directional tuning of the neuron as well as the increased firing rate to coherent motion. To do this, we calculated the average response of the

neuron and kernel to coherent motion of varying strength and direction. Average responses were computed using the 300 msec period before and after the coherent motion began on each trial. We compared the average firing rate predicted by the kernel to the actual response of the neuron. It is important to emphasize that the kernel was estimated based on the response of the neuron to the 0% coherent motion. Thus, we tested the model using coherent motion that was not used to construct the model.

Figures 2*B* and 3*B* show the average firing rates of the two example MT neurons to coherent motion in eight different directions (gray filled triangles). The average firing rates predicted by the kernels (black filled circles) are also shown. The horizontal lines represent the average and predicted response to the 0% coherent motion that preceded the coherent motion in every trial. Although both kernels capture the directional selectivity of the neurons, the one in Figure 3*B* does a better job matching the average firing rate of the cell.

We also compared the response predicted by the kernel with the response of the cell to increasing coherent motion strength. Figures 2*C* and 3*C* show the average response of both the kernel and cell to the three levels of coherent motion in the preferred direction of the cell. The kernel in Figure 2*C* matched the firing rate of the cell as the strength of the coherent motion increased, whereas in Figure 3*C*, the kernel exhibits a slight reduction in gain. The solid horizontal line is the average response to the 0% coherent motion. As expected, the kernels predicted the average response to the 0% coherent motion in Figures 2*C* and 3*C* because they were estimated using these data. We also compared the response of the neuron and kernel when the animals withdrew their attention (attend-out), which occurred during the medium strength coherent motion only (Figs. 2*C*, 3*C*, open symbols). For this, we used the kernel derived when the animals were attending to the stimulus outside the RF. Withdrawing attention reduced neuronal responses to the medium coherent motion and to the 0% coherent stimulus (dashed lines).

We quantified the ability of the kernel to account for the neuronal response to the coherent motion stimuli. Using the directional tuning and response gain (Figs. 2*B*, *C*, 3*B*, *C*), we computed the proportion of the variance (r_p^2) in the neuronal response accounted for by the kernels. Twelve points were used (eight directions, three attended coherent motion levels, and one unattended medium coherent motion). We found a wide range of kernel performances as illustrated by the wide distribution of r_p^2 values for all MT neurons (Fig. 4*B*). For comparison, the linear kernels for the example neurons in Figures 2 and 3 had r_p^2 values of 0.68 and 0.84, respectively. Notably, the kernels that had a good S/N tended to have high r_p^2 values (Fig. 4*B*, gray bars) and thus captured the neuronal response to the coherent motion.

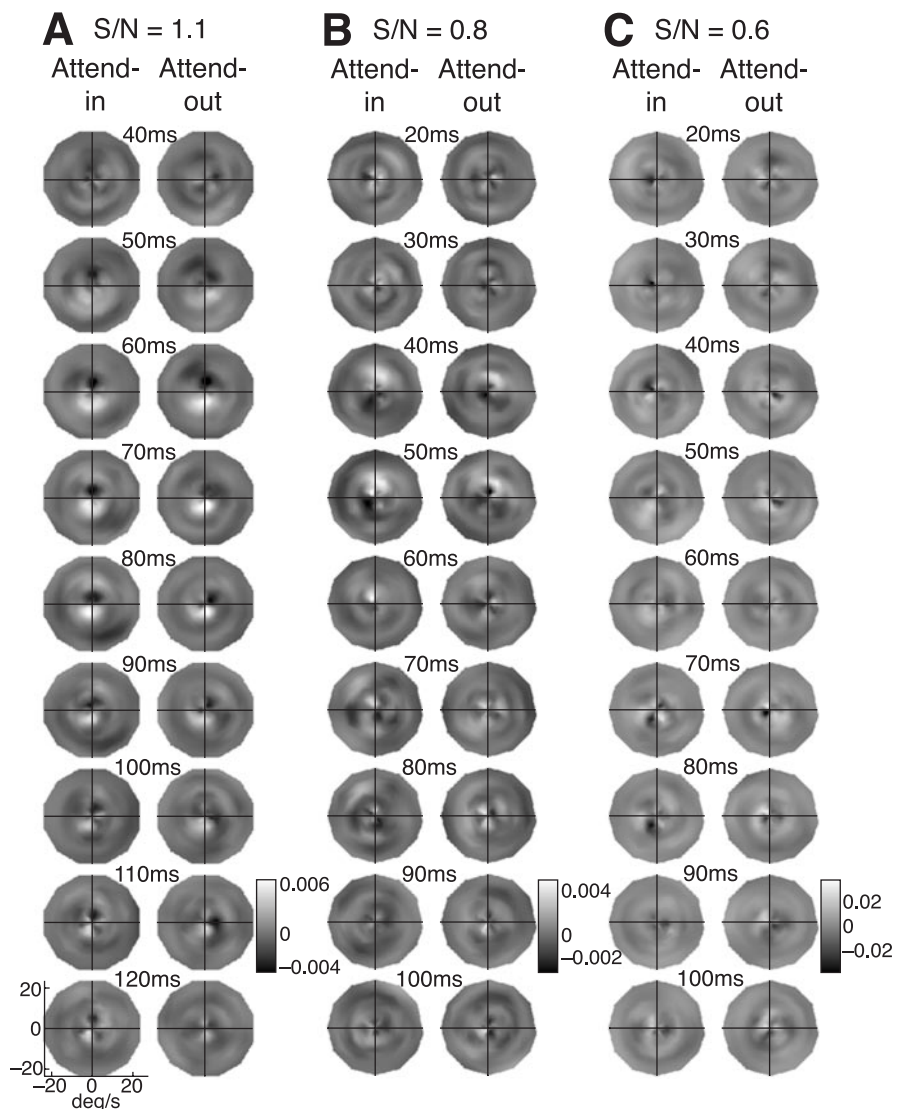


Figure 5. Example of kernels with a progressively lower S/N. *A*, *B*, Example of kernels with an S/N above the 0.75 criterion and included in the analysis. *C*, Example of a kernel with a poor S/N and not included in the analysis.

Figure 4*B* also indicates some kernels with a poor S/N also had high r_p^2 values. These cells did not have well defined regions of integration, yet accounted for the mean firing rate to the coherent motion. Visual inspection of these cells revealed that many had either weak directional tuning to begin with or, in most cases, had regions of integration that were overwhelmed with noise.

MT neurons usually increase their response when the animals orient their attention to the stimulus located in the RF of the neuron (Seidemann and Newsome, 1999; Treue and Maunsell, 1999; Cook and Maunsell, 2002a). The amount of AM of the average firing rate in response to the 0% coherent motion (in a ratiometric form) is shown for our population of MT neurons in Figure 4*C*, and as this histogram illustrates, the amount of modulation varies widely between cells. The mean AM to the 0% coherent motion for cells that produced kernels with a good and poor S/N was 11 and 21% (median, 7 and 16%), respectively. However, this difference in modulation was not significant ($p = 0.12$; two-sample t test).

Why did some neurons produce kernels with regions of well defined integration (as measured by the S/N), whereas others did not? Possible explanations include failure of our description of

the motion stimulus to capture the true input driving the cell and the assumption that MT neurons are linear. The quality of the kernel can also be affected by the limited amount of data available for neurons that had very low firing rates or for which fewer trials were sampled. These last two possibilities were analyzed by computing the partial correlations (ρ) between the number of spikes recorded (n_s), the S/N of the kernel, and the amount of AM. We found that the number of spikes recorded correlated with the S/N (ρ between n_s and S/N = 0.70). Thus, neurons for which we had more data produced kernels with a higher S/N and better defined regions of integration. We found, however, no strong correlations between the kernel S/N and AM or between the number of spikes and AM (ρ between S/N and AM = -0.02 ; ρ between n_s and AM = -0.15). Other studies that have investigated temporal integration of motion by MT neurons did not quantify the quality of their kernels (Bair et al., 1997; Livingstone et al., 2001; Borghuis et al., 2003), and thus we do not know whether the observed range of quality in our estimated kernels is typical.

In addition to the difference in the amount of AM between cells that produce good and poor S/N kernels, these two populations also exhibited substantially different response dynamics to visual stimulus. Figure 4, *D* and *E*, shows the average response of neurons to the onset of the coherent motion and to a flash of a small stationary target in the center of the RF. Normalized population averages are shown for the neurons with good S/N kernels (thick gray) and poor S/N kernels (thin dashed black). For each cell, we normalized by the peak neuronal response, allowing each neuron to contribute equally to the average responses shown. The responses to the coherent motion in Figure 4*D* came from directional tuning trials that used the same coherent motion level (50%) for all cells. The responses shown are for motion in the preferred direction (as determined from the directional tuning trials). The responses to the target in Figure 4*E* came from memory saccade trials performed for most cells (see Materials and Methods). In this case, the animals were fixating when a small target first appeared in the center of the RF (target on).

It is clear from Figure 4, *D* and *E*, that the two populations show different temporal dynamics in response to the same stimuli. For the coherent motion, neurons with good S/N kernels show strong, transient responses each time the coherent motion was updated (every 26.6 msec). Although neurons with poor S/N kernels demonstrate a comparable average response to the coherent motion, their dynamics exhibit much less temporal modulation. For the contrast change produced by the target onset in Figure 4*E*, cells with good S/N kernels had robust transient responses, whereas the cells with poor S/N kernels responded weakly. One possibility for this difference is that neurons that produced kernels with a high S/N have their response properties tightly linked to the visual stimulus. The response properties of neurons with poor S/N kernels may either be influenced by other aspects of the visual input or contain larger extra-retinal components, neither of which would be captured in our kernel estimates. Another possibility is that the integration window of neurons with poor S/N kernels may have exceeded the 90 msec duration defined by our analysis. We do not think this is likely because we varied the delay between stimulus and response up to 100 msec to make sure our kernels would overlap with the temporal integration window of the cell.

Effects of attention on Gaussian fits of the kernels

We wanted to know how the attentional state of the animal affected the kernels estimated for our MT neurons. To address this

question, we only analyzed the population of neurons with good S/N kernels because those were the cells that produced kernels with reasonably well defined regions of motion integration that also tended to account for the average response to the coherent motion.

To reveal how attention affected temporal integration, we first parameterized the shape of the estimated kernels. For this, we fit each kernel with a function that was the product of three Gaussians corresponding to direction, speed, and time (Eq. 6). Figures 2, *D* and *E*, and 3, *D* and *E*, show the results of fitting kernels for our two example cells. In Figures 2*D* and 3*D*, we plot as a function of time the coefficients (or weights) of the kernels and the Gaussian fit corresponding to a constant (preferred) direction and speed. This slice of the kernel through time provides a profile of the temporal integration window. Each plot shows four time slices corresponding to the peak excitatory and inhibitory regions for the two attentional conditions. Fits were done for each attentional condition independently. In Figures 2*E* and 3*E*, we plot slices of the kernels through direction (constant speed and time), which shows the directional tuning profiles for both attentional conditions. Spatial attention increased the values of the kernel weights for both cells, although the effect was somewhat larger for the cell in Figure 3.

To examine the effects of attention on the kernels, we parameterized the shape of the kernels using the fits to our Gaussian function. Figure 6*A* shows five of the Gaussian parameters for the kernels with good S/N cells corresponding to the two attentional states: (1) the amplitude of the kernels (G); (2) the time of the peak excitatory and inhibitory integration regions (μ_e); (3) the SD of the temporal width of the kernels (σ_t); (4) the preferred direction of motion that corresponds to the peak (μ_d); and (5) the SD of the width of the directional tuning (σ_d). The other three parameters of the Gaussian for speed tuning and offset showed similar relationships with attention. Because the average AM of our neurons with good S/N kernels was relatively weak, we separated out those cells that also had high AM $\geq 10\%$ (Fig. 6*A*, filled circles) ($n = 20$).

In Figure 6*B*, we show the distribution of the AM of each parameter. AM was expressed in a ratiometric form $(P_{in} - P_{out}) / (P_{in} + P_{out})$, where P is the parameter of interest. The exception, however, was for the modulation of the time and direction peaks, which was computed as the difference between the two attentional conditions $(P_{in} - P_{out})$. The filled bars correspond to good S/N kernels that also experienced high AM. The median modulation of each parameter for good S/N neurons and good S/N with high-attention neurons is indicated by the open and filled triangles, respectively.

Although there is a relatively large amount of variability in the data in Figure 6, only the kernel magnitude (G) tends to be consistently larger when the animals directed their attention to the RF of the neuron. The median AM of the kernel magnitude is 7 and 17% for the kernels with good S/N and good S/N plus high attention, respectively. This modulation was significantly different than zero for the high-attention group ($p = 0.21$ and 0.004 for the good S/N and good S/N plus high attention, respectively; Wilcoxon signed rank test). The median modulation of the other parameters was not significantly different than zero for either kernels with good S/N or good S/N plus high-attention groups (Wilcoxon signed rank test). Overall, these data suggest attention scaled the Gaussian fits of the kernels with no appreciable shift in peak or change in shape.

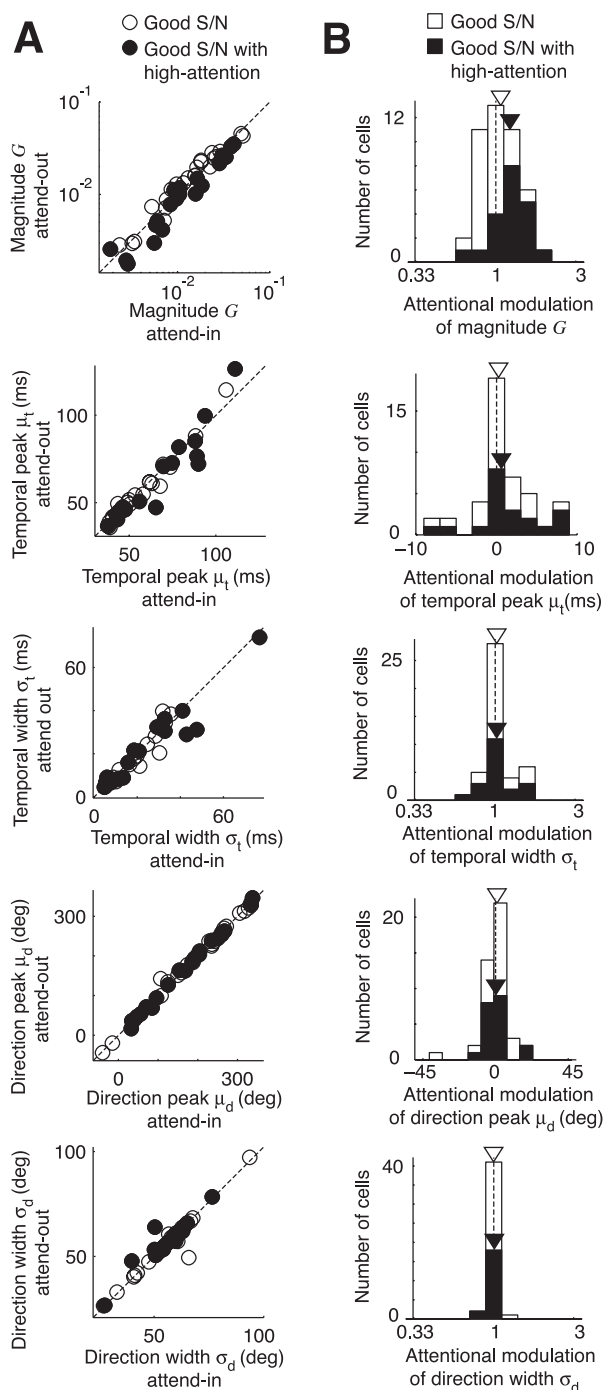


Figure 6. Comparison of Gaussian fit parameters (Eq. 6) for the two attentional conditions for the 44 good S/N kernels. *A*, Each parameter is from the fit to the kernel corresponding to the attend-in (*x*-axis) versus attend-out (*y*-axis) conditions. Kernels from each attentional condition were fit independently. Top to bottom: kernel gain (G), the center (μ_t) and width (σ_t) of the temporal integration window, and the preferred direction (μ_d) and direction tuning width (σ_d) for the attend-in and attend-out conditions. The filled symbols correspond to those neurons with good S/N kernels that also experienced $\geq 10\%$ AM of the firing rate. The dashed line in each plot is unity slope. *B*, Distribution of AM of the Gaussian parameters. AM was computed using a ratiometric form $(P_{in} - P_{out}) / (P_{in} + P_{out})$ and plotted with the *x*-axis labeled using the corresponding ratio (P_{in} / P_{out}) . For the parameters corresponding to time of peak (μ_t) and preferred direction (μ_d), the effects of attention are expressed as the difference between attend-in and attend-out. The filled bars correspond to good S/N kernels with AM of the firing rate $\geq 10\%$ and the triangles correspond to the median of each parameter for the good S/N (open) and high-attention subpopulation (filled). The vertical dashed line corresponds to no AM.

Effects of attention on population kernels

To further examine the effects of attention on the linear model, we compared the average population of good S/N kernels for each condition. For each cell, we normalized the kernel weights to the peak value in the attend-in condition. We then computed the population kernel by averaging the normalized kernels from each cell. Before averaging, we aligned both the temporal peak (at time 0 msec) and the speed peak (at 8.5°/sec) for each kernel based on the Gaussian fit to the attend-in condition. We then rotated the kernel so that the preferred direction was up. Because we used only Gaussian parameters for the attend-in condition to align the kernels, this analysis would reveal any systematic shifts or changes in kernel shape that occurred in the attend-out condition. Figure 7*A* shows the average population kernels for each attentional condition for our neurons with a good S/N. We chose to align the kernels at the peak to reveal the average integration period.

The average population kernels for the 44 good S/N kernels are very similar for both attentional conditions. Figure 7*C* shows a comparison of the average normalized firing rate to the coherent motion for the cells and that predicted by the kernels. On average, the kernels capture the directional selectivity and increased response to the coherent motion. Figure 7*C* reveals, however, that the kernels predicted a persistent reduction in response amplitude at higher coherence levels. Because individual cells were recorded using different levels of coherent motion, responses were grouped into low, medium, and high coherent levels in Figure 7*C*. The average normalized response in the unattended condition shows a small amount of modulation at both 0% coherent (dashed line) and medium coherent motion (open symbols) for both the cells and kernels.

To better reveal the effects of attention on the kernels, we further subdivided the 44 good S/N neurons by computing an average population kernel using only cells with high AM $\geq 10\%$ (Fig. 7*B*). The average amount of AM of the spike rate for this group was 33% (median, 27%). For these cells, attention did not significantly alter the shape of the kernels between the attend-in and attend-out conditions. Figure 7*D* shows the average normalized responses for the cells and predicted response for the kernels corresponding to the high-attention group. Note that the average predicted response did not match the gain of the neurons at high coherence levels. However, for the high-attention group of cells, the effect of withdrawing attention is readily seen in the normalized population response to the 0% coherent motion (dash line) and medium coherent motion (open symbols) for both the cells and their corresponding kernels.

To examine the effect of attention on the integration of the motion, we fit the average population kernels in Figure 7, *A* and *B*, with the product of the Gaussian model of Equation 6 (fitting was done in the same manner as for the individual kernels). Figure 8, *A* and *B*, show the time, direction, and speed profiles of our population kernels and the optimal fits for both attentional conditions. For the temporal profiles, we show slices through both the peak excitatory and inhibitory integration regions. It should be noted that the reason some of the fits do not appear optimal in Figure 8 is because they were optimized on the entire set of kernel coefficients and not just the subset of data points shown.

For the fits to the population kernels, attention produced a 10 and 34% modulation in the amplitude parameter (G) in the good S/N and good S/N plus high-attention kernels, respectively, and was similar to the AM of the spike rate for the two populations of neurons. The width of the temporal integration window did not appreciably change for the high-attention population (attend-in

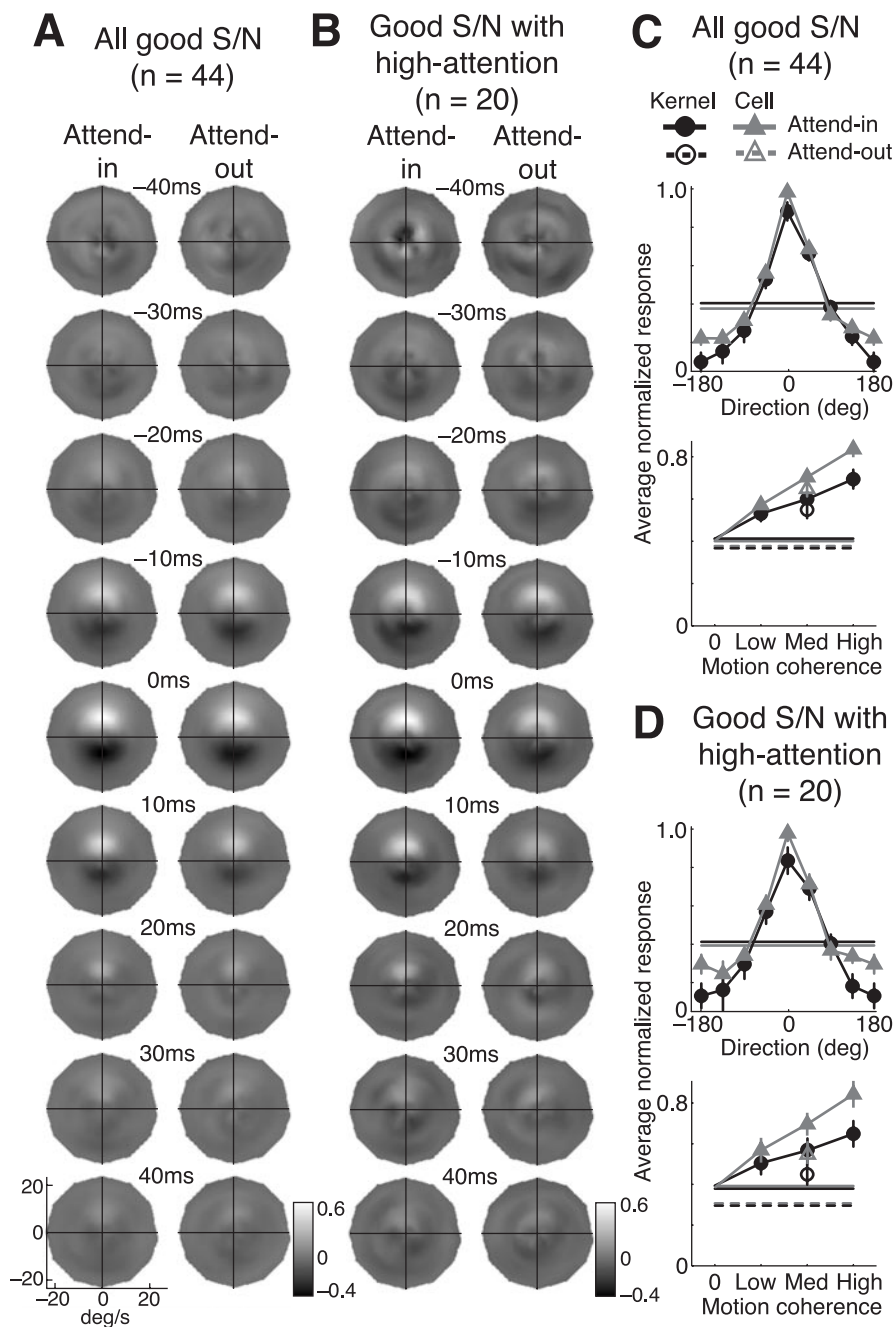


Figure 7. Average of kernels normalized by peak weight, oriented with preferred direction (μ_d) pointing up, and with temporal peak (μ_t) aligned at 0 msec and speed peak (μ_s) aligned at 8.5°/sec. Alignment of both the attend-in and attend-out kernels was based on parameters from the Gaussian fits for the attend-in condition only. *A*, Average population kernels from the 44 good S/N cells. Time points are relative to the alignment of the peaks. *B*, Average population kernels from the 20 good S/N cells that also had high AM ($\geq 10\%$). *C*, *D*, Comparison of the average response predicted by the kernels with the actual average response for each subpopulation of cells. Averages were computed using the normalized responses during the 300 msec before and after the coherent motion began. The format shown is the same as that in Figure 2, *B* and *C*. The error bars are SEM.

vs attend-out: $\sigma_t = 15.0$ vs 14.0 msec for all good S/N and $\sigma_t = 16.2$ vs 17.4 msec for good S/N plus high attention). Comparing the two population temporal integration profiles suggests that cells with larger AM tended to have wider temporal integration windows [across our 44 good S/N cells, there was a weak correlation between temporal integration width in the attend-in condition (σ_t) and the amount of AM of 0.24].

The directional tuning widths of the population kernels were also similar between the two attention conditions (attend-in vs

attend-out: $\sigma_d = 55.6$ vs 58.4° for all good S/N and $\sigma_d = 57.7$ vs 59.2° for good S/N plus high attention). Similar effects of attention were observed for the speed profile (attend-in vs attend-out: $\sigma_s = 4.6$ vs 4.7°/sec for all good S/N and $\sigma_s = 4.0$ vs 3.9°/sec for good S/N plus high attention). At the higher speeds, however, our Gaussian fits underestimated the data and may reflect the skewed shape of MT neuron speed tuning expressed on a linear scale (Lagae et al., 1993). Overall, attention scaled the average population kernels with no appreciable change in shape.

AM of kernels and spike rate

Our results thus far suggest that attention scaled the kernels with no appreciable change in shape. However, not all neuronal spike rates were equally modulated by attention. Some neurons experienced more modulation of their spike rate than others. We therefore wanted to know how changes in kernel amplitude corresponded with changes in the neuronal spike rate attributable to attention.

In Figure 9A, we plot the AM of the fitted Gaussian amplitude versus the AM of the spike rate in response to the 0% coherent motion. AM is expressed as a ratio-metric index and, as in previous analyses, we only show the 44 neurons classified with good S/N kernels. Although there is scatter, this analysis suggests that as the modulation of the spike rate increases, so does the modulation of the fitted Gaussian amplitude ($r^2 = 0.46$). To reveal how the noise in our kernels affected our analysis, we separated out those neurons with the highest S/N kernels ($S/N \geq 1.3$; $n = 15$) (Fig. 9A, filled circles). Cells with the highest S/N kernels show good correlation between the fitted Gaussian amplitude and spike rate modulation ($r^2 = 0.52$). Thus, for neurons for which we had the most reliable kernel estimates, AM of kernel amplitude and spike rate were closely matched.

Model-free effects of attention on the kernels

Until now, we have used the parameters of the Gaussian fits to examine the effects of attention on the motion kernels in MT.

Although most kernels were well described by our product of Gaussian model, it is possible some kernels may have experienced changes attributable to attention that were not captured by the Gaussian fit. To address this, we used a model-free analysis to examine how our kernels changed with attention.

If attention multiplicatively scales the kernels, then it should be possible to estimate this attentional scale factor (β) with no prior assumptions on kernel shape. This can be expressed as $K_{IN} = \beta K_{OUT}$, where K_{IN} and K_{OUT} are the attend-in and attend-

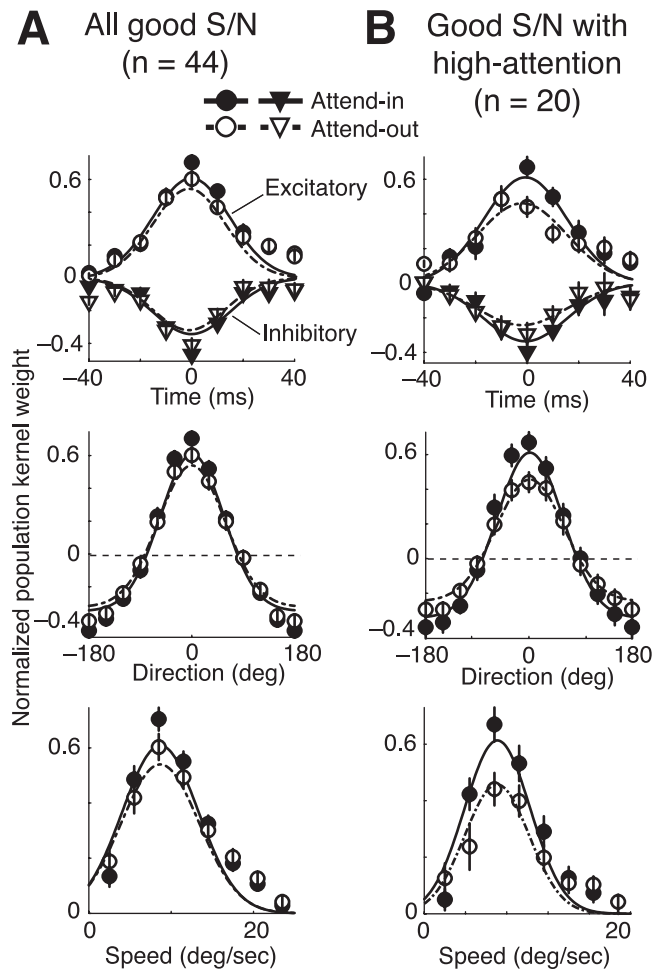


Figure 8. Temporal, directional, and speed profiles for the average normalized population kernels shown in Figure 7, *A* and *B*. *A*, Good S/N kernels. *B*, Good S/N plus high attention. The curves are the best fit Gaussian function (Eq. 6) using all kernel coefficients from Figure 7, *A* and *B*. The error bars are SEM.

out kernels, respectively. To compute β , we nonlinearly scaled the attend-out kernels to match the attend-in kernels (Eq. 7). Figure 9*B* shows the scale factor, β , for each cell plotted against the AM of spike rate. Both β and spike rate modulation are expressed in a ratiometric form. For our 44 good S/N neurons, β and spike rate modulation weakly covaried (all points; $r^2 = 0.13$). However, for the 15 neurons with the most reliable kernel estimates, the multiplicative scaling factor, β , was strongly related to the modulation in spike rate (filled points; $r^2 = 0.76$).

The multiplicative scaling of the kernels with attention can also be computed using the estimated kernel variance σ_{kernel}^2 . If the kernels are multiplicatively scaled copies of each other, then the scale factor can also be computed as the ratio of the SDs of the kernels (Eq. 8). Figure 9*C* shows this scale factor, γ , plotted against the AM of spike rate (both expressed as an equivalent ratiometric) and is nearly identical to the nonlinearly estimated scale factor β . Importantly, the two model-free scale factors, β and γ , are very similar to the scaling predicted by the Gaussian fits in Figure 9*A*. This is especially true for the kernels with the highest S/N. Thus, for cells for which we were able to estimate the most reliable kernels, the multiplicative scaling of the kernels with attention, regardless of how it was computed, did a good job of predicting the AM in average firing rates.

Computing a kernel scale factor does not address whether

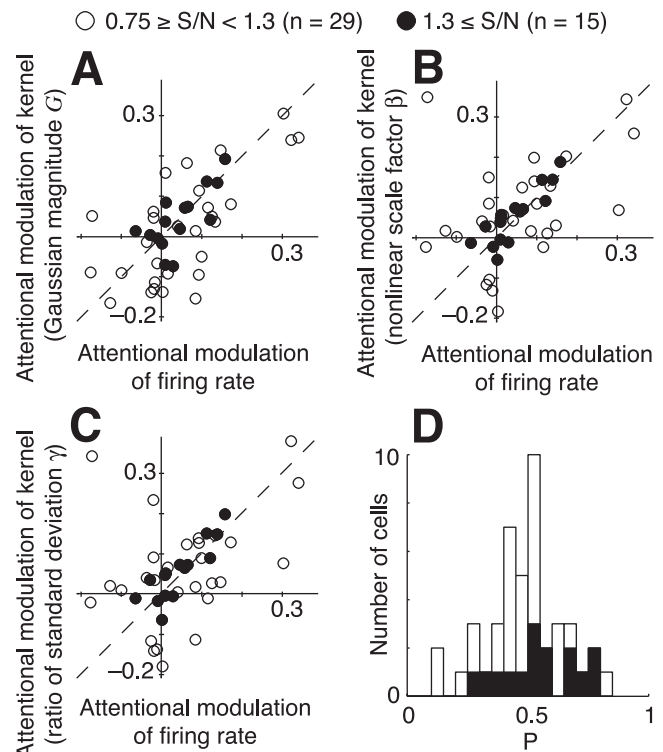


Figure 9. Attentional effects on the kernels and firing rates. *A*, AM of the amplitude (G) of the Gaussian fits versus the AM of firing rates for the 0% coherent motion. AM is expressed as a ratiometric $(P_{\text{in}} - P_{\text{out}})/(P_{\text{in}} + P_{\text{out}})$, where P_{in} and P_{out} correspond to the parameter of interest for the attend-in and attend-out conditions, respectively. Shown are all good S/N kernels. The filled symbols are kernels with a high S/N ≥ 1.3 . The dashed line is unity slope. *B*, The kernel scaling factor (β) expressed in an equivalent ratiometric form versus the AM of the firing rate. β was computed by nonlinearly scaling the kernel derived from the attend-out condition into the kernel corresponding to the attend-in condition. *C*, The kernel scaling factor (γ) expressed in an equivalent ratiometric form versus the AM of the firing rate. γ was computed by dividing the attend-in kernel SD by the attend-out kernel SD. *D*, Distribution of p values for the χ^2 test that the variance in the residual kernel, σ_{res}^2 , is greater than the variance predicted by a multiplicative scaling of the noise, σ_0^2 , or $H_0: \sigma_{\text{res}}^2 \leq \sigma_0^2$; $H_A: \sigma_{\text{res}}^2 > \sigma_0^2$. The filled bars correspond to those kernels with a high S/N. No kernel had a p value that was statistically significant at 0.05.

attention produced a multiplicative scaling alone or also produced a change in kernel shape. To examine whether kernels underwent a change in shape with attention, we computed the residual kernel $K_{\text{res}} = K_{\text{IN}} - \beta K_{\text{OUT}}$, using the estimated multiplicative scaling factor, β . If K_{IN} and K_{OUT} are multiplicatively scaled copies of each other, then K_{res} would be zero. Alternatively, non-zero elements in K_{res} would indicate a change in kernel shape with attention. One difficulty in determining whether K_{res} is flat (or contains all zeros), is that noise in our kernel estimates also contributes to non-zero elements in K_{res} . To determine whether the non-zero elements in K_{res} are attributable to noise or a shape change in our kernels, we used the estimates of kernel noise (σ_{noise}^2) to predict the variance in K_{res} (referred to as σ_{res}^2) we would expect from a pure multiplicative scaling with attention. If σ_{res}^2 is greater than the expected variance (σ_0^2), this would suggest that attention produced a change in kernel shape.

We used a χ^2 distribution to test whether the variance in K_{res} was greater than expected given the variance of the noise in K_{IN} and K_{OUT} (Eq. 11). For this analysis, we used the scaling factor β to compute the expected variance, σ_0^2 , in K_{res} . The distribution of p values for this test are shown in Figure 9*D* for our 44 good S/N kernels (one-tailed χ^2 for the null hypothesis $\sigma_{\text{res}}^2 \leq \sigma_0^2$).

Most p values in Figure 9D fall well above 0.05, which suggests no kernels underwent a significant change in shape. Because noise in the kernels reduces the power of this test to detect a shape change, neurons with the best S/N estimates are highlighted in Figure 9D (solid bars). Even for kernels with the lowest amounts of noise, the p values are well above significance, suggesting these kernels underwent a pure multiplicative scaling with attention. This result agrees with the population averages in Figure 7 that also suggest AM does not change kernel shape in a systematic way.

Discussion

We examined the effects of attention on the estimated linear transfer function of MT neurons. From a subpopulation of neurons that produced well defined kernels, we found that attention scaled the directional, speed, and temporal profiles of the linear transfer function without an appreciable change in shape. This suggests that AM does not alter the period of temporal integration of single neurons in MT. In addition, we found that neurons that produced the best kernels with the lowest noise also had their response properties to the coherent motion well described by the kernels and demonstrated strong transient response dynamics.

For kernels with good S/N that also had high AM, the effect of attention was also a multiplicative scaling of the kernel. Attention scaled these kernels by 34% (Fig. 8B), approximately the same amount the average firing rate was modulated by attention. Thus, even strong AM did not overtly change the shape of the integration window or shift the window in time. This agrees with previous observations that suggest attention does not alter the stimulus selectivity of neurons in the visual cortex (McAdams and Maunsell, 1999; Treue and Martinez-Trujillo, 1999; Recanzone and Wurtz, 2000). In these studies, the tuning curves of neurons to various stimulus dimensions (e.g., orientation, motion direction) were multiplicatively scaled by attention. Multiplicative interactions have also been observed between different stimulus dimensions (Tolhurst, 1973; Tolhurst and Movshon, 1975; Holub and Morton-Gibson, 1981; Albrecht and Hamilton, 1982; Sclar and Freeman, 1982; Skottun et al., 1987; Hamilton et al., 1989; Friend and Baker, 1993; McLean and Palmer, 1994; Geisler and Albrecht, 1997), suggesting that attentional and sensory inputs may be processed in a similar manner (McAdams and Maunsell, 1999). That attention does not change the shape of the neuronal kernel is supported by a recent study that found attention had a multiplicative effect on the psychophysically measured perceptual filter in humans performing a visual detection task (Eckstein et al., 2002).

Recently, two abstracts have described the effect of attention on spatiotemporal RFs (STRFs) of neurons in V1 (McAdams and Reid, 2003) and V4 (David et al., 2002). The V1 study reported AM of the peak of the estimated STRFs but did not address whether there were corresponding changes in shape. In contrast, the V4 study, which was computed under free-viewing conditions, reported changes in the spatial and temporal tuning of the estimated STRFs. It has also been suggested that attention produces subtle effects on spike timing without changes in the average rate (Niebur et al., 2002). Because of our 10 msec sampling interval, small changes in response latency or the length of the integration window would not be resolved by our kernel estimates.

Before we began this study, there was no *a priori* reason to think that spatial attention would multiplicatively scale temporal integration of motion. Previous studies used the number of spikes produced over hundreds of milliseconds to describe the

multiplicative effects of attention on stimulus selectivity in cortical neurons. Because changes in the synaptic conductance associated with attention could theoretically alter the membrane time constant of neurons (Rall, 1959), it was possible that attention could have dramatically affected the time course of integration, yet still have a multiplicative effect on spike counts. Other models that include the effects of voltage-dependent mechanisms suggest that the membrane time constant can be maintained under changes in synaptic input (Cook and Johnston, 1999). However, these models do not explain how multiplication, which plays a central role in models of sensory processing, is implemented by neurons (Albrecht and Geisler, 1991; Andersen et al., 1997; Sun and Frost, 1998; Taylor et al., 2000; Pena and Konishi, 2001). Although various mechanisms have been proposed to underlie multiplicative interactions in neurons (Srinivasan and Bernard, 1976; Heeger, 1992; Koch and Poggio, 1992; Mel, 1993; Salinas and Abbott, 1996; Hahnloser et al., 1999; Chance et al., 2002), the biophysical mechanisms remain unclear (but see Gabbiani et al., 2002).

The structure of our MT kernels is equivalent to those estimated by Livingstone et al. (2001). In that study, the authors used the motion vectors between two randomly moving dots located in the RF and compensated for small changes in eye position during fixation. Given the relatively large size of the RFs in our population (3.9–10.7°; median, 7.4°), we do not think that small changes in fixation added much noise to our kernel estimates.

Our average temporal profile of our kernels was similar in width to those computed by Borghuis et al. (2003) and Bair et al. (1997) using reverse correlation with random motion sequences. However, many of our best neurons that produced very high S/N kernels had narrow integration windows compared with these previous studies (such as the neuron in Fig. 2). We do not know the origin of this difference. In fact, neurons that produced well defined kernels with a good S/N tended to have shorter temporal integration (the correlation between the kernel S/N and the width of temporal integration was -0.27). Interestingly, neurons with higher AM tended to have slightly longer temporal integration windows (Fig. 8). The reason for this is not known.

It is also possible that some neurons may have been sensitive to motion occurring at speeds higher than that defined by our kernels. We do not think this was likely because we intentionally selected neurons with low speed preferences. This was because the animals were not trained with faster motion stimuli and had difficulty detecting the coherent motion at speeds greater than $\sim 15^\circ/\text{sec}$. Using the peak plus 1 SD along the speed dimension of the Gaussian fits ($\mu_s + \sigma_s$) as a measure of maximum speed sensitivity, for the good S/N kernels, the median and maximum were 12.1 and $23.2^\circ/\text{sec}$, which are below the speed cutoff of $23.5^\circ/\text{sec}$ used in our analysis.

Our average population kernels (Fig. 7A,B) show no trailing temporal inhibition, although several individual cells showed hints of this feature. In the cross-correlation between neuronal responses in MT and a one dimensional moving stimulus, Bair et al. (1997) showed that some MT cells have kernels with weak trailing inhibition. Borghuis et al. (2003) reported approximately half their kernels showed trailing inhibition of some kind. The reason for this discrepancy is unknown. However, our random motion vectors are a more “broadband-like” stimulus compared with the full-field movement impulses used by both the previous studies. Although we assumed the relationship between stimulus and neuronal responses in MT are stationary, recent results suggest that stimulus attributes can affect the estimated linear transfer function of neurons (Bair et al., 1997; Theunissen et al., 2000;

Blake and Merzenich, 2002), and this may explain the difference in our results. In addition, the use of the Gaussian function to parameterize the estimated kernel shapes was an approximation and not suggestive that the kernel profiles are true Gaussian.

Our analysis likely selected neurons that were driven by the global motion as described by our motion vectors. MT neurons respond to other stimulus attributes such as visual disparity (DeAngelis et al., 1998) and changes in contrast such that occurred during the replotting of the dots in our stimulus (flicker). We further assumed that MT neurons integrated the motion vectors uniformly across the RF. Therefore, our model would not account for non-uniform motion integration because of “hot spots” of integration in the RF. In addition, it has been suggested that MT neurons use a power-law summation when presented with multiple stimuli in the RF (Britten and Heuer, 1999), which would not be captured by our linear transfer function.

Livingstone et al. (2001) showed that some MT neurons can integrate motion across several previous stimulus frames, although the contribution of this integration was weak. Such interactions were not included in our model. Other factors, such as the number of spikes available (affected by the duration of recording and responsiveness to the 0% coherent motion), also contributed to whether the kernel of a cell could be estimated. It has also been shown that visual cortical neurons, including MT, respond slower to stimuli that increase firing rates than those that reduce firing rates (Bair et al., 2002). This asymmetry in latency is not featured in our linear model.

Despite these limitations, the kernels did a good job of reproducing the directional selectivity of the 44 good S/N neurons in our population (Fig. 7C). Thus, our estimated kernels captured a large portion of how the neurons processed the motion stimulus. The nonlinear components of the neuronal processing that were not captured by our kernels are evident in the fact the kernels underestimated responsiveness to coherent motion. It is possible that a static nonlinearity could be added to the output of the linear kernel to eliminate these discrepancies (Hunter and Korenberg, 1986). Although it is unlikely that attention would affect the nonlinear components of neuronal processing in such a way as to change the attentional effects on the linear kernels reported here, this is an important topic for future studies.

References

- Albrecht DG, Geisler WS (1991) Motion selectivity and the contrast-response function of simple cells in the visual cortex. *Vis Neurosci* 7:531–546.
- Albrecht DG, Hamilton DB (1982) Striate cortex of monkey and cat: contrast response function. *J Neurophysiol* 48:217–237.
- Andersen RA, Snyder LH, Bradley DC, Xing J (1997) Multimodal representation of space in the posterior parietal cortex and its use in planning movements. *Annu Rev Neurosci* 20:303–330.
- Bair W, Cavanaugh JR, Movshon JA (1997) Reconstructing stimulus velocity from neuronal responses in area MT. In: *Advances in neural information processing systems* (Mozer MC, Jordan MI, Patsche T, eds). Cambridge, MA: MIT.
- Bair W, Cavanaugh JR, Smith MA, Movshon JA (2002) The timing of response onset and offset in macaque visual neurons. *J Neurosci* 22:3189–3205.
- Barlow H, Tripathy SP (1997) Correspondence noise and signal pooling in the detection of coherent visual motion. *J Neurosci* 17:7954–7966.
- Blake DT, Merzenich MM (2002) Changes of AI receptive fields with sound density. *J Neurophysiol* 88:3409–3420.
- Borghuis BG, Perge JA, Vajda I, van Wessel RJ, van de Grind WA, Lankheet MJ (2003) The motion reverse correlation (MRC) method: a linear systems approach in the motion domain. *J Neurosci Methods* 123:153–166.
- Braun J, Koch C, Davis JL (2001) Visual attention and cortical circuits. Cambridge, MA: MIT.
- Britten KH, Heuer HW (1999) Spatial summation in the receptive fields of MT neurons. *J Neurosci* 19:5074–5084.
- Chance FS, Abbott LF, Reyes AD (2002) Gain modulation from background synaptic input. *Neuron* 35:773–782.
- Cook EP, Johnston D (1999) Voltage-dependent properties of dendrites that eliminate location-dependent variability of synaptic input. *J Neurophysiol* 81:535–543.
- Cook EP, Maunsell JHR (2002a) Attentional modulation of behavioral performance and neuronal responses in middle temporal and ventral intraparietal areas of macaque monkey. *J Neurosci* 22:1994–2004.
- Cook EP, Maunsell JHR (2002b) Dynamics of neuronal responses in macaque MT and VIP during motion detection. *Nat Neurosci* 5:985–994.
- Cook EP, Maunsell JHR (2003) Attentional modulation of motion integration by single neurons in MT of monkeys. *Soc Neurosci Abstr* 29:385.8.
- David SV, Mazer JA, Gallant JL (2002) Pattern specific attentional modulation of V4 spatiotemporal receptive fields during free-viewing visual search. *Soc Neurosci Abstr* 28:559.2.
- DeAngelis GC, Ohzawa I, Freeman RD (1995) Receptive-field dynamics in the central visual pathways. *Trends Neurosci* 18:451–458.
- DeAngelis GC, Cumming BG, Newsome WT (1998) Cortical area MT and the perception of stereoscopic depth. *Nature* 394:677–680.
- Desimone R, Duncan J (1995) Neural mechanisms of selective visual attention. *Annu Rev Neurosci* 18:193–222.
- DiCarlo JJ, Johnson KO, Hsiao SS (1998) Structure of receptive fields in area 3b of primary somatosensory cortex in the alert monkey. *J Neurosci* 18:2626–2645.
- Eckstein MP, Shimozaki SS, Abbey CK (2002) The footprints of visual attention in the Posner cueing paradigm revealed by classification images. *J Vis* 2:25–45.
- Friend SM, Baker CL Jr (1993) Spatio-temporal frequency separability in area 18 neurons of the cat. *Vision Res* 33:1765–1771.
- Fries P, Reynolds JH, Rorie AE, Desimone R (2001) Modulation of oscillatory neuronal synchronization by selective visual attention. *Science* 291:1560–1563.
- Gabbiani F, Krapp HG, Koch C, Laurent G (2002) Multiplicative computation in a visual neuron sensitive to looming. *Nature* 420:320–324.
- Geisler WS, Albrecht DG (1997) Visual cortex neurons in monkeys and cats: detection, discrimination, and identification. *Vis Neurosci* 14:897–919.
- Hahnloser R, Douglas RJ, Mahowald M, Hepp K (1999) Feedback interactions between neuronal pointers and maps for attentional processing. *Nat Neurosci* 2:746–752.
- Hamilton DB, Albrecht DG, Geisler WS (1989) Visual cortical receptive fields in monkey and cat: spatial and temporal phase transfer function. *Vision Res* 29:1285–1308.
- Heeger DJ (1992) Normalization of cell responses in cat striate cortex. *Vis Neurosci* 9:181–197.
- Holub RA, Morton-Gibson M (1981) Response of visual cortical neurons of the cat to moving sinusoidal gratings: response-contrast functions and spatiotemporal interactions. *J Neurophysiol* 46:1244–1259.
- Hunter IW, Korenberg MJ (1986) The identification of nonlinear biological systems: Wiener and Hammerstein cascade models. *Biol Cybern* 55:135–144.
- Jones JP, Palmer LA (1987) The two-dimensional spatial structure of simple receptive fields in cat striate cortex. *J Neurophysiol* 58:1187–1211.
- Judge SJ, Richmond BJ, Chu FC (1980) Implantation of magnetic search coils for measurement of eye position: an improved method. *Vision Res* 20:535–538.
- Koch C, Poggio T (1992) Multiplying with synapses and neurons. In: *Single neuron computation* (McKeena T, Davis JL, Zornetzer S, eds), pp 315–345. Orlando, FL: Academic.
- Lagae L, Raiguel S, Orban GA (1993) Speed and direction selectivity of macaque middle temporal neurons. *J Neurophysiol* 69:19–39.
- Livingstone MS, Pack CC, Born RT (2001) Two-dimensional substructure of MT receptive fields. *Neuron* 30:781–793.
- Martinez-Trujillo J, Treue S (2002) Attentional modulation strength in cortical area MT depends on stimulus contrast. *Neuron* 35:365–370.
- McAdams CJ, Maunsell JHR (1999) Effects of attention on orientation-tuning functions of single neurons in macaque cortical area V4. *J Neurosci* 19:431–441.
- McAdams CJ, Reid RC (2003) Effects of attention on the spatial and temporal structure of receptive fields in macaque primary visual cortex. *Soc Neurosci Abstr* 29:551.9.

- McLean J, Palmer LA (1994) Organization of simple cell responses in the three-dimensional (3-D) frequency domain. *Vis Neurosci* 11:295–306.
- Mel BW (1993) Synaptic integration in an excitable dendritic tree. *J Neurophysiol* 70:1086–1101.
- Niebur E, Hsiao SS, Johnson KO (2002) Synchrony: a neuronal mechanism for attentional selection? *Curr Opin Neurobiol* 12:190–194.
- Pena JL, Konishi M (2001) Auditory spatial receptive fields created by multiplication. *Science* 292:249–252.
- Posner MI (1980) Orienting of attention. *Q J Exp Psychol* 32:3–25.
- Press WH, Flannery BP, Teukolsky SA, Vetterling WT (1986) Numerical recipes: the art of scientific computing. Cambridge, UK: Cambridge UP.
- Rall W (1959) Branching dendritic trees and motoneuron membrane resistivity. *Exp Neurol* 1:491–527.
- Recanzone GH, Wurtz RH (2000) Effects of attention on MT and MST neuronal activity during pursuit initiation. *J Neurophysiol* 83:777–790.
- Reynolds JH, Pasternak T, Desimone R (2000) Attention increases sensitivity of V4 neurons. *Neuron* 26:703–714.
- Robinson DA (1963) A method of measuring eye movement using a scleral search coil in a magnetic field. *IEEE Trans Biomed Eng* 10:137–145.
- Salinas E, Abbott LF (1996) A model of multiplicative neural responses in parietal cortex. *Proc Natl Acad Sci USA* 93:11956–11961.
- Sclar G, Freeman RD (1982) Orientation selectivity in the cat's striate cortex is invariant with stimulus contrast. *Exp Brain Res* 46:457–461.
- Seidemann E, Newsome WT (1999) Effect of spatial attention on the responses of area MT neurons. *J Neurophysiol* 81:1783–1794.
- Skottun BC, Bradley A, Sclar G, Ohzawa I, Freeman RD (1987) The effects of contrast on visual orientation and spatial frequency discrimination: a comparison of single cells and behavior. *J Neurophysiol* 57:773–786.
- Srinivasan MV, Bernard GD (1976) A proposed mechanism for multiplication of neural signals. *Biol Cybern* 21:227–236.
- Sun H, Frost BJ (1998) Computation of different optical variables of looming objects in pigeon nucleus rotundus neurons. *Nat Neurosci* 1:296–303.
- Taylor WR, He S, Levick WR, Vaney DI (2000) Dendritic computation of direction selectivity by retinal ganglion cells. *Science* 289:2347–2350.
- Theunissen FE, Sen K, Doupe AJ (2000) Spectral-temporal receptive fields of nonlinear auditory neurons obtained using natural sounds. *J Neurosci* 20:2315–2331.
- Theunissen FE, David SV, Singh NC, Hsu A, Vinje WE, Gallant JL (2001) Estimating spatio-temporal receptive fields of auditory and visual neurons from their responses to natural stimuli. *Network* 12:289–316.
- Tolhurst DJ (1973) Separate channels for the analysis of the shape and the movement of moving visual stimulus. *J Physiol (Lond)* 231:385–402.
- Tolhurst DJ, Movshon JA (1975) Spatial and temporal contrast sensitivity of striate cortical neurones. *Nature* 257:674–675.
- Treue S (2003) Visual attention: the where, what, how and why of saliency. *Curr Opin Neurobiol* 13:428–432.
- Treue S, Martinez-Trujillo JC (1999) Feature-based attention influences motion processing gain in macaque visual cortex. *Nature* 399:575–579.
- Treue S, Maunsell JHR (1999) Effects of attention on the processing of motion in macaque middle temporal and medial superior temporal visual cortical areas. *J Neurosci* 19:7591–7602.
- Van Essen DC, Maunsell JHR, Bixby JL (1981) The middle temporal visual area in the macaque: myeloarchitecture, connections, functional properties and topographic organization. *J Comp Neurol* 199:293–326.
- White JM, Sparks DL (1986) Saccades to remembered targets: a behavioral and neurophysiological study. *Invest Ophthalmol Vis Sci [Suppl]* 27:155.
- Zar JH (1999) Biostatistical analysis, Ed 4. Upper Saddle River, NJ: Prentice Hall.