Spatial Summation, End Inhibition and Side Inhibition in the Middle Temporal Visual Area (MT)

Leo L. Lui, James A. Bourne, and Marcello G. P. Rosa
Department of Physiology, Monash University, Clayton, Victoria, Australia

Submitted 25 September 2006; accepted in final form 9 November 2006

Lui LL, Bourne JA, Rosa MG. Spatial summation, end inhibition and side inhibition in the middle temporal visual area (MT). J Neurophysiol 97: 1135–1148, 2007. First published November 15, 2006; doi:10.1152/jn.01018.2006. We investigated the responses of single neurons in the middle temporal area (MT) of anesthetized marmoset monkeys to sine-wave gratings of various lengths and widths. For the vast majority of MT cells maximal responses were obtained on presentation of gratings of specific dimensions, which were typically asymmetrical along the length and width axes. The strength of end inhibition was dependent on the width of the stimulus, with many cells showing clear end inhibition only when wide gratings were used. Conversely, the strength of side inhibition was dependent on stimulus length. Furthermore, for over one third of MT cells length summation properties could not be defined without consideration of stimulus width and vice versa. These neurons, which we refer to as “length–width inseparable” (LWI) cells, were rare in layer 4. The majority of LWI neurons was strongly inhibited by wide-field stimuli and responded preferentially to gratings that were elongated, along either the length or width dimensions. However, rather than forming a homogeneous and entirely distinct group, LWI cells represented the upper end of a continuum of complexity in spatial summation response properties, which characterized the population of MT cells. Only a minority of MT neurons (22.3%) showed no evidence of inhibition by wide-field stimuli, with this type of response being common among layer 5 cells. These results demonstrate distinct patterns of spatial selectivity in MT, supporting the notion that neurons in this area can perform various roles in terms of grouping and segmentation of motion signals.

INTRODUCTION

As well as their traditionally accepted roles in the processing of direction of motion and speed (Albright 1984; Dubner and Zeki 1971; Maunsell and Van Essen 1983), neurons in the middle temporal area (MT; also known as V5) are involved in the integration and segmentation of motion signals (Bradley and Anderson 1998; Movshon et al. 1985). It was previously proposed that a large proportion of cells in MT have receptive fields that include antagonistic motion-sensitive surrounds (Allman et al. 1985; Tanaka et al. 1986), such that the magnitude of the elicited responses is modulated by the relative direction of motion of patterns presented in the “classical” excitatory receptive field and in its surround region. Functionally, cells with antagonistic motion surrounds are likely to code local motion contrasts, such as those experienced during the motion of individual objects in extrapersonal space, whereas cells that lack such suppressive surrounds respond optimally to wide-field motion, such as that experienced during one’s self-motion (Born and Tootell 1992).

Ideas regarding what constitutes the receptive field of an MT neuron continue to be refined, with new evidence pointing to increasingly complex spatial profiles. Contrary to early proposals, there are now data to suggest that a model consisting of concentric excitatory center and inhibitory surround regions cannot adequately describe the structure of a substantial proportion of MT receptive fields. In many cells the spatial structure of the inhibitory surround is not homogeneous, perhaps endowing neurons with the ability to make local motion comparisons (Raiguel et al. 1995; Xiao et al. 1995). In addition, a subpopulation of MT cells shows complex preference for motion contrasts, whereby optimal stimuli consist of a motion discontinuity, irrespective of exactly where this is present over the extent of the receptive fields (Born 2000). MT cells can also respond to behaviorally relevant stimuli in locations remote from their classical receptive field (Zaksas and Pasternak 2005) and their summation properties change as a function of contrast in a way that cannot be predicted by linear summation (Heuer and Britten 2002; Pack et al. 2005). In summary, the evidence to date indicates that strict spatial linearity may not apply to MT receptive fields: the response to two simultaneous stimuli cannot always be predicted by the sum of the responses to the same two stimuli, presented separately (Britten and Heuer 1999).

The receptive field structure of MT neurons was previously investigated with two-dimensional (2D) stimuli, in particular those consisting of drifting random dot patches, alone or in combination with single bars (Born 2000; Perge et al. 2005; Raiguel et al. 1995; Xiao et al. 1995). These studies either varied the diameter of circular stimuli or presented small stimuli at different locations within the receptive field. In contrast, much of our modern understanding of center-surround interactions in other visual areas, including the primary visual cortex (V1), comes from studies using combinations of luminance-defined gratings (e.g., Angelucci et al. 2002; Bair et al. 2003; Cavanaugh et al. 2002; Levitt and Lund 2002; Lui et al. 2006; Sceniak et al. 1999, 2002; Xu et al. 2005). In the current study, we determined the responses of MT neurons to sine-wave gratings of optimal direction of motion, spatial frequency, and temporal frequency, presented within rectangular windows centered in the cell’s receptive field. Like most natural objects, the elements of such gratings have oriented boundaries defined by first-order cues; it is of interest to determine how MT cells respond to such patterns as a function of variations in the size of their angular projection in the retina. The stimuli were varied independently in terms of length (i.e.,
parallel to the orientation of the elements of the grating; Fig. 1A) and width (parallel to the direction of motion of the elements of the grating). Previous work in V1 indicates that responses can be modulated in different ways by stimuli presented along these dimensions (e.g., Born and Tootell 1991; Cavanaugh et al. 2002; DeAngelis et al. 1994; Kapadia et al. 2000), particularly in terms of strength of end inhibition (i.e., inhibition caused by antagonistic zones located adjacent to, and aligned with, the receptive field’s “length” axis) and side inhibition (inhibition caused by antagonistic zones adjacent to, and aligned with, the receptive field’s “width” axis). However, no comparable information exists for MT, despite earlier work in which first-order length selectivity was investigated using single moving bars (e.g., Cheng et al. 1994; Felleman and Kaas 1984; Olavarria et al. 1992). We found that the length and width summation properties of MT neurons are not immutable; rather, in many cells length selectivity depends on the width of the stimulus and vice versa. These observations add support to the argument that the receptive fields of MT cells are complex structures and, in the light of other recent studies, indicate that the responses of MT cells are likely to involve nonlinear spatial summation of signals.

M E T H O D S

Animals

Single-unit recordings were obtained from MT of 12 adult New World monkeys (Callithrix jacchus, the common marmoset). Marmosets are small (about 350 g) diurnal primates, with eyes characterized by a well-developed fovea and steep central-peripheral gradients of photoreceptors and ganglion cells (Wild et al. 1996). The boundaries, topographic organization, and connections of MT have been well studied in this species (Krubitzer and Kaas 1990; Palmer and Rosa 2006). Moreover, it was demonstrated that, similar to Old World monkeys, the vast majority of marmoset MT neurons are direction selective (Lui et al. 2005; Rosa and Elston 1998). Experiments were conducted in accordance with the Australian Code of Practice for the Care and Use of Animals for Scientific Purposes and all procedures were approved by the Monash University Animal Ethics Experimentation Committee. Some of these animals were also part of parallel anatomical investigations of the visual and auditory cortices, having received tracer injections in the contralateral hemisphere.

Surgical preparation

The methods for preparation and procedures for recording and visual stimulation were previously described in detail (Bourne and Rosa 2003). Anesthesia was induced with ketamine (50 mg/kg; Parnell, Sydney, Australia) and xylazine (3 mg/kg; Illium, Sydney, Australia), allowing a tracheotomy, cannulation of the saphenous vein, and a craniotomy to be performed. The dura mater overlying the dorsal cortical surface was covered with a thin layer of silicone oil to prevent desiccation. After all surgical procedures were completed, the animal was administered an intravenous infusion of pancuronium bromide (0.1 mg · kg⁻¹ · h⁻¹; Organon, Sydney, Australia), combined with sufentanil (6 μg · kg⁻¹ · h⁻¹; Janssen-Cilag, Sydney, Australia) and dexamethasone (0.4 mg · kg⁻¹ · h⁻¹; David Bull, Melbourne, Australia), in a saline/glucose solution, which induced muscular paralysis while maintaining anesthesia. The animal was artificially ventilated with a gaseous mixture of nitrous oxide and oxygen (7:3). The electrocardiogram (SpO2 level) and the level of cortical spontaneous activity were continuously monitored. Administration of atropine (1%) and phenylephrine hydrochloride (10%) eye drops (Sigma Pharmaceuticals, Melbourne, Australia) resulted in mydriasis and cycloplegia. Appropriate focus and protection of the corneas from desiccation were achieved by means of hard contact lenses selected by streak retinoscopy. These lenses brought into focus the surface of a computer monitor located 40 cm in front of the animal. Visual stimuli were presented to the eye contralateral to the hemisphere from which the neuronal recordings were obtained (neurons in marmoset MT either respond equally well to input from the two eyes or show a slight contralateral bias).

Electrophysiological recordings

Parylene-coated tungsten microelectrodes with exposed tips of 10 μm were moved in 5- to 10-μm steps until a unit with amplitude that was clearly above the background was detected. Amplification and filtering of the electrophysiological signal were achieved by a Model 1800 Microelectrode AC amplifier (AM systems, Everett, WA) and a 50-Hz eliminator (HumBug, Quest Scientific, Vancouver, Canada). The processed signal was fed into a waveform discrimination system (SPS-8701, Signal Processing Systems, Adelphia, New Jersey) run with the aid of a Pentium PC, allowing the isolation of single-unit signals by means of a template-matching algorithm. The neural activity was continuously monitored by means of loudspeakers, an oscilloscope (raw signal), and computer displays (processed signal, corresponding to the unit under investigation). For quantitative analyses, the spike trains processed by the SPS-8701 system were collected by a high-fidelity interface (ITC-16, Instrutech, Great Neck, NY) into a Macintosh computer, which also controlled the visual stimulus generation and displayed the accumulated peristimulus time histograms (PSTHs) in real time.

Area MT was initially located by stereotaxic coordinates obtained as part of the study by Rosa and Elston (1998) and by visualization of the posterior tip of the lateral sulcus. The provisional attribution of recording sites to MT during the experiment was based on mapping of multunit receptive fields, using electrodes that penetrated vertically. The dorsoventral trajectory of the electrodes resulted in a gradual shift of MT receptive field center positions, from the lower quadrant toward the upper quadrant, and a gradual decrease in the eccentricity of receptive fields (Rosa et al. 2000). These topographic trends allowed us to estimate the dorsal and ventral borders of MT during the recording session; the location of the recording sites within this area

![Diagram](http://jn.physiology.org/Downloadedfrom/10.220.33.5/10.220.33.5/files/1136_L.L.LUI_J.A.BOURNE_M.G.P.ROSA)
was later confirmed by histological examination of the electrode tracks.

The initial exploration of the receptive field boundaries was conducted using hand-held stimuli (such as rulers of various lengths and widths) moved at various speeds, orientations, and directions of motion across the screen of a 20-in. monitor (refresh rate 75 Hz, resolution 1,024 × 768 pixels). The estimates of receptive field borders were then refined by presenting computer-generated, small, mouse-controlled moving bars or gratings of near-optimal orientation and direction of motion, while listening to the cell’s activity. The receptive fields mapped in this way were similar in size and shape to those reported by previous studies in marmoset MT (Rosa and Elston 1998). Finally, to ensure that the stimuli used for quantitative tests were appropriately centered on the receptive fields, we used small stimuli (such as a 1° crosshair or square) flashed or moved at different points in the receptive field, to determine the point in the visual field that elicited the maximal response. Reflecting previous observations in the macaque (Raiguel et al. 1995), we found that the point of maximal response usually coincided with the geometrical center of the excitatory receptive field. Repeated estimations of the receptive field center of the same cell typically yielded estimates within 1° of each other and never exceeding 2°. This margin of error is small relative to the dimensions of the minimum response fields in the explored region of MT (mean square root of receptive field area = 9.4°, range 5–16°; Rosa and Elston 1998).

After determination of the receptive field center, neuronal response properties were studied quantitatively using computer-controlled stimuli. The stimuli for quantitative tests consisted of rectangular patches of drifting gratings presented against a uniform gray background, which corresponded to the grating’s average luminance (2.6 cd/m²). Each trial started with a 0.5-s presentation of the gray screen, during which measurements of spontaneous activity were obtained. Drifting gratings were then presented for 2 s, with the phase varied from trial to trial, at constant speed. An intertrial interval of 4 s, during which the gray screen was presented, separated trials.

Tests were performed to determine optimal values of direction of motion, spatial frequency (range tested 0.08–2.4 cycles/deg), and temporal frequency (0.18–10.9 Hz) for each neuron. Thus stimuli of near-optimal spatiotemporal characteristics were used for the size-selectivity tests. References to the length of the grating indicate the dimension that had constant luminance, whereas along the width of the grating luminance changed according to a sine-wave function. The direction of motion was always parallel to the width of the grating (Fig. 1A). Five values of length and width were used, including dimensions that were smaller than the receptive fields of most MT neurons (2°, 4°), others within the range of excitatory receptive field sizes we typically observed (8°, 16°) and others that were larger than most receptive fields (30°). The length and width of the stimuli were varied independently, resulting in 25 stimulus conditions that were presented a minimum of eight times, in randomized sequence (Fig. 2).

Histology

At the end of the experiment the animal was administered an overdose of sodium pentobarbitone (Rhone-Merieux, Brisbane, Australia) and perfused transcardially with 0.9% saline, followed by 4% paraformaldehyde in 0.1 M phosphate buffer (pH 7.4). After cryoprotection by increasing concentrations of sucrose and sectioning, alternate slides (40 μm) were stained for Nissl substance and myelin, allowing for reconstruction of electrode tracks relative to histological borders. Electrode tracks were reconstructed with the aid of small electrolytic lesions (4 μA, 10 s), which were placed at various sites

FIG. 2. Sets of peristimulus time histograms (50-ms bin width) illustrating neuronal activity in response to gratings of different dimensions, centered on the receptive field. Gray highlight indicates the time during which the stimuli were presented. Numbers above each histogram denote the cell’s mean response level as a ratio of its maximum response. A: neuron without significant end inhibition or side inhibition (spatial frequency = 0.15 cycles/deg, temporal frequency = 4.8 Hz). B: cell displaying complex inhibition in both dimensions (spatial frequency = 0.30 cycles/deg; temporal frequency = 9.6 Hz). Significant end inhibition was observed in response to the wide patches (8, 16, and 30°), whereas this was not the case for narrower patches (2 and 4°). Side inhibition was significant for the longest patch (30°), but not for shorter patches.
during the experiment. Only cells confirmed as belonging to MT, on the basis of the pattern of myelination observed in sections stained with the method of Gallyas (1979), were included in the present report. The laminar distribution of the recorded units was also determined based on histological analysis (e.g., Bourne and Rosa 2006).

Data analysis

The responses of each cell were converted into PSTHs with a 10-ms bin width, which formed the basis of all subsequent analyses. A single trial response was computed as the mean firing rate over the entire duration of stimulus presentation at 66% contrast. Spontaneous activity was calculated from the mean firing rate during the 500 ms before stimulus onset. The results reported here include only cells that responded to near-optimal stimuli at a level at least 2 SDs above the mean spontaneous activity. Size-tuning curves, both one and two dimensional, were fitted using the Matlab function “lsqcurvefit” (The MathWorks, Natick, MA); the best fit for each neuron was obtained by minimizing the sum-squared error between the neuronal response and the values obtained by the function. Curve fitting was based on the entire matrix of single trial responses, rather than the mean responses to each stimulus condition. The fittings were constrained by the requirement that the curves cross the level of spontaneous activity at length and width zero (e.g., Bourne et al. 2004). However, data points corresponding to spontaneous activity were not included in the calculation of $r^2$ values. Nonparametric statistical tests were used in most analyses because of nonnormal distributions of data (see Results for details).

RESULTS

We characterized the responses of 157 neurons located in MT, with receptive fields centered at an average eccentricity of 10.6 ± 5.2°. Eighty-eight percent of these cells (138/157) were classified as strongly direction selective using a quantitative criterion based on vector sum analysis (Lui et al. 2006).

“Classic” one-dimensional analysis of size selectivity: end inhibition and side inhibition

The strength of end inhibition demonstrated by many MT cells was dependent on the width of the grating stimulus. For the purposes of this analysis neurons were classified as end inhibited if, using stimuli of constant width, the response to the longest grating in our set of stimuli (30°) was >2 SEs lower than the response evoked by a shorter, near-optimal grating. Overall, 45% of the cells in MT (71 units) showed no evidence of end inhibition, irrespective of the width of the stimulus (e.g., Fig. 2A). On the other hand, only 13% of our sample (21/157 units) showed consistent end inhibition, here defined as end inhibition that was statistically significant in tests using at least four of the five grating widths tested. Thus in the majority of MT cells end inhibition was present, but depended on stimulus width. In particular, a substantial subpopulation of MT cells showed end inhibition when stimulated with the wider gratings in our stimulus set, but not with the narrow ones (e.g., Fig. 3, A and B). This point is illustrated by the fact that, taking the entire population of MT cells into consideration, estimates of the percentage of end inhibited cells increased significantly ($r^2 = 0.94, P = 0.016$) as a function of the width of the test stimulus (Fig. 1B).

Conversely, we also observed that, in the majority of MT cells, side inhibition is a response property that is conditional on stimulus length (Fig. 1C). For this analysis, neurons were considered side inhibited if the response to the 30°-wide grating was >2 SEs lower than the activity evoked by a narrower grating of similar length. Forty-eight percent of the MT cells (76/157) showed no evidence of side inhibition. However, only 9% of cells (14/157) displayed consistent side inhibition, evident in tests using at least four of the five grating lengths tested. Estimates of the proportion of cells manifesting side inhibition were positively correlated with the length of the grating presented ($r^2 = 0.97, P = 0.005$; see Fig. 1C), mainly reflecting a relatively common pattern of response whereby side inhibition occurred only on presentation of long gratings (Figs. 2B and 3, C and D). Interestingly, conditional end and side inhibition properties were not mutually exclusive, with 18 units (11% of the sample) displaying complex behavior in both dimensions (e.g., see unit illustrated later in Fig. 9A).

FIG. 3. Complex patterns of spatial selectivity of 4 middle temporal area (MT) neurons. In A and B, the x-axis represents the change in grating length and the data symbols refer to the grating width. Significant end inhibition was observed for only the widest 2 patches. In C and D the x-axis represents the changes in width and the different symbols refer to different grating lengths. Significant side inhibition was observed for only the longest grating patches. Neuronal responses that did not show significant end or side inhibition were fitted with error functions (integral of Gaussian functions) and responses showing significant inhibition were fitted with difference of error functions.
As demonstrated in Fig. 3, neurons with conditional size-selectivity properties in MT did not form a homogeneous group, in terms of either length or width selectivity. For example, the cell illustrated in Fig. 3A responded best to gratings that were short and wide. Even though the responses to long gratings were always weak, the strongest responses to short gratings, and the most evident end inhibition, occurred on the presentation of wide patches. In contrast, the optimal stimulus length for the unit shown in Fig. 3B depended on the stimulus width. Figure 3, C and D represents observations with respect to the occurrence of side inhibition and optimal width at different stimulus lengths.

Finally, for comparison with earlier studies, in which stimuli varied simultaneously along the length and width dimensions of MT receptive fields (e.g., those in which circular gratings were used), we also examined “symmetrical” inhibition, by considering the responses to the five grating patches of equal length and width in our stimulus set. Using a criterion analogous to that applied to the unidimensional analyses (i.e., response to the 30° × 30° grating being more than 2 SEs lower than the activity evoked by a smaller square grating), we found that only 58 cells (37%) in our sample showed evidence of inhibition by stimuli that were larger than an optimal value. Thus analyses restricted to stimuli that grow symmetrically in any dimension. Thus although acknowledging that additional parameters may be needed to encapsulate the full range of response properties found among MT cells, we did not exclude cells with low \( r^2 \) values from the analyses detailed below.

**End inhibition and side inhibition according to the 2D model**

We used the 2D model to classify cells in terms of their length and width selectivity characteristics. Confidence intervals for parameter estimates were computed from the Jacobian matrix and the residuals using the Matlab function “nlparci.” If the 95% confidence interval prediction for the parameter \( l_{opt} \) (optimal length) did not overlap with 30° (maximum patch length tested), the cell was classified as end inhibited; conversely, if the 95% confidence interval prediction for parameter \( w_{opt} \) (optimal width) did not extend to 30°, the cell was classified as side inhibited. Using these criteria, 44 cells (28.0% of our sample) were classified as showing inhibition in both the length and width dimensions (e.g., Fig. 4, A and B). A similar proportion of MT neurons (29.3%; 46 cells) showed end inhibition but no side inhibition (Fig. 4C), whereas 20.4% (32 cells) showed only side inhibition (Fig. 4D); the difference between the proportions of end and side inhibited cells was not statistically significant [\( \chi^2(1) = 2.513, P = 0.11 \)]. The remaining 35 cells (22.3%) were not inhibited along either dimension, thus qualifying as true “wide-field” neurons (Born 2000; Born and Tootell 1992; see Fig. 4, E and F). Although the responses of the unit illustrated in Fig. 4E summed up to the maximum dimension of our stimuli, those shown in Fig. 4F saturated at estimated values of \( l_{opt} = 26.1° \) and \( w_{opt} = 20.2° \); however, the 95% confidence interval prediction of these parameters extended to 30°. The occurrences of significant end and side inhibition appeared to be statistically independent. Cells showing end inhibition were not any more likely to be side inhibited than cells without end inhibition; conversely, whether a cell was end inhibited could not be reliably predicted by side inhibition [\( \chi^2(1) = 0.02, P > 0.05 \)].

The majority of wide-field cells defined on the basis of the 2D model (30%5; 86%) also showed no evidence of inhibition along either the length or width dimension, according to the one-dimensional (1D) analyses, which were based on different (threshold-based) statistical criteria. In particular, none of the wide field cells showed evidence of inhibition when only the responses to symmetrical stimuli were considered. Of the 90 cells showing end inhibition according to the 2D analysis, 75
(84% were also considered to be end inhibited according to the 1D analysis. A similar relationship was observed with respect to side inhibition (83%, or 63/76 of the units). In summary, the 2D model provides a satisfactory account of “classical” end and side inhibition among MT cells, while also allowing for interactions between these dimensions, as described later.

Preferred length and width

Figure 5 is a summary of the optimal lengths and widths as estimated by parameters \( l_{\text{opt}} \) and \( w_{\text{opt}} \). The median preferred length in the entire sample was found to be 13.8°, whereas the median optimal width was 21.3°. These values were significantly different (Wilcoxon signed-rank test, \( z = 2.14, P = 0.032 \)), indicating that the excitatory summation zones of MT cells tend to be elongated in the direction parallel to the preferred axis of motion. This difference persisted when we considered the optimal lengths for end inhibited cells alone, in comparison with the optimal widths for side inhibited cells (median \( l_{\text{opt}} = 7.10° \), median \( w_{\text{opt}} = 10.20° \), Wilcoxon rank-sum test, \( z = 2.38, P = 0.017 \); see arrows in Fig. 5). Thus the difference in optimal lengths and widths cannot be attributed solely to the larger number of non-side inhibited cells, in comparison to end inhibited cells. No correlation was found between optimal lengths and widths (Spearman’s correlation, \( r^2 = 0.07, P > 0.05 \)) and 61% (96 units) of the cells in our sample preferred stimulus dimensions that were clearly asymmetrical. Figure 6 illustrates the distribution of optimal grating aspect ratios (optimal length/optimal width) for the entire MT population. For 37 cells (23.6% of the sample) the optimal

FIG. 4. Examples of neuronal selectivity to length and width of a grating patch centered on the receptive field. Spontaneous activity was subtracted from the responses, after which the responses were normalized with respect to the mean peak. Illustrated 3-dimensional (3D) surfaces were fitted using the “oriented” version of the 2-dimensional (2D) model, as explained in RESULTS. A and B: neurons that were inhibited in both dimensions. C: neuron showing end inhibition only. D: neuron showing side inhibition only. E and F: cells showing no significant inhibition (wide-field). Calculation of the \( r^2 \) values (shown for all fits) did not include the points corresponding to length and width zero (i.e., level of spontaneous activity), although these were used to constrain the surfaces.

FIG. 5. A, bottom left: combinations of length and width eliciting the maximal responses among the neurons in our sample. Classification of the neurons in terms of type of inhibition displayed is also indicated by different symbols. B and C: distributions of preferred lengths and widths, respectively. Arrows indicate the medians for the subpopulation of cells showing significant inhibition by stimuli larger than an optimal value, according to the 2D analysis. Black bars indicate cells in which no end or side inhibition was detected.
stimuli had aspect ratios >2:1, whereas for 59 cells’ (37.6%) the optimal stimuli aspect ratios were <1:2. As expected from the analysis of optimal lengths and widths, significantly more cells preferred gratings that were short and wide, rather than long and narrow [23.6%; \chi^2(1) = 5.04, \( P = 0.023 \)].

The distributions of the preferred lengths and widths (Fig. 5), when considered independently, were not unimodal (length: Dip = 0.112, \( P < 0.001 \); width: Dip = 0.055, \( P = 0.002 \); Hartigan and Hartigan 1985). Indeed, Fig. 5 indicates that the distributions of length and width selectivity in MT can be reasonably described as bimodal, with one mode reflecting cells that display inhibition along a particular dimension, whereas the other mode represents cells that are not inhibited along that dimension. One obvious question that arises from this conclusion is whether there is also evidence of a bimodal distribution when both dimensions are considered in a dependent manner, perhaps reflecting the distinction between “local motion” and “wide-field motion” units reported by earlier studies (Born and Tootell 1992). As a test of whether this is the case, in Fig. 7 we illustrate the ratios between the responses to the 30° grating and the response to a grating of optimal dimensions, based on the 2D model. The pattern here is certainly not as clear: although a second peak appears to occupy the upper end of the distribution (corresponding to “wide-field” cells), this distribution does not deviate significantly from a unimodal distribution (Dip = 0.033, \( P = 0.268 \)). At least to a first approximation, these results indicate that neurons in MT may form a continuum in terms of response properties, from those showing specificity for stimulus size and strong antagonistic surrounds, to “wide-field” units.

**Interaction effects between length and width**

The 1D analyses described above suggested that neuronal selectivities to length and width must be viewed, in some cases, as mutually interdependent. To study this phenomenon in a more rigorous manner, we investigated the effect of constraining parameter \( \theta \), which in the 2D model allows for a rotation of the Gaussian around its peak, thereby modeling interactions between length and width selectivity. In a “nonoriented” version of the 2D model, parameter \( \theta \) was constrained to 0°. Whereas Eq. 1 remains the same, Eqs. 2 and 3 are substituted by Eqs. 4 and 5, respectively.

\[
R(l, w) = \left[ \exp\left(-\frac{l^2}{\sigma_l^2}\right)\right] \left[ \exp\left(-\frac{w^2}{\sigma_w^2}\right)\right]
\]

where

\[
l' = \log\left(\frac{1 + l_{off}}{l_{opt} + l_{off}}\right)
\]

and

\[
w' = \log\left(\frac{w + w_{off}}{w_{opt} + w_{off}}\right)
\]

The responses of units showing an interaction effect between length and width selectivity are better modeled by the oriented version of the model (Eqs. 1–3), whereas those in which the preferred length and/or width are invariant are equally well described by the nonoriented model. To detect cells in these categories, we compared the residuals to the oriented and nonoriented fits using a sequential F-test (DeAngelis and Uka 2003; Nover et al. 2005), where the degrees of freedom were adjusted according to the number of free parameters. The distribution of F-values and the corresponding levels of significance are illustrated in Fig. 8A. One hundred units (64% of the sample) had an F-value >3.87 (\( P = 0.05 \)) and, among those, 78 cells (50%) had an F-value >6.73 (\( P = 0.01 \)), suggesting that interdependent length and width selectivities represent a relatively common characteristic of MT cells. However, given that the sequential F-test is best applied to functions that are linear in their parameters, we adopted a much more stringent criterion (\( P < 0.001 \); \( F = 11.06 \)), which provided a conservative estimate of the proportion of units showing an interaction between length and width selectivities.

Figure 8B compares the \( r^2 \) values yielded by the fits to the oriented and nonoriented models. As expected, the oriented model yielded a higher median \( r^2 \) value (0.81) than the nonoriented model (0.77; Wilcoxon signed-rank test: \( z = -10.82, \ P < 0.001 \)). For 54 neurons (34% of the sample) the oriented model provided a significantly better description (at a criterion level of \( P < 0.001 \)) of the observed selectivity for stimulus dimensions: the length and width selectivity properties of these cells were not separable. Three examples of such cells are illustrated in Fig. 9, with the results yielded by both the oriented and the nonoriented models displayed. It is clear from

![Figure 6](http://jn.physiology.org/.../1141SPATIAL_SELECTIVITY_IN_VISUAL_AREA_MT)

![Figure 7](http://jn.physiology.org/.../1141SPATIAL_SELECTIVITY_IN_VISUAL_AREA_MT)
Figs. 9 and 10 that the population of “length–width inseparable” (LWI) cells included a variety of spatial integration properties, including cells showing strong inhibition by large gratings (e.g., Figs. 9A and C, and 10A) as well as wide-field cells (Figs. 9E and 10B). For example, the responses of the unit illustrated in Fig. 10A were such that the preferred length decreased, whereas the magnitude of end inhibition increased, as a function of increased patch width. For wide-field cells (Fig. 10B), the response tended to saturate more rapidly as a function of length as wider gratings were used. Overall, 18 of the 54 LWI cells (33%) were end inhibited only, 10 cells (18.5%) showed only side inhibition, 10 cells were inhibited in both dimensions (18.5%), and 16 cells (30%) were wide-field. These proportions were not different from those found among other MT cells [i.e., those that were not classified as LWI; \( \chi^2(3) = 2.79, P > 0.05 \)]. Although the LWI subpopulation contained a variety of cell types, it was homogeneous in the fact that the axis of rotation of the model was always perpendicular to the axis of stimulus symmetry; for example, no cells were found for which the optimal length increased as wider gratings were used, or vice versa.

In an analysis that excluded wide-field cells, it was apparent that LWI cells tended to show a stronger degree of inhibition by large stimuli, compared to other MT cells. The ratio of the responses elicited by the largest patches (30° × 30°) and optimal gratings size was significantly lower (0.37) for LWI cells than for non-LWI cells (0.56; Wilcoxon rank-sum test, \( z = -2.30, P = 0.022 \)). However, the distributions of optimal
lengths and widths were not different (length: Wilcoxon rank-sum test, \(z = -1.13, P > 0.05\); width: \(z = 0.35, P > 0.05\)). It was also evident that, with the exception of the wide-field subpopulation, LWI cells rarely preferred symmetrical stimuli, with only 4/38 units (10.5%) having optimal stimulus aspect ratios between 2:1 and 1:2.

Laminar distribution

Table 1 shows the laminar distribution of cells with different types of size selectivity in MT. Given the shallow angle of our penetrations, cells in the infragranular layers (in particular, layer 6) were underrepresented in our sample. The data indicate that LWI cells were relatively rare in layer 4, in comparison to the remaining layers \(\chi^2(1) = 4.89, P = 0.027\), and that wide-field cells were significantly less common in the supragranular layers than in the granular or infragranular layers \(\chi^2(1) = 42.44, P < 0.001\). In particular, the proportion of wide-field cells was greater in layer 5 compared with the remainder of the population \(\chi^2(1) = 31.13, P < 0.001\). We are confident that these differences in size summation properties cannot be attributed simply to differences in the mean eccentricity (and thus overall size) of the receptive fields sampled (see Table 1).

There is additional evidence to suggest that other response properties change as a function of cortical depth (Fig. 11). Typically, cells showing the strongest inhibition, as revealed by the lowest ratios of the modeled responses to the largest \(30^\circ\) and optimally sized gratings, were located in the superficial layers (Fig. 11A). A significant difference was found in this respect [Kruskal–Wallis test: \(\chi^2(4) = 34.72, P < 0.001\], with post hoc analysis revealing that the median ratio was significantly larger for cells in layer 5 than for cells in layers 1–3 \((P < 0.05)\). Optimal stimulus lengths (Fig. 11B) and widths (Fig. 11C) also varied significantly as a function of cortical layer [length: Kruskal–Wallis test: \(\chi^2(4) = 20.61, P < 0.001\); width: \(\chi^2(4) = 18.85, P = 0.001\)]. Post hoc analyses indicated that cells in layer 5 layers preferred longer gratings compared with cells in all supragranular layers and wider gratings than cells in layer 3 \((P < 0.05)\).

**DISCUSSION**

We have described the spatial summation properties of neurons in MT using first-order, sine-wave gratings as stimuli. As well as having characterized basic properties such as end inhibition and side inhibition, we found a substantial proportion of neurons for which the selectivities to stimulus length...
and width were interdependent. In addition, differences in spatial selectivity were described between neurons in different layers. These results provide new evidence regarding the receptive fields of MT neurons, contributing to the idea that their spatial structure is relatively complex.

Modeling size summation

The basis of our 2D model came from the work of Perrone and Thiele (2001) on the interaction effects of spatial frequency and temporal frequency, where a linear version of the product of two Gaussians model was used. The transformation from the product of the two-Gaussians model to the product of the two log-Gaussians model was performed using the method of Nover et al. (2005). We recognize that, in theory, size summation curves should take the form of the integral of the receptive field model. Because the spatial structures of MT receptive fields were previously described by Gaussian (wide-field cells) or difference of Gaussians (center-surround, or local motion cells) functions, MT size summation curves have generally been fitted either by error (erf) or difference of error (DoE) functions, respectively (DeAngelis and Uka 2003). Following from this, 2D receptive fields can be modeled by the product of two such functions, depending on whether the cell is inhibited by larger than optimal stimuli and in which dimension(s). However, this is unwieldy because the model applying to different cells would need to be based on different combinations of functions. One advantage of the approach used in the present study is that it enabled us to fit the same seven-parameter model to all data, resulting in good descriptions of the essential elements being examined (i.e., the length and width of optimal stimuli, presence or absence of end and side inhibitions, and the extent to which integration over the length and width dimensions is independent). As is apparent from Fig. 4 the fitted surfaces represent these elements of the data well, even in cases when the \( r^2 \) values obtained were comparatively low. Finally, this model does not rely on the assumption that receptive fields necessarily follow either a Gaussian or a difference-of-Gaussians function, an assumption that can be questioned on the basis of the results of the present and previous studies (Born 2000; Raiguel et al. 1995). The resulting accurate descriptions of MT response properties are likely to provide a strong empirical basis for additional modeling of MT receptive field structure. Among its potential weaknesses, the present model does not allow for inhibition below the level of spontaneous activity. However, this is a very rare occurrence in MT, at least in conditions in which the neuronal responses are probed with stimuli moving in the optimal direction. Although in the present study complete inhibition was observed in four units, in no case was the activity recorded on the presentation of the largest gratings significantly below the level of spontaneous activity.

End inhibition and side inhibition properties

Given that end inhibition and side inhibition were neither correlated nor mutually exclusive, four groups of neurons could be identified in terms of spatial integration properties. It makes sense to propose that cells showing both end and side inhibition are useful for the coding of local motion contrasts, whereas cells inhibited in neither dimension are most adapted for wide-field coding (Born and Tootell 1992). As well as coding for local motion discontinuities, cells showing inhibition in only one dimension are well suited for extracting information concerning the 3D structure (Buracas and Albright 1996). Cells with asymmetric inhibitory or bilaterally symmetric surrounds were reported by previous studies of MT cells, which used moving random dot patterns as stimuli (Xiao et al. 1997). The reported combined prevalence of these types of cells is similar to the proportion of cells that showed inhibition along only one dimension (length or width) in the present sample.

Neurons showing exclusively end inhibition or side inhibition may be distinct in terms of function, if one considers what constitutes their optimal stimulus, and how their responses vary as a function of changes in stimulus size. Side inhibited–only cells can segregate long contours from backgrounds with greater spatiotemporal resolution than that of end inhibited–only cells (which cannot differentiate between stimuli of different characteristics on the axis parallel to the optimal direction of motion). On the other hand, under our experimental conditions, cells showing end inhibition but lacking side inhibition (nearly one third of our sample) responded maximally to gratings consisting of many short bars moving in file; thus they may appear at first to prefer multiple numbers of small objects moving in the same direction. However, these neurons could also be integrating motion signals from different elements of the same object on the basis of coherent motion, reflecting grouping according to the Gestalt principle of “common fate.” It was recently demonstrated that MT neurons are capable of coding for “camouflaged” objects defined solely by virtue of coherent motion of their textural elements (Dobiecki et al. 2006; see also Bourne et al. 2004). The diversity of response properties we observed in the current study is in agreement with the hypothesis that MT plays a variety of roles in the integration and segmentation of motion signals.

It could be argued that the choice of a rectangular aperture for the presentation of gratings in the present study introduced abrupt luminance differences at the edges of the stimuli, with the “width” edges (i.e., those parallel to the grating bars) appearing to flash at the stimulus temporal frequency. This flashing edge could constitute a powerful stimulus near the receptive field center, but would become progressively less important as the stimulus width increased. However, if this were the case, we would expect to see an unusually high number of cells preferring narrow gratings and, in particular, long and narrow gratings. In fact, the data illustrated in Figs. 5 and 6 indicate that a majority of MT neurons preferred wide gratings. Moreover, the optimal stimuli for most MT neurons were rarely narrower than the classical receptive fields at the corresponding eccentricities (Fig. 5). Finally, a two-way ANOVA indicated that the neurons we classified as having an antagonistic surround tended to prefer higher spatial frequencies in comparison to wide-field cells (mean optimal spatial frequency 0.23 vs. 0.15 cycles/deg; \( P = 0.015 \)), while revealing no significant main effect related to whether the neurons had LWI characteristics (\( P = 0.069 \)). Given that the “flashing” would be maximal at the lowest spatial frequencies, these results argue against the idea that this effect could have affected our conclusions, either by inflating our estimates of the proportion of cells with inhibitory surrounds or by generating false “LWI-like” properties.
Complex spatial integration properties

Our results show that the optimal grating stimuli for MT neurons are often asymmetrical, in contrast with the results of qualitative mapping of minimum response fields (Albright and Desimone 1987; Allman and Kaas 1971; Rosa and Elston 1998). Quantitative assessments of receptive field shapes using random dot patches presented in different locations revealed that a large number of macaque MT receptive fields are elongated (Raiguel et al. 1995), and that this elongation occurs mostly along the preferred axis of motion. Although this finding is compatible with our observation that MT cells tend to prefer gratings that are larger in terms of width, the average elongation determined with this method was only 30%, with very few cells exceeding 70%. This is considerably less than the 2:1 optimal stimulus aspect ratio displayed by many of the cells in our sample (Fig. 6). The optimal stimulus for a large proportion of MT cells may not match the shape of the classical receptive field, indicating that complex excitatory/inhibitory interactions are in operation.

The strongest evidence for complex spatial integration by MT receptive fields in the current study was found among the subpopulation of LWI cells that showed strong inhibition by wide-field stimuli. As well as having a strong response to stimuli of optimal dimensions, these cells were generally activated by moving stimuli occupying elongated sectors of the visual field, either parallel or perpendicular to the preferred axis of motion (Figs. 9, A and B, and 10A). In theory, a simple receptive field model that could account for these responses would be one based on the linear summation of inputs along a cross-shaped excitatory center and inhibitory “corners” (Fig. 12). In this case elongated stimuli centered on the receptive field, oriented either parallel or perpendicular to the axis of motion, would produce the maximal response, whereas a larger square patch would also stimulate the inhibitory subunits and thus produce a weaker response. Although our results alone cannot directly rule out this type of receptive field structure, previous studies in which the receptive fields of MT cells were probed with small moving patches in multiple locations did not report cross-shaped receptive field configurations (Perge et al. 2005; Raiguel et al. 1995; Xiao et al. 1997). Thus it is unlikely that our results reflect simple linear summation. It is also conceivable that a similar pattern of response could be generated by single inhibitory “corners” (for example, an asymmetrical receptive field with the inhibitory region offset from the preferred axis of motion by $45^\circ$), without necessarily implying symmetrical inhibitory zones around the receptive field. Yet, the inhibitory subregions within MT receptive fields do not appear to have sufficient spatial specificity to produce such responses. As demonstrated by Xiao et al. (1997; see their Fig. 12), the surround regions that yield the strongest inhibition are distributed at all possible locations around the cell’s axis of motion, rather than specifically at locations oblique to it, even when the subpopulation of cells with asymmetrical surrounds is considered separately. Thus receptive fields with the required configuration are unlikely to account for the relatively large proportion of LWI-inhibited cells that we observed. Moreover, asymmetrical surrounds in MT take the form of a relatively mild gradient that encompasses regions all around the excitatory receptive field, rather than being restricted to a specific position (e.g., Fig. 11 in Xiao et al. 1997). In summary, rather than the linear summation of excitatory and inhibitory inputs from retinotopic subregions of the receptive field, it is likely that the responses we observed among MT cells reflect the capacity of moving patterns in the same location to produce either excitation or inhibition, depending on context. Receptive fields of this type cannot be accurately modeled using a difference of two Gaussians model, which is spatially static. One possibility is that suppression arises from a peripheral mechanism with nonlinear spatial summation: although long narrow gratings and short wide gratings do not bring this mechanism to threshold, gratings that are larger in two dimensions do.

There are other lines of evidence indicating complex, nonlinear spatial summation by neurons in MT. For example, Raiguel et al. (1995) found that roughly 20% of macaque MT neurons have a classical receptive field that is $\geq 2.5$-fold larger than the optimal stimulus, and one would expect that LWI neurons to overlap with this group (see also Cheng et al. 1994). Although LWI cells were not reported by previous studies, this subpopulation may overlap to some extent with the complex motion contrast cells described by Born (2000) in the owl monkey. The mechanism suggested by Born (2000) for generating complex motion contrasts can also apply to this subpopulation, with gain normalization being generated by “OR”-gating inputs from many neurons with receptive fields showing antagonistic surrounds. This type of “OR”-gating may apply more generally, contributing to the shaping of the responses of MT cells in a variety of situations. For example, the response to two stimuli presented simultaneously tends to be consistently lower than the algebraic sum of the responses to stimuli presented individually (Britten and Heuer 1999). However, our data indicate that complex spatial summation is not an all-or-none phenomenon, restricted to a subpopulation of MT neurons. Rather, cells with complex motion contrasts may be part of a continuum (Fig. 8A), in which some receptive fields may be sufficiently described by linear summation, whereas others

![Diagram of receptive field configuration](image-url)
may be better accounted for by “OR”-gating gain normalization mechanisms (Raiguel et al. 1995).

Global motion processing in MT

We found a proportion of truly “global motion” or “wide-field” cells (22.3%) that corresponds to the lower end of the range reported by previous studies (10–60%; Allman et al. 1985; Born 2000; Bradley and Anderson 1998; Cheng et al. 1994; DeAngelis and Uka 2003; Raiguel et al. 1995; Tanaka et al. 1986). Conceptually, a major difference between our analysis and many of the previous studies lies in the criteria for classification of cells as local motion (inhibited) or wide field. For example, DeAngelis and Uka (2003), who reported the incidence of wide-field responses to be 42%, adopted a criterion whereby cells were classified as having an inhibitory surround only if their responses were better fitted with a DoE model than with an erf model. Here the onus was shifted, with cells being considered to be selective for local motion unless the 95% confidence interval estimates for preferred length and width overlapped with 30 × 30°, the largest stimulus we could generate. Our results are compatible with those of Raiguel et al. (1995; ≈10% wide field), who considered cells to be wide field if the response to the largest stimulus was ≥85% of the response to a smaller, near-optimal stimulus. It is possible that the lower proportion of wide-field cells in our sample is linked to laminar biases because most of the studied neurons were located in the supragranular layers. However, this is unlikely to be the only factor involved. Considering that previous studies used circularly symmetrical stimuli, it is questionable whether optimal stimulus conditions were achieved for a substantial proportion of the population. This could, in turn, have led to an underestimation of the incidence of inhibition in the population. For example, the neuronal responses illustrated in Fig. 4, C and D show clear evidence of inhibition along one dimension; however, if one considered only the five data points along the axis corresponding to stimulus symmetry, inhibition would not necessarily become obvious. Indeed, our results revealed that the proportion of cells classified as wide field by 2D analysis (22.3%) was substantially lower than that revealed by the 1D method using only the responses to five symmetrical stimuli (63%).

Visual processing in area MT

Although direction selectivity is a defining characteristic of MT neurons, quantitative characteristics of neuronal responses, such as the strength and bandwidth of directional tuning, appear to reflect those already represented in its inputs from V1 (Movshon and Newsome 1996). Some of the main perceptual functions attributed to MT are related to the further integration and segmentation of motion signals (Born and Bradley 2005; Britten 2003). Complex patterns of spatial selectivity, as demonstrated by the presence of LWI cells and conditional end and side inhibitions, are likely to represent emergent response properties of cells in MT. Our finding that LWI cells were less common in layer 4 also supports this idea.

The apparent dichotomy between wide-field and local motion cells could, in principle, be the result of computations occurring primarily within MT, with the first group apparently receiving excitatory connections from cells representing other regions of space within MT and the second group receiving strong inhibition. However, the present results point to a more complex situation. The conditional nature of end and side inhibitions we observed in many MT units, as well as of that of length and width preference, together suggest that the responses of neurons in this area cannot be conceptualized on the basis of simple center-surround organizations. Instead, neuronal responses in MT may be defined by a combination of factors involving feedforward, feedback, and intrinsic connections, with intrinsic mechanisms involving inhibition being dominant for local motion cells, and perhaps intercortical projections from other extrastriate dorsal stream areas having a greater influence on wide-field cells. Wide-field cells appear to be common in the medial superior temporal area (MST; Saito et al. 1986) and the dorsomedial area (DM; Lui et al. 2006), which form two of the most numerous contingents of cortical projections to MT (of “feedback” and “lateral” natures, respectively; Palmer and Rosa 2006). In contrast, the majority of MT cells show end inhibition, side inhibition, or a combination of these, depending on the exact stimulus configuration. Considering the spatial extent of the summation zones we observed, this inhibition could be conveyed by horizontal connections between cells within MT, through GABAergic interneurons (Dhar et al. 2001). Indeed, MT is characterized by a particularly rich network of parvalbumin-immunoreactive interneurons, a characteristic that indicates high levels of sustained inhibitory activity in comparison with that of most other visual areas (Bourne et al. 2007). In addition, intrinsic connections within MT link cells with similar orientation and direction preferences (Malach et al. 1997); thus increasing the dimensions of a grating is likely to engage such inputs to a greater extent. This could explain why the percentage of cells showing end inhibition tends to increase when MT cells are probed with very wide stimuli and why the proportion of side inhibition tends to increase when very long stimuli are used. Corticocortical projections originating from specific areas often form periodically arranged patches, perhaps introducing variations in the relative strength of the excitation and inhibition elicited by large moving patterns; this could explain the periodical organization of MT revealed in 2-deoxyglucose experiments (Born and Tootell 1992).

In summary, our observations support the concept that area MT plays a prominent role in the integration and segmentation of motion signals. Our results can be best explained by nonlinear patterns of spatial summation in MT, supporting the notion that MT receptive fields are considerably more complex than previously described.

ACKNOWLEDGMENTS

The authors thank R. Tweedale for correcting the grammar and style. Prof. Paul Martin for assistance with the Hartigan’s Dip statistic, and two anonymous referees for comments and suggestions, which substantially improved this manuscript.

GRANTS

This work was supported by research grants from the National Health and Medical Research Council and from the Australian Research Council. Items of equipment purchased with funds provided by the Clive and Vera Ramaciotti Foundation and by the Australia and New Zealand Banking Group Limited Charitable Trust were crucial for the completion of this study.
REFERENCES


