

# Diverse Suppressible Influences in Area MT and Selectivity to Complex Motion Features

Yuwei Cui,<sup>1</sup> Liu D. Liu,<sup>2</sup> Farhan A. Khawaja,<sup>2</sup> Christopher C. Pack,<sup>2</sup> and Daniel A. Butts<sup>1</sup>

<sup>1</sup>Department of Biology and Program in Neuroscience and Cognitive Science, University of Maryland, College Park, Maryland 20742, and <sup>2</sup>Montreal Neurological Institute, McGill University, Montreal, Quebec, Canada H3A 2B4

Neuronal selectivity results from both excitatory and suppressive inputs to a given neuron. Suppressible influences can often significantly modulate neuronal responses and impart novel selectivity in the context of behaviorally relevant stimuli. In this work, we use a naturalistic optic flow stimulus to explore the responses of neurons in the middle temporal area (MT) of the alert macaque monkey; these responses are interpreted using a hierarchical model that incorporates relevant nonlinear properties of upstream processing in the primary visual cortex (V1). In this stimulus context, MT neuron responses can be predicted from distinct excitatory and suppressible components. Excitation is spatially localized and matches the measured preferred direction of each neuron. Suppression is typically composed of two distinct components: (1) a directionally untuned component, which appears to play the role of surround suppression and normalization; and (2) a direction-selective component, with comparable tuning width as excitation and a distinct spatial footprint that is usually partially overlapping with excitation. The direction preference of this direction-tuned suppression varies widely across MT neurons: approximately one-third have overlapping suppression in the opposite direction as excitation, and many other neurons have suppression with similar direction preferences to excitation. There is also a population of MT neurons with orthogonally oriented suppression. We demonstrate that direction-selective suppression can impart selectivity of MT neurons to more complex velocity fields and that it can be used for improved estimation of the three-dimensional velocity of moving objects. Thus, considering MT neurons in a complex stimulus context reveals a diverse set of computations likely relevant for visual processing in natural visual contexts.

## Introduction

Our ability to perceive natural scenes relies on efficient extraction of higher-order structure, which permits the decomposition of the visual scene into objects and surfaces. The extrastriate cortex of primates is devoted to such higher-order processing of visual signals (Maunsell and Newsome, 1987; Orban, 2008). Consequently, characterizing the relationship between extrastriate cortical activity and complex visual scenes can reveal a great deal about the neuronal computations underlying sensory processing. However, such stimuli often have very high dimensionality and necessarily involve complicated spatiotemporal correlations, making it difficult to isolate those potentially complex aspects of the stimulus that are driving neuronal responses.

One of the most thoroughly studied regions of the extrastriate cortex is the middle temporal (MT) area. MT is somewhat unusual among extrastriate regions in that the vast majority of its

neurons are highly selective within a low-dimensional stimulus feature space. Specifically, MT neurons are selective for motion, which in natural vision occurs in the context of optic flow, comprising the spatiotemporal stimuli observed during translations and rotations of objects in the environment relative to the observer. However, most previous explanations of MT stimulus selectivity have focused on a subset of optic flow patterns that corresponds to the two-dimensional velocities of objects translating within a single depth plane (but see Lagae et al., 1994; Simoncelli and Heeger, 1998; Lisberger and Movshon, 1999; Perone and Thiele, 2001; Rust et al., 2006; Nishimoto and Gallant, 2011). Although this is a very important aspect of MT responses, such studies implicitly neglect the effect of different velocities at different positions within the visual field, which is an important component of optic flow.

There has been much evidence that MT is also selective for complex motion patterns, given the observation of the powerful suppressible surrounds (Allman et al., 1985; Born, 2000; Tsui and Pack, 2011). Previous studies have isolated various aspects of suppression, including its contrast sensitivity (Pack et al., 2005; Hunter and Born, 2011), direction tuning (Allman et al., 1985), and spatial structure (Xiao et al., 1995, 1997), but these experiments have focused on a subset of properties of suppression, often using highly tailored stimuli. Thus, it has been left unclear how the multiple forms of suppression combine with excitation in more natural contexts—in which stimuli driving each element are related because of the statistics of optic flow—to potentially result in higher-order selectivity.

Received Jan. 15, 2013; revised Aug. 17, 2013; accepted Sept. 9, 2013.

Author contributions: C.C.P. and D.A.B. designed research; Y.C., L.D.L., and F.A.K. performed research; Y.C. analyzed data; Y.C. and D.A.B. wrote the paper.

This work is supported by National Science Foundation Collaborative Research in Computational Neuroscience Grant IIS-0904430/CNS-103331 (Y.C., D.A.B., C.C.P.), Canadian Institutes of Health Research Grant MOP-115178 (C.C.P.), Canadian Graduate Scholarship-Doctoral Grant CGSD-121719 (L.D.L.), Natural Sciences and Engineering Research Council of Canada Grant CGS M-408975-2011 (L.D.L.), and Le Fonds de la Recherche en Santé du Québec Dossier 13159 (F.A.K.).

Correspondence should be addressed to Yuwei Cui, Department of Biology, 1210 Biology-Psychology Building 144, University of Maryland, College Park, MD 20742. E-mail: ywui@umd.edu.

DOI:10.1523/JNEUROSCI.0203-13.2013

Copyright © 2013 the authors 0270-6474/13/3316715-14\$15.00/0

Here, we use a continuously varying optic flow stimulus to measure the combination of excitation and multiple forms of suppressive tuning, using a nonlinear modeling approach that can be fit to the recorded neuronal spike trains. Our analysis reveals the full spatial and temporal structure of excitatory and suppressive influences for MT neurons and demonstrates the diversity of computation in area MT from this perspective. Using simulations, we show that these nonlinear properties of MT receptive fields are functionally useful for extracting information about the three-dimensional velocities of moving objects and hence directly facilitate additional motion processing in higher cortical areas (Zemel and Sejnowski, 1998; Mineault et al., 2012).

## Materials and Methods

**Electrophysiology recordings and behavioral task.** Data were recorded from two adult rhesus macaque monkey (one female and one male, referred to as M1 and M2 hereafter), prepared using standard surgical techniques that have been described previously (Mineault et al., 2012). Animals were trained to fixate within 2° of a small fixation point on a computer monitor in return for a liquid reward. Eye movements were monitored at 500 Hz by an infrared eye tracker (EyeLink II; SR Research). Extracellular recordings were performed on 102 well-isolated single units in area MT, which were located using exterior cranial landmarks, anatomical images from magnetic resonance imaging, and/or physiological properties. Data were recorded using either single electrodes ( $n = 63$ , M1;  $n = 18$ , M2) or a multisite linear electrode array ( $n = 21$ , M1). Signals were amplified, bandpass filtered, sorted online, and resorted offline, using spike-sorting software (Plexon) to identify single units. All aspects of the experiments were approved by the Animal Care Committee of the Montreal Neurological Institute and were conducted in compliance with regulations established by the Canadian Council on Animal Care.

**Visual stimuli.** During isolation of a single MT unit, we measured the direction tuning, speed tuning, and size tuning of the neuron using random-dot motion stimuli. We then presented a continuously varying optic flow stimulus composed of moving dots whose velocity varied over space as well as time (Mineault et al., 2012). The velocity field was generated as a random combination of six optic flow components: horizontal/vertical translation, expansion, rotation, and horizontal/vertical shears. The magnitude of each optic flow components varied independently based on low-pass filtered Gaussian noise with a cutoff of 2 or 5 Hz. For the majority of cells ( $n = 84$ ), the stimulus was displayed in a slowly moving aperture with a diameter ranging from 8° to 20° (depending on size of the receptive field), the position of which was determined by another pair of low-pass filtered Gaussian noise with a cutoff of 0.05–0.10 Hz, with its mean at the center of the receptive field, and its SD ranging from 3 to 10°, depending on the size of the receptive field. The other neurons ( $n = 18$ , recorded from monkey M2) used the same optic flow stimulus, although it remained centered on the receptive field of the neurons (location estimated from hand-mapping), and there was no moving aperture; instead, stimuli were displayed either on the full screen or in a very large static aperture with 30° diameter. The stimuli were presented on a LCD monitor (Dell 2707WFP) with a display resolution of 1600 × 1000 pixels (49° × 36° of visual field at a distance of 50 cm) and a refresh frame rate of 60 Hz ( $n = 73$ , M1) or 75 Hz ( $n = 11$ , M1;  $n = 18$ , M2). There were no notable differences for cells recorded with different refresh rate. In all cases, the dots were 0.1° in diameter against a black background. The luminance of white dots was 194 cd/m<sup>2</sup>, and the luminance of black background was 0.2 cd/m<sup>2</sup>.

During the experiment, the stimulus was displayed in 6 or 8 min blocks until the animal stopped behaving or the unit was lost. The stimulus presented in each block was different, and the data were thus combined to form a longer continuous stimulus with a median length of ~18 min/unit. The stimuli would revisit a given spatial position an average of 36 times. For a subset of recordings ( $n = 20$ ), we showed a repeated short segment of the stimulus (5 s) that had its aperture centered on the receptive field of the cell to measure the response reliability and to calculate the explained variance ( $R^2$ ) of the model.

**Data analyses.** The measured spike trains were binned at 25 ms temporal resolution to obtain the observed response rate  $r_{\text{obs}}(t)$ . We excluded data from 100 ms before fixation breaks (when the animal's gaze location deviated by >1.5° from the fixation point) to 500 ms after the recovery of fixation. To avoid any saccade-related effects or transients, only periods with fixations that were longer than 1 s were used. The total recording of each neuron was broken into 10 s segments, which were randomly divided into two groups: (1) 80% of these segments were used to estimate model parameters ("training set"); and (2) model performance was evaluated on the remaining 20% of the data ("cross-validation set"). The use of cross-validation ensures that the improved model performance is truly attributable to an increased ability to capture the relationship between the stimulus and response rather than a tendency to characterize random fluctuations (noise).

Of the 102 units recorded, we excluded eight units for the following reasons. From monkey M1, two units were excluded because they did not have a consistent stimulus-dependent response, which meant that a stimulus-independent "null model" (that only predicted the average firing rate) outperformed all stimulus-dependent models tested. We also excluded six units recorded from monkey M2, in which the measured receptive field did not align with the hand-mapped receptive field center (<30% of the excitatory weights were within 7.5° of the receptive field center). Although we did obtain significant model fits for these neurons, the spatial elements of the model were hard to interpret because the stimuli were not centered on the receptive field. The remaining 94 units were included for this study.

Finally, because units recorded from the same electrode often had very similar relationships between excitation and suppression, we only included a single neuron from each multisite electrode experiment ( $n = 5$  of 21) in the figures relating to the overall distribution of MT neuron properties (see Fig. 5 F, G) to avoid sampling bias.

We also carefully controlled for the effects of fixational eye movements on the 53 recordings for which there were eye signals of sufficient quality, including both fixational drift and microsaccades (Martinez-Conde et al., 2013). The speed of fixational drift was small during fixation (median, 1.515 ± 0.090°/s) compared with the typical speed of the stimulus (~20°/s), which had an undetectably small effect on the neuronal response (data not shown). Microsaccades were detected using the algorithm proposed previously (Engbert and Mergenthaler, 2006) and were observed to occur with a frequency of 1.77 ± 0.10 Hz during the experiment. The impact of microsaccades on the neural response was direction dependent as reported previously (Bair and O'Keefe, 1998). However, we found that this effect was primarily unrelated to the stimulus-dependent terms of the model and had no impact on the results presented. As a result, we did not include these additional analyses here.

**Modeling of MT neurons.** To understand how MT units respond to the complex motion stimuli, we developed a hierarchical modeling framework. We assumed that MT neuron responses are generated by an inhomogeneous Poisson process with an instantaneous rate  $r(t)$ . The log-likelihood of the model is then given by (up to an additive constant):

$$LL [r_{\text{obs}}(t), r(t)] = \sum_t [r_{\text{obs}}(t) \log r(t) - r(t)], \quad (1)$$

where  $r_{\text{obs}}(t)$  is the measured neuronal response, and  $r(t)$  is the model predicted firing rate (Paninski, 2004).

All models we consider use a fixed spiking nonlinearity  $F[\cdot]$  that acts on the stimulus-dependent terms of the model, which we refer to as the generating signal  $g(t)$ ,

$$r(t) = F[g(t) - b], \quad (2)$$

where  $b$  is the spiking threshold, and we choose the spiking nonlinearity function  $F[\cdot]$  to be of the form  $\log[1 + \exp(\cdot)]$ . This functional form resembles a familiar rectified-linear function and additionally facilitates well behaved model optimization (Paninski, 2004; McFarland et al., 2013). To validate the use of this parametric form of  $F[\cdot]$ , we also measure the spiking nonlinearity using nonparametric histogram method (Chichilnisky, 2001; Paninski, 2004). In all cases, we find that the chosen form of nonlinear function gives a good description of the measured spiking nonlinearity.

Other nonlinear functions, such as the power law transformation (Nykamp and Ringach, 2002; Ghose and Baur, 2010) can fit the measured spiking nonlinearity equally well (data not shown) but have additional parameters and are not as well behaved for parameter optimization (Paninski, 2004).

The model acts on the continuously varying optic flow stimulus, which is described by a local motion speed  $\rho(t, x, y)$  and direction  $\theta(t, x, y)$ , sampled at a spatial resolution of  $2^\circ$ . This local motion signal is first processed by a set of subunits that are described by a speed-tuning function  $f_v[\cdot]$  and a direction-tuning function  $f_\theta[\cdot]$ . The subunit output is thus given by the following:

$$\text{subunit}(t, x, y) = f_v[\rho(t, x, y)] f_\theta[\theta(t, x, y)]. \quad (3)$$

To test the idea that primary visual cortex (V1) neurons are detectors of one-dimensional velocity, we also implemented an alternative formulation for the subunit, in which velocity was first projected onto the preferred direction of the subunit before being processed by the subunit nonlinearity:

$$\text{subunit}(t, x, y) = f[\rho(t, x, y) \cos(\theta(t, x, y) - \theta^{\text{pref}})]. \quad (4)$$

This formulation is consistent with the assumptions of various models of MT (Simoncelli and Heeger, 1998). However, because direction tuning and speed tuning are entangled in this formulation, the resulting models were more difficult to fit and interpret. Thus, because we did not find any significant difference in model prediction for these two formulations, we used the first subunit model (Eq. 3) for the majority of this work.

The generating signal of the model is computed by integrating the subunit outputs over space and time. In the motion–opponency model (MO model), this is given as follows:

$$g^{\text{MO}}(t) = \sum_{\tau} k_T(\tau) \left[ \sum_{x, y} w(x, y) \text{subunit}(t, x, y) \right], \quad (5)$$

where  $k_T$  is the temporal kernel, and  $w$  is the spatial weighting function. With a fixed speed-tuning function, the temporal kernel and the spatial kernel can be efficiently estimated with the GLM framework (Paninski, 2004). The nonlinear speed-tuning function can also be efficiently optimized by expressing it as a linear combination of basis functions  $f_v(x) = \sum_n \alpha_n \xi_n(x)$ , which were chosen to be overlapping tent-basis functions (Ahrens et al., 2008b; Butts et al., 2011; McFarland et al., 2013). We chose the center of the tent basis to be equally spaced on a logarithmic scale.

The nonlinear direction-tuning functions can be similarly optimized by expressing them using the tent-basis functions. However, in practice, we found that it was more reliable to assume a parametric form for the direction-tuning functions. To mimic the local motion–opponency mechanism (Qian and Andersen, 1994), we used the von Mises function as the direction-tuning function:

$$f_{\theta(\phi)} = \exp[b \cos(\phi - \phi_p)], \quad (6)$$

where  $\phi_p$  denotes the preferred direction, and  $b$  controls the direction-tuning width. The cosine function implements local opponency, because a nonpreferred stimuli will lead to suppressed output. These functions are always rectified (positive) because of the use of the exponential function. Speed tuning, direction tuning, spatial weights, and temporal kernels can be alternatively optimized with an appropriate choice of initial guesses.

For the excitation–suppression model (ES model), the stimulus processing is performed separately by excitatory and suppressive components, each using their own direction-tuning functions and rectified spatial weighting functions. Here, the superscripts  $E$  and  $S$  denote the excitatory and direction-selective-suppressive (DS-Sup) components, respectively. The generating signal is then given as follows:

$$\begin{aligned} g^{\text{ES}}(t) &= g^E(t) + g^S(t) \\ &= \sum_{\tau} k_T^E(\tau) \sum_{x, y} w^E(x, y) f_v^E[\rho(t - \tau, x, y)] f_\theta^E[\theta(t - \tau, x, y)] \\ &\quad + \sum_{\tau} k_T^S(\tau) \sum_{x, y} w^S(x, y) f_v^S[\rho(t - \tau, x, y)] f_\theta^S[\theta(t - \tau, x, y)], \end{aligned} \quad (7)$$

where  $k_T^E$  and  $k_T^S$  are the temporal kernels,  $w^E$  and  $w^S$  the spatial weighting functions,  $f_v^E[\cdot]$  and  $f_v^S[\cdot]$  the speed-tuning functions, and  $f_\theta^E[\cdot]$  and  $f_\theta^S[\cdot]$  the direction-tuning functions. For some models, an additional non-direction-selective suppression (NS-Sup) is included, which takes the following form:

$$g^{\text{NS}}(t) = \sum_{\tau} k_T^{\text{NS}}(\tau) \sum_{x, y} w^{\text{NS}}(x, y) f_v^{\text{NS}}[\rho(t - \tau, x, y)]. \quad (8)$$

Because this component lacks a direction-tuning function, its output is independent of stimulus direction.

This model is thus structured as a multilinear model of functions of the stimulus (Ahrens et al., 2008a), and each set of parameters can be efficiently optimized while holding the others constant. Because the overall multilinear optimization procedure can get stuck in local optima, it was important to choose an appropriate initialization of these parameters for the optimization procedure. The excitatory component was first optimized as the only model component, and then different types of suppressive components were added to the model to improve its performance. For the excitatory component, the temporal kernel was initialized to unity over the range of 50–150 ms and was zero for all other time lags based on the typical response latency of MT cells. The direction-tuning parameter  $b$  (Eq. 6) was initialized to be 1, corresponding to an SD of  $42.6^\circ$  for the direction-tuning function (similar to measured tuning width using random dot patterns; Snowden et al., 1992:  $46.5^\circ$ ), and the speed-tuning function was initialized to be linear over a logarithmic scale. The direction preference was initialized to be one of eight possible directions, equally spaced from  $0^\circ$  to  $360^\circ$ . Spatial weighting functions were optimized for each direction, and the best one was selected for additional refinement. We then alternatively optimized the direction-tuning parameters, the temporal kernel, the speed-tuning functions, and the spatial weights.

Similar procedures were performed when a DS-Sup component was added to the model: the suppressive direction that could most improve the model was selected, and then other model components were refined alternatively. For models with NS-Sup components, NS-Sup was simultaneously optimized with the other model components. To determine which components to include in the model of a given neuron, four separate models were fit and compared for each cell: models with (1) only excitation; (2) excitation and DS-Sup; (3) excitation and NS-Sup; and (4) excitation and both types of suppression. We selected the model with the best cross-validated performance.

Each component of the model has 225 parameters ( $15 \times 15$ ) for the spatial weighting function, 40 parameters for the temporal kernel, 10 parameters for the speed-tuning function, and two parameters for the direction-tuning function (preferred direction and tuning width). There is an additional parameter for the spiking nonlinearity for each model. The MO model thus contains 307 parameters, an NS-Sup component adds an additional 306 parameters, and an NS-Sup component adds an additional 304 parameters. A model with both DS-Sup and NS-Sup thus has 917 parameters. Note that such models are well constrained by the median stimulus duration of 18 min, which corresponds to 43,200 different stimuli at a sampling rate of 25 ms. Note that the evaluation of model performance using a cross-validation dataset avoids bias toward models with more parameters.

Because of the complexity of the model, regularization techniques were used to prevent overfitting. A penalty term was added to the log-likelihood proportional to the second derivative or the Laplacian of the temporal kernel and the speed-tuning function, e.g.,  $-\lambda \sum_i \partial^2 k_T / \partial \tau^2$ . For the two-dimensional spatial weighting function, we used both “smoothness” and “sparseness” penalties. The smoothness penalty was proportional to the sum of the squared slope relative to four nearest neighbors (up, down, left, and right), and the sparseness penalty was proportional to the sum of absolute value of the weighting function. The use of regularization techniques imposes certain prior distributions on model parameters and reduces the “effective” number of free parameters. For example, although we used 225 parameters to describe each spatial weighting function to make it flexible enough, the usage of smoothness and sparseness regularization ensure that only a small fraction of the

parameters will be nonzero, and these nonzero weights usually vary smoothly across space.

Because the temporal kernels and the speed-tuning functions usually had a stereotyped shape, we used the same regularization parameters for these functions across all cells. In contrast, the smoothness regularization parameters were adjusted for each cell and each model component individually using a nested cross-validation scheme. Specifically, 20% of the fitting data were randomly selected and reserved during the optimization of the spatial weights (note that this was different from the cross-validation data, which was never used during model estimation). The regularization parameters that gave best performance on the reserved data were used for the final fits to the data.

**Measurement of the properties of suppressive components.** Properties of the DS-Sup components were analyzed when detectable ( $n = 62$  of 94). The spatial profiles, direction-tuning strengths, and temporal dynamics of DS-Sup and NS-Sup components were also compared when both were detectable ( $n = 33$  of 94).

Following the fitting procedure described above, we investigated our assumptions about the direction tuning of each component by refitting the direction-tuning curves using tent-basis functions (Ahrens et al., 2008b; Butts et al., 2011), which yields a nonparametric estimate of each direction-tuning function. The direction-tuning strength was then measured using circular variance (Ringach et al., 2002), defined as  $\text{Var} = 1 - |R|$ , where  $R$  is given by the following:

$$R = \frac{\sum_k f_{\theta}(\rho_k) e^{i\rho_k}}{\sum_k f_{\theta}(\rho_k)}, \quad (9)$$

where  $\rho_k$  are the centers of the tent basis, and angles are expressed in radians. The value of circular variances ranges from 0 to 1, with lower values indicating tighter clustering around a single mean. A circular variance close to 1 indicates no direction tuning.

Three indexes were calculated to describe spatial profiles of suppression. First, we calculated the center of mass of the excitatory weights and the suppressive weights. The weighted distance between suppression and excitation was then calculated as follows:

$$D_{ES} = - \sum_{x,y} w_s(x,y) \sqrt{(x - x_c^E)^2 + (y - y_c^E)^2}, \quad (10)$$

where  $(x_c^E, y_c^E)$  is the center of excitation. This distance reflects how far suppression is from the receptive field center. A second index was calculated to reflect the dispersion of suppression:

$$\text{Dispersion} = - \sum_{x,y} w_s(x,y) \sqrt{(x - x_c^S)^2 + (y - y_c^S)^2}, \quad (11)$$

where  $(x_c^S, y_c^S)$  is the center of suppression. Both indexes were separately calculated for DS-Sup and NS-Sup components. Finally, an overlap index was calculated as follows:

$$\text{Overlap} = - \frac{\sum_{xy} w_E(x,y) w_S(x,y)}{\|w_E\| \|w_S\|}. \quad (12)$$

Note that a minus sign is introduced in Equations 10–12 because the suppressive weights  $w_s(x,y)$  are negative.

The temporal kernel was specified at 25 ms resolution. To more accurately measure the latency of a given model component, we used cubic spline interpolation with five points around the peak of the temporal kernel and used the peak of the cubic spline function as the latency of the component.

**Measurement of the selectivity of MT neurons for nontranslational optic flow.** To gauge the selectivity of MT neurons to nontranslational optic flow, we calculated the correlation coefficient  $\rho$  between each optic flow component and the neural response. For a given optic flow component, we computed the difference between the measured correlation coefficient and that predicted by the model, which reveals how well the models predicted the actual optic flow selectivity of the neuron.

We also compared the responses of the MO and ES models directly to optic flow stimuli, in which each model was presented with a randomly

selected combination of optic flow components centered on its receptive field, and the fraction of nontranslational stimuli was systematically varied from 0 to 1, with 200 random stimulus samples drawn at each level. For each simulation, we calculated the correlation coefficient between model outputs and also calculated the correlation coefficient between model response and the translational optic flow component.

**Population decoding simulations.** To simulate an MT population response to different stimuli, we “cloned” each model fit by replicating it across space (translating its position) and direction (rotating its spatial features and direction selectivity). For each of the 33 cells for which both types of suppression were detected, the MO model and the ES model fits were shifted on a  $5 \times 5$  grid with spacing of  $5^\circ$  and rotated to include the antipreferred direction and both perpendicular directions. Therefore, we generated  $25 \times 4 = 100$  virtual models for each cell. We pooled all virtual models together to create an entire virtual population of 3300 cells.

We simulated the response of this population to velocity field induced by 3D motion in different directions. For these simulations, an object covering the central  $20^\circ$  of the visual field was simulated undergoing 3D motion in 200 directions sampled from a spherical distribution randomly. For the purposes of the simulation, we assumed that the distance from the observer to the object was 5 m, and the speed of the object was uniformly distributed in the range of 0.87 to 2.62 m/s, matching the speed range we explored with the optic flow stimuli ( $10\text{--}30^\circ/\text{s}$ ). The spiking threshold of different model types was readjusted to give the same average firing rate (20 Hz) across all stimulus patterns.

We compared the capacity of an optimal linear decoder to extract information relevant to behavior, which were the physical parameters of the stimulus [3D velocity  $(v_x, v_y, v_z)$ ], given the output of a population of model MT cells. Specifically, we computed a weight vector  $\mathbf{w}$  to minimize the mean squared error between the parameter to decode (e.g.,  $v_x$ ) and the estimated parameter of the decoder  $\mathbf{X}\mathbf{w}$ , where  $\mathbf{X}$  is the matrix with one row for each stimulus and one column for each simulated MT response. At each repeat, we randomly selected 200 virtual models from the entire population or from a subset of the population and evaluated the reconstruction error as the ratio between the root mean squared error and the range of the physical parameter. The procedure was repeated 20 times to estimate the error of our evaluation.

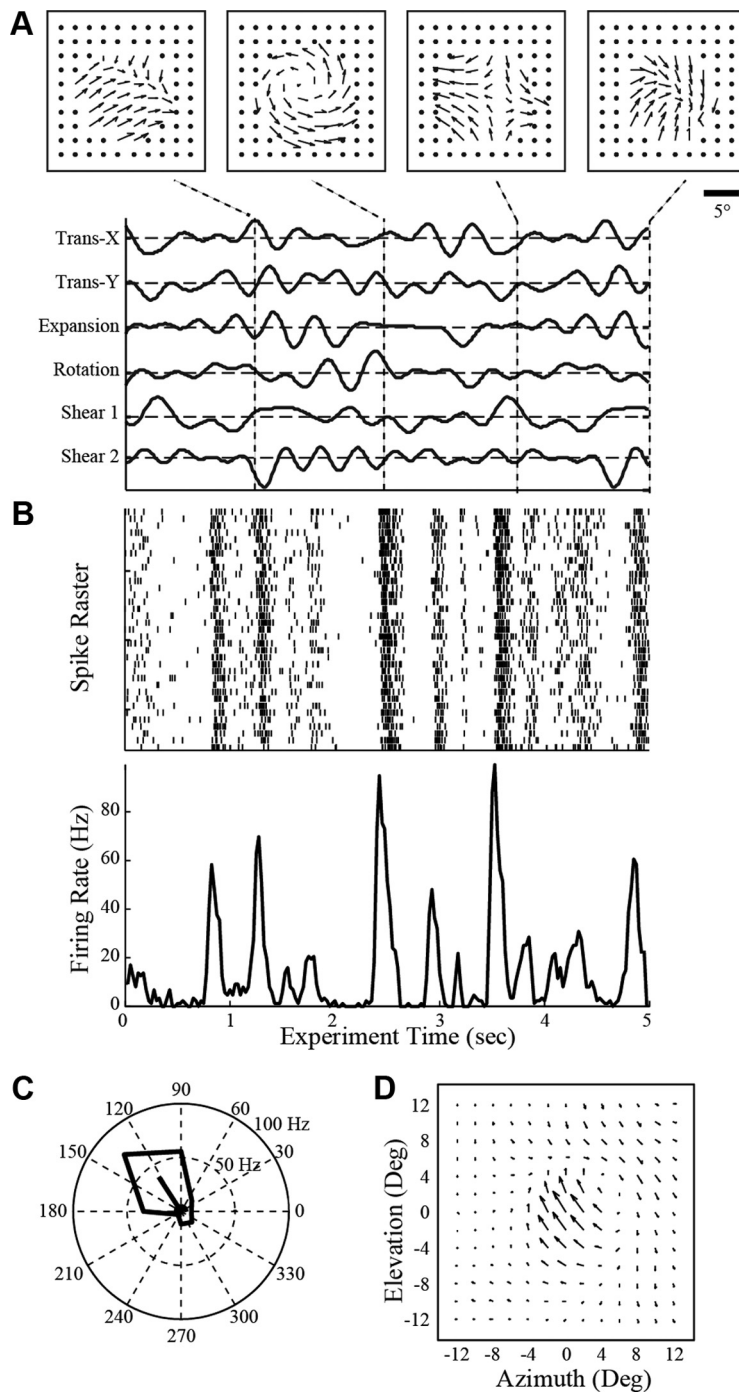
## Results

To explore the selectivity of MT neurons to complex motion stimuli, we recorded from single units in area MT during the presentation of a continuous optic flow stimulus composed of a random-dot field with the velocity field specified by a random combination of six optic flow dimensions (Fig. 1A; Mineault et al., 2012). MT neurons generally responded very reliably to this stimulus, as demonstrated by the reproducible patterns of spikes in response to multiple repeats of the same stimulus sequence (Fig. 1B).

MT neurons are thought to primarily be selective for the direction of motion stimuli in their receptive field (Maunsell and Van Essen, 1983; Albright, 1984; Mikami et al., 1986), as characterized by their average firing rate as a function of motion direction (Fig. 1C). Such “first-order” tuning is also reflected in more complex stimulus contexts, such as during the continuous optic flow stimulus, which can be demonstrated by fitting a linear model to explain responses in this context (Fig. 1D; Weber et al., 2010; Richert et al., 2013). Measurements of tuning in this stimulus context have the added advantage that the optic flow stimulus can reveal more complex aspects of MT tuning, because the stimulus incorporates different combinations of velocities across space and time. Indeed, as we show, more accurate models of stimulus processing are necessary to capture these details.

### Hierarchical modeling framework for MT neurons

To interpret the neuronal response in this complex motion context, we adopt a hierarchical modeling framework (Fig. 2A). Re-



**Figure 1.** Response of MT neurons to continuous optic flow stimuli. **A**, The naturalistic optic flow stimuli used in this study is composed of six independently varying optic flow components: horizontal and vertical translation (Trans-X and Trans-Y), expansion rotation, and shear along both axes. Each optic flow component is independently specified by low-passed Gaussian noise (bottom) and is displayed over a circular aperture that is moving around slowly to explore the spatial profile of the MT receptive field. The resulting velocity fields explore different types of flow, as shown by four example velocity fields that occur over the 5 s period shown. While stimulus is displayed at high spatial resolution, the models use the velocity field sampled over a uniform grid at a resolution of 2°. **B**, Response of an example MT neuron to the optic flow stimulus shown. The repeated spike responses are represented in a raster plot (top), with each vertical bar indicating a spike, and the peristimulus time histogram is shown at the bottom. **C**, The direction-tuning curve of this example MT neuron, measured as the average firing rate in response to random-dot motion in the receptive field. **D**, Spatial kernel of the linear model fit to the continuous optic flow stimuli for the same MT neuron.

ceptive fields of MT cells are much larger than those at earlier stages of the visual hierarchy and thus likely reflect aggregated responses of a large number of V1 neurons. Therefore, we choose an analysis resolution that gives fine detail for MT spatial recep-

tive fields by dividing them into smaller subunits, each presumably representing pooled responses of V1 neurons from that location. We assume that speed and direction are separately processed within each subunit (Hammond and Reck, 1980; Movshon et al., 1985; Rodman and Albright, 1987), such that subunit output is given by its direction tuning acting on the velocity (magnitude and direction) at each spatial location (Fig. 2B), multiplied by its speed-tuning function (Fig. 2C).

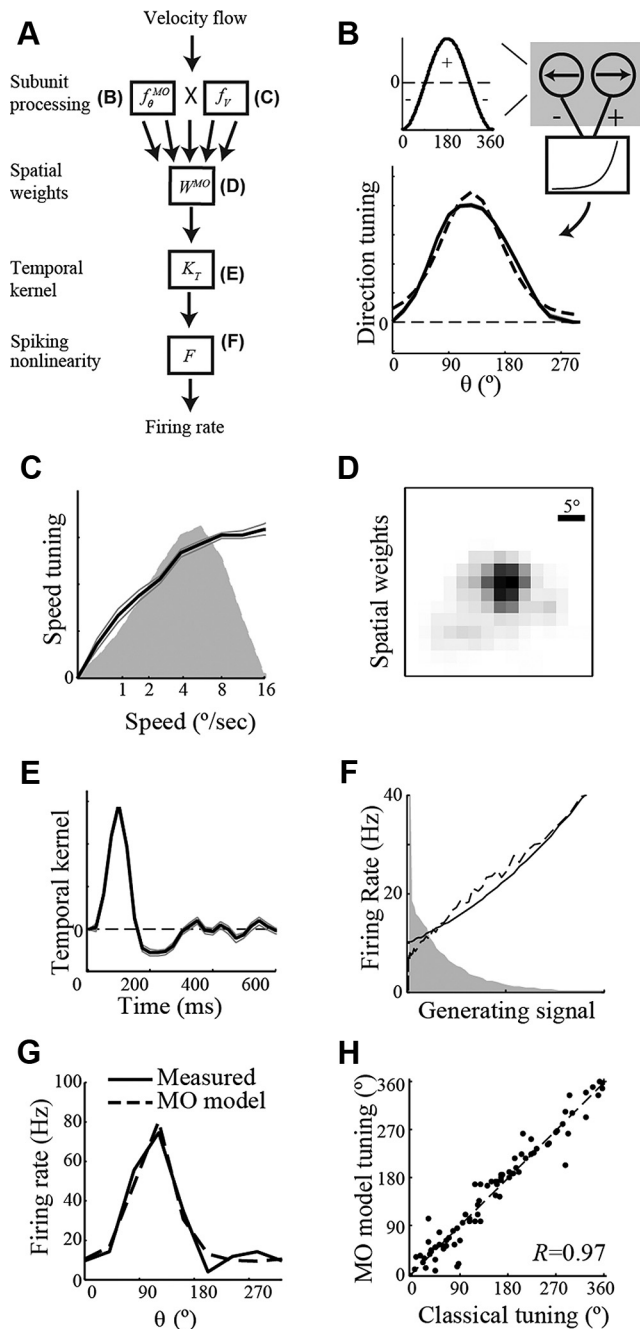
The output of each subunit is then integrated across space with a spatial weighting function (Fig. 2D) and integrated over time with a temporal kernel (Fig. 2E). Our model assumes the same direction and speed processing at all relevant positions across space, in part because we found that relaxing these constraints does not improve model predictions (data not shown, but see Discussion). The final signal after integrating over space is converted into a spike rate through a spiking nonlinearity (Fig. 2F).

Note that the structure of this model follows earlier models of MT responses (Qian et al., 1994), which assumes that the response of each subunit is enhanced by the preferred stimulus and suppressed by the nonpreferred stimulus, followed by a rectified nonlinearity (Fig. 2B, top). This is often called “motion opponency,” because the direction preference of suppression is always opposite from that of excitation (Adelson and Bergen, 1985; Simoncelli and Heeger, 1998); previous studies showed that this opponency-direction suppression is local (Qian and Andersen, 1994; Pack et al., 2006). We thus refer to this model as the MO model.

The parameters of this model structure can be tractably optimized using maximum-likelihood estimation methods, in which parameters representing the stimulus selectivity of individual subunits, the spatial weights, and the temporal kernels can be fit to the observed neuronal response (see Materials and Methods). Alternative formulations of MT processing, such as ones in which V1 cells only respond to the velocity component projected onto their preferred velocities (Movshon et al., 1985), do not substantially change the model performance (Wilcoxon’s rank-sum test on the cross-validated likelihood across the population between models of each type,  $p > 0.1$ ), but the resulting model is more difficult to fit

and interpret because speed tuning and direction tuning become entangled.

In general, the model fitted from data recorded in the context of complex optic flow stimuli has direction preferences that cor-



**Figure 2.** Analysis of MT neurons with the MO model. **A**, MO model schematic: motion is first processed by local subunits with direction and speed-tuning functions (top row) and then is pooled across space (second row), with the resulting output integrated over time (third row). Finally, a spiking nonlinearity is applied on this signal to transform it into a firing rate prediction (bottom row). **B–F**, The model components for the example MT neuron with SD of the fits (indicated as gray shaded area) calculated using bootstrapping techniques (100 repeated resampling with replacement). **B**, The direction-tuning function for the subunits. The model is fit by assuming a von Mises function (dashed), which contains a motion-opponency stage, followed by a rectified nonlinearity (top). This can be validated by nonparametrically fitting the direction tuning directly (solid), which is in close agreement. **C**, The speed-tuning function for the subunits (solid) compared with the distribution of stimulus speed (shaded gray). **D**, Spatial weighting function. **E**, Temporal kernel. **F**, The measured spiking nonlinearity (dashed) is fit parametrically (solid), with the distribution of the generating signal indicated in shaded gray. **G**, The model has a nearly identical direction-tuning curve (generated through simulation; dashed) as that of the neuron (solid). **H**, The model predicted direction preferences are highly correlated with the measured ones across the population of neurons in the study ( $r = 0.97$ ,  $p < 10^{-49}$ ,  $n = 84$ ).

relate with the preferred direction of the neuron measured using the standard approaches (see Materials and Methods; Fig. 2H,  $r = 0.97$ ).

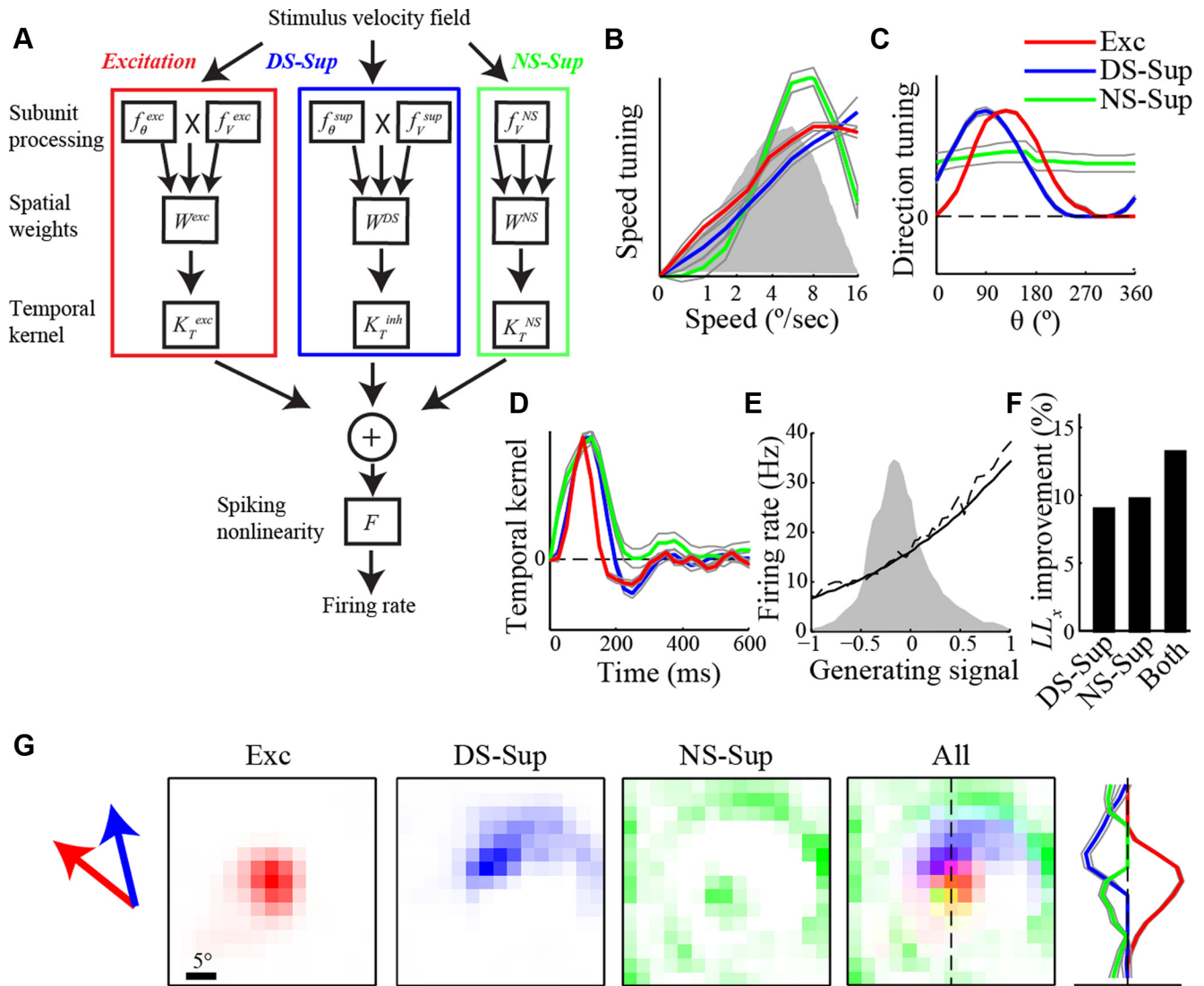
### Extension of the hierarchical modeling framework with suppression

Although the MO model provides a good description of the direction tuning of the MT neurons (Fig. 2G), it only incorporates a single form of suppression, which is always spatially localized with excitation and with opposite direction tuning. In contrast, numerous studies that probed stimulus selectivity with more than one spatially localized motion component have observed suppression outside the classical receptive field—in the “surround”—in a majority of MT neurons (Allman et al., 1985; Xiao et al., 1997; Born, 2000; Tsui and Pack, 2011). This suggests an extension of the MO model to include suppression that is not simply colocalized with excitation.

We thus extend the hierarchical framework to model a spatially distinct suppressive influence (Fig. 3A) in addition to the excitation described above, comprising what we call the ES model. Motivated by detailed studies of suppression in V1, in which suppression can have both an orientation-selective component (Ringach et al., 2002) and a non-selective component (Sengpiel et al., 1997), we incorporate both forms of suppression into the ES model: (1) a DS-Sup component (blue); and (2) an NS-Sup component (green). The DS-Sup component has a direction-tuning function, which is potentially distinct from the excitatory direction tuning, whereas the NS-Sup responds equally to all directions. The excitatory, DS-Sup, and NS-Sup components are pooled across space and integrated over time with their own spatial weighting functions and temporal kernels. The excitatory weights are constrained to be positive and the suppressive weights (both DS and NS) to be negative. As a result, the contribution of each is distinct, and simultaneous optimization of these components is tractable, resulting in a single optimal description of the excitatory and suppressive substructure of the MT receptive field of this form.

Across the population of MT neurons, we find that excitation, DS-Sup, and NS-Sup typically have different spatial profiles and direction preferences and, in some cases, different temporal dynamics. For example, the previously considered neuron (Fig. 2) has a DS-Sup component with similar direction preference as excitation (Fig. 3C) but forms an asymmetric surround structure that is essentially non-overlapping with the excitatory region (Fig. 3F). Furthermore, there is also a distinct NS-Sup component that has spatial weights in two areas: (1) a central area that is essentially overlapping with the excitatory component; and (2) a surround component that is farther away from the receptive field center than the DS-Sup. The temporal kernels of both suppressive components are slightly delayed relative to excitation (Fig. 3D). Although the ES model does not predict a different excitatory tuning of the neuron, it has a much better cross-validated performance than the MO model (Fig. 3E).

To gain additional insight into the degree of model performance improvement, for a subset of the recorded neurons, we also presented repeats of a short segment of the stimulus (Fig. 1B), allowing us to evaluate model performance using both cross-validated log-likelihood ( $LL_x$ ) and more traditional peristimulus histogram-based methods, such as explained variance ( $R^2$ ). This analysis reveals that the models can explain  $34.5 \pm 3.1\%$  of the variance of the response, comparable with the performance of other models of MT processing (Nishimoto and Gallant, 2011). Furthermore, the improvement of model performance after inclusion of suppression is significant using both metrics and corresponds to a median of 25.3% of explained variance and 26.2% of  $LL_x$  (Figs. 4A, B). Moreover, there is a corre-



**Figure 3.** Incorporation of suppressive components into the description of MT processing. **A**, Model schematic for the ES model, with excitation (Exc, red), as well as DS-Sup (blue) and NS-Sup (green). The DS-Sup component has the same computational structure as the excitatory component but with the spatial weights constrained to be negative. The NS-Sup component is like DS-Sup but does not have a direction-tuning function and thus responds equally to all directions. **B**, Each component has its own speed-tuning function, but they are all very similar in this case. Note that the gray lines around each curve show the SD estimated with bootstrapping techniques. The shaded gray area denotes stimulus speed distribution. **C**, Subsequent unconstrained fits of the direction-tuning functions for each model component validates the forms used in fitting them, where Exc (red) and DS-Sup (blue) have very similar (antagonistic) tuning and NS-Sup component (green) is not selective for direction. **D**, Temporal kernels of Exc (red), DS-Sup (blue), and NS-Sup (green), demonstrating a slight delay of suppressive components. **E**, The measured spiking nonlinearity (dashed) and the corresponding parametric fit (solid) relative to the distribution of the generating signal (shaded gray). **F**, The improvement of cross-validated likelihood of ES models with each (or both) suppressive components added over the M0 model, expressed as a fraction of the performance of the model with both DS-Sup and NS-Sup. **G**, Spatial footprints of Exc (red), DS-Sup (blue), and NS-Sup (green). The arrows on the left indicate direction preferences of excitation and DS-Sup. A vertical slice of the weighting function is shown on the right, with the gray lines indicating the SE of each function over multiple model fits.

lation between log-likelihood and  $R^2$  (Fig. 4C), indicating a consistency of the two metrics. Because it provides a much more reliable metric of model performance in this case and does not require repeated stimulus presentations, we will only use cross-validated log-likelihood to measure accuracy of model predictions in the rest of the study.

**Properties of suppressions in MT**

Across the population of cells from which we recorded, the addition of one or both types of suppression improves model prediction for most of the neurons (81 of 94). For one-third of the neurons, the best model has both DS-Sup and NS-Sup (33 of 94,  $LL_x$  improvement =  $14.6 \pm 3.2\%$ ). For the rest of the neurons in the population, the best model either only has DS-Sup (29 of 94,

$LL_x$  improvement =  $11.9 \pm 5.3\%$ ) or only has NS-Sup (19 of 94,  $LL_x$  improvement =  $15.8 \pm 9.1\%$ ). Because a model with more components requires more data to fit and is more susceptible to overfitting, the fraction of cells that have both suppressive components might be underestimated using our criteria of requiring better cross-validated likelihood. For the neurons with both DS-Sup and NS-Sup components, we verified that there was a significant improvement in model performance over models with only one suppressive component (Fig. 5A), suggesting that DS-Sup and NS-Sup are describing different aspects of neuronal selectivity.

The two forms of suppression thus appear to be distinct contributions to the computation performed by the neuron. To explore this further, we first tested whether we correctly assumed

the forms of DS-Sup and NS-Sup tuning by relaxing constraints on the tuning functions and fitting these functions using nonparametric methods directly (see Materials and Methods). The resulting direction tuning of the NS-Sup components are indeed flat in almost every case, reflected as a circular variance very close to unity (Fig. 5B). In contrast, the direction tuning of DS-Sup fit in the same way has a circular variance comparable with that of the excitatory component (Fig. 5B).

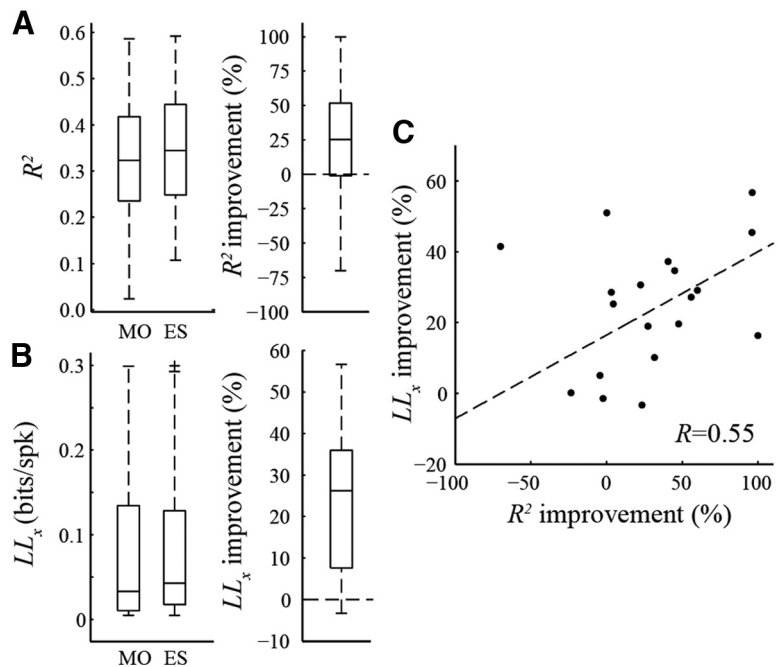
The two components also typically have different spatial profiles (Fig. 3F), with the spatial weights of NS-Sup usually located farther from the excitatory center (Fig. 5D) and with more dispersion than that of the DS-Sup (Fig. 5E). Both DS-Sup and NS-Sup temporal integration is significantly delayed relative to the excitatory components (relative latency of DS-Sup =  $12.6 \pm 4.0$  ms, relative latency of NS-Sup =  $25.5 \pm 6.9$  ms). The latency of DS-Sup is consistent with previous observations of MT surrounds ( $16 \pm 10$  ms; Perge et al., 2005),

In summary, the NS-Sup shares many characteristics of the previously described MT surround (Born and Bradley, 2005; Hunter and Born, 2011), which is farther from the center, covers a larger area, and is very broadly tuned and delayed. In contrast, DS-Sup is more likely to influence stimulus selectivity directly within or close to the receptive field. However, the different direction selectivity and spatial profiles of DS-Sup and excitation can lead to local tuning differences in subregions of the receptive field, seen for example in the study by Richert et al. (2013). Below, we will thus focus on the properties of MT neurons imparted by DS-Sup.

### Diversity of MT selectivity reflected in DS-Sup

The combination of their excitatory and suppressive influences can potentially give MT neurons selectivity to complex motion. To understand such selectivity, we first characterize the relationship between excitation and DS-Sup with two parameters: (1) the differences of direction preferences; and (2) the overlap extent of spatial weights (Fig. 5F). Note that here we only consider cells with DS-Sup detected and only include a single neuron from each multisite electrode experiment (see Materials and Methods), resulting in a population size of 55. It is most common for MT neurons to have suppression and excitation with matching tuning (antagonistic suppression; green), as well as with opposite preference (opponent suppression; red). Nevertheless, this distribution is continuous—with no apparent holes—and there are also many cells with suppression orthogonal to excitation (orthogonal suppression; blue). However, the distribution also shows a certain degree of diversity, especially for cells with orthogonal suppression. To highlight examples from different parts of this distribution and further investigate the diversity of the neuronal computation pictured, we divide cells into three groups based on their position within this two-dimensional distribution (Fig. 5F). Example model fits for cells in each group are shown in Figure 6.

Cells with antagonistic suppression represent the majority of our population (23 of 55). Spatially, DS-Sup has an asymmetric

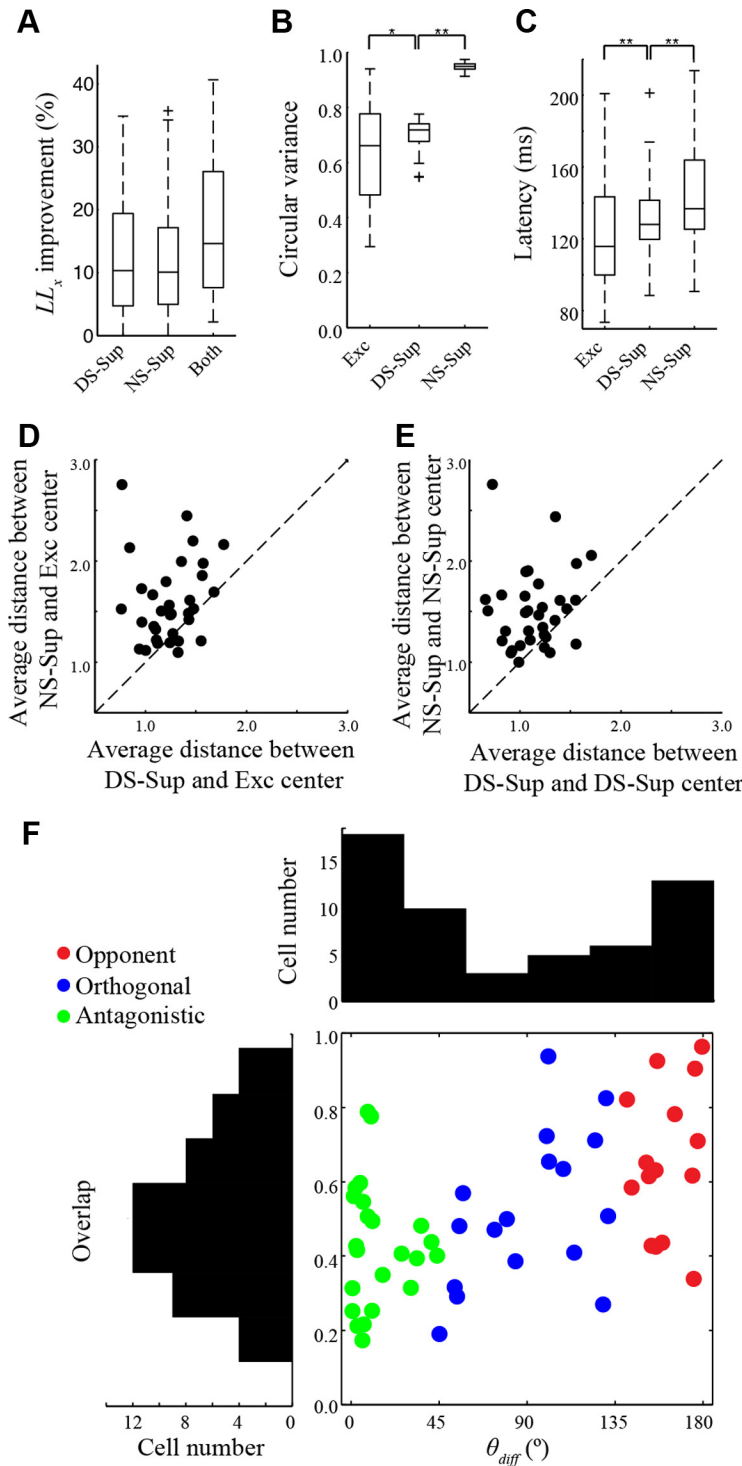


**Figure 4.** Model prediction accuracy. **A**, Fraction of explained variance ( $R^2$ ) for the MO and ES models (left). Paired comparisons (right) demonstrate significant improvement:  $22.0 \pm 11.7\%$  ( $p < 0.05$ , Wilcoxon's signed-rank test). **B**, Cross-validated likelihood ( $LL_x$ ) for the MO and ES models, showing a similar trend as in **A**. Percentage improvement is  $21.7 \pm 5.0\%$  ( $p < 0.05$ , Wilcoxon's signed-rank test). **C**, A comparison between the two metrics shows the correlation between them ( $r = 0.55$ ,  $p < 0.05$ ).

surround structure around excitation (Fig. 6A). Note that the lack of observed neurons in the top left corner of Figure 5F, corresponding to overlapping antagonistic suppression, is likely attributable to the inability of the model to detect overlapping suppression that has similar selectivity as excitation, because adding excitation and suppression with identical tuning at the same position will have no effect on the model output in the stimulus contexts we studied. In contrast, other studies have observed overlapping, antagonistic suppression using tailored stimuli (DeAngelis et al., 1992; Cavanaugh et al., 2002). In most cases, NS-Sup can also be detected at a farther distance from the center, suggesting that MT surrounds can be decomposed into a direction-selective and a more distant non-selective component. In principle, both suppressive components could contribute to classically measured “size tuning” (Allman et al., 1985). However, we find that size tuning is not limited to cells with antagonistic suppression, and only the strength of NS-Sup is significantly correlated with the extent of size tuning ( $r = 0.50$ ,  $p < 0.05$ ). This is also consistent with previous reports of broader direction tuning of the surround than that of the center (Born and Bradley, 2005; Hunter and Born, 2011). Conversely, the spatial footprint of DS-Sup resembles the asymmetric antagonistic surround seen in other studies (Xiao et al., 1995; Orban, 2008) and may contribute to selectivity for complex motion features, such as speed gradients and surface orientation (Xiao et al., 1997; Gautama and Van Hulle, 2001; Sanada et al., 2012).

Orthogonal suppression is found in a proportion of MT neurons (17 of 55; Fig. 6B). This type of suppression is not as well documented as the other type of suppressions, presumably because previous studies of MT surround usually restrict motion in the same or the opposite direction as the preferred direction. The spatial footprint of orthogonal DS-Sup is generally very different from excitation and exhibits a large degree of diversity across neurons with orthogonal suppression. Usually this suppression is confined to one side of the receptive field. In addition to image





**Figure 5.** Properties of the suppressive components. **A**, Box plots showing the improvement of cross-validated likelihood ( $LL_x$ ) over the MO model, applied to MT neurons with detectable DS-Sup and NS-Sup components ( $n = 33$ ). Including only DS-Sup on average increases the  $LL_x$  by 12.4% (left); the improvement is 12.1% for a model with only NS-Sup (middle) and 20.7% for a model with both DS-Sup and NS-Sup (right). **B**, Circular variance, which measures direction-tuning width, of the direction-tuning curve for excitation (left), DS-Sup (middle), and NS-Sup (right). DS-Sup has comparable tuning variance as excitation, whereas NS-Sup is almost completely non-selective to motion direction (circular variance close to 1).  $*p < 0.05$ ,  $**p < 10^{-6}$ . **C**, Both types of suppression are delayed relative to excitation, with latency relative to excitation for DS-Sup (left;  $12.6 \pm 4.0$  ms) and NS-Sup (right;  $25.5 \pm 6.9$  ms). NS-Sup is further delayed relative to DS-Sup ( $*p < 0.05$ ,  $**p < 0.001$ ). **D**, Average distance from the receptive field center for DS-Sup ( $x$ -axis) and NS-Sup ( $y$ -axis), in units of receptive field size. NS-Sup is farther away from the center than DS-Sup ( $p < 0.001$ ). **E**, NS-Sup is more dispersed than DS-Sup, as demonstrated by measuring the distance from the centroid of DS-Sup ( $x$ -axis) and NS-Sup ( $y$ -axis) ( $p < 0.001$ ). **F**, The population of MT neurons demonstrates a diversity of relationship between excitation and suppression, as shown by plotting the direction difference of each neuron between excitation and DS-Sup (horizontal) and amount of spatial overlap between them (vertical) ( $n = 55$ ). Each dot shows an individual neuron; although the

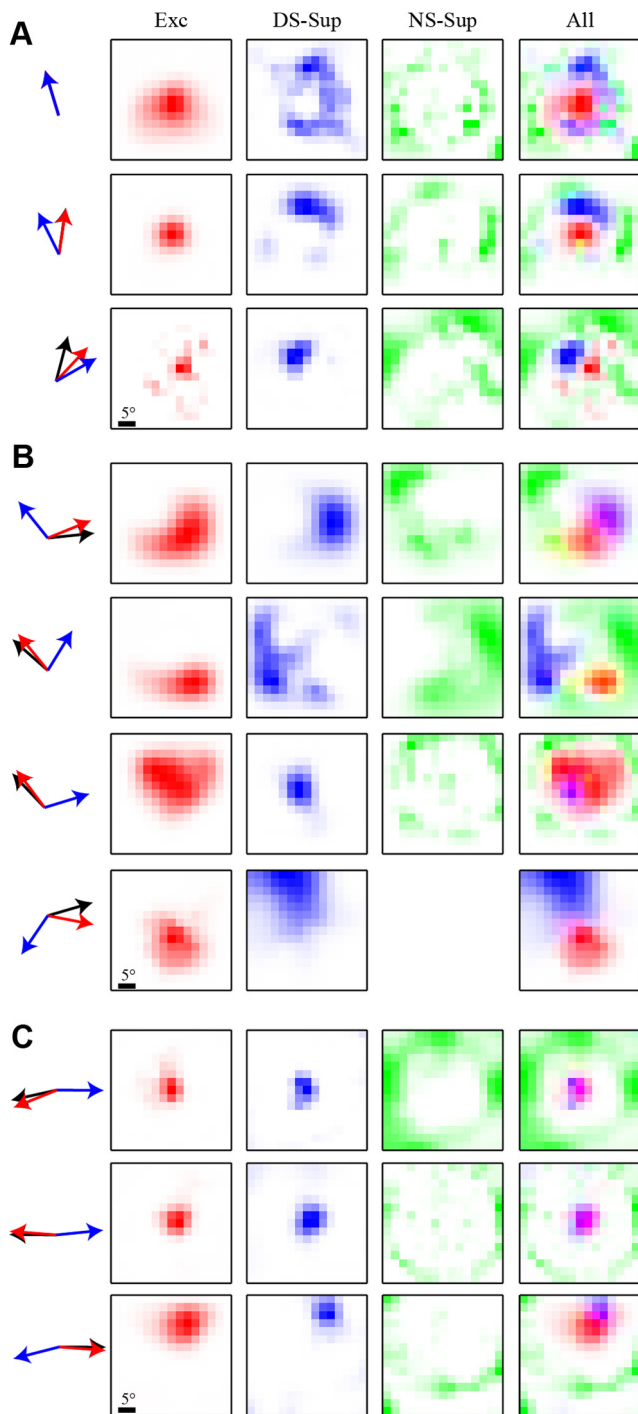
discontinuities, this arrangement may provide selectivity to curvature in the motion field, which is an important aspect of natural motion.

Finally, cells with opponent suppression comprise the rest of MT neurons in the population in which we detected DS-Sup (15 of 55). DS-Sup is typically colocalized with excitation for these cells (Fig. 6C), resembling the organization of the motion-opponent subunit we used in the MO model (Qian et al., 1994; Fig. 2A). However, the ES model still represents a different computation from the MO model because of the rectified nonlinearity in the subunit. As a result, the subunits will not give any response for nonpreferred stimuli and the MT neuron will respond at its spontaneous level, whereas for models with opponent suppression, the neuronal response will be truly suppressed below the baseline for nonpreferred stimuli. Indeed, we can observe this in the measured direction-tuning curve, in which we calculated the amount of suppression relative to the baseline firing rate in the nonpreferred direction. We find that the amount of suppression below baseline is significantly larger for cells with opponent suppression compared with the rest of the cells in the population ( $p < 0.05$ , Mann–Whitney  $U$  test).

**Selectivity and coding of complex optic flow**

The diversity of both excitatory and suppressive tuning direction and their spatial footprints suggest that MT neurons might be specifically tuned to velocity fields with different directions of motion at different spatial locations. Such visual stimuli are quite common in more natural settings in which velocity fields are attributable to motion of the observer and with objects at different depths. A simple way to reveal optic flow selectivity is to ask to what extent the firing rate of the neuron is modulated by a given flow component, which we gauge using the correlation coefficient  $\rho$  between each of the six optic flow components and (1) the actual neural response, (2) the predicted response of the MO model, and (3) the predicted responses of the ES model (i.e., with both types of suppression). Because MT re-

distribution is continuous, we draw color-coded distinctions to analyze different regimens of tuning (green, antagonistic suppression; blue, orthogonal suppression; red, overlapping opponent suppression). Marginal distributions over direction difference and overlapping extent are shown at the bottom and the left, respectively.



**Figure 6.** Example model fits for different types of tuning. **A–C**, Example fits are shown for different types of tuning: **A**, antagonistic suppression; **B**, orthogonal suppression; and **C**, overlapping opponent suppression. Each row shows (from left to right) direction preferences of excitation (Exc, red), DS-Sup (blue) and the measured preferred direction (black), spatial weights for excitation, DS-Sup, NS-Sup, and overlapping of them.

sponses are significantly correlated with nontranslational flow components in many cases, the models with suppressive components predicted their correlation with each optic flow component better than the MO model (Fig. 7A). Therefore, optic flow selectivity in MT relies at least in part on the suppressive influences we characterized previously.

We can also gauge the effect of adding suppression by comparing the responses of MO and ES models fit to the same neu-

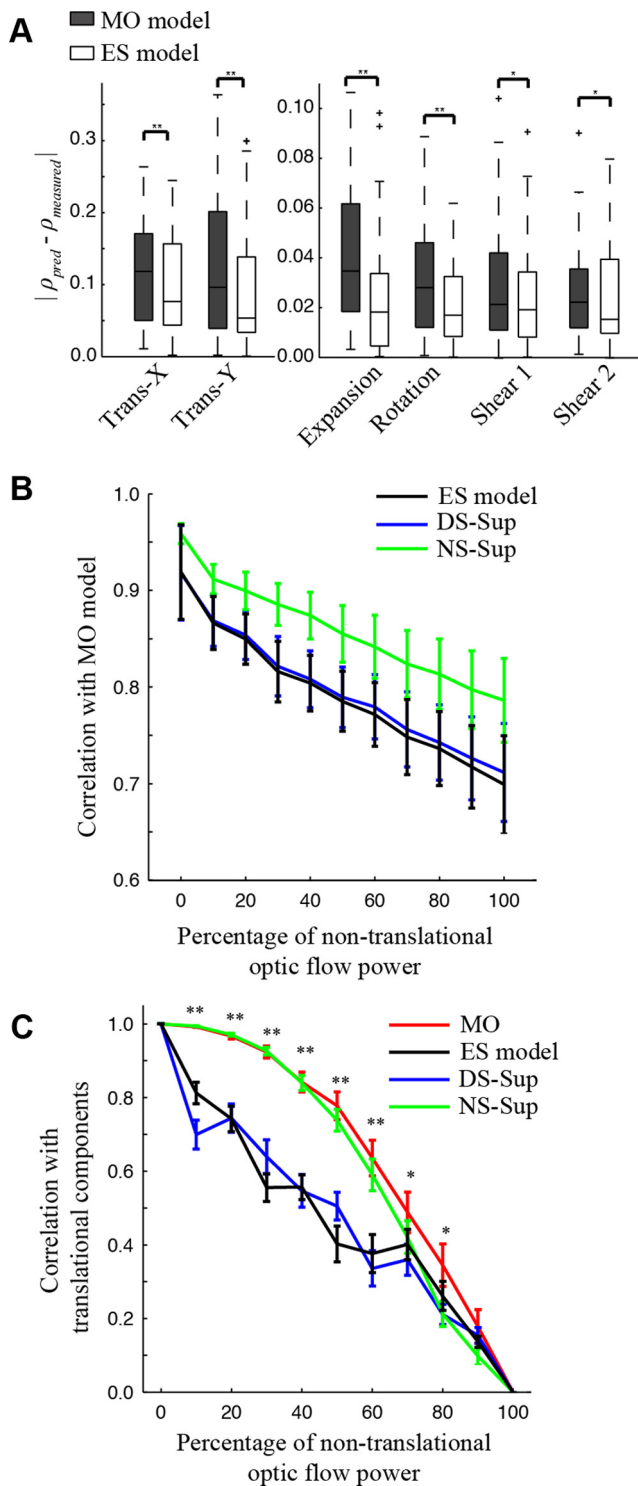
ron, as different amounts of nontranslational optic flow are presented. First, for stimuli that consist of only translation, the outputs of the MO and ES models are highly correlated ( $r = 0.91 \pm 0.06$ ), which is expected because both models can predict the standard direction-tuning properties of the cell. However, the correlation between the two models decreases as more nontranslational optic flow is introduced into the stimulus (Fig. 7B). When the stimulus is composed of only nontranslational components, the average correlation coefficient is only 0.67, with an SE of 0.06. This implies that, as expected, DS-Sup results in different responses to more complex motion stimuli. Such selectivity to complex optic flow appears to arise from the DS-Sup component rather than NS-Sup, because models with excitation and DS-Sup alone (Fig. 7B, blue) are similar to the full ES models, whereas models with only excitation and NS-Sup (red) are generally more correlated to the MO models.

Next we examine which components of the stimulus are most correlated with the predicted response for models with and without suppression. In general, the MO model outputs are much more correlated with the translational components than the ES models (Fig. 7C). The difference is most significant when a moderate amount of complex optic flow is introduced (10–40%). Consistent with Figure 7B, we find that DS-Sup alone is enough to explain this difference, whereas models with only NS-Sup mostly resemble behaviors of the MO models. This suggests that the difference between model predictions is attributable to the fact that MO models are only selective to translational motion, whereas the ES models exhibit additional selectivity to more complex optic flow stimuli. When more complex optic flow is introduced, both types of models are driven by the nontranslational flow components, but the responses of the two models generally depend on the stimulus in different ways, as reflected by the decreased correlation in Figure 7B.

Thus, the suppressive components of the ES model contribute to selectivity to nontranslation components of optic flow stimuli. Does such selectivity allow for a better representation of such stimuli by MT neurons? Previous work has shown that individual MT neurons exhibit fairly weak tuning for optic flow patterns, such as expansion and rotation, and that such tuning is highly dependent on stimulus position (Lagae et al., 1994). Here we ask whether a population of MT neurons could encode three-dimensional motion patterns and to what extent this encoding depended on the presence of nonlinear surround suppression.

To address these points, we applied an optimal linear decoding framework (DiCarlo and Cox, 2007; Mineault et al., 2012) to recover the velocity of a simulated object moving with different 3D velocities at a specific position relative to the observer (Fig. 8A). This task is also related to decoding egomotion, in which one typically assumes that the visual scene is static and the observer is moving. In this case, the “object” is the entire visual for egomotion decoding, because we are not examining the more complex case of recovering egomotion during simultaneous rotations of the eyes, head, or body. The results we report here are not sensitive to the size of the object.

For each simulation, we calculate the velocity pattern generated by 3D object motion in different directions (Fig. 8A, B) and then simulate a population of MT neuron responses to these stimuli (see Materials and Methods). An optimal linear decoder is trained to infer the 3D motion direction based on the simulated neuronal response using either the MO models or the ES models. The reconstruction performance of the decoder reflects a lower bound on the information that would be available for a downstream brain area (DiCarlo and Cox, 2007).



**Figure 7.** Suppression enhances selectivity to complex optic flow. **A**, To gauge the ability of the MO and ES models to capture selectivity of MT neurons to different components of optic flow, we calculate the correlation coefficient  $\rho$  between each flow component and the neuron response and compare that predicted by each model across the population of neurons with measured suppression ( $n = 33$ ). There are significant improvements for all six optic flow components with incorporation of suppression into the model ( $*p < 0.05$ ,  $**p < 0.001$ , Wilcoxon’s signed-rank test). **B**, As a second way to gauge this selectivity, we simulate the response of MO and ES models of the same neuron to different combinations of optic flow and measure how correlated the responses are as a function of the amount of nontranslational optic flow present in the stimulus. Each data point shows the average correlation coefficient between the responses of the MO and ES models over the 33 MT neurons that have both types of suppression. Responses of the two models, with both components included (black), are highly correlated for

The results of this simulation show that decoding performance using the ES model responses is significantly greater (i.e., less decoding error; Fig. 8C) compared with that using the MO model, whereas additional incorporation of NS-Sup does not further improve the performance. Moreover, the decoding performance depends on properties of suppression: although significant improvement is observed if we only use cells with antagonistic suppression or orthogonal suppression for the decoding task, such differences are not observed for cells with opponent suppression (Fig. 8D). This is likely because spatial offsets between excitatory and suppressive components may be instrumental in creating nontranslational optic flow sensitivity, but opponent suppression tends to have high degree of overlap with excitation and thus little spatial offset (Fig. 5F). Note that the performance of both models significantly degrades when stimuli are not presented in the center of the receptive field, in marked contrast to the same decoding task applied in the medial superior temporal (MST) area (Mineault et al., 2012), suggesting that computations described in MT represent one element of what is likely a hierarchical computation. That is, the initial selectivity developed in MT is further refined and generalized across spatial positions in higher-level areas such as area MST, as has been suggested for analogous computations in other areas (Riesenhuber and Poggio, 2000).

**Discussion**

In this study, we recorded the responses of MT cells in the context of naturalistic optic flow stimuli. To interpret the neuronal responses in this rich stimulus context, we constructed a hierarchical modeling framework that describes MT processing as integration over excitatory and both DS-Sup and NS-Sup inputs. Most previous studies of MT have focused on isolating either the excitatory tuning of MT neurons in the receptive field center or individual suppressive influences, such as the surround suppression (Xiao et al., 1995; Born 2000) or motion opponency (Snowden et al., 1991; Qian et al. 1994). For example, experimental studies of suppression usually use the preferred stimulus to drive the center and gauge the effects of suppression in this simplified context (Snowden et al., 1991; Xiao et al., 1995, 1997). At the same time, previous modeling studies have primarily focused on explaining how MT computes velocity from local motion signals (Qian et al., 1994; Simoncelli and Heeger, 1998; Rust et al., 2006; Nishimoto and Gallant, 2011) within its receptive field.

Here, by fitting a hierarchical ES model with both excitatory and suppressive components, our results provide the most complete picture of how different types of suppressive influences interact with excitation to impart selectivity to higher-order

translational stimuli ( $r = 0.92 \pm 0.05$ ) but decrease progressively for stimuli that contain more nontranslational optic flow components (black). This trend is highly significant (Spearman’s rank correlation coefficient,  $\rho = -0.38$ ,  $p < 10^{-11}$ ). DS-Sup appears to contribute to this selectivity to nontranslational optic flow, as demonstrated by including only the DS-Sup term with excitation (blue) or NS-Sup term with excitation (green). Models with only NS-Sup are more correlated with the MO model in general ( $p < 0.05$  when percentage of nontranslational optic flow is not 0,  $t$  test), whereas models with only DS-Sup are similar to the full ES model (Spearman’s rank correlation coefficient,  $\rho = -0.36$ ,  $p < 10^{-11}$ ). **C**, To measure to what extent model output is determined by translational component of the stimulus, we also report the correlation coefficients between model outputs and the translational components of the optic flow stimulus in the same context as in **B** as a function percentage of nontranslational optic flow. Significant differences between the MO model (red) and ES model (black) are revealed when the stimuli contain moderate amount of complex optic flow ( $*p < 0.05$ ,  $**p < 0.001$ ,  $t$  test). This difference is also observed for models with only DS-Sup (blue) but is absent for models with only NS-Sup (green).

motion stimuli. In particular, we find that suppression can be divided into a direction-selective component, which exhibits diverse structure and imparts functionally useful higher-order selectivity, and a non-selective component, which seems to play the role of surround suppression and normalization.

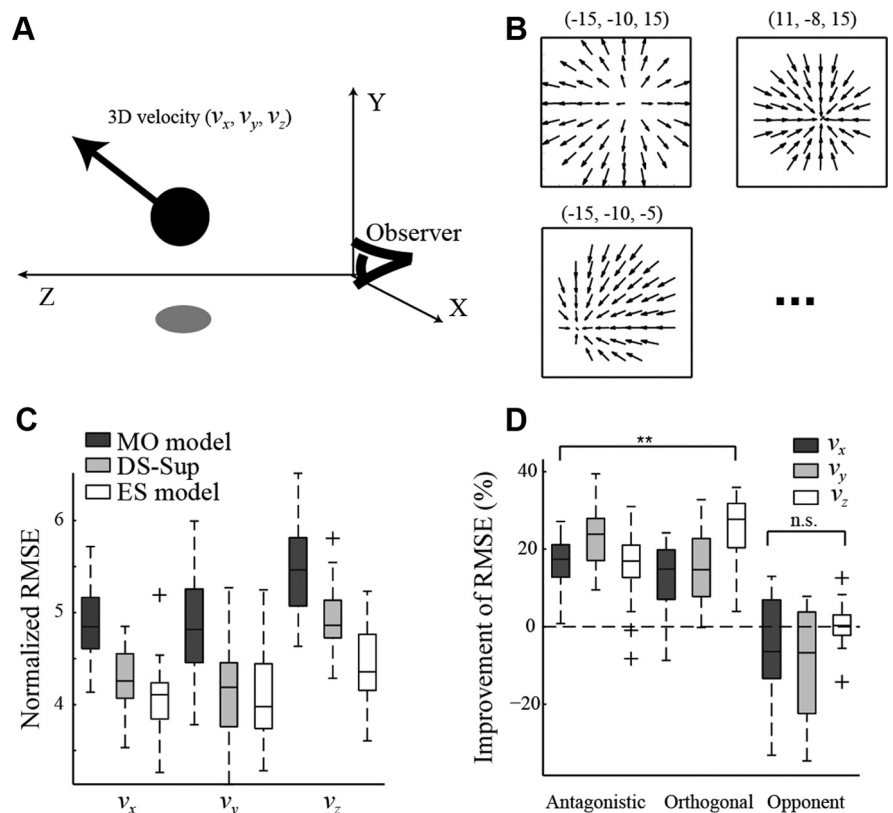
### The use of complex motion stimuli to probe area MT

The majority of work on area MT has explored its role in estimating the velocity of a rigidly translating object (Simoncelli and Heeger, 1998; Lisberger and Movshon, 1999; Rust et al., 2006). The idea dates back to the discovery that response of MT neurons is primarily dependent on the motion direction within a two-dimensional plane (Zeki, 1974). However, when more complex motion fields are used, selectivity to higher-order features, such as speed gradients (Treue and Andersen, 1996; Xiao et al., 1997) and surface orientation (Nguyenkim and DeAngelis, 2003; Sanada et al., 2012), has been observed, raising the question of what features of natural vision are represented by MT neurons. Our results show that, although excitatory contributions dictate direction selectivity, they are not sufficient to explain responses to motion stimuli with different velocities across visual space: suppressive contributions with different spatial profiles than excitation also significantly modulate MT responses and thus impart selectivity to complex motion features.

Here, we focus on understanding the role of spatial heterogeneity in natural motion fields and thus use a stimulus—a random-dot field—that produces an unambiguous velocity signal as a function of space. This allows us to map the excitatory and suppressive influences on an MT neuron across space and purposefully avoids the complexities associated with extracting velocity from texture patterns, which is another known aspect of MT processing (Pack and Born 2001; Rust et al., 2006; Jazayeri et al., 2012). In particular, one recent modeling study (Nishimoto and Gallant, 2011) extended such texture-based processing to explain complex motion stimuli, although as a result did not focus on the specific roles of spatially distributed suppression in processing such stimuli. We thus regard this approach as being orthogonal to understanding motion estimation for textures patterns and expect that the models investigated here might be consistent with and/or ultimately combined with models that address those complexities (Bradley and Goyal, 2008; Nishimoto and Gallant, 2011).

### Different forms of suppression in MT

Several forms of suppression have been documented in the literature. For example, neurons in area MT are often suppressed by motion in the antipreferred direction (Mikami et al., 1986). This is often termed “motion-opponent suppression,” which is likely



**Figure 8.** Role of suppressive surround revealed by population decoding of 3D velocity. **A**, Schematic of the 3D motion population decoding task. We calculate the optic flow pattern generated by object motion in a 3D space, with the motion direction and speed randomly selected (see Materials and Methods). The resulting velocity field is processed by a population of MT models, and a linear decoder is fitted to reconstruct the 3D motion-based outputs of 200 randomly selected models. **B**, Example velocity patterns used in the decoding task, with the 3D motion direction labeled above each pattern. **C**, Performance of the decoder based on input from the MO model (black) and ES model with only DS-Sup (gray) and ES model with both type of suppression (white). Results are quantified as the ratio of root mean squared error to the range of the parameter; smaller values indicate better performance. Using outputs of ES models gives better reconstruction performance along all three dimensions than using the MO models population. ( $p < 0.05$ ,  $t$  test). **D**, Percentage improvement of reconstruction performance between the MO model and the full ES model for model cells with antagonistic suppression (left), orthogonal suppression (middle), and opponent suppression (right). Significant improvements of reconstruction are observed for cells with antagonistic suppression and orthogonal suppression but not for cells with opponent suppression (\*\* $p < 0.01$ , n.s.  $p > 0.05$ ,  $t$  test).

related to similar phenomena that have been reported in psychophysical studies (Levinson and Sekuler, 1975; Qian et al., 1994), single-unit recordings (Mikami et al., 1986; Rodman and Albright, 1987), and functional magnetic resonance imaging (Heeger et al., 1999).

Surround suppression is another well studied property of MT receptive fields. Most neurons in area MT have receptive fields with antagonistic surrounds (Allman et al., 1985; Tanaka et al., 1986; Born, 2000; Tsui and Pack, 2011), and the classical view is that maximal suppression occurs when the surround stimulus moves in the same direction as that in the center (Bradley and Andersen, 1998; Born and Bradley, 2005). Other studies have shown that the suppressive surrounds could be quite complex relative to the center, exhibiting such properties as asymmetric spatial organization (Xiao et al., 1997), different contrast sensitivity (Pack et al., 2005), and less direction selectivity (Hunter and Born, 2011).

Although different suppressive mechanisms are often separately studied using targeted stimuli, our modeling approach provides a unified framework for characterizing suppression. We incorporate the motion-opponency assumption at the local unit level and allow for both DS-Sup and NS-Sup. Interestingly, for the majority of the cells we studied, the best model has both selective and non-selective

suppression but with different spatial profiles. The DS-Sup is closer to the receptive field center and appears to directly modulate motion selectivity, whereas the NS-Sup is more like to be the previously described surround suppression.

Although our model assumes a feedforward structure, the source of different forms of suppression remains unclear. Given the longer latency of suppression, it may be that these suppressive contributions actually reflect horizontal connections within area MT or feedback from higher areas, such as area MST. The contribution of different types of suppression may also depend on the stimulus (Huang et al. 2007). Although the goal of this study is to reveal types of suppression that are functionally relevant to processing of naturalistic optic flow, other types of suppression, such as those that are spatially and directionally aligned with excitation (DeAngelis et al., 1992; Cavanaugh et al., 2002), may not be revealed by our method.

### Spatial heterogeneity of MT processing

Although the classical view of MT receptive fields is that preferred directions in the center are essentially homogeneous and the surround is antagonistic and circularly symmetric (Tanaka et al. 1986), a handful of studies have focused on spatial heterogeneity of processing, using stimuli that either separately drove center and surround regions (Xiao et al., 1995; Orban, 2008) or separately mapped direction preferences (Richert et al., 2013) or sensitivity (Britten and Heuer, 1999) within the classical receptive field.

In our model, the spatial heterogeneity is reflected by the different spatial profiles and direction preferences of excitation and suppression. Although we assume that the subunit selectivity is the same for each model component, different direction selectivity and sensitivity could emerge within the receptive field because of the different combinations of excitation and suppression across space attributable to their different spatial footprints. This effect is most prominent for cells with orthogonal oriented suppression. The complex surround of area MT has been shown to play important roles in 3D shape estimation and motion segmentation (Buracas and Albright, 1996; Xiao et al., 1997; Gautama and Van Hulle, 2001). The additional heterogeneity in direction preferences may further contribute to selectivity to curvature in the motion field, which is an important aspect of natural motion. Indeed, the improvement of accuracy in 3D velocity estimation is most significant when orthogonal suppression is introduced to the model (Fig. 8D).

### The role of area MT in visual motion processing

From a computational perspective, the goal of visual motion processing is much broader than estimating 2D velocity. For a behaving animal, motion stimulus not only depends on object motion and visual depth but also results from optic flow patterns imparted by egomotion and eye movements. Multiple problems are thus involved in motion processing, such as detection of independently moving objects, egomotion estimation, 3D velocity estimation, and structure from motion (Beauchemin and Barron, 1995; Bradley et al., 1998; Pauwels et al., 2010; Sanada et al., 2012). Although other cues (e.g., disparity; DeAngelis et al., 1998) are also important, some of these problems involve recognition of complex motion patterns (Fermüller and Aloimonos, 1995) and discontinuities of the motion field (e.g., motion segmentation), which likely require selectivity to higher-order motion features. Although motion processing is certainly not complete at the stage of MT, our results suggest that such higher-order selectivity is already present in feedforward MT processing and can support behaviorally relevant tasks, such as estimation of the 3D velocity of a moving object.

MT neurons project to area MST (Ungerleider and Desimone, 1986; Tanaka et al., 1993), which is thought to calculate the heading direction of the observer (Perrone and Stone, 1994, 1998) and to estimate 3D velocity (Zemel and Sejnowski, 1998). MST neurons are more selective to optic flow components than MT neurons (Lagae et al., 1994) and are more invariant to the stimulus shape (Geesaman et al., 1997) and position (Duffy and Wurtz, 1995). This selectivity can be partly explained in a hierarchical framework using model MT neurons as inputs (Mineault et al., 2012), and our results here suggest that processing in MT neurons might more directly facilitate these ultimate goals of motion processing.

### References

- Adelson EH, Bergen JR (1985) Spatiotemporal energy models for the perception of motion. *J Opt Soc Am A* 2:284–299. [CrossRef Medline](#)
- Ahrens MB, Linden JF, Sahani M (2008a) Nonlinearities and contextual influences in auditory cortical responses modeled with multilinear spectrotemporal methods. *J Neurosci* 28:1929–1942. [CrossRef Medline](#)
- Ahrens MB, Paninski L, Sahani M (2008b) Inferring input nonlinearities in neural encoding models. *Network* 19:35–67. [CrossRef Medline](#)
- Albright TD (1984) Direction and orientation selectivity of neurons in visual area MT of the macaque. *J Neurophysiol* 52:1106–1130. [Medline](#)
- Allman J, Miezin F, McGuinness E (1985) Stimulus specific responses from beyond the classical receptive field: neurophysiological mechanisms for local-global comparisons in visual neurons. *Annu Rev Neurosci* 8:407–430. [CrossRef Medline](#)
- Bair W, O'Keefe LP (1998) The influence of fixational eye movements on the response of neurons in area MT of the macaque. *Vis Neurosci* 15:779–786. [Medline](#)
- Beauchemin SS, Barron JL (1995) The computation of optical flow. *ACM Computing Surveys* 27:433–466. [CrossRef](#)
- Born RT (2000) Center-surround interactions in the middle temporal visual area of the owl monkey. *J Neurophysiol* 84:2658–2669. [Medline](#)
- Born RT, Bradley DC (2005) Structure and function of visual area MT. *Annu Rev Neurosci* 28:157–189. [CrossRef Medline](#)
- Bradley DC, Andersen RA (1998) Center-surround antagonism based on disparity in primate area MT. *J Neurosci* 18:7552–7565. [Medline](#)
- Bradley DC, Goyal MS (2008) Velocity computation in the primate visual system. *Nat Rev Neurosci* 9:686–695. [CrossRef Medline](#)
- Bradley DC, Chang GC, Andersen RA (1998) Encoding of three-dimensional structure-from-motion by primate area MT neurons. *Nature* 392:714–717. [CrossRef Medline](#)
- Britten KH, Heuer HW (1999) Spatial summation in the receptive fields of MT neurons. *J Neurosci* 19:5074–5084. [Medline](#)
- Buracas GT, Albright TD (1996) Contribution of area MT to perception of three-dimensional shape: a computational study. *Vision Res* 36:869–887. [CrossRef Medline](#)
- Butts DA, Weng C, Jin J, Alonso JM, Paninski L (2011) Temporal precision in the visual pathway through the interplay of excitation and stimulus-driven suppression. *J Neurosci* 31:11313–11327. [CrossRef Medline](#)
- Cavanaugh JR, Bair W, Movshon JA (2002) Nature and interaction of signals from the receptive field center and surround in macaque V1 neurons. *J Neurophysiol* 88:2530–2546. [CrossRef Medline](#)
- Chichilnisky EJ (2001) A simple white noise analysis of neuronal light responses. *Network* 12:199–213. [Medline](#)
- DeAngelis GC, Robson JG, Ohzawa I, Freeman RD (1992) Organization of suppression in receptive fields of neurons in cat visual cortex. *J Neurophysiol* 68:144–163. [Medline](#)
- DeAngelis GC, Cumming BG, Newsome WT (1998) Cortical area MT and the perception of stereoscopic depth. *Nature* 394:677–680. [CrossRef Medline](#)
- DiCarlo JJ, Cox DD (2007) Untangling invariant object recognition. *Trends Cogn Sci* 11:333–341. [CrossRef Medline](#)
- Duffy CJ, Wurtz RH (1995) Response of monkey MST neurons to optic flow stimuli with shifted centers of motion. *J Neurosci* 15:5192–5208. [Medline](#)
- Engbert R, Mergenthaler K (2006) Microsaccades are triggered by low retinal image slip. *Proc Natl Acad Sci U S A* 103:7192–7197. [CrossRef Medline](#)
- Fermüller C, Aloimonos Y (1995) Direct perception of three-dimensional motion from patterns of visual motion. *Science* 270:1973–1976. [CrossRef Medline](#)

- Gautama T, Van Hulle MM (2001) Function of center-surround antagonism for motion in visual area MT/V5: a modeling study. *Vision Res* 41:3917–3930. [CrossRef Medline](#)
- Geesaman BJ, Born RT, Andersen RA, Tootell RB (1997) Maps of complex motion selectivity in the superior temporal cortex of the alert macaque monkey: a double-label 2-deoxyglucose study. *Cereb Cortex* 7:749–757. [CrossRef Medline](#)
- Ghose GM, Bearn DW (2010) Attention directed by expectations enhances receptive fields in cortical area MT. *Vision Res* 50:441–451. [CrossRef Medline](#)
- Hammond P, Reck J (1980) Influence of velocity on directional tuning of complex cells in cat striate cortex for texture motion. *Neurosci Lett* 19:309–314. [CrossRef Medline](#)
- Heeger DJ, Boynton GM, Demb JB, Seidemann E, Newsome WT (1999) Motion opponency in visual cortex. *J Neurosci* 19:7162–7174. [Medline](#)
- Huang X, Albright TD, Stoner GR (2007) Adaptive surround modulation in cortical area MT. *Neuron* 53:761–770. [CrossRef Medline](#)
- Hunter JN, Born RT (2011) Stimulus-dependent modulation of suppressive influences in MT. *J Neurosci* 31:678–686. [CrossRef Medline](#)
- Jazayeri M, Wallisch P, Movshon JA (2012) Dynamics of macaque MT cell responses to grating triplets. *J Neurosci* 32:8242–8253. [CrossRef Medline](#)
- Lagae L, Maes H, Raiguel S, Xiao DK, Orban GA (1994) Responses of macaque STS neurons to optic flow components: a comparison of areas MT and MST. *J Neurophysiol* 71:1597–1626. [Medline](#)
- Levinson E, Sekuler R (1975) The independence of channels in human vision selective for direction of movement. *J Physiol* 250:347–366. [Medline](#)
- Lisberger SG, Movshon JA (1999) Visual motion analysis for pursuit eye movements in area MT of macaque monkeys. *J Neurosci* 19:2224–2246. [Medline](#)
- Martinez-Conde S, Otero-Millan J, Macknik SL (2013) The impact of microsaccades on vision: towards a unified theory of saccadic function. *Nat Rev Neurosci* 14:83–96. [CrossRef Medline](#)
- Maunsell JH, Newsome WT (1987) Visual processing in monkey extrastriate cortex. *Annu Rev Neurosci* 10:363–401. [CrossRef Medline](#)
- Maunsell JH, Van Essen DC (1983) Functional properties of neurons in middle temporal visual area of the macaque monkey. II. Binocular interactions and sensitivity to binocular disparity. *J Neurophysiol* 49:1148–1167. [Medline](#)
- McFarland JM, Cui Y, Butts DA (2013) Inferring nonlinear neuronal computation based on physiologically plausible inputs. *PLoS Comput Biol* 9:e1003143. [CrossRef Medline](#)
- Mikami A, Newsome WT, Wurtz RH (1986) Motion selectivity in macaque visual cortex. I. Mechanisms of direction and speed selectivity in extrastriate area MT. *J Neurophysiol* 55:1308–1327. [Medline](#)
- Mineault PJ, Khawaja FA, Butts DA, Pack CC (2012) Hierarchical processing of complex motion along the primate dorsal visual pathway. *Proc Natl Acad Sci U S A* 109:E972–E980. [CrossRef Medline](#)
- Movshon J, Adelson E, Gizzi M, Newsome W (1985) The analysis of moving visual patterns. In: *Pattern recognition mechanisms* (Chagas C, Gattass R, Gross C, eds), pp 54, 117–151. Rome: Vatican.
- Nguyenkim JD, DeAngelis GC (2003) Disparity-based coding of three-dimensional surface orientation by macaque middle temporal neurons. *J Neurosci* 23:7117–7128. [Medline](#)
- Nishimoto S, Gallant JL (2011) A three-dimensional spatiotemporal receptive field model explains responses of area MT neurons to naturalistic movies. *J Neurosci* 31:14551–14564. [CrossRef Medline](#)
- Nykamp DQ, Ringach DL (2002) Full identification of a linear-nonlinear system via cross-correlation analysis. *J Vis* 2:1–11. [CrossRef Medline](#)
- Orban GA (2008) Higher order visual processing in macaque extrastriate cortex. *Physiol Rev* 88:59–89. [CrossRef Medline](#)
- Pack CC, Born RT (2001) Temporal dynamics of a neural solution to the aperture problem in visual area MT of macaque brain. *Nature* 409:1040–1042. [CrossRef Medline](#)
- Pack CC, Hunter JN, Born RT (2005) Contrast dependence of suppressive influences in cortical area MT of alert macaque. *J Neurophysiol* 93:1809–1815. [Medline](#)
- Pack CC, Conway BR, Born RT, Livingstone MS (2006) Spatiotemporal structure of nonlinear subunits in macaque visual cortex. *J Neurosci* 26:893–907. [CrossRef Medline](#)
- Paninski L (2004) Maximum likelihood estimation of cascade point-process neural encoding models. *Network* 15:243–262. [CrossRef Medline](#)
- Pauwels K, Kruger N, Lappe M, Woergoetter F, Van Hulle MM (2010) A cortical architecture on parallel hardware for motion processing in real time. *J Vis* 10:18. [CrossRef Medline](#)
- Perge JA, Borghuis BG, Bours RJ, Lankheet MJ, van Wezel RJ (2005) Dynamics of directional selectivity in MT receptive field centre and surround. *Eur J Neurosci* 22:2049–2058. [CrossRef Medline](#)
- Perrone JA, Stone LS (1994) A model of self-motion estimation within primate extrastriate visual cortex. *Vision Res* 34:2917–2938. [CrossRef Medline](#)
- Perrone JA, Stone LS (1998) Emulating the visual receptive-field properties of MST neurons with a template model of heading estimation. *J Neurosci* 18:5958–5975. [Medline](#)
- Perrone JA, Thiele A (2001) Speed skills: measuring the visual speed analyzing properties of primate MT neurons. *Nat Neurosci* 4:526–532. [CrossRef Medline](#)
- Qian N, Andersen RA (1994) Transparent motion perception as detection of unbalanced motion signals. II. Physiology. *J Neurosci* 14:7367–7380. [Medline](#)
- Qian N, Andersen RA, Adelson EH (1994) Transparent motion perception as detection of unbalanced motion signals. I. Psychophysics. *J Neurosci* 14:7357–7366. [Medline](#)
- Richert M, Albright TD, Krekelberg B (2013) The complex structure of receptive fields in the middle temporal area. *Front Syst Neurosci* 7:2. [CrossRef Medline](#)
- Riesenhuber M, Poggio T (2000) Models of object recognition. *Nat Neurosci [Suppl]* 3:1199–1204. [CrossRef Medline](#)
- Ringach DL, Shapley RM, Hawken MJ (2002) Orientation selectivity in macaque V1: diversity and laminar dependence. *J Neurosci* 22:5639–5651. [Medline](#)
- Rodman HR, Albright TD (1987) Coding of visual stimulus velocity in area MT of the macaque. *Vision Res* 27:2035–2048. [CrossRef Medline](#)
- Rust NC, Mante V, Simoncelli EP, Movshon JA (2006) How MT cells analyze the motion of visual patterns. *Nat Neurosci* 9:1421–1431. [CrossRef Medline](#)
- Sanada TM, Nguyenkim JD, DeAngelis GC (2012) Representation of 3D surface orientation by velocity and disparity gradient cues in area MT. *J Neurophysiol* 107:2109–2122. [CrossRef Medline](#)
- Sengpiel F, Sen A, Blakemore C (1997) Characteristics of surround inhibition in cat area 17. *Exp Brain Res* 116:216–228. [CrossRef Medline](#)
- Simoncelli EP, Heeger DJ (1998) A model of neuronal responses in visual area MT. *Vision Res* 38:743–761. [CrossRef Medline](#)
- Snowden RJ, Treue S, Erickson RG, Andersen RA (1991) The response of area MT and V1 neurons to transparent motion. *J Neurosci* 11:2768–2785. [Medline](#)
- Snowden RJ, Treue S, Andersen RA (1992) The response of neurons in areas V1 and MT of the alert rhesus monkey to moving random dot patterns. *Exp Brain Res* 88:389–400. [CrossRef Medline](#)
- Tanaka K, Hikosaka K, Saito H, Yukie M, Fukada Y, Iwai E (1986) Analysis of local and wide-field movements in the superior temporal visual areas of the macaque monkey. *J Neurosci* 6:134–144. [Medline](#)
- Tanaka K, Sugita Y, Moriya M, Saito H (1993) Analysis of object motion in the ventral part of the medial superior temporal area of the macaque visual cortex. *J Neurophysiol* 69:128–142. [Medline](#)
- Treue S, Andersen RA (1996) Neural responses to velocity gradients in macaque cortical area MT. *Vis Neurosci* 13:797–804. [CrossRef Medline](#)
- Tsui JM, Pack CC (2011) Contrast sensitivity of MT receptive field centers and surrounds. *J Neurophysiol* 106:1888–1900. [CrossRef Medline](#)
- Ungerleider LG, Desimone R (1986) Cortical connections of visual area MT in the macaque. *J Comp Neurol* 248:190–222. [CrossRef Medline](#)
- Weber F, Machens CK, Borst A (2010) Spatiotemporal response properties of optic-flow processing neurons. *Neuron* 67:629–642. [CrossRef Medline](#)
- Xiao DK, Raiguel S, Marcar V, Koenderink J, Orban GA (1995) Spatial heterogeneity of inhibitory surrounds in the middle temporal visual area. *Proc Natl Acad Sci U S A* 92:11303–11306. [CrossRef Medline](#)
- Xiao DK, Raiguel S, Marcar V, Orban GA (1997) The spatial distribution of the antagonistic surround of MT/V5 neurons. *Cereb Cortex* 7:662–677. [CrossRef Medline](#)
- Zeki SM (1974) Functional organization of a visual area in the posterior bank of the superior temporal sulcus of the rhesus monkey. *J Physiol* 236:549–573. [Medline](#)
- Zemel RS, Sejnowski TJ (1998) A model for encoding multiple object motions and self-motion in area MST of primate visual cortex. *J Neurosci* 18:531–547. [Medline](#)