VOLUME 9 NUMBER 11 NOVEMBER 2006
www.nature.com/natureneuroscience

nature neuroscience

Modeling motion processing in MT

Inhibiting glycolysis reduces seizures

Glutamate transporter currents carry information

nature neuroscience

How MT cells analyze the motion of visual patterns

Nicole C Rust[1,2,4], Valerio Mante[2–4], Eero P Simoncelli[1,2,5] & J Anthony Movshon[2,5]

Neurons in area MT (V5) are selective for the direction of visual motion. In addition, many are selective for the motion of complex patterns independent of the orientation of their components, a behavior not seen in earlier visual areas. We show that the responses of MT cells can be captured by a linear-nonlinear model that operates not on the visual stimulus, but on the afferent responses of a population of nonlinear V1 cells. We fit this cascade model to responses of individual MT neurons and show that it robustly predicts the separately measured responses to gratings and plaids. The model captures the full range of pattern motion selectivity found in MT. Cells that signal pattern motion are distinguished by having convergent excitatory input from V1 cells with a wide range of preferred directions, strong motion opponent suppression and a tuned normalization that may reflect suppressive input from the surround of V1 cells.

Analyzing visual scenes requires the visual system to integrate information from many sources. It is widely believed that the complex of visual areas outside the primary visual cortex (V1) performs such a synthesis, based on a hierarchical elaboration of the information represented by neurons in V1 (ref. 1). The analysis of visual motion is a particularly useful case in point: the computation of complex object motion by integration of simple object signals has been studied in both experiment and theory[2,3].

To analyze visual motion, the visual system first computes the motion of oriented elements in visual scenes, by filtering the image in both space and time[4]. This computation is represented by the activity of neurons in V1 that are selective for both stimulus orientation and direction[5–7]. Knowing the motion of a single component does not, however, reveal the motion of the pattern containing that component, because the components of a moving pattern can move in different directions[8]. The visual system solves this problem by combining motion signals from multiple V1 cells to compute pattern motion, represented by the activity of a population of directionally selective neurons in extrastriate area MT (V5)[2,9,10].

Modeling the computation performed by pattern direction–selective neurons in MT is a significant challenge. These neurons have nonlinear response properties that are evident from a comparison of their responses to single oriented gratings with those to plaids formed from superimposed pairs of gratings. The motion of a plaid can be constructed so that it is quite different from the motions of its component gratings, yet pattern neurons in MT respond with invariant tuning to both kinds of pattern. Furthermore, whereas some cells in MT are 'pattern direction selective', others exhibit a range of behaviors including 'component direction selectivity' as seen in V1 (ref. 2). A satisfactory model would account for this heterogeneity as well as for pattern direction selectivity in its pure form.

Here, we explored a model in which MT cells linearly combine signals from V1, which are themselves nonlinear. This is an elaborated form of a linear-nonlinear ('L-N') model, of a kind that is increasingly useful in sensory neuroscience. The power of L-N models lies in their simplicity, the ease with which they can be fit to data, and their ability to describe stimulus selectivity for a wide variety of neurons (see ref. 11). For example, the responses of V1 cells are well captured by a model that passes a stimulus through a set of one or more suitable linear filters, followed by a nonlinearity[12–14].

Pattern direction selectivity in MT cells, however, is not a tractable problem for an L-N model—linear models fail to account for pattern direction selectivity[3,15]. Moreover, we know that the V1 cells that provide inputs to MT are themselves nonlinear, so it is not reasonable to lump them with their projection to MT into a single linear stage. However, we do not need to abandon the entire L-N framework, but instead can evaluate a model in which MT cells are described by a linear filter that acts on signals from a population of direction-selective V1 complex cells, rather than directly on the visual stimulus[3,15]. The responses of such a cascaded L–N model reflect both the computations performed by the MT cell and the computations performed by its afferents. To test this model, we developed a new technique that allowed us to fit the cascaded L-N model to data from individual MT cells. We show that the model captures the variety of response properties observed in MT cells and provides an accurate and parsimonious account of the motion computations performed by cells signaling either component or pattern motion. The success of this enterprise suggests that such cascaded models may prove useful in accounting for the properties of other sensory neurons that are too far removed from the input stimulus to be described using more conventional approaches.

[1]Howard Hughes Medical Institute and [2]Center for Neural Science, New York University, New York, New York 10003, USA. [3]Institute of Neuroinformatics, University of Zurich, and Federal Institute of Technology, CH-8057, Zurich, Switzerland. [4]Present addresses: McGovern Institute for Brain Research, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA (N.C.R.) and Department of Neurobiology, Stanford University School of Medicine, Stanford, California 94305, USA (V.M.). [5]These authors contributed equally to this work. Correspondence should be addressed to N.C.R. (rust@mit.edu).

Received 15 August; accepted 18 September; published online 15 October 2006; doi:10.1038/nn1786

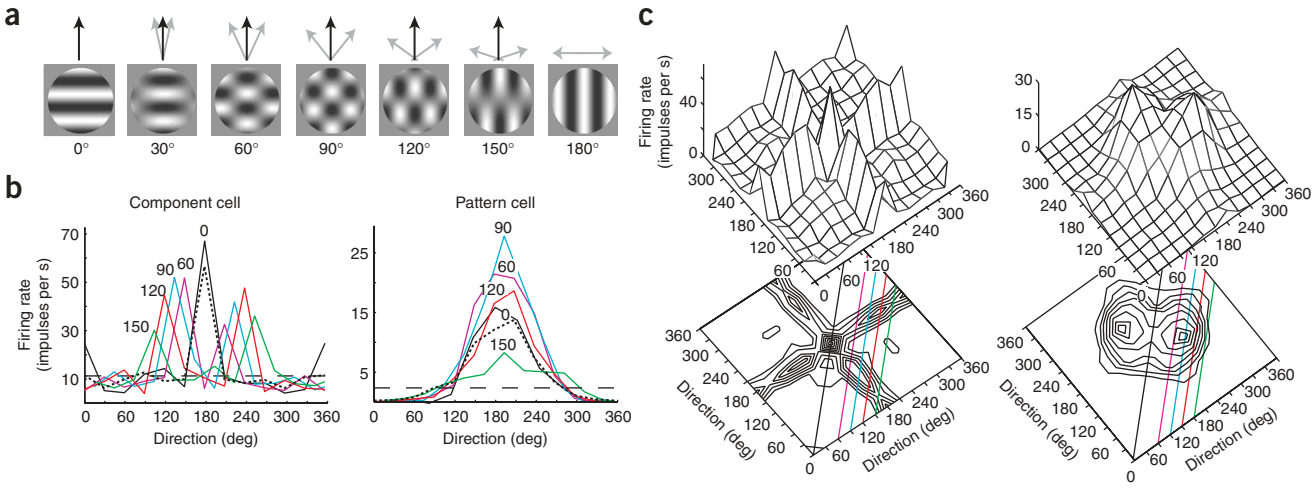

Figure 1 Component and pattern MT cell responses to plaid stimuli. (**a**) Example plaid stimuli. All plaids were constructed by superimposing two sinusoidal gratings of equal contrast (contrast = 1/6), moving at the same spatial and temporal frequency. Gray arrows show the motion directions of the two gratings; black arrows show the motion direction of the plaid[8]. The angular separation between the two gratings (the plaid angle) is given below each stimulus. When the plaid angle is zero, the gratings form a single grating with twice the contrast. (**b**) Direction tuning curves for example component (left) and pattern (right) cells, collected for plaids with different plaid angles. Each colored tuning curve represents the response to a particular plaid. For the component cell on the left, the functions have two peaks, displaced from one another by the angle of the plaid; for the pattern cell on the right, all have similar shapes and preferred directions. Also shown is the half-contrast grating tuning curve (black dotted line). (**c**) Surface and contour plots of response as a function of the direction of the two grating components. The colored lines in **c** indicate the loci of the particular plaids whose responses are shown in the same colors in **b**. Each direction-interaction plot is symmetrical about the main diagonal (a plaid constructed from gratings 0° and 120° gratings is equivalent to a plaid constructed from 120° and 0° gratings). In the contour plots, contours begin at 20% of the maximum firing rate and subsequent contours indicate 10% increments.

RESULTS

Pattern and component direction selectivity

The challenge for an MT model can be understood by considering the responses of MT cells to a set of stimuli consisting of sinusoidal gratings moving in all directions and all the plaids that can be constructed by additive combinations of these gratings (**Fig. 1a**). We measured direction tuning curves for the responses of cells to gratings and to a range of possible plaids (shown for two example cells in **Fig. 1b**). For the component cell, the functions have two peaks, displaced from one another by the angle of the plaid; for the pattern cell, all have similar shapes and preferred directions. These responses are drawn from full direction-interaction surfaces (shown in **Fig. 1c** for the same two cells). These surfaces fully reveal the characteristic differences between component and pattern cells. For the component cell, responses are confined to a plus-shaped locus, representing all stimuli for which at least one component is a grating moving in a direction close to the cell's optimum (here and hereafter rotated to 180° for clarity). For the pattern cell, responses are aligned parallel to a line with a slope of –1. Such a line represents a family of plaids that share an axis of motion but are composed of different combinations of grating directions (that is, the family of stimuli in **Fig. 1a**). The component cell surface is roughly what one would expect from the additive combination of the responses to single gratings. The pattern cell surface cannot be approximated by any additive combination of grating responses. Tracing the loci of particular plaids across the surface plots reveals that the double peak of the component cell tuning curves arises as the stimulus crosses the two arms of the plus, whereas the single peak in each pattern cell tuning curve corresponds to the unimodal core of the pattern cell's response surface (**Fig. 1c**).

To quantify this behavior, we compared each cell's plaid responses to idealized component and pattern predictions, by computing partial correlations between actual and observed responses[2]. **Figure 2a** shows a scatterplot of these correlations for our cells, transformed to Z-scores using Fisher's r-to-Z transformation[16]. Z_p is the figure of merit for pattern direction selectivity, and Z_c is for component direction selectivity. We took the difference between these scores as a 'pattern index', which ranges from negative values indicating component-selective behavior to positive values indicating pattern-selective behavior. The

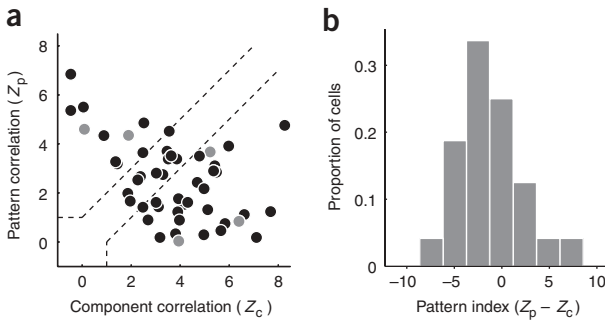

Figure 2 Quantification of neuronal responses to plaids. We compared each cell's response to idealized component and pattern predictions (Methods). Briefly, the plaid response is quantified by computing the partial correlation between the actual response of the cell and the component and pattern predictions and taking its Z-score, Z_p and Z_c. Component predictions are constructed by taking the linear superposition of two half-contrast grating tuning curves shifted by an amount corresponding to the plaid angle and subtracting the baseline response. For all plaid angles, the pattern prediction is the half-contrast grating tuning curve. For most cells ($n = 39$ of 50), Z_p and Z_c were computed as the mean Z_p and Z_c for 60°, 90°, 120° and 150° plaids. For the remaining cells, Z_p and Z_c are computed based on the responses to gratings and 120° plaids. (**a**) Z_p plotted against Z_c for 50 cells. Data for the five cells used as examples in subsequent figures are drawn in gray. (**b**) The distribution of the difference between Z_p and Z_c (hereafter the pattern index). The center bin of the histogram includes all cells classified as "unclassed"; cells with larger values are classified as pattern direction selective and cells with smaller values as component direction selective.

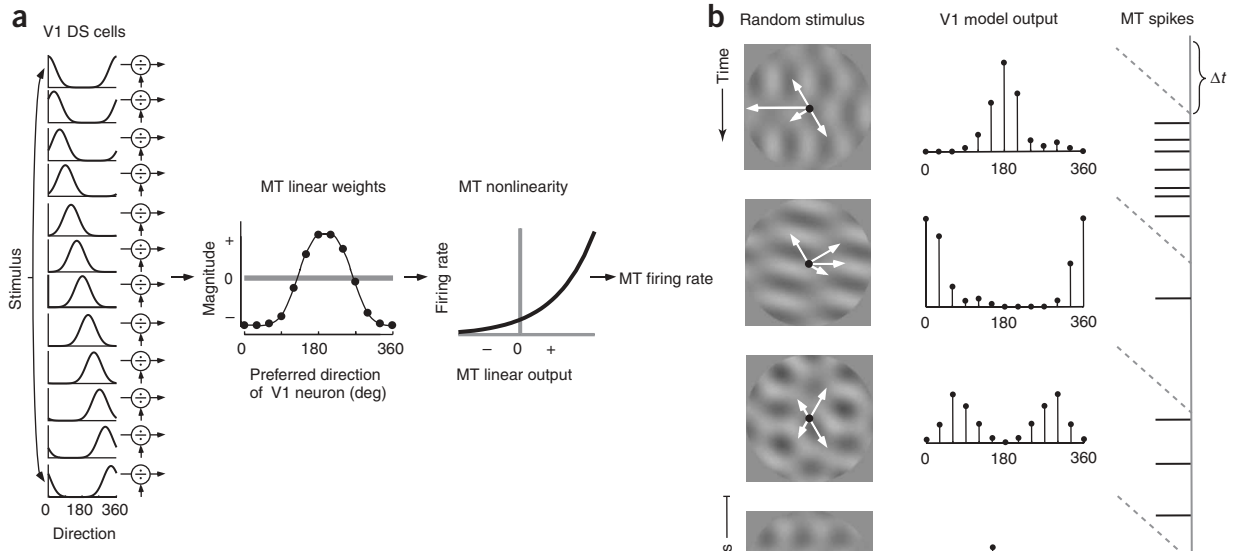

Figure 3 The cascade model and its characterization. (**a**) A stimulus is first passed through a population of 12 direction tuning curves with direction preferences spaced in 30° steps. The instantaneous response of each V1 neuron is divisively normalized. V1 parameters for each neuron include the bandwidth of the V1 direction tuning curve and strength of two normalization components, one tuned for direction and one untuned. In the MT stage, the outputs of the model V1 cells are linearly combined according to the MT linear weighting function. Linear weights can take on positive or negative values, corresponding to excitatory or inhibitory influences of the V1 neuron's response on the MT cell. The signal is then transformed into a firing rate via an instantaneous nonlinear function. (**b**) To recover the components of the model, we presented a sequence of stimuli containing random combinations of 6 gratings, chosen with replacement from a pool of 12 gratings drifting in different directions. Arrows indicate the grating components randomly selected on a particular trial; longer arrows indicate the selection of more than one grating drifting in the same direction. After 160 ms, another set of 6 randomly selected gratings was immediately presented. The second column shows the trial-by-trial V1 population response to the random grating stimulus. Plots show the responses of the set of 12 V1 neurons, ordered according to preferred direction. The third column displays a hypothetical spike train of an MT cell in response to the random grating stimulus. The spikes were shifted by the latency of the cell's response (Δt) and the number of spikes occurring in a 160-ms bin counted. We then fit the cascade model to these spike count data, using procedures described in Methods.

distribution of the pattern index is unimodal (**Fig. 2b**), suggesting that, by this measure, MT cells form a single heterogeneous population.

The cascade model

To investigate the computation of component and pattern selectivity in MT cells, we fit a relative of a previously published model[3] to the responses of individual MT neurons. The model (**Fig. 3a**) describes how a stimulus comprised of one or more superimposed gratings is transformed into the firing rate response of an MT cell. In the first stage of the model, the stimulus is passed through a population of directionally selective V1 neurons with equally spaced preferred directions. We describe the direction tuning curves of these neurons with a set of von Mises functions (circular Gaussians) of equal (but unknown) width. The response of each model V1 neuron is divisively normalized, both by the summed responses of the V1 population and by their own response (the latter is equivalent to applying a compressive nonlinearity). Each element is controlled by a variable that determines the strength of the normalization. We refer to these as the untuned and tuned components of normalization. As we consider in the Discussion section, these components may correspond, respectively, to the suppression that arises within the receptive fields of V1 cells[17] and the suppression from the surround[18]. The MT cell then computes a weighted sum of V1 responses; weights can be positive (excitatory) or negative (inhibitory). The result is then transformed into a firing rate by an instantaneous nonlinear function that captures the effect of spike threshold and other nonlinear operations that take place after summation.

To specify the model for each MT neuron, we presented 'hyperplaids' made by random combinations of six gratings chosen (with replacement) from a pool of 12 gratings of equispaced directions (**Fig. 3b**, left column; **Supplementary Video 1** online). Each hyperplaid was displayed for 160 ms and was immediately followed by another. Presenting this random selection of directions in a compound stimulus is akin to using a random selection of pixel values in a conventional spike-triggered analysis[6]. To measure the neuron's response, we counted the number of spikes that occurred in a 160-ms window, delayed from the stimulus transitions to take account of the neuron's latency (Methods). We then optimized the model parameters for each cell to maximize the likelihood of the observed responses. We first selected an initial set of values for the three V1 parameters and used these to compute the responses of the V1 population to each stimulus (**Fig. 3b**, center). We then determined the MT weighting function by linear regression of the MT cell's response against the model V1 responses. This procedure is similar to reverse correlation procedures that are commonly used to characterize L-N models. Finally, we recovered the MT nonlinearity, which describes the relationship between the output of the MT linear weighting function and the spikes. This process was repeated, searching over the three V1 parameters and recalculating the MT weights and MT nonlinearity.

Model predictions of responses to plaids and gratings

We used the optimized model for each cell, without further parameter adjustment, to predict the results of the independently measured

responses to gratings and plaids. **Figure 4** shows the actual and predicted responses for five example cells, ordered from the most component selective (**Fig. 4a**) to the most pattern selective (**Fig. 4e**). Comparing the actual responses with the predictions reveals that the cascade model captures all the main features of these responses. The model predicts the plus-shaped direction-interaction surfaces for the component and intermediate cells (**Fig. 4a–c**), and the attenuation or absence of these responses from the pattern cells (**Fig. 4d,e**). The asymmetric component response seen in some cells (for example, in **Fig. 4b**) is correctly captured. The model predicts the negative-diagonal alignment of the 'core' of the response surface for pattern cells, which indicates that they respond only to plaids whose pattern motion direction is similar to the preferred direction of the cell. Within this core response, the pattern cells were tuned for the same direction at all plaid angles and the direction tuning curves collected at different plaid angles had the same bandwidth; the cascade model has this behavior as well (**Fig. 4d,e**). Finally, the cascade model captures the two peaks in the surface plot for the intermediate and pattern cells (**Fig. 4c–e**); these indicate a more vigorous response to some plaids than to preferred gratings of the same contrast.

We then compared the pattern index observed for each of our cells with the index predicted by the cascade model (**Fig. 5**). As suggested by the examples (**Fig. 4**), the cascade model is a good predictor of the full observed range of selectivity for pattern motion.

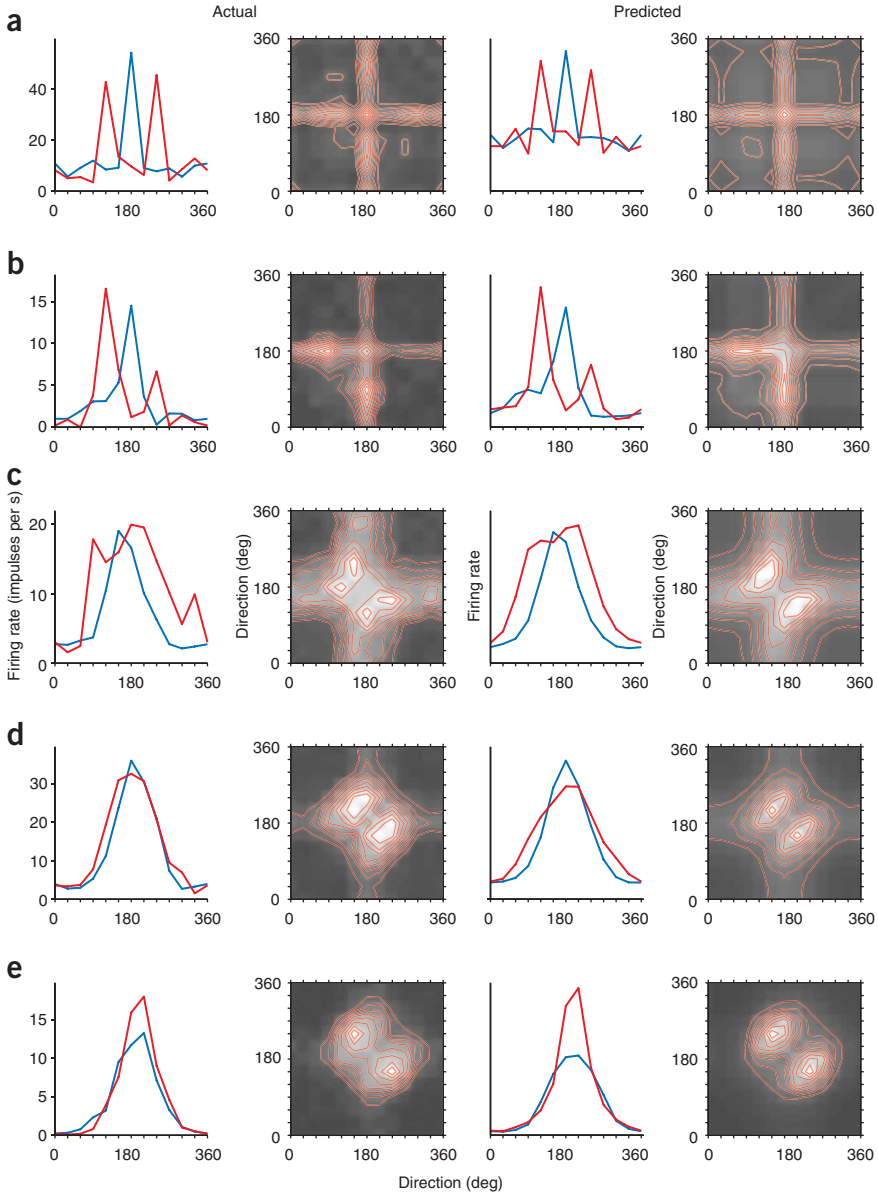

Figure 4 Comparison of actual and model-predicted responses to plaids for five example cells. (**a**–**e**) Neurons are ordered according to their pattern index, from the most component-like (**a**) to the most pattern-like (**e**). The first column shows the actual responses of the cell to gratings of 1/6 contrast (blue) and 120° plaids (red). The second column shows the direction-interaction surface. Here and hereafter, response magnitude for the direction-interaction surfaces is indicated by a combination of a gray-level image whose lightness indicates response and a set of overlaid contour lines. The right two columns show the responses predicted by the cascade model. The cell whose data are shown in **a** is the example component cell shown in the left column of **Figure 1b,c**. The cell whose data are shown in **e** is the example pattern cell shown in the right column of **Figure 1b,c**. In columns 2 and 4, the lowest contour corresponds to 20% of the difference between the minimum and maximum firing rates, and subsequent contours correspond to increases of 10%.

Recovered elements of the cascade model

The success of the cascade model's predictions suggests that this relatively simple model captures the computation of pattern motion by MT cells, and also the variety of behaviors shown by different cells. Which aspects of the model determine where a cell will fall on the continuum from component to pattern selectivity? We examined the internal elements of the recovered cascade model (**Fig. 6**) for our five example cells. The direction tuning curves (first column) indicate the selectivity determined by the model for the V1 stage. We visualize normalization strength (second column) as a function of the direction of a grating for a cell preferring 180°, and resolve it into an untuned component (dashed line) and into the contributions of both the tuned and untuned components together (solid line). We also obtained the recovered MT linear weighting functions for each cell (third column) and the output nonlinearities for the MT stage (fourth column).

Comparison of the models for the different cells (**Fig. 6**) suggests that some elements of the cascade model covary with the pattern index. The tuning bandwidth of the V1 input cells shows no consistent covariation with the pattern index (**Fig. 6**, first column). The overall strength of normalization at the preferred direction does not vary systematically across the example cells, but the relative strength of the tuned component of normalization increases with pattern index (**Fig. 6**, second column). The distribution of the excitatory MT linear weights is narrow for component cells, and becomes broader and develops more profound inhibition for cells with higher pattern index

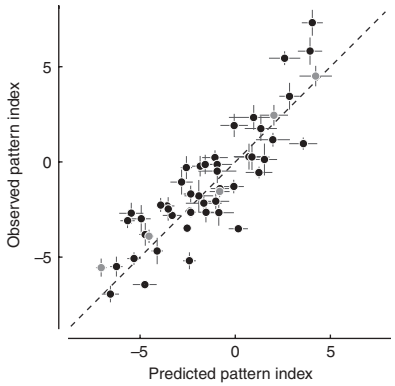

Figure 5 Comparison of the pattern index observed and predicted by the cascade model. The diagonal indicates identity. Data from the five example cells from **Figure 4** are drawn in gray. We computed predicted pattern indices by simulating the same number of trials acquired during data collection with a Poisson spiking mechanism. Error bars for the actual pattern indices were computed by bootstrap resampling of the data (Methods). Error bars for the predicted pattern indices were computed by bootstrap resampling of the responses to hyperplaids, refitting the data for each sample and computing the pattern index arising from the resulting model (*Methods*). Error bars indicate ± one s.d. of the bootstrap distributions.

(**Fig. 6**, third column). Finally, the MT nonlinearity seems similar across all the examples (**Fig. 6**, fourth column).

We examined the variation of model parameters with pattern index for the population. Three parameters determine the properties of the V1 stage (**Fig. 7a–c**); the other two represent the fractions of the V1 cell population that send strong excitatory and inhibitory weights to MT (**Fig. 7d,e**). Inspection of these scatter diagrams suggests that three of the parameters—the V1 tuned normalization component (**Fig. 7c**) and the MT excitatory and inhibitory weights (**Fig. 7d,e**)—are related to the pattern index. We used multiple linear regression to establish the reliability of these correlations. The five model parameters together account for 58% of the variance in the pattern index. The most significant predictors were the fraction of excitatory ($P < 0.0001$) and inhibitory ($P = 0.004$) weights. The strength of tuned normalization approached, but did not achieve, significance as a predictor ($P = 0.069$), and the V1 bandwidth and untuned normalization showed no reliable relationship with the pattern index ($P = 0.286$ and $P = 0.686$, respectively). When we restricted the regression analysis to the three V1 parameters (**Fig. 7a–c**), they accounted for 17% of the variance (a significant fraction: $P = 0.033$), with the strength of tuned normalization being the only individually significant predictor ($P = 0.023$). We did not observe a systematic relationship between the pattern index and any aspect of the nonlinear function that converts the output of the MT linear stage into a firing rate (data not shown).

We conclude that, in a statistical sense, variations in the distribution of the feedforward MT weights in the model carried the greatest weight, but that variation in the parameters of the V1 stage made a significant contribution to the model's ability to account for the range of behavior of MT cells.

The increase in the fraction of excitatory weights with the pattern index (**Fig. 7d**) suggests that the convergence of excitatory input from V1 cells of different direction preference increases systematically with the pattern index. Consistent with this, we found a positive correlation between direction bandwidth and pattern index for a population of MT cells recorded in this laboratory for a different study ($r = 0.38$, $n = 183$, $P < 0.0001$). The nature of V1 normalization signals also varied with the pattern index: pattern cells relied on a stronger tuned normalization signal in their inputs from V1 than did component cells (**Fig. 7b,c**). Although the distribution of excitation and inhibition differed for component and pattern cells, there was no consistent variation in the total strength of the inhibitory input with pattern index (data not shown). This suggests that inhibition is an important factor in determining the responses of all MT cells.

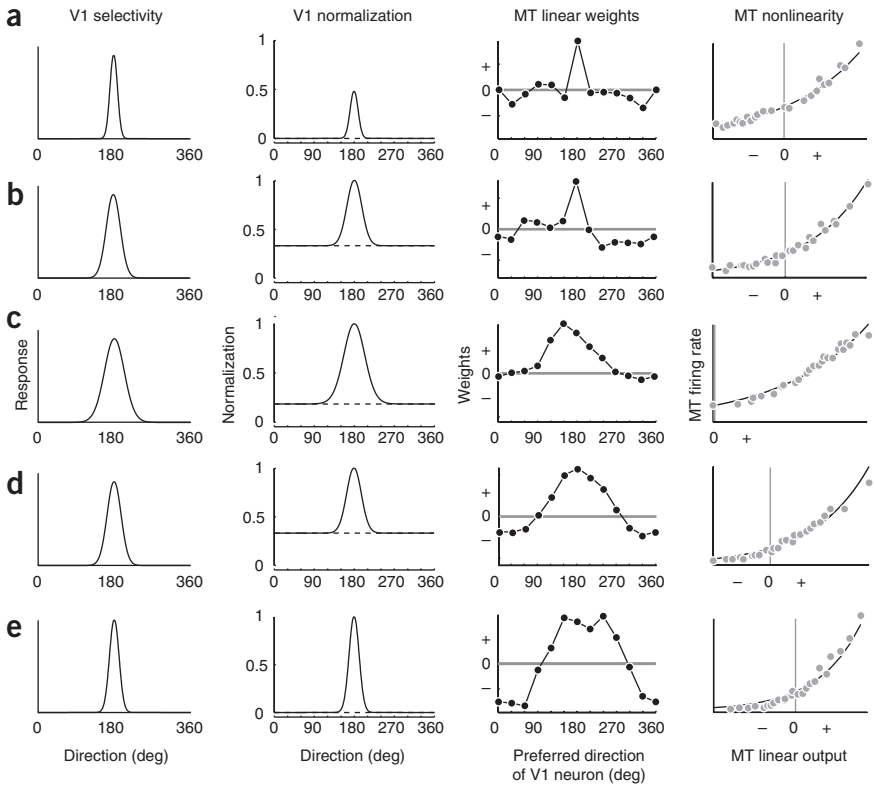

Figure 6 Recovered models for five example MT neurons, corresponding to the same cells whose data are shown in **Figure 4**. (**a–e**) The first column shows the prenormalized direction tuning curve for the V1 neuron with a 180° direction preference (all V1 neurons have the same direction tuning bandwidth). The next column illustrates the strength of the normalization signal, ranging from 0 (no normalization) to 1 (maximal normalization). Solid lines indicate the strength of the concatenated tuned and untuned normalization components ($\alpha_1 + \alpha_2$, Methods). Dashed lines indicate the normalization strength for the untuned normalization component alone (α_2). The third column illustrates the MT linear weights recovered for each cell. The final column shows the nonlinear function that converts the output of the MT linear weights into a firing rate response. The dots indicate computed, binned values for illustration purposes; the dark line indicates the exponential function that was fit to the unbinned data points (Methods). An MT linear output of zero corresponds to balanced excitatory and inhibitory input to the cell.

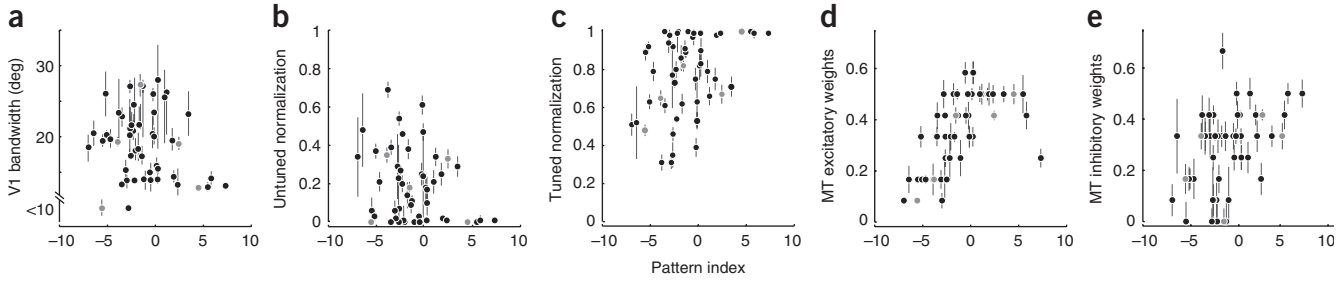

Figure 7 Relationship between the recovered cascade model parameters and pattern index. (**a**) The V1 direction tuning curve bandwidth, before normalization. The break in the axis indicates the model V1 cells whose tuning bandwidths are narrower than the resolution tested by these experiments and thus unconstrained by the fit. (**b**) Strength of the untuned normalization signal, measured by the fit weight α_1 (Methods). (**c**) Strength of the tuned normalization signal, measured by the fit weight α_2 (Methods). (**d**) The fraction of robustly excitatory weights in the recovered MT linear weighting function (those that exceeded 20% of the peak recovered weight). (**e**) The fraction of robustly inhibitory weights in the recovered MT linear weighting function (those whose magnitude exceeded 20% of the peak recovered weight).

Inspection of the standard deviations (error bars in **Fig. 7**) of bootstrap estimates for the model parameters for each cell (Methods) reveals that the parameter values chosen by the model are well constrained by the data. We also asked whether each element of the V1 stage of the model was needed to account for MT response properties. We constrained each model parameter in turn (by fixing the V1 bandwidth at an average value, by removing the untuned normalization, by removing the tuned normalization, and by removing the normalization altogether) and fit each constrained model to the data for all cells. Each constraint decreased the quality of the fit for a majority of cells—the constrained models were unable to capture the full range of variation in the behavior of MT cells.

How the cascade model computes pattern motion

To understand how the different factors in the cascade model combine to create pattern direction selectivity, we examined the direction-interaction surfaces at partial stages of the computation (**Fig. 8a–e**). The response surface for an un-normalized V1 cell—a simple motion energy detector[4]—has the basic 'plus' configuration of component cells and varies only in breadth of tuning (shown by the thickness of the 'arms': **Fig. 8a–e**, first column). When the two normalization components are included in the V1 response (**Fig. 8a–e**, second column), the peak response at the intersection of the plus—the response to a double contrast grating drifting in the preferred direction—is reduced relative to that at the arms (which contain one grating drifting in the preferred direction and a second in a different direction). The difference in the two sets of plots (column 1 versus column 2) can be attributed almost entirely to the tuned normalization component—the untuned component primarily rescales the response. Note that although this saturation in V1 is crucial in conferring the response properties of MT cells (see also **Fig. 8f**), normalization in V1 does not by itself produce pattern-like behavior at the V1 stage. The distribution of excitatory weights in the forward projection to MT also influences the response (**Fig. 8a–e**, third column). The plus configuration can still be discerned in all five plots, but the broad tuning of the pattern cells (**Fig. 8d,e**) shows where the core of the final response surface will lie. Adding the inhibitory weights in the forward projection to MT further shapes pattern selectivity (**Fig. 8a–e**, fourth column). This has little influence on the shape of the response surface in the component and intermediate cells (**Fig. 8a–c**), but attenuates or abolishes the arms of the plus in the pattern cells, more or less strictly confining responses to the central region of the direction-interaction surface.

The role of normalization in V1 in conferring the response properties of MT pattern cells is not easily seen in this representation. We therefore explored (for the example cell of **Fig. 8e**) a different dissection designed to illustrate the contribution of each of the three factors that determine pattern direction selectivity (**Fig. 8f**). The excitation that arises jointly from the un-normalized V1 tuning and the convergence of excitatory feedforward projections together combine to give the range of excitatory input for the cell (first column). The inhibition creates selectivity for preferred plaids by eliminating responses to plaids moving in other directions (second column). This tuning surface shows many of the characteristics of pattern cells, but crucially fails to be invariantly tuned for direction—observe, for example, that tuning for gratings (along the unity diagonal) is broader than tuning for plaids (along other diagonals). The invariance is conferred by tuned normalization signals in V1, which have a negative diagonal structure (third column)—this surface represents the denominator in the normalization equation. The effect of division is thus to suppress responses along this main diagonal, reducing the tuning width for gratings and creating an enhanced response to plaids over gratings (fourth column). Recall that this normalization signal within V1 fails to produce pattern-like responses from V1 cells (**Fig. 8a–e**, second column), because the tuning bandwidths of excitation and inhibition in V1 are matched and the normalization can therefore only modify the height of the response peak, but not its shape.

DISCUSSION

MT cells respond selectively to the direction of moving stimuli. Some respond selectively to pattern motion, whereas others represent the motion of stimulus components. We have shown that this range of behavior can be captured by a cascade model consisting of two L-N stages representing V1 and MT. Unlike models of visual processing that represent the transformations from stimulus to input in a single linear stage, the cascade model explicitly represents an intermediate nonlinearity that represents the behavior of V1 neurons. The properties of this nonlinear stage are consistent with the properties of V1 cells that project to MT (ref. 10), and the model as a whole captures the main variations in response that are seen in MT cells.

We have also achieved an important practical advance by developing a new method to fit the cascade model to data from individual neurons. A number of other researchers have developed methods to fit L-N models to the outputs of a preprocessing stage based on spectral energy[19–21], but these are limited by the preprocessing stage, whose

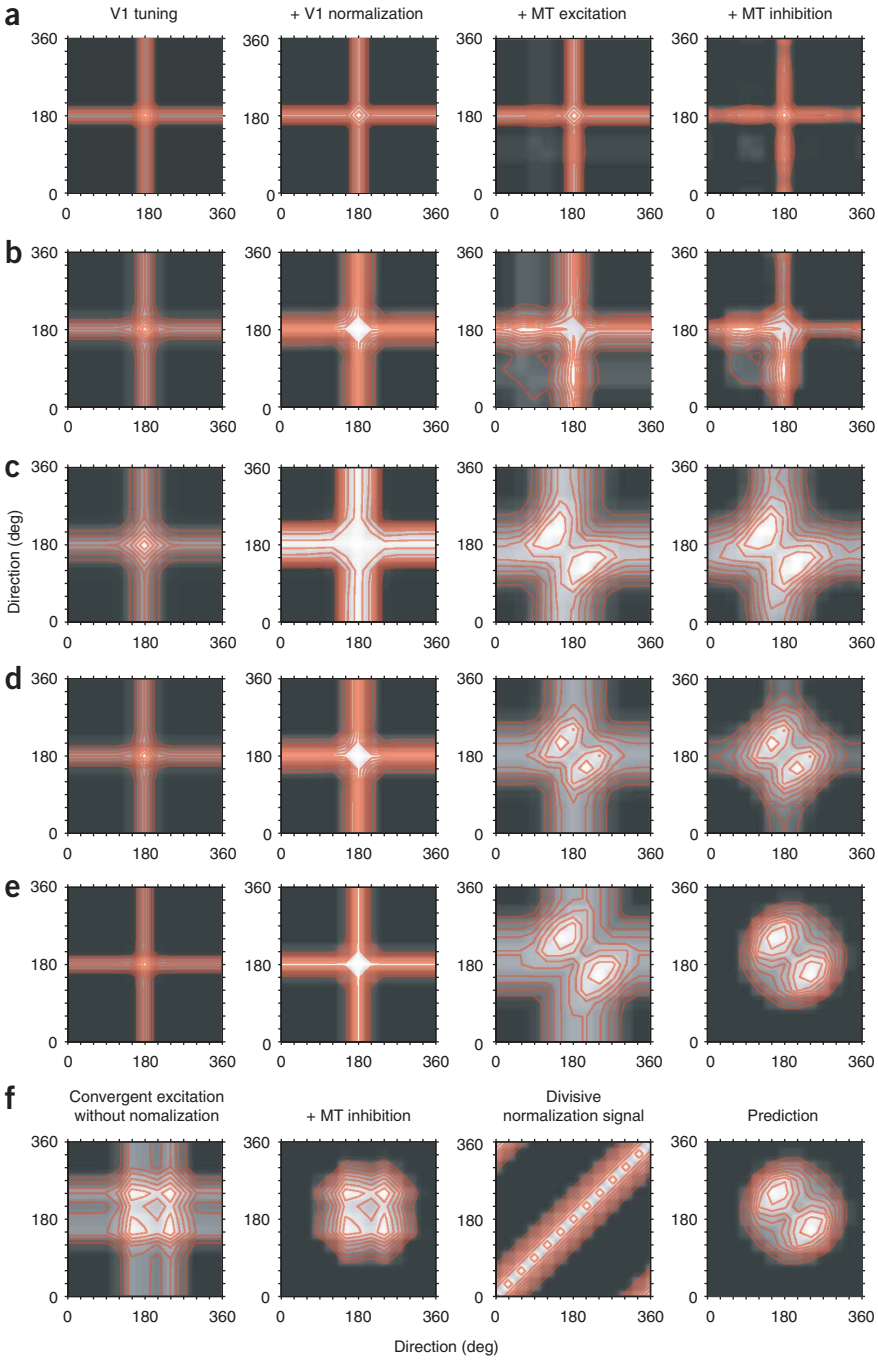

Figure 8 Dissection of the elements of the cascade model that create particular kinds of selectivity for the motion of plaids. (**a**–**e**) Each column shows the direction-interaction surface response at one stage of the cascade model, for the same five example cells of **Figures 4** and **6**, ordered by pattern index. The first column shows the response of a version of the V1 stage of the model, without normalization, equivalent to the square root of the response of an energy model[4]. The second column shows the response of the V1 stage when normalization is included. The third column shows the response that would be obtained if this V1 stage projected to the MT stage using only the excitatory feedforward weights. The fourth column shows the effect of adding the feedforward inhibitory weights. The full model predictions (which differ from the fourth column by the addition of the MT nonlinear processing stage) are shown for these cells in **Figure 4**. (**f**) Plots illustrating the role of V1 normalization in the computation of pattern motion for the fifth example cell. The first panel shows the combined effects of the V1 tuning bandwidth and the distribution of MT excitatory weights, without normalization. The second panel demonstrates the effect of including the MT inhibitory weights. The third panel illustrates the distribution of the V1 normalization signal across the direction-interaction surface. The final panel shows the effect of adding the normalization signal at the V1 stage. This is like dividing the response due to convergent excitation and inhibition (panel 2) by the normalization signal (panel 3). Only in this final panel is direction-invariant pattern selectivity seen.

form is fixed. Our model is of course also limited, but we tried to ensure that our preprocessing stage reflected the behavior of direction-selective V1 complex cells as accurately as necessary, by including two forms of divisive normalization in the preprocessing. More significantly, we did not assume a fixed form, but explicitly fit the properties of the preprocessing stage in recovering the model. This had a considerable impact on our conclusions—analysis showed that variations in the properties of MT neurons depend importantly on the particular properties of the neurons that provide their inputs. For example, analysis of the cascade model confirms an earlier proposal[15] that the tuned component of V1 normalization is essential for generating the pure form of pattern direction selectivity (**Fig. 8f**).

Interpreting the cascade model

What physiological mechanisms do the elements of the cascade model represent? Throughout this paper, we have freely used terms with specific anatomical and physiological meanings to name elements of the cascade model. And yet our model is, at heart, a functional model that does not aspire to complete anatomical and physiological accuracy. It is rather intended to capture, in the least complex possible way, the essential neural computations that lead to the observed selectivity of MT cells.

The first stage of the model is of V1. Its essential elements—direction selectivity and normalization—are well-known characteristics of V1 neurons, and we have made the model faithful to these properties of V1 cells. But we simulate V1 with only 12 neurons, which represent the many thousands of V1 neurons that project to any particular site in MT (ref. 22). We also recognize, of course, that MT receives inputs from other cortical areas, notably V2 and V3, as well as from a number of subcortical sources[22,23]. Little is known about the properties of these input neurons, though neither in V2 nor in V3 is pattern direction selectivity more prevalent than in V1 (refs. 24,25). Moreover, it seems that most or all nonstriate inputs to MT depend on V1 for their function[26,27], suggesting that encapsulating all our inputs in a V1 stage is not unreasonable. We also recognize that some of the properties important to the creation of

pattern direction selectivity that we attribute to V1 may be represented even more strongly in other extrastriate areas providing input to MT.

We attribute two normalization mechanisms to V1, an untuned component that is widely accepted[17,28,29] and a tuned component that may seem unfamiliar. Each has an important role in the cascade model's account of the variation of properties among MT cells (**Fig. 6**). As a confirmation of the physiological plausibility of our model, we computed the contrast of a preferred-direction grating that produced the half-maximal response for each model V1 neuron, as a measure of the combined strength of the tuned and untuned normalization components. The simulated values fell within the range previously reported for V1 cells (reported range 0.01–0.4 (ref. 30), mean 0.193; model values: 0.01–0.26, mean 0.03). We used cross-orientation suppression to assess whether the relative contributions of the tuned and untuned normalization components were reasonable. Cross-orientation suppression is the reduction in response that occurs when an orthogonal grating is superimposed on a preferred grating. In normalization models like ours, this behavior arises because the orthogonal grating reduces the response by adding to the untuned normalization signal. We quantified this suppression as the ratio between the response to a plaid and that to a preferred grating. Simulated values from our model V1 neurons (0.56–1.0, mean 0.90) matched the range of the values we obtained by analysis of unpublished data for the cells collected in a study of V1 cells projecting to MT (ref. 10; range 0.56–1.04, mean 0.79).

We justify the unfamiliar tuned normalization mechanism by recalling that our stimuli were all large, matched in size to the receptive fields of the MT cells studied. Such stimuli would engage the suppressive surround mechanisms that have been well documented in V1; these act in a divisive fashion and are selective for both orientation and direction[31,32]. Moreover, selective surround suppression is often strong in directionally selective V1 neurons[33]. We therefore suggest that our tuned normalization mechanism represents the action of this surround mechanism. This interpretation is also consistent with the suggestion that pattern selective neurons in MT get input from "end-stopped" cells in V1 (refs. 34–38), because these are the cells with the strongest surround suppression and our tuned normalization mechanism is strongest in pattern cells (**Figs. 6** and **7**).

The feedforward excitation in the cascade model has a direct interpretation in terms of the projections of V1 cells to MT, which are excitatory[39]. The individual direction tuning bandwidths fit to the V1 cells in our model are within the ranges of bandwidths based on data from antidromically identified MT-projecting V1 cells (fit 10–29°; data 11.2–73.8°). A robust response to a large angle plaid (for example, 150°) requires excitatory input in the direction of each of the component gratings (for example, ± 75°). This is not likely to arise from V1 cells tuned for the cell's preferred direction, but rather from the convergence of V1 cells tuned for different directions. However, broadly tuned excitatory input alone would result in broadly tuned plaid direction tuning curves, which are not seen in pattern cells. In our model, this is prevented by strong inhibitory feedforward input, which contributes directly to pattern selectivity (**Fig. 8**). As there is no known anatomical substrate for feedforward inhibition, we must postulate that it arises either from inhibitory local-circuit neurons in MT or from motion-opponent inhibition in V1 (N.C.R., N.J. Majaj, E.P.S. & J.A.M., *Soc. Neurosci. Abstr.* 657.10, 2002). We did not include directional inhibition in the V1 stage of the cascade model, and this omission might manifest itself in a physiologically implausible but functionally reasonable way.

More generally, we included only mechanisms required to predict the steady-state responses of component and pattern cells to plaid stimuli, and omitted a number of mechanisms known to exist in both V1 and MT, including adaptation[40], spatial integration (N.J. Majaj, M. Carandini, M.A. Smith & J.A.M., *Soc. Neurosci. Abstr.* 674, 1999), dynamical modulation[41] and MT contrast gain control[42,43]. Our model here seeks to identify the physiological mechanisms that have a role in computing pattern motion, without necessarily providing a robust account of all the response properties of these cells.

Cell classes and model continua

The first accounts of pattern direction– and component direction–selective cells in MT suggested that the two cell types might be distinct[2,44,45], though the evidence on that issue was equivocal. Viewed through the lens of the cascade model, we see that the diversity of MT cell behavior can be captured by a single parametric model, variations of which generate everything from pure component to pure pattern selectivity. This unification is reminiscent of recent re-examinations of the classical distinction between simple and complex cells in V1 (ref. 5). It emerges that conceptually similar variations in the parameters of a single underlying model can account for the range of behaviors from purely simple to purely complex[46–48]. Such models serve a valuable purpose in accounting for the range of response properties in a particular area with a single underlying computational form. They do not invalidate the idea that separate classes of cells may be identified on other grounds, such as anatomical location or projection, but they give support to the idea that seeking a unified account of the functions of individual cortical areas is a worthy and attainable goal.

METHODS

Recording methods. We prepared adult macaque monkeys (*Macaca fascicularis* and *Macaca nemestrina*) for recording as described previously[32]. We maintained anesthesia with an intravenous infusion of 4–30 μg per kg (body weight) per h of sufentinil citrate in lactate dextrose-saline (4–10 ml per kg body weight per h). We also infused vecuronium bromide (Norcuron, Organon: 0.15 μg per kg body weight per h) to prevent eye movements. The monkey was artificially ventilated, and body temperature was maintained with a thermostatically controlled heating pad. We continuously monitored vital signs (heart rate, lung pressure, electroencephalogram (EEG), electrocardiogram (ECG), body temperature, urine flow and osmolarity, and end-tidal P_{CO2}). Gas-permeable contact lenses protected the corneas, and supplementary lenses chosen by direct ophthalmoscopy made the retinas conjugate with a screen 80–180 cm distant. At the end of the experiment, the monkey was killed with an overdose of sodium pentobarbital, and perfused with 0.1% phosphate-buffered saline (PBS) followed by 4% paraformaldehyde. Confirmation that recording sites lay within MT was made through histological identification of electrolytic lesions in Nissl- and myelin-stained frozen 40-μm sections. We conducted all experiments in compliance with the US National Institutes of Health Guide for the Care and Use of Laboratory Animals and with the guidelines established by the New York University Animal Welfare Committee.

We recorded from well-isolated single neurons in MT with quartz-glass microelectrodes (Thomas Recording). We identified activity in MT from the vigorous direction-selective response of isolated neurons and unresolved background activity. All receptive fields were centered between 2° and 20° from the fovea. Recorded signals were amplified, band-pass filtered, and fed into a time-amplitude window discriminator. Spike arrival times and stimulus synchronization pulses were stored with a resolution of 0.1 ms.

Visual stimulation. Stimuli were presented on a gamma-corrected monitor with a refresh rate of 100 Hz and a mean luminance of 33 cd m^{-2}. Stimuli were generated by an Apple Macintosh G5 workstation and presented to each cell's preferred eye; the other eye was covered. On encountering a cell, we first optimized sinusoidal grating stimuli for direction, spatial frequency, temporal frequency (drift rate) and size. Thereafter, we presented all stimuli at optimal spatial and temporal frequency within a circular window confined to the classical receptive field[32]. All stimuli were presented in randomly interleaved blocks.

Analysis of neuronal response. To measure visual latency, we presented stimuli in a continuous stream, in which each randomly selected stimulus was displayed for 160 ms and was immediately followed by the next stimulus. Responses to these stimuli were estimated by counting spikes in a 160-ms window shifted in time from the stimulus stream by the cell's response latency (Δt, **Fig. 3b**). We estimated latency for each cell by choosing the latency that maximized the modulation of the direction tuning curve, as measured by the variance of its 12 constituent values[16]. For the 50 cells reported here, we recovered latencies of 54–147 ms with a mean of 88 ms. Our latency values were slightly larger than those reported earlier[16] because we used a lower contrast (0.16 versus 0.5).

To quantify MT cell responses to plaids, we used standard methods to compute the partial correlation of the actual response to plaids with the predictions of idealized models of pattern and component direction selectivity (r_p and r_c, respectively)[2,16]. All grating components were displayed at 0.16 contrast. The predicted response of the pattern model was the grating tuning curve at 0.16 contrast. The predicted response of the component model was computed as the sum of two direction tuning curves generated in response to a 0.16-contrast grating, each shifted by an amount appropriate for the plaid angle, with the baseline subtracted. To stabilize the variance of these correlations and to permit us to combine them across conditions, we converted the values to *Z*-scores[16]. For all cells, we computed Z_p and Z_c—the Z-transforms of r_p and r_c—from the responses to gratings and 120° plaids. For 39 of the 50 cells, we collected full direction-interaction surfaces like the ones shown in **Figure 1c,** and we computed Z_p and Z_c as the mean Z_p and Z_c for 60°, 90°, 120° and 150° plaids. We computed a pattern index as $Z_p - Z_c$. For model-predicted responses, we computed the pattern index from Poisson-simulated spike counts for each trial, with total number of trials equal to the number acquired experimentally. This ensured that any bias in the pattern index resulting from trial-to-trial variability was approximately matched for actual and model-predicted estimates.

To estimate the variability in the pattern index for each cell, we performed a bootstrap analysis[49]. On each iteration of the bootstrap, we estimated the pattern index for a set of trials randomly resampled, with replacement, from the full set of experimental conditions, and equal in number to the full set. For each cell, we performed 100 such bootstrap estimates and took the standard deviation of the estimates as the variability of the pattern index.

To test the cascade model, we fit a simplified two-stage feedforward model to data from individual MT neurons (**Fig. 2a**). The model describes the transformation from a stimulus of superimposed drifting gratings (of varying orientation, but with the same spatial and temporal frequency) into the firing rate of an MT cell. The model presented is based on a previous model of motion processing in MT (the "SH model", ref. 3), but differs in several important ways. First, the SH model is designed to operate on spatiotemporal image intensities, whereas the current model is restricted to inputs that are mixtures of 12 sinusoidal gratings of a fixed spatial and temporal frequency. Second, the SH model includes (untuned) divisive normalization in the MT stage, whereas the present model has only a static nonlinearity. Finally, the SH model includes only untuned normalization in V1.

To characterize the model, we used a modified 'reverse correlation' approach. We stimulated MT neurons with hyperplaids made from the sum of six gratings randomly selected (with replacement) from a pool of 12 gratings, each with a spatial and temporal frequency chosen to match the preferences of the cell and drifting in 1 of 12 directions. The initial phase of each grating was randomized. Each hyperplaid was displayed for 160 ms after which the next stimulus was immediately displayed. Single-frame images of example stimuli are shown on the left of **Figure 3b**.

We parameterize the stimulus with a matrix, $S(\theta_m, t_i)$, whose ith row is a vector whose entries indicate the contrast of each of 12 gratings (at orientations θ_m) that constitute the stimulus displayed during the interval starting at time t_i.

These stimuli were first processed with a population of 12 model V1 cells, each with a direction tuning curve centered on one of the grating directions. The direction tuning curve for the nth model V1 neuron is described by a von Mises function:

$$d_n(\theta_m) = e^{b*\cos(\theta_m - p_n)},$$

where p_n (equal to $30n$ deg) is the direction preference of the nth model V1 neuron and b controls the the direction bandwidth (common to all model V1 neurons). The direction tuning curves are then normalized to unit area:

$$d_n'(\theta_m) = \frac{d_n(\theta_m)}{\sum_k d_n(\theta_k)}$$

The trial-by-trial linear response of each model V1 neuron is computed by taking the inner product of the stimulus and the neuron's direction tuning curve:

$$L_n(t_i) = \sum_m d_n'(\theta_m) S(\theta_m, t_i)$$

This linear V1 response is then subjected to two stages of instantaneous gain control. The linear responses are squared and normalized by the pooled, squared linear responses of all cells[3,28]:

$$P_n(t_i) = \frac{L_n(t_i)^2}{\sum_k L_k(t_i)^2 + \sigma_1^2}$$

We refer to this as the 'untuned' normalization, but this component does in fact have a small 'tuned' component that arises when the fringes of neighboring V1 tuning curves overlap, thereby contributing doubly to the untuned normalization term[15]. In our initial explorations of the model, we found that this untuned normalization stage was not sufficient to account for the plaid responses of strongly pattern-selective cells. Specifically, a model with only this form of normalization did not exhibit stronger response to plaids than gratings, such as seen in **Figure 4c–e**. We therefore incorporated an additional tuned normalization component in the V1 response[15] via a self-normalization stage (equivalent to a static nonlinearity):

$$V_n(t_i) = \frac{P_n(t_i)}{P_n(t_i) + \sigma_2}$$

The tuned and untuned normalization may be combined algebraically to describe V1 responses as resulting from a single stage of weighted normalization with three positive constants (see also **Fig. 5**, column 2):

$$V_n(t_i) = \frac{L_n(t_i)^2}{\alpha_1 L_n(t_i)^2 + \frac{\alpha_2}{12}\sum_k L_k(t_i)^2 + \alpha_3 \bar{L}},$$

where $\bar{L}$ is the mean squared contrast of the hyperplaid stimuli. The response of a model MT cell is based on a linear combination of the responses of the 12 V1 neurons:

$$Q(t_i) = \sum_k w_k V_k(t_i)$$

Finally, the output of this linear stage is transformed into a firing rate via a static nonlinear function: $M(t_i) = f(Q(t_i))$.

Recovering the cascade model for individual cells. In total, the cascade model has 17 parameters: 3 V1 parameters, 12 MT linear weights and 2 parameters for the MT nonlinearity. For each individual cell, we fit these parameters by maximizing the likelihood of the observed responses. Rather than search over all the parameters simultaneously, we performed the optimization in a nested fashion. At top level, we searched over the set of V1 parameters. For each choice of V1 parameters, we optimized the MT linear weights, and for the combined choice of V1 parameters and MT weights, we optimized the MT nonlinearity.

We found it advantageous for fitting purposes to re-parameterize the V1 normalization constants using spherical coordinates:

$$\alpha_1 = \cos^2(\phi_1)\cos^2(\phi_2)$$

$$\alpha_2 = \cos^2(\phi_1)\sin^2(\phi_2)$$

$$\alpha_3 = \sin^2(\phi_1) + \varepsilon$$

where ϕ_1 represents the elevation and ϕ_2 represents the azimuth. The squaring ensures that all three constants are positive, and the small positive constant ε is chosen to set a lower bound of 0.01 on the semisaturation contrast for cells in V1.

For each set of V1 parameters, we solved for the MT linear weights w_k using multivariate linear regression (this is equivalent to reverse-correlation after whitening the V1 afferent inputs to the MT linear stage). Specifically, we minimized the error expression:

$$E(\vec{w}) = \|V\vec{w} - R\|^2 + \lambda\|\vec{w}\|^2$$

Here, we use vector/matrix notation: $\vec{w}$ is a vector containing the MT linear weights w_k, V is a matrix whose rows contain the V1 population output at each time after the mean V1 response is subtracted, and R is a vector containing the recorded MT responses (spike counts) over time, after subtracting the mean spike count. The second term is a regularizer, with λ a small constant chosen to avoid instability that could occur when portions of the input space are not well-covered by the V1 responses (sometimes referred to as "ridge regression"). The optimal linear weights can be computed in closed form using the expression:

$$\vec{w} = (V^T V + \lambda I)^{-1} V^T R$$

where T indicates the transpose, -1 indicates the inverse, I the identity.

For the MT nonlinearity, $f()$, we assumed an exponential form for the nonlinear function $M(t_i) = Ae^{BQ(t_i)}$ where $Q(t_i)$ is the output of the MT linear stage. A model of this form guarantees that the regression solution used for the linear weights is also the solution that maximizes the likelihood[50]. In addition, we found empirically that an exponential nonlinearity provided a good description of the data. The parameters $[A,B]$ of the exponential nonlinearity are chosen to maximize the likelihood of the spiking responses given the MT linear responses $Q(t_i)$.

We used a simplex algorithm (the Matlab function 'fminsearch') to search the space of V1 parameters $[b,\phi_1,\phi_2]$, to minimize the negative log likelihood (NLL) of the observed responses under a Poisson spiking model with rate $M(t_i)$. Specifically, we minimized the expression:

$$NLL(b,\phi_1,\phi_2) = -\sum_i R(t_i)\log(M(t_i)) - M(t_i)$$

Although the error surfaces corresponding to different model parameter combinations were quite smooth, we were concerned that our fits might converge to local rather than global minima. To reduce this possibility, we sampled the error surface for each cell in evenly spaced increments and used the global minimum as the starting point for each search. We also compared this fit with fits started from three other arbitrarily chosen parameter combinations. Fits converged to different solutions for only a few cells; in all cases, starting from the global minimum of the estimated error surface produced the smallest error.

To estimate the variability in the model parameters and predictions resulting from the fit for each cell, we performed a bootstrap analysis similar to that described for the calculation of the pattern index. On each iteration of the bootstrap, we fit the model to a data set resampled, with replacement, from the original dataset, preserving the total number of trials. We used the fit of the intact data set as the initial starting point for the fits on each bootstrap iteration. Finally, we computed the model-predicted pattern index, as described above, from the simulated data arising from the parameters of each fit. We performed 500 such bootstrap estimates for each cell and estimated the variability of each parameter as the standard deviation about the mean.

Note: Supplementary information is available on the Nature Neuroscience website.

ACKNOWLEDGMENTS

We are grateful to M. Carandini and N. Majaj for helpful discussions. This work was supported by the Howard Hughes Medical Institute through an Investigatorship to E.P.S. and by a grant from the National Eye Institute to J.A.M. (EY02017). V.M. was supported by a grant to M. Carandini from the Swiss National Science Foundation.

COMPETING INTERESTS STATEMENT

The authors declare that they have no competing financial interests.

Published online at http://www.nature.com/natureneuroscience
Reprints and permissions information is available online at http://npg.nature.com/reprintsandpermissions/

1. Felleman, D.J. & Van Essen, D.C. Distributed hierarchical processing in the primate cerebral cortex. *Cereb. Cortex* **1**, 1–47 (1991).
2. Movshon, J.A., Adelson, E.H., Gizzi, M.S. & Newsome, W.T. The analysis of moving visual patterns. in *Pattern Recognition Mechanisms (Pontificiae Academiae Scientiarum Scripta Varia)* Vol. 54 (eds. Chagas, C., Gattass, R. & Gross, C.) 117–151 (Vatican Press, Rome, 1985).
3. Simoncelli, E.P. & Heeger, D.J. A model of neuronal responses in visual area MT. *Vision Res.* **38**, 743–761 (1998).
4. Adelson, E.H. & Bergen, J.R. Spatiotemporal energy models for the perception of motion. *J. Opt. Soc. Am. A* **2**, 284–299 (1985).
5. Hubel, D.H. & Wiesel, T.N. Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *J. Physiol. (Lond.)* **160**, 106–154 (1962).
6. DeAngelis, G.C., Ohzawa, I. & Freeman, R.D. Spatiotemporal organization of simple-cell receptive fields in the cat's striate cortex. I. General characteristics and postnatal development. *J. Neurophysiol.* **69**, 1091–1117 (1993).
7. Emerson, R.C., Bergen, J.R. & Adelson, E.H. Directionally selective complex cells and the computation of motion energy in cat visual cortex. *Vision Res.* **32**, 203–218 (1992).
8. Adelson, E.H. & Movshon, J.A. Phenomenal coherence of moving visual patterns. *Nature* **300**, 523–525 (1982).
9. Zeki, S.M. Functional organization of a visual area in the posterior bank of the superior temporal sulcus of the rhesus monkey. *J. Physiol. (Lond.)* **236**, 549–573 (1974).
10. Movshon, J.A. & Newsome, W.T. Visual response properties of striate cortical neurons projecting to area MT in macaque monkeys. *J. Neurosci.* **16**, 7733–7741 (1996).
11. Carandini, M. *et al.* Do we know what the early visual system does? *J. Neurosci.* **25**, 10577–10597 (2005).
12. Movshon, J.A., Thompson, I.D. & Tolhurst, D.J. Spatial summation in the receptive fields of simple cells in the cat's striate cortex. *J. Physiol. (Lond.)* **283**, 53–77 (1978).
13. Movshon, J.A., Thompson, I.D. & Tolhurst, D.J. Receptive field organization of complex cells in the cat's striate cortex. *J. Physiol. (Lond.)* **283**, 79–99 (1978).
14. Jones, J.P., Stepnoski, A. & Palmer, L.A. The two-dimensional spectral structure of simple receptive fields in cat striate cortex. *J. Neurophysiol.* **58**, 1212–1232 (1987).
15. Mante, V. *Testing Models of Cortical Area MT.* Thesis, Institute of Neuroinformatics, ETH, University of Zurich (2000).
16. Smith, M.A., Majaj, N.J. & Movshon, J.A. Dynamics of motion signaling by neurons in macaque area MT. *Nat. Neurosci.* **8**, 220–228 (2005).
17. Carandini, M., Heeger, D.J. & Movshon, J.A. Linearity and normalization in simple cells of the macaque primary visual cortex. *J. Neurosci.* **17**, 8621–8644 (1997).
18. Lennie, P. & Movshon, J.A. Coding of color and form in the geniculostriate visual pathway. *J. Opt. Soc. Am. A Opt. Image Sci. Vis.* **22**, 2013–2033 (2005).
19. Ringach, D.L., Sapiro, G. & Shapley, R. A subspace reverse-correlation technique for the study of visual neurons. *Vision Res.* **37**, 2455–2464 (1997).
20. Theunissen, F.E., Sen, K. & Doupe, A.J. Spectral-temporal receptive fields of nonlinear auditory neurons obtained using natural sounds. *J. Neurosci.* **20**, 2315–2331 (2000).
21. David, S.V., Vinje, W.E. & Gallant, J.L. Natural stimulus statistics alter the receptive field structure of V1 neurons. *J. Neurosci.* **24**, 6991–7006 (2004).
22. Maunsell, J.H. & van Essen, D.C. The connections of the middle temporal visual area (MT) and their relationship to a cortical hierarchy in the macaque monkey. *J. Neurosci.* **3**, 2563–2586 (1983).
23. Sincich, L.C., Park, K.F., Wohlgemuth, M.J. & Horton, J.C. Bypassing V1: a direct geniculate input to area MT. *Nat. Neurosci.* **7**, 1123–1128 (2004).
24. Levitt, J.B., Kiper, D.C. & Movshon, J.A. Receptive fields and functional architecture of macaque V2. *J. Neurophysiol.* **71**, 2517–2542 (1994).
25. Gegenfurtner, K.R., Kiper, D.C. & Levitt, J.B. Functional properties of neurons in macaque area V3. *J. Neurophysiol.* **77**, 1906–1923 (1997).
26. Rodman, H.R., Gross, C.G. & Albright, T.D. Afferent basis of visual response properties in area MT of the macaque. I. Effects of striate cortex removal. *J. Neurosci.* **9**, 2033–2050 (1989).
27. Collins, C.E., Lyon, D.C. & Kaas, J.H. Responses of neurons in the middle temporal visual area after long-standing lesions of the primary visual cortex in adult new world monkeys. *J. Neurosci.* **23**, 2251–2264 (2003).
28. Heeger, D.J. Normalization of cell responses in cat striate cortex. *Vis. Neurosci.* **9**, 181–197 (1992).
29. Geisler, W.S. & Albrecht, D.G. Cortical neurons: isolation of contrast gain control. *Vision Res.* **32**, 1409–1410 (1992).
30. Albrecht, D.G. & Hamilton, D.B. Striate cortex of monkey and cat: contrast response function. *J. Neurophysiol.* **48**, 217–237 (1982).
31. Cavanaugh, J.R., Bair, W. & Movshon, J.A. Selectivity and spatial distribution of signals from the receptive field surround in macaque V1 neurons. *J. Neurophysiol.* **88**, 2547–2556 (2002).
32. Cavanaugh, J.R., Bair, W. & Movshon, J.A. Nature and interaction of signals from the receptive field center and surround in macaque V1 neurons. *J. Neurophysiol.* **88**, 2530–2546 (2002).
33. Sceniak, M.P., Hawken, M.J. & Shapley, R. Visual spatial characterization of macaque V1 neurons. *J. Neurophysiol.* **85**, 1873–1887 (2001).

34. Livingstone, M.S., Pack, C.C. & Born, R.T. Two-dimensional substructure of MT receptive fields. *Neuron* **30**, 781–793 (2001).
35. Pack, C.C., Livingstone, M.S., Duffy, K.R. & Born, R.T. End-stopping and the aperture problem: two-dimensional motion signals in macaque V1. *Neuron* **39**, 671–680 (2003).
36. Pack, C.C., Gartland, A.J. & Born, R.T. Integration of contour and terminator signals in visual area MT of alert macaque. *J. Neurosci.* **24**, 3268–3280 (2004).
37. van den Berg, A.V. & Noest, A.J. Motion transparency and coherence in plaids: the role of end-stopped cells. *Exp. Brain Res.* **96**, 519–533 (1993).
38. Tinsley, C.J. *et al.* The nature of V1 neural responses to 2D moving patterns depends on receptive-field structure in the marmoset monkey. *J. Neurophysiol.* **90**, 930–937 (2003).
39. Rockland, K.S. Bistratified distribution of terminal arbors of individual axons projecting from area V1 to middle temporal area (MT) in the macaque monkey. *Vis. Neurosci.* **3**, 155–170 (1989).
40. Kohn, A. & Movshon, J.A. Neuronal adaptation to visual motion in area MT of the macaque. *Neuron* **39**, 681–691 (2003).
41. Bair, W. & Movshon, J.A. Adaptive temporal integration of motion in direction-selective neurons in macaque visual cortex. *J. Neurosci.* **24**, 7305–7323 (2004).
42. Sclar, G., Maunsell, J.H. & Lennie, P. Coding of image contrast in central visual pathways of the macaque monkey. *Vision Res.* **30**, 1–10 (1990).
43. Britten, K.H. & Heuer, H.W. Spatial summation in the receptive fields of MT neurons. *J. Neurosci.* **19**, 5074–5084 (1999).
44. Rodman, H.R. & Albright, T.D. Coding of visual stimulus velocity in area MT of the macaque. *Vision Res.* **27**, 2035–2048 (1987).
45. Rodman, H.R. & Albright, T.D. Single-unit analysis of pattern-motion selective properties in the middle temporal visual area (MT). *Exp. Brain Res.* **75**, 53–64 (1989).
46. Chance, F.S., Nelson, S.B. & Abbott, L.F. Complex cells as cortically amplified simple cells. *Nat. Neurosci.* **2**, 277–282 (1999).
47. Priebe, N.J., Mechler, F., Carandini, M. & Ferster, D. The contribution of spike threshold to the dichotomy of cortical simple and complex cells. *Nat. Neurosci.* **7**, 1113–1122 (2004).
48. Rust, N.C., Schwartz, O., Movshon, J.A. & Simoncelli, E.P. Spatiotemporal elements of macaque V1 receptive fields. *Neuron* **46**, 945–956 (2005).
49. Efron, B. & Tibshirani, R.J. *An Introduction to the Boostrap* (Chapman & Hall/CRC Press, Boca Raton, Florida, 1993).
50. Paninski, L. Maximum likelihood estimation of cascade point-process neural encoding models. *Network* **15**, 243–262 (2004).