Behavioral/Systems/Cognitive

Disparity-Based Coding of Three-Dimensional Surface Orientation by Macaque Middle Temporal Neurons

Jerry D. Nguyenkim and Gregory C. DeAngelis

Department of Anatomy and Neurobiology, Washington University School of Medicine, St. Louis, Missouri 63110

Gradients of binocular disparity across the visual field provide a potent cue to the three-dimensional (3-D) orientation of surfaces in a scene. Neurons selective for 3-D surface orientation defined by disparity gradients have recently been described in parietal cortex, but little is known about where and how this selectivity arises within the visual pathways. Because the middle temporal area (MT) has previously been implicated in depth perception, we tested whether MT neurons could signal the 3-D orientation (as parameterized by tilt and slant) of planar surfaces that were depicted by random-dot stereograms containing a linear gradient of horizontal disparities. We find that many MT neurons are tuned for 3-D surface orientation, and that tilt and slant generally have independent effects on MT responses. This separable coding of tilt and slant is reminiscent of the joint coding of variables in other areas (e.g., orientation and spatial frequency in V1). We show that tilt tuning remains unchanged when all coherent motion is removed from the visual stimuli, indicating that tilt selectivity is not a byproduct of 3-D velocity coding. Moreover, tilt tuning is typically insensitive to changes in the mean disparity (depth) of gradient stimuli, indicating that tilt tuning cannot be explained by conventional tuning for frontoparallel disparities. Finally, we explore the receptive field mechanisms underlying selectivity for 3-D surface orientation, and we show that tilt tuning arises through heterogeneous disparity tuning within the receptive fields of MT neurons. Our findings show that MT neurons carry high-level signals about 3-D surface structure, in addition to coding retinal image velocities.

Key words: visual cortex; extrastriate; stereopsis; binocular disparity; surface; tilt; slant

Introduction

A typical visual environment contains a variety of surfaces at different three-dimensional (3-D) orientations relative to one's line of sight. For planar surfaces, the 3-D orientation can be described in terms of "tilt" and "slant" (see Fig. 1A). Accurate information about 3-D surface orientation is important for visual navigation and object manipulation, as well as for object recognition itself. Many visual cues can be used to judge the tilt and slant of surfaces, including texture gradients, velocity gradients, shading, perspective, and binocular disparity gradients (Sedgwick, 1986; Howard and Rogers, 2002). Disparity gradients are quantitatively related to 3-D surface orientation, knowing only the positions of the two eyes, whereas interpretation of other cues requires additional knowledge about object structure, observer motion, lighting, etc. Thus, disparity gradients provide robust information about 3-D surface orientation.

Recent physiological studies have described neurons in parietal and temporal cortex that are sensitive to 3-D structure defined by disparity gradients. Taira et al. (2000) and Tsutsui et al. (2002) have reported that neurons in the caudal intraparietal (CIP) area signal surface tilt defined by disparity gradients. Meanwhile, Janssen et al. (1999, 2000, 2001) have shown that inferotemporal neurons signal 3-D shape via disparity gradients. Although these studies establish the presence of disparity gradient signals at the upper levels of the dorsal and ventral processing streams, the origins and mechanisms of gradient selectivity in visual cortex remain unknown.

Computational and psychophysical studies (Gibson, 1950; Marr, 1982; Nakayama, 1996) suggest that 3-D surface structure should be computed early in the visual pathways. It seems unlikely that disparity gradients could be effectively coded in areas V1 or V2 because of the small size of receptive fields in these areas. We reasoned that the middle temporal area (MT) might participate in gradient computations because the receptive fields of MT neurons are several-fold larger than those of their primary inputs from V1/V2 (Albright and Desimone, 1987; Maunsell and Van Essen, 1987). In addition, recent studies have shown that MT contains strong disparity signals (Maunsell and Van Essen, 1983a; DeAngelis and Uka, 2003), that disparity-selective MT neurons are organized topographically (DeAngelis and Newsome, 1999), and that electrical stimulation of MT influences depth perception (DeAngelis et al., 1998). We therefore tested whether MT neurons signal the tilt and slant of 3-D surfaces defined solely by disparity gradients.

Two critical factors to control in these experiments are vergence eye movements and stimulus centering on the receptive field. Systematic changes in vergence angle with tilt or slant could give rise to artifactual tuning for surface orientation. Similarly,

improper centering of the gradient stimulus on the receptive field can give a false impression of tilt selectivity unless tilt tuning is shown to be invariant to changes in the mean disparity of the gradient stimulus. These factors have not been rigorously controlled in previous studies (Janssen et al., 1999; Taira et al., 2000), whereas our experiments and analyses were designed specifically to account for them.

We show that many MT neurons exhibit robust tuning for the tilt and slant of disparity-defined surfaces. Our results complement previous studies of the responses of MT neurons to speed gradients (Treue and Andersen, 1996; Xiao et al., 1997) and suggest that MT neurons use multiple cues for computing 3-D surface orientation. These findings offer additional evidence that area MT plays important roles in 3-D vision.

Materials and Methods

Two male rhesus monkeys (*Macaca mulatta*), weighing between 5 and 7 kg, performed a standard fixation task during extracellular recording experiments. A detailed description of our methods has recently appeared (DeAngelis and Uka, 2003); here, we briefly review these procedures, focusing on those aspects most relevant to the present study. All experimental procedures were approved by the Institutional Animal Care and Use Committee at Washington University and conformed to National Institutes of Health guidelines.

Visual stimuli. Stereoscopic visual stimuli were presented using frame alternation (at 100 Hz) on a 22 inch flat-face monitor that subtended $40 \times 30°$ at the viewing distance of 57 cm. Random-dot stereograms were generated by an OpenGL accelerator board (3 Dlabs Oxygen GVX1) and were viewed by the monkey through ferroelectric liquid crystal shutters that were synchronized to the monitor refresh. Stereo crosstalk was ~3%. Some of the later experiments (including monocular controls) (see Fig. 5) were performed using a stereoscopic projector (Christie Digital Mirage 2000; image subtense: $56 \times 46°$) that had no measurable stereo crosstalk; similar results were obtained using both display devices.

Stereograms consisted of red dots (~0.1° diameter) presented on a black background. Dot density was generally 64 dots per square degree per second, and dots were presented within a circular aperture. Precise disparities and smooth motion were achieved by plotting dots with subpixel resolution using hardware anti-aliasing under OpenGL. Except where noted in the text, dots moved coherently at the preferred direction and speed of each MT neuron and wrapped around when they reached the edge of the aperture.

In these experiments, 3-D surface orientation was varied by applying linear gradients of horizontal disparities to the random-dot stereograms. It is important to note that the disparity gradient was the only useful cue to surface orientation in these stimuli: there were no corresponding speed or texture gradients in the stimulus as would typically occur for a real slanted surface in a natural scene. Note, however, that application of a disparity gradient does produce very subtle variations in dot density along the axis of the gradient. We thus performed monocular controls (described below) (see Fig. 5) to exclude the possibility that these subtle monocular density cues account for tilt tuning.

Task and data collection. Monkeys were required to maintain their conjugate eye position within a 1.5° diameter fixation window that was centered on the fixation point. Fixation began 300 msec before presentation of the random-dot stereogram and had to be maintained throughout the 1.5 sec stimulus presentation to receive a liquid reward. Only data from successfully completed trials were analyzed. Movements of both
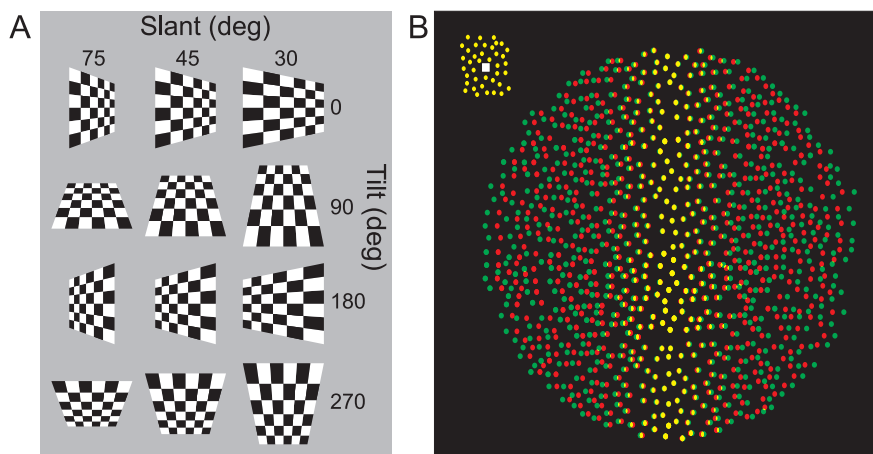
Figure 1. Schematic illustration of the 3-D orientation of planar surfaces, parameterized by tilt and slant. *A*, Tilt refers to the axis around which the plane is rotated away from frontoparallel, and slant defines the amount by which the plane is rotated (Gibson, 1950; Stevens, 1983). Zero slant corresponds to a frontoparallel surface for which the tilt is undefined. In this illustration, tilt and slant are defined by perspective and texture gradient cues. In our experiments, surface orientation was defined solely by the direction and magnitude of a linear gradient of horizontal disparity. *B*, An example of a 0 degree tilt stimulus depicted using a red–green anaglyph. Dots in our experiments were all red, and stereoscopic presentation was accomplished using ferroelectric liquid crystal shutters synchronized to the monitor refresh.

eyes were measured in all experiments using eye coils that were sutured to the sclera; eye position signals were stored to disk at 250 Hz.

Tungsten microelectrodes were introduced into the cortex through a transdural guide tube, and area MT was recognized based on the following criteria: the pattern of gray and white matter transitions along electrode penetrations, the response properties of single units and multiunit clusters (direction, speed, and disparity tuning), retinal topography, the relationship between receptive field size and eccentricity, and the subsequent entry into gray matter with response properties typical of the medial superior temporal area. All data included in this study were taken from portions of electrode penetrations that were confidently assigned to area MT. Raw neural signals were amplified and bandpass filtered (500–5000 Hz) using conventional electronic equipment. Action potentials of single MT units were isolated using a dual voltage–time window discriminator (Bak Electronics) and time-stamped with 1 msec resolution.

Experimental protocol. The receptive field (RF) of each isolated MT neuron was initially explored using a mapping program to carefully estimate the RF location and size, preferred velocity, and preferred disparity. We subsequently performed the following series of quantitative tests on each MT neuron (each condition below represents a separate block of trials). (1) A direction-tuning curve was obtained by presenting moving random-dot patterns at eight directions of motion, 45° apart. (2) A speed-tuning curve was obtained by presenting random-dot patterns at speeds of 0, 0.5, 1, 2, 4, 8, 16, and 32°/sec, with direction fixed to the optimal value. (3) Horizontal disparity tuning was measured by presenting moving random dots at nine disparities typically ranging from −1.6 to 1.6° in steps of 0.4°. These parameters were adjusted as necessary based on the initial RF exploration. (4) The receptive field was mapped quantitatively by presenting small (<0.25 × RF size) rectangular patches of moving dots at 16 locations on a 4 × 4 grid that covered the receptive field. A two-dimensional Gaussian was fit to this RF map to determine the center location of the receptive field. (5) A size-tuning curve was obtained by presenting moving random dots in circular apertures having sizes of 0, 1, 2, 4, 8, 16, and 32°. Results of this test were used to quantify the extent (percent) of surround inhibition exhibited by each neuron (DeAngelis and Uka, 2003). (6) Tilt tuning was assessed by presenting stereograms containing a linear gradient of horizontal disparities across the circular aperture (Fig. 1*B*). The stimuli depicted surfaces at eight tilt angles, 45 degrees apart (see Fig. 1*A* for convention). Each tilt angle was presented at three to five different mean disparities that typically flanked the peak in the disparity tuning curve of the neuron (see Fig. 3*A,B,D*). The magnitude of the disparity gradient was 0.15°/° for most experi-

ments, corresponding to a surface slanted ~67 degrees away from fronto-parallel. We chose a steep slant to maximize our chances of observing tilt tuning in this test (similar to Xiao et al., 1997).

Stimulus size for the tilt-tuning measurements was chosen based on the results of the receptive-field mapping and size-tuning experiments. For neurons that did not show any surround inhibition in the size-tuning test, stimulus size was chosen to encompass the entire classical receptive field (including the weakest flanks) as mapped using the 4×4 grid described above. For neurons with clear surround inhibition, stimulus size was chosen to be two or three times larger than the stimulus that elicited a maximal response, so that the stimulus encompassed a large portion of the nonclassical inhibitory surround. In some cases of exceptionally strong surround inhibition, however, a stimulus two or three times the optimal size elicited little or no response from the neuron. In these instances, stimulus size was reduced until the neuron gave an approximately half-maximal response. Because we found no overall correlation between tilt selectivity and the strength of surround inhibition (see Fig. 10), our population analyses were done by combining data across neurons regardless of the presence of surround inhibition.

Because the relationship between disparity and depth is nonlinear, our linear disparity gradients depict surfaces that are not exactly planar (although this departure is generally not evident to human observers). Slant is not constant across space when the stimulus is large, and slant also varies a bit with the mean disparity of the gradient stimulus. However, given that tilt tuning was observed across a variety of stimulus sizes and that tilt tuning is generally invariant to changes in both mean disparity and slant, the subtle deviations from planarity in our stimuli cannot explain our results.

The above set of tests was performed on all 97 neurons included in the present study. For some neurons, we also performed one or more of the following additional tests. (1) The interaction between tilt and slant was examined for 29 neurons by presenting eight tilts at each of five to seven different slants chosen from the following set of gradient magnitudes: 0.001, 0.002, 0.01, 0.02, 0.05, 0.1, 0.15, 0.2, or 0.25°/°. These correspond to slants of 0.75, 1.5, 7, 15, 35, 54, 67, 73, and 76 degrees, measured at the center of a stimulus with zero mean disparity. (2) The effect of removing coherent motion from the stimulus was assessed by testing 10 neurons with stereograms in which dots were either stationary or randomly re-plotted every fourth video frame (0% motion coherence). (3) For 15 neurons, monocular tilt-tuning controls were obtained by turning off the dots presented to either the left or right eye while the image to the other eye was presented intact. (4) For some neurons with strong tilt selectivity, we probed the 3-D substructure of the receptive field by presenting pairs of circular patches of random dots. One member of the pair was always centered on the classical RF, and the other member of the pair was chosen from six locations surrounding the center stimulus (see Fig. 8C). The disparity of the center patch was held fixed at the optimal value, whereas the disparity of the surrounding patches varied from −2 to 2° in steps of 0.5°. This allowed us to measure a disparity-tuning curve for each of the six surrounding locations (see Fig. 8D). For neurons without surround inhibition, the entire array of seven patches was presented within the classical RF, such that the center patch was approximately one-third the size of the RF. When surround inhibition was present, the center patch was set to the optimal size (from the size-tuning curve), and the six surrounding patches extended into the inhibitory surround. Thus, our experiments probed for heterogeneous disparity tuning within either the classical RF or the nonclassical inhibitory surround (when present). In most cases, the center patch had the same dimensions as each of the six surrounding patches, but sometimes the size of the center patch was reduced to enhance the response modulations produced by varying the disparities of the six surrounding patches.

Data analysis. The response to each stimulus presentation was quantified as the average firing rate over the 1.5 sec stimulus period. Each different stimulus was typically presented five times in blocks of randomly interleaved trials. Tuning curves were constructed by plotting the mean ± SE of the response across repetitions of each different stimulus.

Each tilt-tuning curve was fit with a modified sinusoid having the following form:

$$R(\theta) = A \times G(\sin(f\theta + \psi)) + R_o, \tag{1}$$

where

$$G(x) = \frac{e^{nx} - 1}{n}. \tag{2}$$

θ denotes the tilt angle, and A, f, ψ, R_o, and n are free parameters. $G(x)$ is an exponential function that can distort the sinusoid such that the peak is taller than the trough or vice versa. We found that this distortion of the sinusoid was necessary to fit the tilt-tuning curves of some MT neurons (see Figs. 8E, 9A). The best fit of this function to the data was achieved by minimizing the sum squared error between the responses of the neuron and the values of the function, using the constrained minimization tool, "lsqcurvefit", in Matlab (Mathworks). To homogenize the variance of the neural responses across different stimulus values, we minimized the difference between the square root of the neural responses and the square root of the function (Prince et al., 2002). Curve fits were generally quite good, accounting for 85% (median across all neurons) of the variance in MT responses. Additional details about our fitting procedures are described elsewhere (DeAngelis and Uka, 2003).

The frequency, f, of the modified sinusoid was constrained to lie within a range from 0.4 to 1.6. Although most of the fitted values of f were very close to unity, the fits for a minority of neurons were significantly improved when the frequency was allowed to differ from unity. This could present a problem if we were using the phase parameter, ψ, of the fits to characterize the stimulus preference. However, tilt preferences were always computed by finding the actual peak of the modified sinusoid, such that there is no difficulty associated with frequencies that depart somewhat from unity.

To test if tilt tuning was sensitive to changes in slant or mean disparity (depth), we analyzed the data using two different models. In the first model, we fit the tilt-tuning curve for each different slant or mean disparity with an independent sinusoid given by Equation 1. We then computed the total sum-squared error of the independent fits. In the second model, we fit all tilt-tuning curves simultaneously while forcing the phase (ψ) and frequency (f) parameters of the sinusoids to be shared (constrained fits). The remaining parameters had independent values for each curve. This second model constrains the fitted curves to have identical peak and trough locations (i.e., a constant preferred tilt) while allowing them to have different amplitudes and mean responses. We then compared the total error of the constrained fits to that of the independent fits using a sequential F test (Draper and Smith, 1966), with a significance criterion of $p < 0.05$. If the difference between models is insignificant ($p > 0.05$), we can conclude that the tilt preference is invariant to changes in slant or mean disparity.

Results

We recorded from 203 neurons in two alert rhesus monkeys that performed a standard fixation task. There were no intentional selection criteria for sampling neurons, so the sample should be unbiased. We isolated 97/203 neurons long enough to obtain a complete set of data, which required the monkey to execute at least 486 correct trials (see Materials and Methods).

Figure 2 shows data for an exemplar neuron. This MT unit preferred far (uncrossed) disparities (Fig. 2A) and exhibited powerful surround inhibition when the diameter of the stimulus aperture was increased beyond a few degrees of visual angle (Fig. 2B). After mapping the RF quantitatively (Fig. 2C), we centered a 6° stimulus aperture (dashed circle) over the receptive field. This size was chosen to cover most of the excitatory RF without eliciting too much surround inhibition. In this aperture, we presented stereograms that simulated planar surfaces at eight tilt angles (45 degrees apart) relative to the line of sight; the simulated slant angle was 70 degrees. Figure 2D shows neuronal response
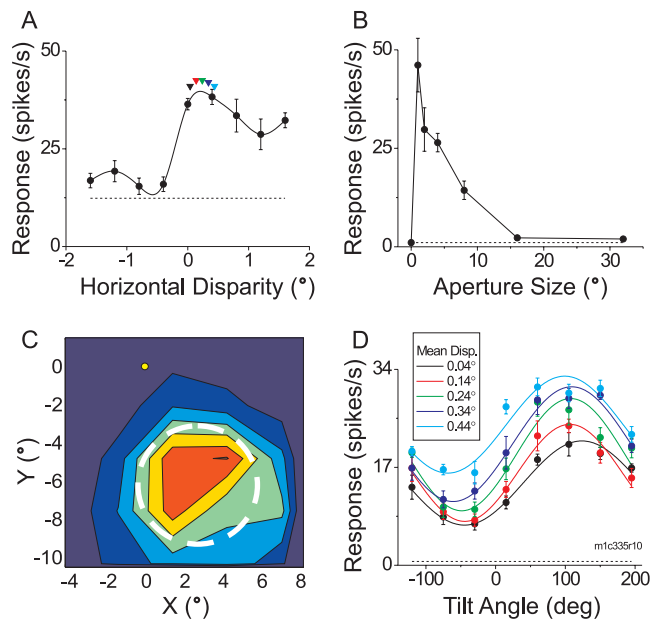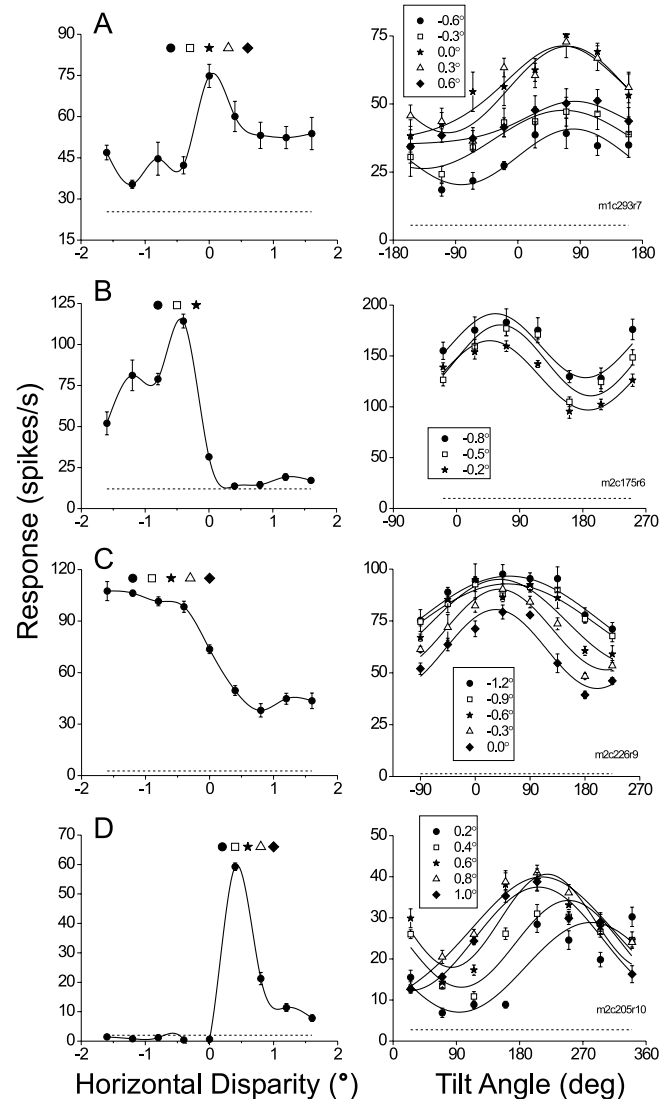
Figure 2. A dataset for an example MT neuron that exhibits tilt selectivity. *A*, A conventional disparity-tuning curve measured using random-dot stereograms (i.e., slant was zero, and different uniform horizontal disparities were applied). Mean responses ± SE are shown for each different stimulus disparity, along with a spline fit. Colored arrowheads indicate the five mean disparities used for the disparity gradients in *D*. *B*, A size-tuning (area summation) curve. A frontoparallel (zero slant) surface was presented at the preferred disparity, and the diameter of the stimulus aperture varied. The response of this neuron was abolished at large sizes, indicating the presence of powerful (96%) surround inhibition. *C*, A quantitative receptive-field map was measured by presenting small (1.3 × 1.3°) patches of dots at 16 spatial locations on a 4 × 4 grid. Response strength is color-coded, from low (dark blue) to high (red); peak response was 45 spikes/sec. The dashed white circle shows the location and size of the stimulus aperture in which disparity gradient stimuli were presented. *D*, Tilt-tuning curves at five different mean disparities (color-coded). Smooth curves indicate the best fits of the modified sinusoid (Eqs. 1, 2). Stimulus parameters were as follows: direction of motion, 105° (convention: rightward, 0°; upward, 90°); speed of motion, 17°/sec; aperture diameter, 6°; eccentricity, 6.8°; and gradient magnitude, 0.2°/°.

Figure 3. Horizontal disparity-tuning curves (left) and tilt-tuning curves (right) for four additional MT neurons. The format is similar to that of Figure 2, *A* and *D*, except that different mean disparities of the gradient stimulus are denoted here by different symbol types. Stimulus parameters were as follows: *A*, direction of motion, 70°; speed of motion, 1°/sec; aperture diameter, 4°; eccentricity, 5.5°; surround inhibition, 29%. *B*, Direction of motion, 161°; speed of motion, 1.5°/sec; aperture diameter, 27°; eccentricity, 10°; surround inhibition, 0%. *C*, Direction of motion, 135°; speed of motion, 12°/sec; aperture diameter, 27°; eccentricity, 15°; surround inhibition, 17%. *D*, Direction of motion, 250°; speed of motion, 8°/sec; aperture diameter, 24°; eccentricity, 11°; surround inhibition, 0%.

plotted as a function of tilt angle, with each curve corresponding to a different mean disparity of the gradient, ranging from 0.04 to 0.44°. Smooth curves are the best fits of a modified sinusoid (Eqs. 1, 2). Note that the response of the neuron is well tuned for surface tilt and that the shape of the tilt-tuning curves varies little over the range of mean disparities tested.

For a slanted plane viewed through a fixed aperture, moving the surface in depth is equivalent to shifting it within a frontoparallel plane. For this example neuron, the range of mean disparities (i.e., depths) that we tested is equivalent to shifting the center of the gradient over a range of 2° relative to the center of the RF. This allows for a considerable amount of error in centering the stimulus on the receptive field.

Figure 3 shows data from four additional MT neurons that were tested across broader ranges of mean disparities. For the neurons in Figure 3, *A* and *B*, mean disparities were chosen to straddle the peak in the disparity-tuning curve (left panels). If tilt tuning were an artifact of mis-centering the stimulus over the receptive field, then the tilt-tuning curve should undergo a phase shift of ~180 degrees for mean disparities on opposite sides of the peak. Clearly, this is not the case for either of these neurons: the shape of the tilt-tuning curve is consistent across mean disparities, although the amplitude and baseline levels of the curves vary somewhat. A similar result is seen in Figure 3*C* for a neuron that was broadly tuned to near (crossed) disparities. These neurons

provide consistent signals about 3-D surface orientation across a large range of depths.

Figure 3*D* shows data that is characteristic of other neurons that we recorded (see also Fig. 8*E*). This neuron exhibits strong tilt selectivity, but the tilt-tuning curve shifts horizontally with changes in mean disparity. Although tilt preference is not invariant to changes in mean disparity, the effect is much more subtle than the 180 degrees phase shift that one would expect to see if tilt tuning were the result of poorly centering the stimulus over the receptive field of a non-tilt-selective neuron. Thus, neurons like those in Figures 3*D* and 8*E* can still provide useful signals about surface orientation. Many other MT neurons had no tilt selectivity at all (quantified below), and presumably cannot contribute to discrimination of surface orientation.

Population analyses

To quantify the strength of tilt tuning, we equated the average response of an MT neuron to all mean disparities by vertically shifting the individual tilt-tuning curves. We then combined the data across mean disparities to create a single "grand" tilt-tuning curve. Note that this allows tilt tuning to cancel across mean disparities when the preferred tilts differ by close to 180 degrees. Thus, neurons with inconsistent tilt preferences across mean disparities will have weak tuning in the grand curve. For each neuron, we computed two metrics from this grand curve: a modulation index and a discrimination index:

$$\text{Tilt modulation index} = \frac{R_{max} - R_{min}}{R_{max} - S} \qquad (3)$$

$$\text{Tilt discrimination index} = \frac{R_{max} - R_{min}}{R_{max} - R_{min} + 2\sqrt{SSE/(N - M)}}. \qquad (4)$$

R_{max} and R_{min} denote the mean firing rates of the neuron at the tilt angles that elicited maximal and minimal responses, respectively. S denotes spontaneous activity. SSE is the sum-squared error around the mean responses, N is the total number of observations (trials), and M is the number of distinct tilt values. Note that the denominator of the discrimination index incorporates a metric of response variability, whereas the modulation index does not. We present both metrics because they provide complementary information (Prince et al., 2002; DeAngelis and Uka, 2003).

Figure 4A shows a scatter plot of the discrimination and modulation indices for all 97 neurons in our sample, with marginal distributions along the edges of the plot. Filled symbols denote neurons for which response depended significantly on tilt ($p <$ 0.05), as assessed using a two-way ANOVA with tilt angle and mean disparity as factors. By this criterion, 72% (70/97) of MT neurons are significantly tuned for surface tilt. It should be noted, however, that tilt tuning in MT is generally much weaker than either direction or disparity selectivity. The mean modulation/discrimination indices for tilt (0.29/0.42) in our sample are significantly smaller than the mean modulation/discrimination indices for both direction (0.98/0.78) and disparity (0.81/0.71) (paired t test, $p \ll 0.0001$ for all comparisons). Some of this difference may be attributable to the fact that the slant was not optimized for each MT neuron and that tilt-tuning curves were combined across mean disparities, but we expect these factors to account for only a small portion of the weaker tuning to surface orientation. By varying only the disparity gradient in our stimuli, we have placed this cue to surface orientation in conflict with other cues such as texture and velocity gradients. Thus, it is also possible that tilt tuning is muted in our experiments by this cue conflict, a possibility that we cannot address at this time. In our present data set, many MT neurons exhibit clear tilt tuning, but this property is much less prominent than either direction or disparity tuning.

To quantify the consistency of tilt tuning across different mean disparities, we computed the magnitude of the difference in preferred tilt, $|\Delta\text{Pref. Tilt}|$, between all unique pairings of mean disparities for which there was significant tilt tuning (ANOVA, $p < 0.05$). For this analysis, preferred tilts were determined from the peaks of the independent sinusoid fits. Figure 4B shows the $|\Delta\text{Pref. Tilt}|$ values for each neuron plotted as a function of the tilt discrimination index (TDI). Most neurons contribute multiple points to this plot (aligned vertically), and the largest value of
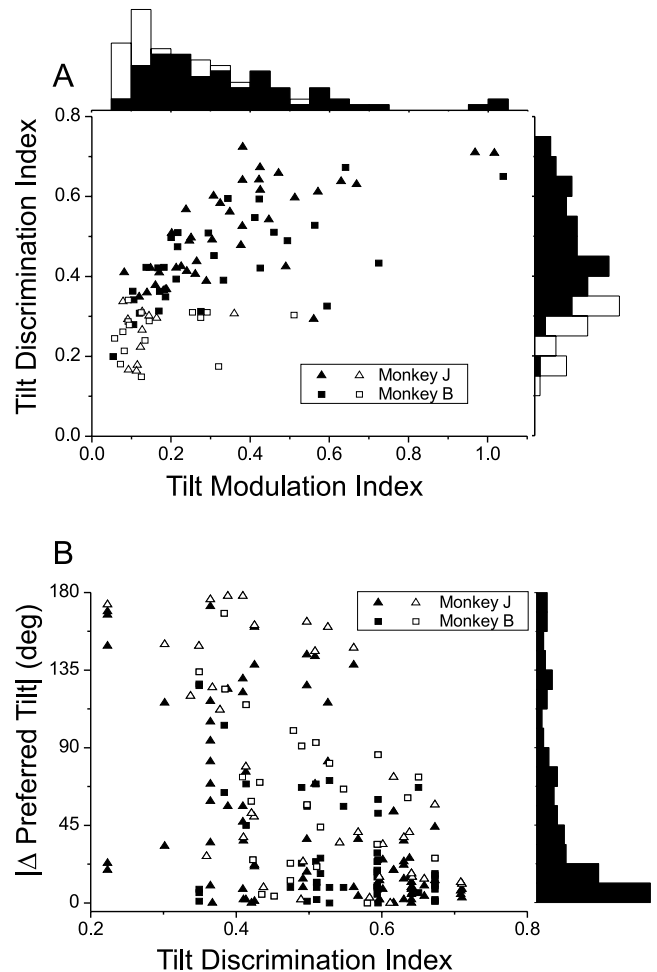
Figure 4. Population summary of tilt selectivity. *A,* Summary of tuning strength. The tilt discrimination index (Eq. 4) is plotted against the tilt modulation index (Eq. 3) for 97 MT neurons. Squares and triangles indicate data from monkeys B and J, respectively. Filled symbols indicate neurons for which the main effect of tilt was significant (two-way ANOVA; p < 0.05). Histograms along the margins of the scatter plot give distributions of the discrimination and modulation indices. *B,* Summary of the consistency of preferred tilts across mean disparities. For each neuron, we computed the absolute value of the difference in tilt preference, $|\Delta\text{Pref. Tilt}|$, between all unique pairs of mean disparities for which there was significant tilt tuning (ANOVA, p < 0.05). The scatter plot shows 219 $|\Delta\text{Pref. Tilt}|$ values (from 64 neurons) plotted against the tilt discrimination index for each neuron. There are up to 10 data points for each neuron, aligned vertically. Open symbols show the largest value of $|\Delta\text{Pref. Tilt}|$ for each neuron. The histogram (right) shows the marginal distribution of $|\Delta\text{Pref. Tilt}|$.

$|\Delta\text{Pref. Tilt}|$ for each neuron is indicated by an open symbol. For neurons with large values of TDI, $|\Delta\text{Pref. Tilt}|$ values are generally less than our sampling interval of 45°, indicating that tilt tuning was quite consistent across mean disparities. Correspondingly, the largest $|\Delta\text{Pref. Tilt}|$ value for these well tuned neurons is also quite small. Overall, the marginal distribution in Figure 4B shows that 62% (135/219) of all data points correspond to $|\Delta\text{Pref. Tilt}|$ values smaller than 45°. However, some neurons with low values of TDI exhibited large differences between preferred tilts at different mean disparities. The presence of $|\Delta\text{Pref. Tilt}|$ values near 180° suggests that some of these neurons exhibit tilt tuning (at individual mean disparities) that is an artifact of mis-centering the visual stimulus over the receptive field. This highlights the importance of analyzing responses to multiple mean disparities straddling the peak of the disparity-tuning curve.

To determine if tilt preference was truly invariant to changes

in mean disparity, we fit the data from each neuron with two models (see Materials and Methods): one in which the tilt preference (determined by the phase and frequency of the fitted sinusoid) was allowed to vary with mean disparity, and one in which the tilt preference was constrained to be identical across mean disparities. For 25/64 neurons in Figure 4 *B*, there was no significant difference between these two models (sequential *F* test, $p > 0.05$), indicating that tilt preference was invariant to changes in mean disparity over the range tested. For many of these invariant neurons, the range of disparities tested was at least 0.8° and included disparities on both sides of the preferred disparity. We thus conclude that a substantial fraction of MT neurons code surface orientation in a depth-invariant manner.

Tilt tuning in MT cannot be explained as an artifact of vergence eye movements. We measured the mean vergence angle of the monkey for each trial and subjected these vergence data to the same two-way ANOVA as the firing rates. Vergence angle showed a significant dependence on tilt for only 12% (12/97) of neurons. Moreover, when vergence angle was added as a covariate to the analysis of firing rates, the significance of the main effect of tilt on firing rate was unchanged for all but one of our units.

Similarly, tilt tuning does not arise from the subtle monocular dot-density cues that accompany a linear disparity gradient (see Materials and Methods). To exclude this possibility, 15 neurons were tested with left- and right-eye half-images presented separately. If tilt tuning resulted from monocular dot-density cues, then tilt selectivity should still be observed in these monocular controls. Figure 5*A* shows data from one of the neurons tested. This neuron exhibited strong tilt tuning to disparity gradients at three different mean disparities, but no significant tilt selectivity in the monocular controls. Figure 5*B* shows TDI values from monocular measurements plotted against TDI values for binocular stimuli at each of three mean disparities for each neuron (resulting in 45 data points for each eye). Only 13% of the monocular controls yielded significant tuning (ANOVA, $p > 0.05$), and there was no significant correlation between monocular and binocular measurements ($r = 0.06$; $p = 0.73$). Thus, monocular cues cannot account for tilt tuning in MT.

Joint coding of tilt and slant
Population decoding of 3-D surface orientation signals might be more difficult if the tilt preference of single neurons varies substantially with surface slant. An alternative possibility, consistent with the joint coding of variables in other areas (e.g., orientation and spatial frequency in V1), is that tilt and slant have separable influences on the firing rate of single neurons such that slant simply modulates the strength of tuning for tilt. To examine the joint coding of tilt and slant, we obtained tilt-tuning curves at several different surface slants for a subset (29/97) of our neurons. Figure 6*A* shows a typical result. This neuron exhibited significant tilt tuning (ANOVA, $p < 0.01$) across a range of slants (from 35 to 74 degrees), with only small changes in the preferred tilt. There was no significant tilt tuning ($p = 0.14$) at a slant of 3
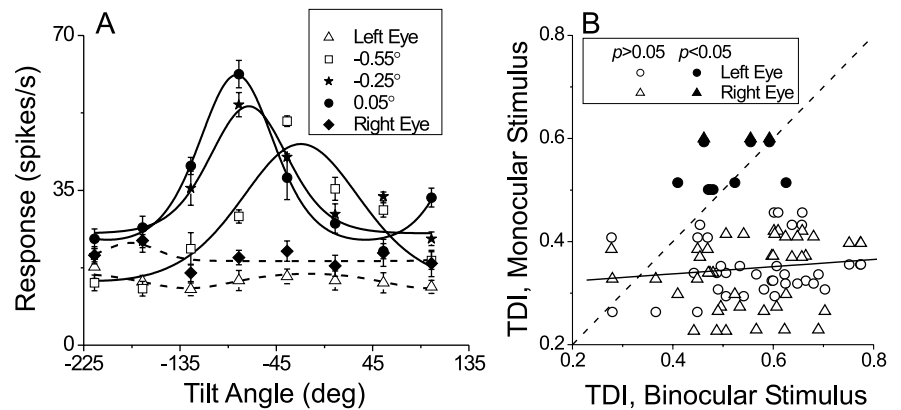

Figure 5. Tilt tuning does not result from monocular cues. *A*, Data from an example neuron. Solid curves show binocular tilt-tuning measurements taken at three mean disparities ranging from −0.55 to 0.05°. Dashed curves show tilt tuning for monocular controls in which only the left or right half-image was presented to the monkey. Direction of motion, 280°; speed of motion, 17°/sec; aperture diameter, 4°; eccentricity, 8.5°; surround inhibition, 72%. *B*, For each neuron tested ($n = 15$), TDI values calculated from the left-eye (circles) and right-eye (triangles) half-images are plotted against TDI values calculated from binocular stimuli at each of three mean disparities. Thus, there are 90 data points in this plot: two eyes × three mean disparities × 15 neurons. Filled symbols denote monocular controls for which tilt tuning was statistically significant (ANOVA; $p < 0.05$). The dashed line has unity slope, and the solid line is the best fit to the data using linear regression.

degrees for this neuron, and tilt tuning was weak even for the 35 degree slant.

We computed a TDI metric at each tested slant for all of the 29 neurons that were studied. Figure 6*B* summarizes how the strength of tilt tuning (TDI) varies with slant; each MT neuron is represented by four to six points in this scatter plot. There is a significant positive correlation ($r = 0.46$; $p < 0.001$) in these data, showing that tilt tuning was generally strong only for large slants. Open symbols in Figure 6*B* indicate the slant at which each neuron showed its maximal TDI. We took this as a measure of the preferred slant of each neuron because we found that peak firing rates generally varied little with slant and, thus, were an unreliable predictor of how slant modulated the tuning for tilt. Although some MT neurons prefer intermediate slants (near 45 degrees), most neurons preferred slants that were close to the largest values tested. These data appear to indicate that MT is insensitive to small slants, but there are two important caveats to be noted. First, these tilt versus slant experiments were usually done only when a neuron displayed clear tilt tuning in the initial tests with a slant of ~67 degrees. We therefore might have missed neurons that were tuned to small slants. Second, as Figure 6*B* indicates, we did not sample small slant values extensively. For these reasons, it is unclear whether there are MT neurons that are strongly tuned to small slants, and further experiments will be needed to clarify this point.

The main purpose of these tests was to determine if tilt preference was independent of slant. In Figure 6*C*, the preferred tilt of each neuron is plotted as a function of slant for all slant values that yielded significant tilt tuning (ANOVA; $p < 0.05$). Most of the curves are quite flat, indicating that there is much greater variance in preferred tilt across neurons than there is across slants for a particular neuron. In fact, neuron identity alone accounts for 90% of the variance in the data of Figure 6*B* (ANOVA), whereas adding slant as a covariate (ANCOVA) accounts for only an additional 1% of variance.

To quantify the dependence of tilt preference on slant for individual neurons, we applied the same fitting methods described above for analyzing effects of mean disparity. For the neuron in Figure 6*A* (Fig. 6*C*, open stars), independent fits to
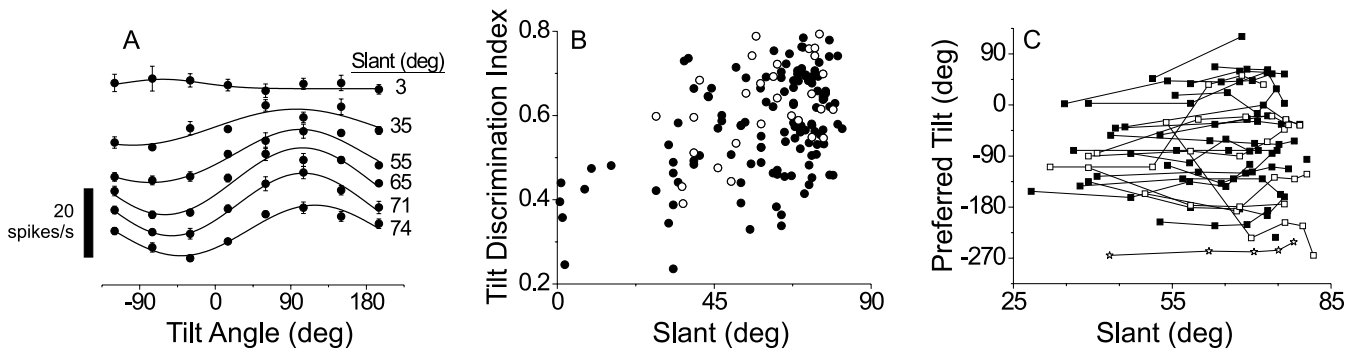
Figure 6. Joint coding of tilt and slant. *A,* Tilt-tuning measurements made at six different slants for the same MT neuron as in Figure 2. From top to bottom, the disparity gradient magnitudes are 0.002, 0.05, 0.1, 0.15, 0.2, and 0.25°/°. The corresponding slants are given along the right side of the plot. Smooth curves are the best fitting sinusoids (Eqs. 1, 2), and have been shifted vertically to minimize overlap and increase clarity. Calibration: 20 spikes/sec. *B,* TDI is plotted as a function of slant for all 29 neurons that were tested in the joint tilt–slant experiment. Each data point shows the TDI value for one slant, such that each neuron is represented four to six times in this plot. Open symbols indicate the slant value at which the maximum TDI was obtained for each neuron. *C,* Preferred tilt is plotted as a function of slant for the same 29 neurons. Preferred tilt values are only shown for slants at which the tilt tuning was significant (ANOVA; $p < 0.05$). Filled symbols denote neurons for which the tilt preference was statistically independent of slant (sequential F test; $p > 0.05$). Stars denote data for the example neuron from *A.*
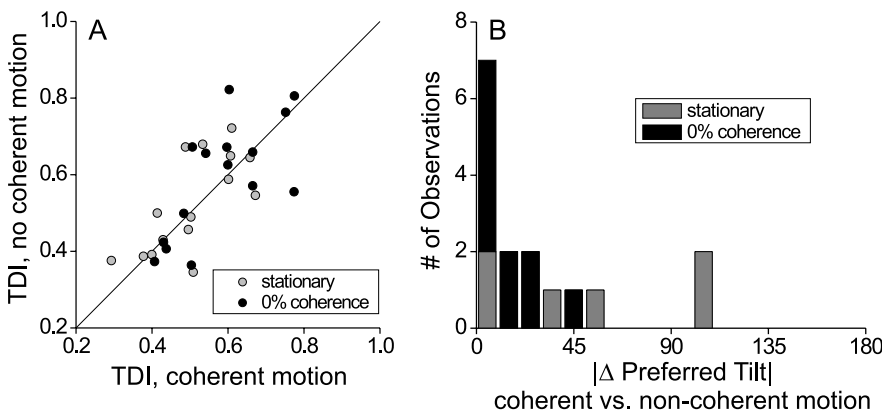
Figure 7. Tilt tuning does not require coherent motion. *A,* Ten MT neurons were tested (at three mean disparities each) using both coherent motion and noncoherent motion. Noncoherent stimuli were either stationary (gray symbols) or 0% coherence (black symbols). For each neuron, TDI values were computed at each mean disparity and for each motion condition; these values are compared across motion conditions in the scatter plot ($n = 30$). The solid diagonal line has unity slope. *B,* For each mean disparity that exhibited significant tilt tuning using both coherent and noncoherent motion (14/30), we computed the absolute difference between the preferred tilts, and these are plotted as a histogram. Gray and black filled bars denote the stationary and 0% coherence cases.

tilt-tuning curves at each slant were slightly, but significantly, better than fits in which the tilt-tuning curve was constrained to have the same peak and trough at each slant (sequential F test; $p = 0.0007$). This example demonstrates the high sensitivity of the sequential F test approach, because the variations in tilt preference across slants in Figure 6*A* are clearly quite modest. Among the 29 neurons that were tested at multiple slants, 21/29 passed the sequential F test ($p > 0.05$). For this majority of neurons (Fig. 6*C,* filled symbols), the tilt preference is statistically invariant with changes in slant. We thus conclude that tilt and slant are coded in a separable manner in MT.

Dependence on coherent motion
In the experiments described above, dots within the MT receptive field always moved with a fixed (preferred) velocity on the display screen. When a disparity gradient is applied, dots appear to stream along an oriented surface in depth. This raises the possibility that tuning for tilt and slant might simply reflect mechanisms in MT for coding 3-D velocity (i.e., motion-in-depth), although previous results from anesthetized monkeys have argued against this possibility (Maunsell and Van Essen, 1983a). To

address this issue, we tested whether tilt and slant tuning are affected when coherent motion is removed from our visual stimuli. This was done either by presenting stationary dots (five neurons) or by randomly replotting the locations of dots every fourth video frame (0% motion coherence, five neurons). If tilt and slant tuning result from sensitivity to specific 3-D trajectories of the moving dots, this surface orientation dependence should be abolished when coherent motion is removed from the display. We found that this was not the case.

Figure 7*A* compares TDI values obtained using both coherent and noncoherent motion. Data are shown for 10 MT neurons, each tested at three mean disparities. Gray and black symbols denote neurons tested with stationary and 0% coherence stimuli, respectively. There is a strong correlation between TDI values for coherent and noncoherent motion (ANCOVA within-cells regression; $r = 0.69$; $p < 0.0001$), with no dependence on the type of non-coherent motion used ($p = 0.95$). Moreover, there is no significant difference between the average TDI values for coherent and noncoherent motion (paired t test; $p = 0.49$). For each mean disparity with significant tilt tuning in both motion conditions, we computed the difference in preferred tilts between the coherent and non-coherent cases. Figure 7*B* shows that the tilt preferences are generally in close agreement.

These analyses show that tilt tuning does not depend on the presence of coherent motion in the receptive fields of MT neurons. Thus, tilt tuning cannot simply be a side effect of selectivity for motion-in-depth based on interocular velocity differences (Cumming, 1994). Further evidence to support coding of surface orientation rather than 3-D velocity is our finding (data not shown) that the preferred tilt axis is not correlated with the preferred (2-D) direction of motion across our population of neurons (randomization test; $p = 0.47$). Thus, it is generally not the case that MT neurons preferred the tilt angle that aligned the 2-D velocity preference with the steep slope of the disparity gradient, as might be expected if these neurons were specialized to signal the 3-D velocity of moving objects.
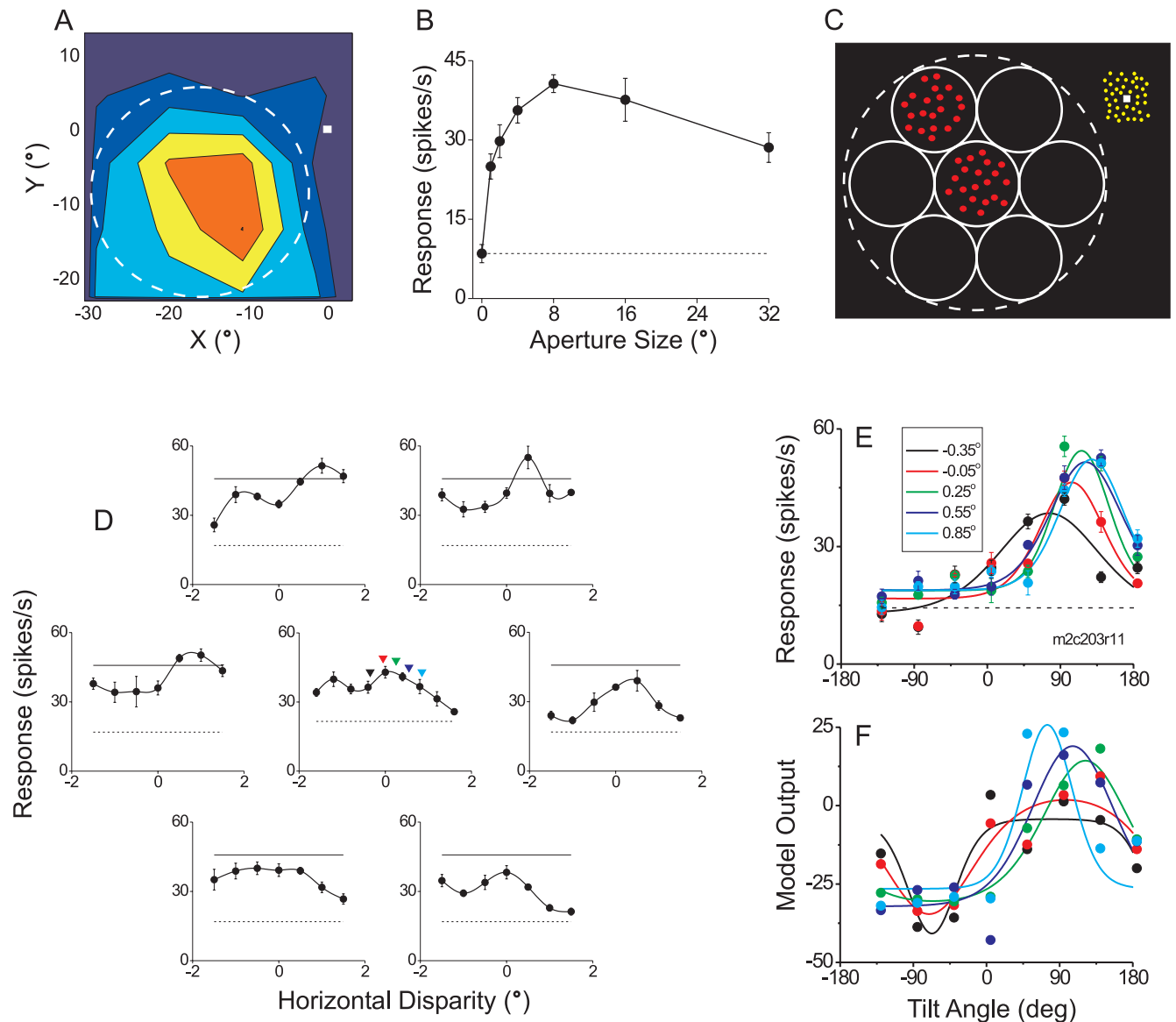
Figure 8. Tilt tuning can be predicted from three-dimensional receptive field substructure. All data in this figure were taken from a single MT neuron. *A,* Receptive field map, conventions as in Figure 2*C*. *B,* Size-tuning curve, conventions as in Figure 2*B*. *C,* Schematic illustration of the stimulus used to probe receptive field substructure. The receptive field (dashed circle, corresponding to that in *A*) was divided into seven subregions: a center patch containing dots at the preferred disparity, and six surrounding patches having variable disparities. A small (2°) patch of zero-disparity dots (yellow) was presented around the fixation point (white square) to help anchor vergence. *D,* Seven disparity-tuning curves are shown, corresponding to the stimulus locations in *C*. The six tuning curves around the perimeter show the disparity tuning of the neuron at each of the locations where the disparity of the surrounding patch was varied. The solid horizontal line in each of these plots shows the response of the neuron to the center patch when presented alone (at the optimal disparity). The dashed horizontal lines denote the level of spontaneous activity. The central disparity-tuning curve shows the measurement obtained with a large patch of dots that covered the entire receptive field (conventions as in Figs. 2 and 3). Arrowheads denote the mean disparities of the gradient stimulus used to test the neuron in *E*. *E,* Tilt tuning tested at five mean disparities ranging from −0.35 to 0.85° (conventions as in Fig. 3). Direction of motion, 230°; speed of motion, 2°/sec; aperture diameter, 24°; eccentricity, 19°; surround inhibition, 25%. *F,* Tilt-tuning curves predicted from a simple model based on linear summation of the responses in *D* (see in text for details).

Receptive field mechanisms

What receptive field mechanisms might underlie the tuning of MT neurons for tilt and slant of 3-D surfaces? One possibility is that tuning for horizontal disparity varies within the classical receptive field and/or within the nonclassical surround. This is quite plausible given that MT neurons have receptive fields several times larger than their primary inputs from V1 and V2 (Albright and Desimone, 1987; Maunsell and Van Essen, 1987), allowing ample opportunity for convergence of heterogeneous disparity-tuned inputs. We therefore probed the 3-D substructure of MT receptive fields and asked whether this substructure could predict the responses to disparity gradients.

Figure 8 shows data from an MT neuron that exhibited weak conventional tuning for frontoparallel disparities (Fig. 8*D*, center panel), moderate surround inhibition (Fig. 8*B*), and strong tuning for tilt (Fig. 8*E*). The 3-D substructure of the receptive field of this neuron was probed with a stimulus array (Fig. 8*C*) consisting of a small center patch of dots presented at the preferred disparity of the neuron and six surrounding patches that had variable disparities. During each trial, the center patch was presented in conjunction with one of the surrounding patches. Because this neuron exhibited clear surround inhibition, the size of the center patch was set to the optimal size from the size-tuning curve (Fig. 8*B*), and the six surrounding patches extended into the nonclas-
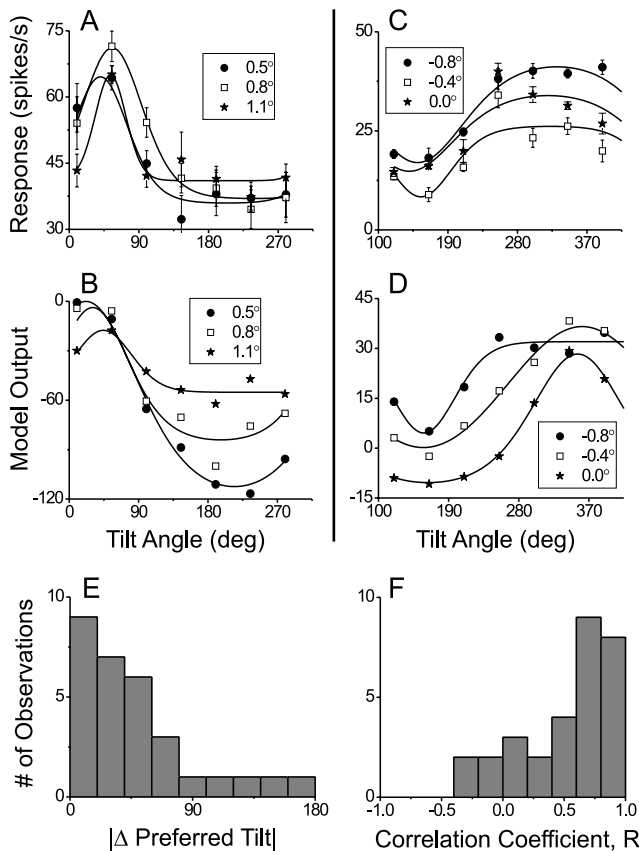
Figure 9. Summary of the quality of model predictions of tilt tuning. *A,* Tilt-tuning curves for an MT neuron taken at three mean disparities ranging from 0.5 to 1.1°. Direction of motion, 100°; speed of motion, 6°/sec; aperture diameter, 14°; eccentricity, 6.9°; surround inhibition, 75%. *B,* Model predictions for the neuron in *A. C,* Tilt tuning of a second MT neuron tested at three mean disparities ranging from −0.8 to 0.0°. Direction of motion, 255°; speed of motion, 8°/sec; aperture diameter, 26°; eccentricity, 11.2°; surround inhibition, 24%. *D,* Model predictions for neuron in *B. E,* Distribution of the absolute value of the difference in tilt preference, $|\Delta \text{Pref. Tilt}|$, between the predicted and the observed responses. Values of $|\Delta \text{Pref. Tilt}|$ are shown for 24 mean disparities (with significant tilt tuning) from nine neurons. *F,* Distribution of correlation coefficients (*R*) between measured tilt-tuning curves and model predictions. One *R* value was computed for each of 24 means disparities from the same nine neurons as in *E*.

sical inhibitory surround. For neurons without any surround inhibition, the entire seven-patch stimulus array was presented within the classical RF (see Materials and Methods for details).

Disparity-tuning curves for each of the six surrounding locations are shown in Figure 8*D*, and it is clear that disparity tuning is not homogeneous throughout the receptive field. Maximal responses were observed at large far (uncrossed) disparities for top-left locations, whereas these disparities elicited near-minimal responses at bottom-right locations. To test whether this heterogeneity underlies tilt tuning, we crudely approximated each different disparity-gradient stimulus by an appropriate combination of disparities in these seven patches. This allowed us to predict responses of the neuron to gradients by linearly summing appropriate portions of the data in Figure 8*D*. Predicted tilt-tuning curves are shown in Figure 8*F*, and it is clear that these curves provide a good first-order prediction of the observed responses. Note, however, that the predicted responses of the model are mainly negative for this example MT neuron. This occurs because the neuron exhibits surround inhibition (Fig. 8*B*), such that responses elicited by the six surrounding patches were generally lower than responses to the center patch presented in isolation (Fig. 8*D*).

Figure 9*A* and *C* shows tilt-tuning curves (at three mean disparities) for two additional MT neurons. Figure 9*B* and *D* shows the corresponding predictions of our model based on data obtained as described in Figure 8, *C* and *D*. Because our model assumes linear summation and contains no normalization mechanisms (Britten and Heuer, 1999), one should not attempt to compare the absolute response levels of the model to those of the MT data. Rather, we emphasize that the basic shapes of the model curves, including the locations of the peaks and troughs, are quite similar for the measured and predicted tuning curves.

To quantify the quality of the model predictions, we computed the difference in preferred tilt, $|\Delta \text{Pref. Tilt}|$, between predicted and measured tilt-tuning curves. This analysis was performed on data from nine neurons that showed both strong tilt selectivity (TDI > = 0.5) and clear disparity selectivity in the seven-patch mapping experiment (average DDI across the six locations ≥0.5). For neurons with weak tilt tuning or weak disparity modulation, we found that model predictions were very noisy. Figure 9*E* shows the histogram of $|\Delta \text{Pref. Tilt}|$ for 24 mean disparities from these nine data sets. Only mean disparities with significant tilt tuning (ANOVA; $p < 0.01$) were included in this analysis. Most of the differences in preferred tilts (60%) were smaller than 45 degrees, and very few were larger than 90 degrees, indicating that tilt preferences were generally well matched between measured and predicted tuning curves. We also calculated the correlation coefficient (*R*) between measured and predicted tuning curves for each mean disparity. Figure 9*F* shows the distribution of these correlation coefficients. Most values are >0.5, indicating that predicted and measured tuning curves typically had quite similar shapes. Together, these results indicate that the tilt selectivity of MT neurons can be primarily explained by variations in local disparity tuning within the MT receptive field.

Involvement of surround inhibition

Previous computational and physiological studies have reported that spatially asymmetric surround inhibition is essential for generating the selectivity of MT neurons to surface orientation defined by speed gradients (Buracas and Albright, 1996; Xiao et al., 1997). Is surround inhibition also necessary for generating the tilt selectivity that we have observed? Among the nine neurons analyzed in Figure 9, *E* and *F*, five showed some surround inhibition, whereas four neurons showed no surround inhibition at all. For the latter neurons, heterogeneous disparity tuning within the classical RF was sufficient to predict tilt preference. This observation suggests that surround inhibition is not a primary determinant of tilt selectivity in our experiments, but this conclusion is tenuous based on only nine neurons. To clarify the role of surround inhibition, we examined how tilt selectivity depends on both the strength and spatial distribution of surround inhibition for our full sample of neurons.

The overall strength of surround inhibition was determined from size tuning curves (Figs. 2*B*, 8*B*) by computing the percentage of surround inhibition:

$$\% \text{ Surround inhibition} = 100 \times \left(\frac{R_{opt} - R_{largest}}{R_{opt} - S} \right), \quad (5)$$

where R_{opt} is the response to the optimal stimulus size, $R_{largest}$ is the response to the largest stimulus, and S denotes the level of spontaneous activity. These values, as well as the statistical significance of surround inhibition, were determined from curve fits to size tuning curves as described elsewhere (DeAngelis and Uka, 2003). Figure 10*A* shows the TDI plotted against percent of sur-

round inhibition for our population of 97 MT neurons. Filled symbols indicate neurons with significant surround inhibition ($p < 0.05$). We find no significant correlation between the strength of surround inhibition and the strength of tilt selectivity ($r = 0.006$; $p = 0.95$), indicating that surround inhibition is not necessary for tilt tuning in MT.

To assess whether tilt selectivity depends on the spatial distribution (i.e., asymmetry) of surround inhibition (Xiao et al., 1997), we analyzed responses from 37 neurons that were tested using the seven-patch stimulus configuration of Figure 8*C*. For each of the six surrounding patch locations, we computed the average response of the MT neuron across disparities, and we plotted a vector having the average response as its length and the location of the patch as its direction. We then computed the vector average across all six patch locations to get an estimate of surround asymmetry. Specifically, we construct a surround asymmetry index as the magnitude of the vector average divided by the average magnitude of the individual vectors. This index will be close to zero if responses to the surrounding patches are symmetric about the receptive field center. Larger values of the index indicate stronger spatial asymmetry in responses to the surrounding patches. Figure 10*B* shows TDI values as a function of the surround asymmetry index for 37 MT neurons. We find no significant correlation between these variables ($r = 0.02$; $p = 0.89$), indicating that tilt tuning does not depend on asymmetric surround effects.

For 44/97 neurons, we measured tilt-tuning curves using two different stimulus sizes, randomly interleaved. The large size was chosen as described in Materials and Methods, whereas the small size was twofold to threefold smaller. Thus, for neurons with surround inhibition, the small size was near-optimal as given by the size-tuning curve. For neurons without surround inhibition, the small size was one-third to one-half the size of the classical RF. For both groups of neurons, TDI values were significantly greater (*t* test; $p < 0.01$) for the large stimulus than for the small stimulus (the percentage difference was 24% for neurons with surround inhibition, 19% for neurons without surround inhibition).

Together, the data of Figures 8–10 indicate that tilt tuning in response to disparity gradients depends mainly on heterogeneity of disparity tuning within the receptive fields of MT neurons (including the nonclassical surround), not on the presence or spatial distribution of surround inhibition. Further work will be necessary to fully understand the 3-D organization of MT receptive fields.

Discussion

Most models of cortical visual processing have focused on the roles that area MT plays in computing motion within frontoparallel planes (Nowlan and Sejnowski, 1995; Wang, 1997; Simoncelli and Heeger, 1998; Koechlin et al., 1999; Perrone and Thiele, 2002) (but see Lappe, 1996; Buracas and Albright, 1996). Recently, it has been demonstrated physiologically that area MT contributes to depth judgments involving frontoparallel surfaces (DeAngelis et al., 1998) and that integration of motion and disparity signals allows MT neurons to signal the perceived depth-ordering of transparent surfaces (Bradley et al., 1995, 1998; Dodd et al., 2001; Grunewald et al., 2002). We now show that MT
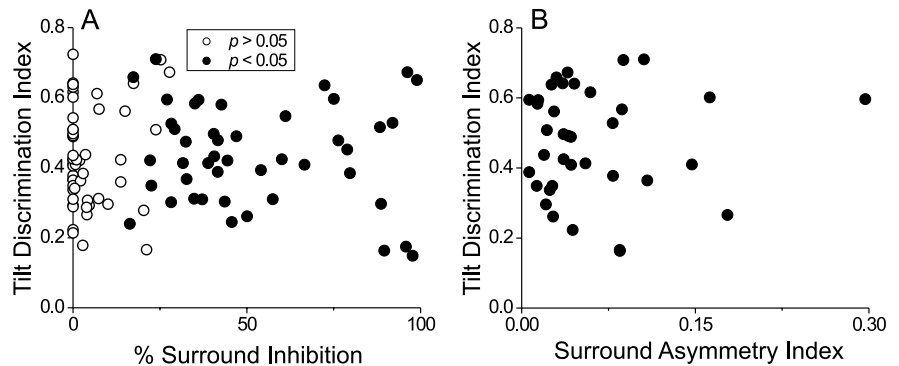
Figure 10. Tilt tuning strength is not correlated with either the strength or spatial asymmetry of surround inhibition. *A,* The TDI is plotted against the percentage of surround inhibition for the 97 MT neurons in our sample. Neurons with significant surround inhibition are indicated by filled symbols ($p < 0.05$). *B,* TDI is plotted against the surround asymmetry index (see text) for 37 MT neurons that were tested with the stimulus configuration of Figure 8*C*.

contains robust, disparity-based signals regarding the 3-D orientation (tilt and slant) of planar surfaces. This tilt selectivity does not result from vergence eye movements or subtle monocular dot-density cues, and approximately one-half of MT neurons respond more strongly to a tilted stimulus (i.e., a nonzero slant) than to any frontoparallel stimulus of the same size (data not shown). In addition, we show that the tilt preference of MT neurons is primarily independent of the mean depth and slant of the surface, properties that may simplify the extraction of 3-D orientation signals from a population of MT neurons. Together, these findings show that the visual representation in MT is more complex than previously thought; it contains information not only about the local velocity of features on the retina, but also about the 3-D structure of the environment from which those velocity signals arise.

Although we have shown that MT neurons carry information about the 3-D orientation of planar surfaces, this selectivity could arise because of other computations in MT. Tilt and slant tuning could be a by-product of selectivity for 3-D velocity (motion-in-depth), which can be computed via either interocular velocity differences or changes in binocular disparity over time (Cumming, 1994). Our control experiments and analyses suggest that this is unlikely, for two main reasons. First, tilt and slant tuning remain unchanged when coherent motion is removed from our stimuli, thus excluding the possibility that tilt tuning reflects the calculation of motion-in-depth based on interocular velocity differences. Second, we found no consistent relationship between the preferred tilt of MT neurons and their preferred 2-D velocity. For an object moving in 3-D space, binocular disparity changes over time along the 3-D vector of the movement. As a result, the direction of maximal slope of the gradient is aligned with the 2-D velocity of the object. If MT neurons were specialized to code 3-D velocity, then we might expect their gradient preference to be similarly aligned with their preferred 2-D velocity (e.g., a neuron preferring rightward 2-D motion would have a preferred tilt of 0 or 180 degrees, whereas a neuron preferring upward motion would prefer a tilt of 90 or 270 degrees) (Fig. 1). Some MT neurons behave this way, but most do not. Thus, our findings suggest that gradient selectivity in MT plays a more general role in the analysis of 3-D scene structure. This conclusion is consistent with that of a previous study in the anesthetized monkey, where a specialization for coding of motion-in-depth was not found in MT (Maunsell and Van Essen, 1983a).

We have shown that the tilt preference of MT neurons can be predicted from heterogeneity of disparity tuning within the clas-

sical receptive field and/or the nonclassical surround. Although the quality of our model predictions is far from ideal, we think their accuracy is striking given the simplicity of the model and the coarseness of our measurements of receptive field substructure. These results suggest that tilt selectivity arises from a combination of inputs with disparity preferences that vary systematically across space within the MT receptive field. The details of the mechanisms that underlie this pooling remain unclear, and our data do not allow us to evaluate whether nonlinear interactions are involved. Our model was based on linear summation of responses to the different stimulus patches (Fig. 8C), but each surrounding patch was always presented in conjunction with the center patch. Thus, our data (Fig. 8D) may include nonlinear interactions between the center and surrounding patches. Further experiments will be needed to clarify the mechanisms underlying tilt selectivity.

Our findings complement and extend those of a few previous studies of disparity-based surface representation. Taira et al. (2000) reported that neurons in the CIP are selective for the tilt of planar surfaces specified by disparity gradients, although they did not sufficiently exclude the possibility that these responses arose through monocular cues, variations in vergence angle, or from inaccurate centering of stimuli over the receptive fields of the neurons. It is also difficult to determine if tilt selectivity is more or less common in CIP than in MT because there is no quantitative summary of tilt selectivity in the Taira et al. (2000) study. In any case, our findings show that disparity-gradient signals arise substantially earlier in the visual hierarchy than the parietal lobe. MT receives direct input from V1 and V2 (Maunsell and van Essen, 1983b), whereas CIP is thought to be two or three synapses removed from these areas (Sakata et al., 1997). Our findings confirm expectations from psychophysical and theoretical considerations that 3-D surface orientation should be coded early in the visual pathways (Gibson, 1950; Marr, 1982; Nakayama, 1996).

Recently, Hinkle and Conner (2002) have reported the presence of 3-D orientation tuning in macaque area V4, indicating that 3-D orientation signals are present midway along both the dorsal and ventral processing streams. A few differences between their study and ours (other than the brain area) are worth noting. First, because Hinkle and Conner (2002) used bar stimuli, tilt was confounded with 2-D orientation in their stimuli. Thus, they clearly demonstrate the presence of slant selectivity in V4, but one cannot draw any conclusions about tilt tuning or about the joint coding of tilt and slant. Second, Hinkle and Conner (2002) did not find slant tuning for textured surface stimuli (like ours) that lacked orientation cues. Thus, V4 neurons do not appear to be coding surface orientation from the gradient of horizontal disparities, but may instead be dependent on orientation disparities. V4 and MT may therefore contain different mechanisms for signaling 3-D orientation.

Our findings dovetail nicely with previous studies showing that the response of MT neurons depends on the spatial orientation of speed gradients (Treue and Andersen, 1996; Xiao et al., 1997), which may also serve as a cue to the tilt and slant of 3-D surfaces. It should be noted, however, that the mean speed of the stimuli was not varied in these studies to control for the possibility that gradient selectivity depends on stimulus centering. Also, the tuning of MT neurons to speed gradients appears to depend on the presence of asymmetric surround inhibition (Buracas and Albright, 1996; Xiao et al., 1997), whereas we did not find any consistent relationship between the strength of tilt tuning and the strength or asymmetry of surround inhibition (Fig. 10). Despite these differences, the combination of these studies suggests that

MT neurons may integrate information from disparity gradients, velocity gradients, and perhaps other cues to provide robust estimates of 3-D surface orientation. We are currently testing this hypothesis.

This work adds to a small, but rapidly growing, body of work on the neural coding of higher-level disparity signals that underlie perception of 3-D structure (Shikata et al., 1996; Bradley et al., 1998; Eifuku and Wurtz, 1999; Janssen et al., 1999, 2000; Taira et al., 2000; von der Heydt et al., 2000; Dodd et al., 2001; Hinkle and Connor, 2002; Thomas et al., 2002). Our results reveal a new aspect of the depth representation found within area MT and provide new support for the idea that MT plays a role in the analysis of 3-D scene structure. Additional studies can now be focused on mapping the 3-D substructure of MT receptive fields, probing for causal links between MT activity and surface perception, and exploring how MT neurons integrate multiple cues to surface orientation. Along with similar studies in other areas, this endeavor should reveal the neural mechanisms that underlie our impressive ability to see the world in three dimensions.

References

Albright TD, Desimone R (1987) Local precision of visuotopic organization in the middle temporal area (MT) of the macaque. Exp Brain Res 65:582–592.

Bradley DC, Qian N, Andersen RA (1995) Integration of motion and stereopsis in middle temporal cortical area of macaques. Nature 373:609–611.

Bradley DC, Chang GC, Andersen RA (1998) Encoding of three-dimensional structure-from-motion by primate area MT neurons. Nature 392:714–717.

Britten KH, Heuer HW (1999) Spatial summation in the receptive fields of MT neurons. J Neurosci 19:5074–5084.

Buracas GT, Albright TD (1996) Contribution of area MT to perception of three-dimensional shape: a computational study. Vision Res 36:869–887.

Cumming BG (1994) Motion-in-depth. In: Visual detection of motion (Smith AT, Snowden RJ, eds), pp 333–366. San Diego: Academic.

DeAngelis GC, Newsome WT (1999) Organization of disparity-selective neurons in macaque area MT. J Neurosci 19:1398–1415.

DeAngelis GC, Uka T (2003) Coding of horizontal disparity and velocity by MT neurons in the alert macaque. J Neurophysiol 89:1094–1111.

DeAngelis GC, Cumming BG, Newsome WT (1998) Cortical area MT and the perception of stereoscopic depth. Nature 394:677–680.

Dodd JV, Krug K, Cumming BG, Parker AJ (2001) Perceptually bistable three-dimensional figures evoke high choice probabilities in cortical area MT. J Neurosci 21:4809–4821.

Draper NR, Smith H (1966) Advanced regression analysis. New York: Wiley.

Eifuku S, Wurtz RH (1999) Response to motion in extrastriate area MSTl: disparity sensitivity. J Neurophysiol 82:2462–2475.

Gibson JJ (1950) The perception of the visual world. Cambridge, MA: Riverside.

Grunewald A, Bradley DC, Andersen RA (2002) Neural correlates of structure-from-motion perception in macaque V1 and MT. J Neurosci 22:6195–6207.

Hinkle DA, Connor CE (2002) Three-dimensional orientation tuning in macaque area V4. Nat Neurosci 5:665–670.

Howard IP, Rogers BJ (2002) Seeing in depth (Vol 2): depth perception. Thornhill, Ontario: I. Porteous.

Janssen P, Vogels R, Orban GA (1999) Macaque inferior temporal neurons are selective for disparity-defined three-dimensional shapes. Proc Natl Acad Sci USA 96:8217–8222.

Janssen P, Vogels R, Orban GA (2000) Three-dimensional shape coding in inferior temporal cortex. Neuron 27:385–397.

Janssen P, Vogels R, Liu Y, Orban GA (2001) Macaque inferior temporal neurons are selective for three-dimensional boundaries and surfaces. J Neurosci 21:9419–9429.

Koechlin E, Anton JL, Burnod Y (1999) Bayesian inference in populations of cortical neurons: a model of motion integration and segmentation in area MT. Biol Cybern 80:25–44.

Lappe M (1996) Functional consequences of an integration of motion and

stereopsis in area MT of monkey extrastriate visual cortex. Neural Comput 8:1449–1461.

Marr D (1982) Vision. New York: Freeman.

Maunsell JH, Van Essen DC (1983a) 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.

Maunsell JH, Van Essen DC (1983b) 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.

Maunsell JH, Van Essen DC (1987) Topographic organization of the middle temporal visual area in the macaque monkey: representational biases and the relationship to callosal connections and myeloarchitectonic boundaries. J Comp Neurol 266:535–555.

Nakayama K (1996) Binocular visual surface perception. Proc Natl Acad Sci USA 93:634–639.

Nowlan SJ, Sejnowski TJ (1995) A selection model for motion processing in area MT of primates. J Neurosci 15:1195–1214.

Perrone JA, Thiele A (2002) A model of speed tuning in MT neurons. Vision Res 42:1035–1051.

Prince SJ, Pointon AD, Cumming BG, Parker AJ (2002) Quantitative analysis of the responses of V1 neurons to horizontal disparity in dynamic random-dot stereograms. J Neurophysiol 87:191–208.

Sakata H, Taira M, Kusunoki M, Murata A, Tanaka Y (1997) The TINS Lecture. The parietal association cortex in depth perception and visual control of hand action. Trends Neurosci 20:350–357.

Sedgwick HA (1986) Space perception. In: Handbook of perception and human performance. Volume 1: sensory processes and perception (Boff KR, Kaufman L, Thomas JP, eds). New York: Wiley.

Shikata E, Tanaka Y, Nakamura H, Taira M, Sakata H (1996) Selectivity of the parietal visual neurones in 3-D orientation of surface of stereoscopic stimuli. NeuroReport 7:2389–2394.

Simoncelli EP, Heeger DJ (1998) A model of neuronal responses in visual area MT. Vision Res 38:743–761.

Stevens KA (1983) Slant-tilt: the visual encoding of surface orientation. Biol Cybern 46:183–195.

Taira M, Tsutsui KI, Jiang M, Yara K, Sakata H (2000) Parietal neurons represent surface orientation from the gradient of binocular disparity. J Neurophysiol 83:3140–3146.

Thomas OM, Cumming BG, Parker AJ (2002) A specialization for relative disparity in V2. Nat Neurosci 5:472–478.

Treue S, Andersen RA (1996) Neural responses to velocity gradients in macaque cortical area MT. Vis Neurosci 13:797–804.

Tsutsui K, Sakata H, Naganuma T, Tanaka Y (2002) Neural correlates for perception of 3-D surface orientation from texture gradient. Science 298:409–412.

von der Heydt R, Zhou H, Friedman HS (2000) Representation of stereoscopic edges in monkey visual cortex. Vision Res 40:1955–1967.

Wang R (1997) A network model of motion processing in area MT of primates. J Comp Neurosci 4:287–308.

Xiao DK, Marcar VL, Raiguel SE, Orban GA (1997) Selectivity of macaque MT/V5 neurons for surface orientation in depth specified by motion. Eur J Neurosci 9:956–964.