

Representation of 3-D surface orientation by velocity and disparity gradient cues in area MT

Takahisa M. Sanada,¹ Jerry D. Nguyenkim,² and Gregory C. DeAngelis¹

¹Department of Brain and Cognitive Sciences, Center for Visual Science, University of Rochester, New York;

and ²Department of Anatomy and Neurobiology, Washington University School of Medicine, St. Louis, Missouri

Submitted 23 June 2011; accepted in final form 3 January 2012

Sanada TM, Nguyenkim JD, DeAngelis GC. Representation of 3-D surface orientation by velocity and disparity gradient cues in area MT. *J Neurophysiol* 107: 2109–2122, 2012. First published January 4, 2012; doi:10.1152/jn.00578.2011.—Neural coding of the three-dimensional (3-D) orientation of planar surface patches may be an important intermediate step in constructing representations of complex 3-D surface structure. Spatial gradients of binocular disparity, image velocity, and texture provide potent cues to the 3-D orientation (tilt and slant) of planar surfaces. Previous studies have described neurons in both dorsal and ventral stream areas that are selective for surface tilt based on one or more of these gradient cues. However, relatively little is known about whether single neurons provide consistent information about surface orientation from multiple gradient cues. Moreover, it is unclear how neural responses to combinations of surface orientation cues are related to responses to the individual cues. We measured responses of middle temporal (MT) neurons to random dot stimuli that simulated planar surfaces at a variety of tilts and slants. Four cue conditions were tested: disparity, velocity, and texture gradients alone, as well as all three gradient cues combined. Many neurons showed robust tuning for surface tilt based on disparity and velocity gradients, with relatively little selectivity for texture gradients. Some neurons showed consistent tilt preferences for disparity and velocity cues, whereas others showed large discrepancies. Responses to the combined stimulus were generally well described as a weighted linear sum of responses to the individual cues, even when disparity and velocity preferences were discrepant. These findings suggest that area MT contains a rudimentary representation of 3-D surface orientation based on multiple cues, with single neurons implementing a simple cue integration rule.

depth; slant; surface; tilt; visual cortex; middle temporal area

THE VISUAL SYSTEM RECONSTRUCTS three-dimensional (3-D) scene structure from images projected onto the two retinas. Many cues, including binocular disparity, relative motion, texture, shading, and perspective, are used to perceive 3-D structure. Most complex surfaces can be approximated by combinations of locally planar surfaces. Thus understanding how planar surfaces are coded in visual cortex may help reveal how complex surface representations are constructed. The 3-D orientation of a plane (tilt and slant) can be specified by gradients of binocular disparity, motion (velocity), or texture. Human perception of 3-D surface orientation from these cues has been well studied, and the findings are often well explained by Bayesian models (Girshick and Banks 2009; Hillis et al. 2004; Jacobs 1999; Knill 2007; Knill and Saunders 2003).

Address for reprint requests and other correspondence: G. C. DeAngelis, Dept. of Brain and Cognitive Sciences, Center for Visual Science, Univ. of Rochester, 315 Meliora Hall, Rochester, NY 14627 (e-mail: gdeangelis@cvs.rochester.edu).

Physiological studies in macaques have identified neurons that signal the 3-D orientation of planar surfaces. In the ventral stream, 3-D orientation tuning has been reported in area V4 for disparity gradients (Hegde and Van Essen 2005) and in inferotemporal (IT) cortex for texture and disparity gradients (Liu et al. 2004). IT neurons also represent surface curvature from disparity cues (Janssen et al. 1999, 2000). In the dorsal stream, neurons in the anterior intraparietal (AIP) area exhibit selectivity for 3-D shapes including slanted and curved surfaces (Srivastava et al. 2009; Verhoef et al. 2010), dorsal medial superior temporal (MSTd) neurons are selective to 3-D orientation based on velocity gradients (Sugihara et al. 2002), and caudal intraparietal (CIP) neurons are tuned for the tilt of planar surfaces defined by perspective cues, texture gradients, or disparity gradients (Taira et al. 2000; Tsutsui et al. 2001, 2002). Moreover, tilt preferences for disparity and texture gradients generally match in CIP (Tsutsui et al. 2002).

Another dorsal stream area that may play a role in computing 3-D surface orientation is the middle temporal (MT) area, which has previously been implicated in perception of depth based on disparity and motion cues (Bradley et al. 1998; Chowdhury and DeAngelis 2008; DeAngelis et al. 1998; DeAngelis and Newsome 2004; Dodd et al. 2001; Krug et al. 2004; Nadler et al. 2008, 2009; Uka and DeAngelis 2003, 2004, 2006). MT neurons are also selective for the tilt of planar surfaces based on velocity (Treue and Andersen 1996; Xiao et al. 1997) and disparity gradients (Nguyenkim and DeAngelis 2003). However, it is not clear whether single MT neurons show tilt selectivity for both velocity and disparity gradients. Furthermore, no previous study has examined responses to combinations of velocity, disparity, and texture gradients. Indeed, only one study has examined neural responses to combinations of 3-D orientation cues, and this involved coding of disparity gradients and figural perspective cues in area CIP (Tsutsui et al. 2001). Thus it is unclear how neurons integrate multiple cues to 3-D orientation. We hypothesize, based on recent findings in area MSTd (Fetsch et al. 2011; Morgan et al. 2008), that responses to combinations of disparity, velocity, and texture gradients may be approximated as a weighted linear sum of individual cue responses.

This study addresses two main questions. First, do single MT neurons signal planar surface orientation defined by multiple gradient cues, and how does tilt tuning compare across cue conditions? Second, can responses of MT neurons to combinations of gradient cues be predicted by a weighted linear sum of single-cue responses? We report that many MT neurons exhibit robust tuning for 3-D surface orientation defined by both disparity and velocity gradients, with relatively

little selectivity for texture gradients. Moreover, responses to combinations of these cues are well approximated by a weighted linear summation model, consistent with predictions of recent theory (Fetsch et al. 2011; Ma et al. 2006).

MATERIALS AND METHODS

Subjects and apparatus. Two male rhesus monkeys (*Macaca mulatta*) served as subjects in this study. A detailed description of our methods has appeared previously (DeAngelis and Uka 2003). All experimental procedures conformed to National Institutes of Health guidelines and were approved by the Institutional Animal Care and Use Committee at Washington University and the University Committee on Animal Resources at the University of Rochester.

Visual stimuli. Three-dimensional visual stimuli were presented to the monkey using a stereoscopic projection system (Christie Digital Mirage 2000). The stimulus display subtended $75^\circ \times 63^\circ$ at the viewing distance of 57 cm. Visual stimuli were generated by an OpenGL accelerator board (Nvidia, Quadro FX1000) and were viewed by the monkey through ferroelectric liquid crystal shutters that were synchronized to the display refresh (100 Hz). There was no noticeable stereo cross talk with this system because the three-chip DLP projector has essentially no persistence.

Planar surface orientation in 3-D was specified by random dot stimuli, and four cue conditions were generated using OpenGL libraries within Visual C++ (Microsoft Visual Studio .Net). Three of the cue conditions depicted planar surfaces defined by isolated gradients of texture, horizontal disparity, or velocity. The final cue condition combined all three gradients in a congruent manner (see Fig. 1B).

In the texture, velocity, and combined conditions, stimuli were generated in a 3-D virtual workspace using the OpenGL libraries. In these cases, the resulting images were determined by the placement of two virtual cameras within the 3-D workspace, to represent the viewpoints of the left and right eyes of the observer. The two cameras were positioned precisely according to the location of the eyes relative to the 3-D coordinate system, with a horizontal separation between the cameras that was determined by the interocular distance of the subject. Once the cameras were positioned, OpenGL calculated all of the appropriate size/density/velocity gradients based on the camera locations and the 3-D orientation of the imaged surface. Images from the two virtual cameras were then presented to the left and right eyes through the shutter glasses. Except where noted below, the visual stimuli contained random dots that drifted in the preferred direction of motion of the MT neuron being recorded, to elicit robust responses. In the 3-D virtual workspace, dot motion was generated as follows. With no slant applied to the surface (frontoparallel), dots were generated such that they moved along the surface in the preferred direction of the neuron. The surface was then rotated in 3-D around a point at the center of the receptive field (RF), to apply the appropriate slant and tilt for each particular trial. Note that dots near the center of the stimulus still move in the neuron's preferred direction following this rotation, although dots closer to the edges of the stimuli will have motion directions (on the screen) that deviate slightly due to the velocity gradient that accompanies surface slant.

When a surface is rendered in a 3-D OpenGL workspace, multiple cues to 3-D surface orientation are generally linked. Thus additional manipulations were required to isolate the individual gradient cues to 3-D surface orientation. For the texture condition, stimuli were presented as dynamic random dots (i.e., 0% coherence), thereby removing any velocity gradient information while still robustly activating neurons. In this condition, square texture elements were used for which size and density varied smoothly across the surface of the planar stimulus. Horizontal disparity information was removed by eliminating the horizontal separation between the two OpenGL cameras (which places the disparity cue in conflict with the texture cue). In the velocity and disparity conditions, the individual elements of the random dot pattern were drawn as points rather than small squares. In

this mode, OpenGL renders the dots with constant size (on the screen) regardless of the location of each point in the 3-D scene; hence, these stimuli did not include any size or perspective cues. In the disparity condition (described further below), this manipulation effectively removes the texture gradient cue. In the Velocity condition, it is also necessary to counter the density gradient that accompanies an oriented surface. For this purpose, 20% of the dots were randomly replotted within the 3-D surface every few video frames, thus limiting the lifetime of the dots. This dramatically reduces (but does not completely eliminate) the density gradient in the velocity condition (this random replotting was not done in the combined condition). Note that any residual density gradient is not likely to have accounted for much of the tilt tuning seen in the velocity condition. If MT neurons were sensitive to density gradients, then they should show strong tilt tuning in the texture condition, but this condition yielded the weakest tuning (see RESULTS). Hence, the strong tilt tuning seen in the velocity condition is very likely attributable to speed gradients and not the residual density gradients.

It was not possible to isolate the disparity gradient cue when rendering the planar stimuli in a 3-D OpenGL context. Hence, the disparity condition was programmed using a different method from the other conditions. To eliminate the velocity gradient, dots in the disparity condition were drifted at a constant velocity on the screen (as opposed to a constant velocity along the oriented 3-D surface). To achieve this, a set of dot locations was chosen randomly in screen coordinates, and each dot was then projected (using a ray tracing procedure) onto an oriented planar 3-D surface. Dots then drifted at a uniform velocity on the screen (in the neuron's preferred direction) and were projected each frame onto the relevant 3-D surface. This allowed us to generate the disparity gradient that would accompany each 3-D surface orientation while eliminating the velocity and texture gradient cues.

Finally, in the combined condition, planar surfaces of various tilts and slants were rendered in a 3-D OpenGL context using all of the natural cues including texture, velocity, and disparity gradients. Note that the gradient of horizontal disparity was computed to simulate a planar 3-D surface. Because depth and disparity are not related linearly, a strictly linear gradient of horizontal disparity would describe a surface that is slightly curved in space, as described previously (Nguyenkim and DeAngelis 2003). Hence, the stimuli used here do not contain strictly linear disparity gradients. Because the texture condition was presented with 0% coherent motion (to eliminate the velocity gradient), it should be noted that the combined stimulus is not a simple combination of the three isolated gradient stimuli, since the combined condition contains coherent motion. However, the texture cue itself is represented in the same manner in the texture and combined conditions.

Task and data collection. Monkeys were required to maintain their conjugate eye position within a 1.5° -diameter fixation window that was centered at the fixation point. Fixation began 300 ms before presentation of the random dot stimulus and had to be maintained throughout the 1.5-s stimulus presentation for the animal to receive a liquid reward. Only data from successfully completed trials were analyzed. Movements of both eyes were measured in all experiments by using eye coils that were sutured to the sclera; eye position signals were stored to a computer disk at a sampling rate of 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 RF size and eccentricity, and the subsequent entry into gray matter with response properties typical of area MST. 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 band-pass filtered (500–5,000 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 a 1-ms resolution. In addition, raw neural signals were digitized and recorded continuously to disk using Spike2 software and a Power 1401 data acquisition system (Cambridge Electronic Design). These raw signals were used for extracting multiunit activity.

Experimental protocol. The experimental protocol used in this study is similar to that described by Nguyenkim and DeAngelis (2003). One practical difference is that depth variables are specified in terms of centimeters in the virtual 3-D environment rather than in degrees of visual angle. This results from the fact that most of the stimuli were generated in a 3-D OpenGL rendering context, rather than a 2-D orthographic projection as used by Nguyenkim and DeAngelis (2003).

The tuning characteristics of each isolated MT neuron were initially estimated qualitatively using a hand mapping program. Estimates of RF center and size, as well as preferred direction, speed, and disparity were gathered. Quantitative measurements of each of these characteristics were then conducted, in separate blocks of trials, as follows (see also DeAngelis and Uka 2003; Nover et al. 2005; Palanca and DeAngelis 2003 for details). 1) A depth tuning curve was measured by presenting random dot stereograms at 8 different depths, ranging from 20 cm in front of the plane of fixation to 20 cm behind the plane of fixation, in steps of 5 cm. These depth values are equivalent to the following binocular disparities (based on an interocular distance of 3.5 cm): -3.80 , -2.51 , -1.50 , -0.68 , 0 , 0.57 , 1.05 , 1.46 , and 1.83 degrees. 2) A speed tuning curve was measured by presenting random dot stereograms (at the preferred depth) with speeds of 0.5 , 1.0 , 2.0 , 4.0 , 8.0 , 16.0 , and 32.0 cm/s. 3) A quantitative map of the RF was obtained by presenting a small patch of random dots ($\sim 25\%$ of the RF diameter) at 1 of 16 locations on a 4×4 grid that covered the estimated RF location. These mapping stimuli were presented at the preferred speed and depth. The data were fit with a 2-D Gaussian to estimate the RF center location and size. 4) A direction tuning curve was next obtained by presenting random dot stereograms moving in 8 directions of motions, 45° apart, while speed and depth were optimized. 5) A size tuning curve was then obtained by presenting moving random dots in circular apertures having sizes of 0 , 1 , 2 , 4 , 8 , 16 , and 32 cm. Results of this test were used to quantify the extent (%) of surround inhibition exhibited by each neuron (DeAngelis and Uka 2003). 6) Tilt tuning curves were then measured for each neuron. The stimulus set consisted of four different cue conditions: velocity, disparity, texture, and combined (see Fig. 1B). In each case, the stimuli depicted 8 different tilts, 45° apart, at a fixed slant of 65° . In addition, each tilt stimulus was presented at three mean depths that were chosen to bracket the peak of the depth tuning curve.

Stimulus size was chosen based on the results of the size tuning curve and quantitative RF mapping. Because we previously found that tilt tuning in response to disparity gradients was not strongly dependent on surround mechanisms (Nguyenkim and DeAngelis 2003), we used stimuli somewhat larger than the receptive field of each neuron. For neurons that displayed clear surround inhibition, stimulus size was chosen to be twofold larger than the optimal size from the size tuning curve. If there was no discernible surround inhibition, the visual stimulus was set to be twofold larger than the classical RF as measured from the quantitative RF map. In some cases of exceptionally strong surround inhibition, however, a stimulus twice the optimal size elicited little or no response from the neuron. In these instances, stimulus size was reduced until the neuron gave a roughly half-maximal response.

Data analysis. The response to each stimulus presentation was quantified as the average firing rate over the 1.5-s stimulus period. Each stimulus was typically presented five times in blocks of randomly interleaved trials. Tuning curves were constructed by plotting the mean $\pm$ SE of the response across repetitions of each stimulus. Each tilt tuning curve was fit with a wrapped Gaussian function of the following form:

$$f(\theta) = A_1 \cdot \left[e^{\frac{-2[1-\cos(\theta-\theta_0)]}{\sigma^2 \kappa}} + A_2 \cdot e^{\frac{-2[1-\cos(\theta-\theta_0-180^\circ)]}{\sigma^2}} \right] + B \quad (1)$$

Because some neurons showed bimodal tilt tuning, Eq. 1 is a sum of two wrapped Gaussian functions, where θ denotes the tilt angle of the stimulus, θ_0 is the location of the primary peak, σ indicates the standard deviation of the Gaussian, A_1 is the overall amplitude, and B is the baseline. The second exponential term in the equation can produce a second peak 180° out of phase with the first, but only if the parameter A_2 is sufficiently large (A_2 is bounded between 0 and 1). The relative widths of the two peaks are determined by the parameter κ , which was bounded between 0 and 3 such that either of the two peaks could be broader than the other. The best fit of this function to the data was achieved by minimizing the sum-squared error between the response of the neuron and the values of the function, using the constrained minimization tool “lsqcurvefit” in Matlab (The Math-Works). Each tilt tuning curve was fitted independently across the different cue conditions and mean depths.

In the above formulation, the two peaks of the wrapped Gaussian function were constrained to lie 180° apart to reduce correlations among variables in the fits. This was justified by the observation that bimodal tuning curves generally showed two peaks that were $\sim 180^\circ$ apart. To confirm this, a subset of 58 tuning curves that were judged to be clearly bimodal by eye were also fit with a sum of two wrapped Gaussians having independent peak locations. For this subset of tuning curves the mean of the distribution of differences in preferred tilts (mean = 179°) was not significantly different from 180° (1-sample *t*-test, $P = 0.603$, $N = 58$). We quantified the extent of bimodality of tuning curves using the following index:

$$\text{Bimodal index} = 1 - \frac{|A_{\text{pref}} - A_{\text{null}}|}{A_{\text{pref}} + A_{\text{null}}} \quad (2)$$

where A_{pref} and A_{null} denote the amplitudes of the primary and secondary peaks in the fitted curve. With respect to the formulation of Eq. 1, $A_{\text{pref}} = A_1$ and $A_{\text{null}} = A_1 \times A_2$. When comparing tilt preferences of a neuron across stimulus conditions (see e.g., Fig. 5), we classified cells as having bimodal tuning when the bimodal index was > 0.6 . In these cases, differences in tilt preferences were computed as the smallest difference between two peaks in different stimulus conditions. This prevented spurious differences in tilt preference close to 180° that could arise when tuning was bimodal in two stimulus conditions but the relative amplitudes of the two peaks varied.

To quantify the strength of tuning, we equated the average response of an MT neuron to all mean depths by vertically shifting the individual tilt tuning curves. We then combined the data across mean depths to create a single “grand” tilt tuning curve. Note that this allows tilt tuning to cancel across mean depths when the tilt preferences differ by close to 180° . Thus neurons with inconsistent tilt preferences across mean depths will have weak selectivity in the grand tilt tuning curve (Nguyenkim and DeAngelis 2003), although this seldom occurs. For each neuron, we quantified strength of tuning by a tilt discrimination index (TDI):

$$\text{TDI} = \frac{R_{\text{max}} - R_{\text{min}}}{R_{\text{max}} + R_{\text{min}} + 2\sqrt{\text{SSE}/(N - M)}} \quad (3)$$

where R_{max} and R_{min} denote the mean firing rates of the neuron (from the grand tuning curve) at the tilt angles that elicited maximal and minimal responses, respectively. 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.

To test whether the responses of MT neurons to combinations of tilt cues can be predicted from the responses to individual cues, we examined whether the combined response could be well approximated as a weighted linear sum of individual cue responses:

$$r_{\text{combined}}(\theta) = w_{\text{velocity}}r_{\text{velocity}}(\theta) + w_{\text{disparity}}r_{\text{disparity}}(\theta) + w_{\text{texture}}r_{\text{texture}}(\theta) + C \quad (4)$$

where θ represents stimulus orientation (tilt) and $r_{\text{combined}}(\theta)$, $r_{\text{velocity}}(\theta)$, $r_{\text{disparity}}(\theta)$, and $r_{\text{texture}}(\theta)$ denote tilt tuning curves for the four cue conditions. The mean response across tilts was subtracted from each of the four tuning curves before fitting such that the model tries to fit the response modulation in the combined condition based on the response modulations in the single-cue conditions (see DISCUSSION). In Eq. 4, w_{velocity} , $w_{\text{disparity}}$, and w_{texture} denote the weights that a neuron applies to each of the three cues, and C is a constant free parameter. Because the stimulus in the texture condition was rendered with 0% motion coherence (unlike the other conditions), the magnitude of the weights for the texture cue may not be easily comparable to the other cues. However, given that texture weights were found to be broadly distributed around zero, this does not substantially limit our conclusions.

To examine how sensitive model predictions were to the presence of each individual cue, we also modeled combined responses as a weighted sum of two of the three cues. All combinations of two-cue models (velocity-disparity, velocity-texture, and disparity-texture) were tested. The statistical significance of the improvement in fit of the three-cue model over the two-cue models was assessed using a sequential F -test. A significant outcome of the sequential F -test ($P < 0.05$) indicates that the three-cue model fits the data significantly better than a particular two-cue model.

To test whether nonlinear interactions between different gradient cues could improve the fits achieved by the model, four nonlinear terms were added:

$$r_{\text{combined}}(\theta) = w_{\text{velocity}}r_{\text{velocity}}(\theta) + w_{\text{disparity}}r_{\text{disparity}}(\theta) + w_{\text{texture}}r_{\text{texture}}(\theta) + w_{\text{velocity,texture}}r_{\text{velocity}}(\theta)r_{\text{texture}}(\theta) + w_{\text{velocity,disparity}}r_{\text{velocity}}(\theta)r_{\text{disparity}}(\theta) + w_{\text{disparity,texture}}r_{\text{disparity}}(\theta)r_{\text{texture}}(\theta) + w_{\text{velocity,disparity,texture}}r_{\text{velocity}}(\theta)r_{\text{disparity}}(\theta)r_{\text{texture}}(\theta) + C \quad (5)$$

where $w_{\text{velocity,disparity}}$, $w_{\text{disparity,texture}}$, $w_{\text{velocity,texture}}$, and $w_{\text{velocity,disparity,texture}}$ denote the weights of the nonlinear response terms. Three of the nonlinear terms correspond to pairwise products of responses to different cues, and the fourth term is a product of all three cues. A sequential F -test was again used to compare fits of the nonlinear model with those of the linear model.

We also compared a nonlinear power law model, without interaction terms (Britten and Heuer 1999), to the models described above, to assess whether an overall nonlinearity would help account for the data:

$$r_{\text{combined}}(\theta) = [w_{\text{velocity}}r_{\text{velocity}}(\theta)^n + w_{\text{disparity}}r_{\text{disparity}}(\theta)^n]^{1/n} + C \quad (6)$$

where θ represents stimulus orientation (tilt) and $r_{\text{combined}}(\theta)$, $r_{\text{velocity}}(\theta)$, and $r_{\text{disparity}}(\theta)$ denote tilt tuning curves. For this analysis, we did not include texture responses, since they account for little of the response in the combined condition. Because the power law model cannot operate on negative firing rates, the mean response was not subtracted from the tuning curves for these fits. Note that the goodness of fit of the power law model was compared with that of the linear and nonlinear models (Eqs. 4 and 5); to allow a fair comparison, those models were also fit to responses without mean response subtraction and without the terms involving the texture cue.

To test for clustering of tilt selectivity, we extracted multiunit (MU) responses from the digitized raw data and analyzed them in the same manner as we did for the single-unit (SU) data. MU responses were extracted from the raw neural signals that were digitized using Spike2 software by setting an amplitude threshold such that the spontaneous event rate for MU activity was 75 impulses/s greater than the spontaneous rate for SU activity. To make the MU signal independent from the SU activity, each SU spike was removed (off-line) from the MU

event train. The success of this manipulation was confirmed by computing cross-correlograms between the SU and MU spike trains (see Chen et al. 2008 and DeAngelis and Newsome 1999 for details).

RESULTS

We recorded from 156 neurons in 2 monkeys and successfully maintained single-unit isolation long enough to complete the experimental protocol for 96 neurons (see MATERIALS AND METHODS). Three-dimensional orientation selectivity was measured using random dot stimuli that depicted planar surfaces at various tilts and slants. Tilt is defined as the axis around which the plane is rotated away from frontoparallel, and slant is defined as the amount by which the plane is rotated (Fig. 1A). Tilt and slant could be specified by isolated gradients of texture, binocular disparity, or image velocity, and all three cues could also be presented together in the combined condition (Fig. 1B). Because tilt tuning for disparity gradients was

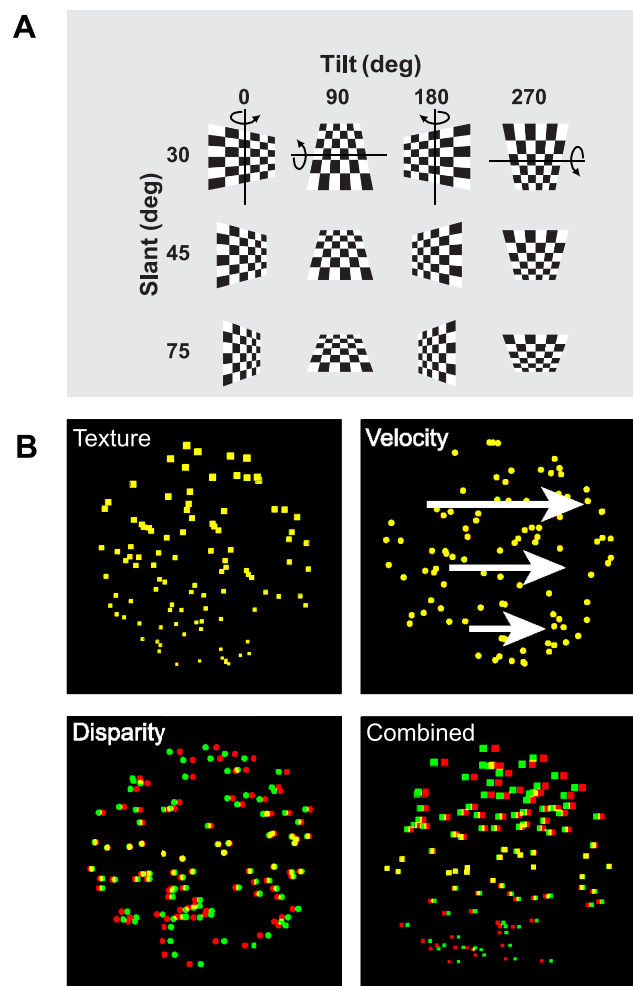


Fig. 1. A: schematic illustration of the 3-dimensional (3-D) orientation of planar surfaces, parameterized by the parameters tilt and slant. Four different tilts (0° , 90° , 180° , 270°) are shown for each of 3 different slants. Tilt refers to the axis around which the plane is rotated away from frontoparallel, and slant defines the amount by which the plane is rotated. B: schematic illustration of the 4 cue conditions used to measure tilt selectivity. In the texture condition, the stimulus incorporated gradients of element size and density. The velocity stimulus was defined by a gradient of speed of dot motion, and the disparity stimulus was defined by a gradient of horizontal disparity. In the combined condition, the stimulus contained congruent texture, velocity, and disparity gradients. Binocular disparities are represented as a red/green anaglyph.

previously found to be consistent across slants (Nguyenkim and DeAngelis 2003), we fixed the slant of the stimulus at 65° for most experiments reported here, except where specifically noted.

For each neuron, we measured tilt tuning for the four cue conditions described above. To control for potential artifacts of mis-centering the stimulus on the RF (Nguyenkim and DeAngelis 2003), we also presented each tilt stimulus at three different mean depths that were chosen to flank the peak of the frontoparallel depth tuning curve (see MATERIALS AND METHODS). Thus 12 tilt tuning curves were obtained for each neuron (4 cue conditions $\times$ 3 mean depths). Example data sets from three representative neurons are illustrated in Fig. 2. We quantified the strength of tilt tuning by calculating a tilt discrimination index (TDI; Eq. 3), which ranges from 0 to 1, with larger values indicating stronger selectivity. For the neuron in Fig. 2A, there was little tilt selectivity in the texture condition (TDI = 0.39), moderate tilt selectivity in the disparity condition (TDI = 0.66), and strong tilt tuning in the velocity condition (TDI = 0.83). Tilt preferences of this neuron in the disparity and velocity conditions were well matched, and thus tilt tuning was also robust and similar in the combined condition (TDI = 0.72). Within each cue condition, this neuron exhibited tilt selectivity that was fairly consistent across the three mean depths tested, suggesting that selectivity was not an artifact of mis-centering the stimulus on the RF (Nguyenkim and DeAngelis 2003) (but see Bridge and Cumming 2008).

The example neuron in Fig. 2B shows a qualitatively different pattern of results. This neuron shows clear tilt selectivity in the disparity (TDI = 0.58) and velocity conditions (TDI = 0.73), but the tilt preference differs markedly between these conditions. Clear tilt tuning is also observed in the combined

condition (TDI = 0.75); however, the tuning is slightly broader and the tilt preference is similar to the velocity preference but shifted slightly toward the disparity preference. Again, tilt tuning was weak in the texture condition (TDI = 0.44). As discussed further below, some neurons with discrepant tilt preferences for disparity and velocity cues exhibited tilt tuning in the combined condition that was dominated by either the disparity or the velocity cue (like that in Fig. 2B), whereas other neurons showed combined tuning that was intermediate.

The third example neuron, in Fig. 2C, shows robust tilt tuning in the velocity (TDI = 0.87) and texture conditions (TDI = 0.65) but little tuning in the disparity condition (TDI = 0.33). Tuning in the combined condition was robust (TDI = 0.79) and similar to that of the velocity condition. This neuron showed the strongest tuning for texture gradients that we observed in the population. Note that the mean firing rate of this neuron was substantially lower for texture than for other cues, perhaps because the texture gradient was presented without coherent motion (0% coherence, see MATERIALS AND METHODS).

Population summary of tilt selectivity. Using TDI as a metric of selectivity, we investigated quantitatively how tilt tuning depends on cue conditions across our population of neurons. Marginal histograms in Fig. 3 show distributions of TDI for each of the four cue conditions. Responses of each neuron were subjected to a two-way ANOVA (tilt $\times$ mean depth), and filled bars in the marginal histograms indicate neurons with significant tilt tuning (main effect of tilt, $P < 0.05$). Mean TDI values are 0.56, 0.53, 0.43, and 0.38 for the combined, velocity, disparity, and texture conditions, respectively, with the corresponding percentages of selective neurons being 90.1, 76.0, 70.0, and 58.0%. In terms of conjunctions of selectivities, 59%

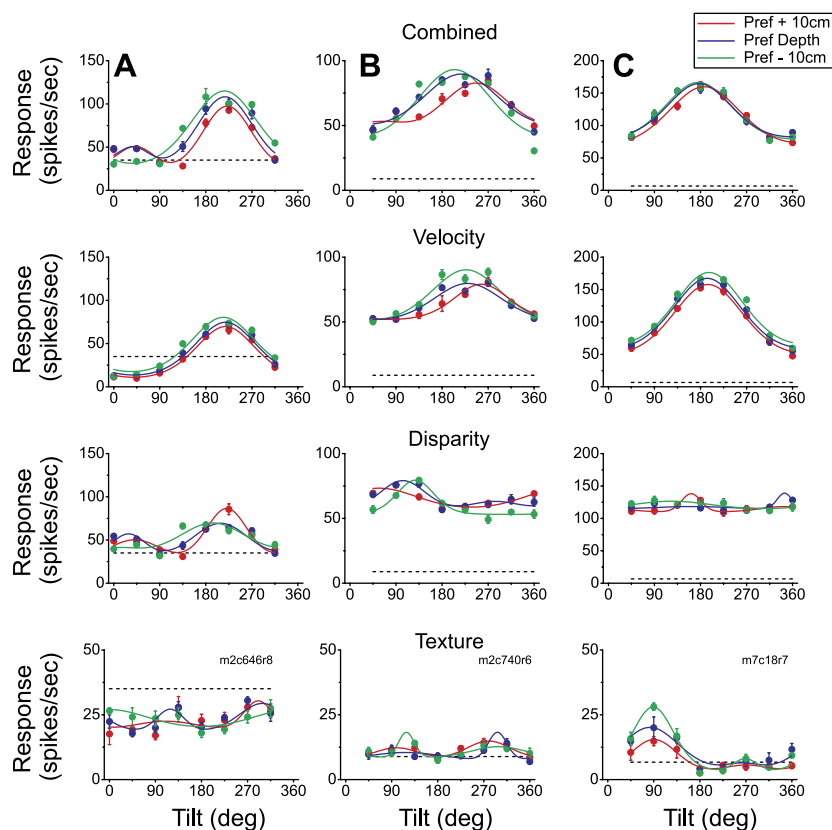
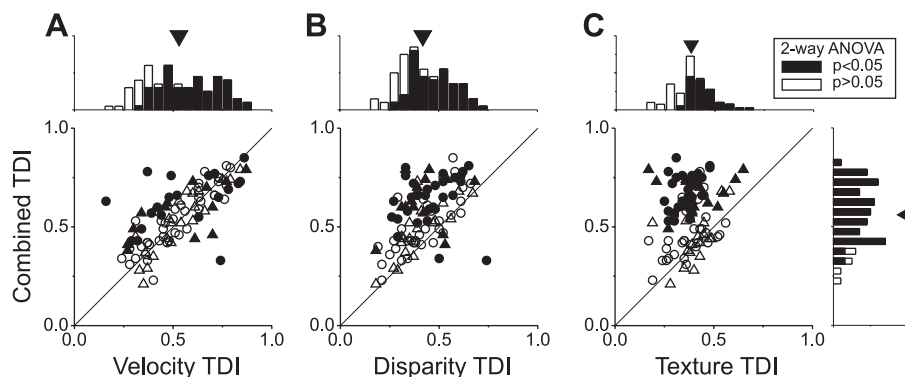


Fig. 2. A–C: tilt tuning curves for 3 representative neurons. Tuning curves at 3 different mean depths (color coded) are plotted for each of the 4 different stimulus conditions. From top to bottom, data from the combined, velocity, disparity, and texture conditions are shown. Solid lines show best fits of a wrapped Gaussian function. Dashed horizontal line indicates spontaneous activity. Error bars are SEs of the mean. Pref, preferred.

Fig. 3. Summary of the strength of tilt selectivity for the population of middle temporal area (MT) neurons. Marginal histograms show distributions of the tilt discrimination index (TDI) for each of the 4 cue conditions, with filled bars denoting a significant main effect of tilt (2-way ANOVA, $P < 0.05$). A: scatter plot of TDI for the combined vs. velocity conditions. Neurons with significantly different TDIs between the 2 cue conditions are plotted as filled symbols (bootstrap test, $P < 0.05$). Each datum represents 1 neuron, and data from monkey J and monkey M are shown as circles and triangles, respectively. B: scatter plot of TDI for the combined vs. disparity conditions. C: scatter plot of TDI for the combined vs. texture conditions.



of MT neurons showed significant tilt tuning in the combined, velocity, and disparity conditions, 15% were selective in only the combined and velocity conditions, 10% were selective in only the combined and disparity conditions, and 2% showed selectivity in only the combined and texture conditions.

We next compared TDI values for the combined condition with those from each of the single-cue conditions (scatter plots of Fig. 3). TDI in the velocity condition was strongly correlated with that in the combined condition (Fig. 3A; $r = 0.72$, $P < 0.001$, $N = 96$). For one-third of the neurons (33/96), TDI was significantly different between the velocity and combined conditions (bootstrap test, $P < 0.05$; filled symbols in Fig. 3A), with the combined TDI being larger for 22/33 neurons. Overall, the average TDI for the velocity condition was slightly, but significantly, less than that for the combined condition (paired t -test, $P = 0.025$, $N = 96$). Note, however, that more data points are above the unity-slope diagonal in Fig. 3A when TDI for the velocity condition is low. We separated the neurons into two groups according to the median TDI in the velocity condition, and we found that the increase in TDI in the combined condition was highly significant for the group of neurons with weaker tilt tuning in the velocity condition (paired t -test, $P < 0.001$). In contrast, for the other half of neurons with strong velocity-based selectivity, mean TDI values were not significantly different between the combined and velocity conditions (paired t -test, $P > 0.1$). This indicates that the addition of disparity and texture cues to a velocity gradient stimulus substantially improves tilt selectivity when the velocity cue by itself does not produce very strong selectivity.

Comparing the combined and disparity conditions, we found that TDI values were again significantly correlated across conditions (Fig. 3B; $r = 0.52$, $P < 0.001$, $N = 96$). Moreover, TDI values in the combined condition were systematically greater than those for the disparity condition (paired t -test, $P < 0.001$, $N = 96$), indicating that the addition of velocity and texture cues to disparity gradients substantially improved tilt selectivity. For the texture condition, TDI values were only marginally correlated with those in the combined condition (Fig. 3C; $r = 0.21$, $P = 0.038$, $N = 96$), and the average TDI value was much higher in the combined condition (paired t -test, $P < 0.001$).

To examine the possibility that differences in tilt selectivity across stimulus condition might be confounded with differences in response strength, Fig. 4 shows TDI plotted as a function of firing rate (square root transformed to reduce skew in the distribution). As illustrated by the example neuron of Fig. 2C, firing rates were generally lower in the texture con-

dition than in the other conditions. However, there was no significant correlation between TDI and firing rate across all conditions ($P = 0.84$, main effect of firing rate, ANCOVA), and there was no significant interaction between firing rate and stimulus condition ($P = 0.99$, ANCOVA), indicating that the relationship between TDI and firing rate did not differ significantly across stimulus conditions. Thus it is clear that the lower average TDI values observed in the texture and disparity conditions were not simply the result of weaker responses in these conditions.

Together, these results suggest that tilt selectivity in area MT increases as multiple cues to surface orientation are combined, with velocity and disparity gradient cues providing the strongest inputs and texture having a weaker contribution. Cue combination only fails to improve selectivity when neurons have very strong tuning for the velocity stimulus by itself, suggesting that tilt tuning is dominated by velocity gradients for these neurons. Data from the two monkeys were consistent (circles and triangles in Fig. 3) and have been combined in subsequent analyses.

One might expect that neurons with congruent tilt preferences in the single-cue conditions would contribute most to the enhancement of TDI values in the combined condition. Indeed, we found that the difference in TDI between the combined and velocity conditions was negatively correlated with the absolute difference in tilt preference between the disparity and velocity conditions ($r = -0.32$, $P = 0.02$, $N = 52$). Thus neurons with congruent tilt preferences for disparity and velocity cues showed the largest increases in TDI in the combined condition relative to the velocity condition. A similar trend was present

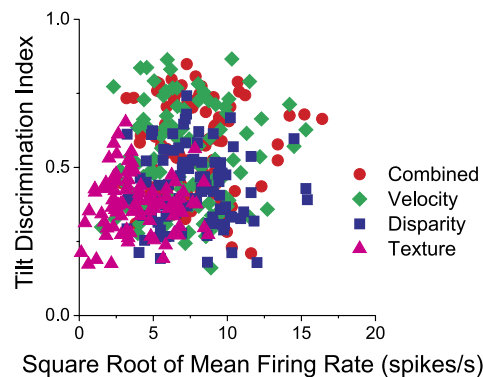


Fig. 4. Relationship between tilt selectivity (TDI) and firing rate for each cue condition. The square root of mean firing rate is plotted on the abscissa to roughly normalize the distribution of response rates. Stimulus conditions are represented as different colors.

for the difference in TDI between the combined and disparity conditions, but the effect did not quite reach significance ($r = -0.24$, $P = 0.08$, $N = 52$).

Previously, tilt selectivity based on velocity gradients was described in MT by Treue and Andersen (1996) and Xiao et al. (1997), and Xiao et al. reported a robust correlation between tilt selectivity and surround suppression. In contrast, tilt selectivity based on disparity gradients was not found to correlate with the strength of surround suppression in a previous study (Nguyenkim and DeAngelis 2003), and we did not find any significant correlations between tilt selectivity and surround suppression in the present study for any of the stimulus conditions (combined: $r = -0.028$, $P = 0.79$; velocity: $r = 0.0038$, $P = 0.97$; disparity: $r = -0.14$, $P = 0.18$; texture: $r = -0.05$, $P = 0.63$, $N = 96$). Thus it does not appear that surround suppression is generally linked to tilt selectivity. Although the reasons for the difference between our results and those of Xiao et al. (1997) are not clear, one potentially relevant methodological difference is that Xiao et al. measured size tuning curves monocularly, whereas we measured size tuning using stereoscopic stimuli in our experiment.

Tilt preferences across cue conditions. We now consider the similarity of tilt preferences across cue conditions. We estimated the tilt preference of each neuron by fitting its tuning curve with a wrapped Gaussian function (see MATERIALS AND METHODS). Because a small percentage of neurons show bimodal tilt tuning curves with two peaks roughly 180° apart, responses were fit with the sum of two wrapped Gaussian functions having peaks that were constrained to lie 180° apart. To obtain a single tilt preference for each neuron in each cue condition, tuning curves were fit after averaging responses across the three mean depths. The tilt preference was defined as the tilt for which the fitted curve had its largest peak (but see MATERIALS AND METHODS for calculation of differences in tilt preferences when tuning is bimodal). Figure 5A compares tilt preferences for the combined and velocity conditions, with each symbol representing a neuron that showed significant tuning in both cue conditions. Because tilt angle is a circular variable, data points in this scatter plot are constrained to lie

within the dashed lines that define a 180° difference between tilt preferences for the two cues. There is a strong correlation between tilt preferences in the combined and velocity conditions (circular correlation coefficient, $r = 0.76$, $P < 0.001$, $N = 61$), and most cells (72%) have tilt preferences that differ by $<30^\circ$ between conditions (Fig. 5D). A similar result was observed for the disparity cue. Tilt preference in the disparity condition was significantly correlated with that in the combined condition (circular correlation coefficient, $r = 0.51$, $P < 0.001$, $N = 60$), and 57% of cases showed tilt preferences within 30° (Fig. 5E). In Fig. 5, D and E, the distribution of differences in tilt preference was significantly different from uniform (Fig. 5D: Rayleigh test, $P < 0.001$, $N = 61$, Fig. 5E: $P < 0.001$, $N = 60$).

Perhaps surprisingly, tilt preferences in the velocity and disparity conditions were not significantly correlated with each other (circular correlation coefficient, $r = 0.22$, $P = 0.14$, $N = 52$; Fig. 5C), and the distribution of differences in preferred tilt between these conditions was not significantly different from uniform (Fig. 5F; Rayleigh test, $P = 0.20$, $N = 52$). Thus, although tilt preferences in the combined condition were correlated with those in the velocity and disparity conditions, there was no consistent relationship between tilt preferences based on velocity and disparity alone. In addition, tilt preferences in the texture condition were not significantly correlated with those for the other cue conditions (texture-combined: $r = 0.13$, $P = 0.38$, $N = 42$; texture-velocity: $r = 0.03$, $P = 0.88$, $N = 39$; texture-disparity: $r = 0.12$, $P = 0.46$, $N = 37$).

To better understand this somewhat puzzling pattern of results, we further analyzed the relationships between differences in tilt preference among cue conditions. When the difference in tilt preference between combined and velocity conditions is plotted as a function of the difference in preference between velocity and disparity conditions (Fig. 6A), three main groups of cells become apparent: those for which the combined tilt preference is consistent with both the velocity and disparity preferences (red symbols), those for which the combined tilt preference is dominated by velocity (green symbols), and those for which the combined preference is domi-

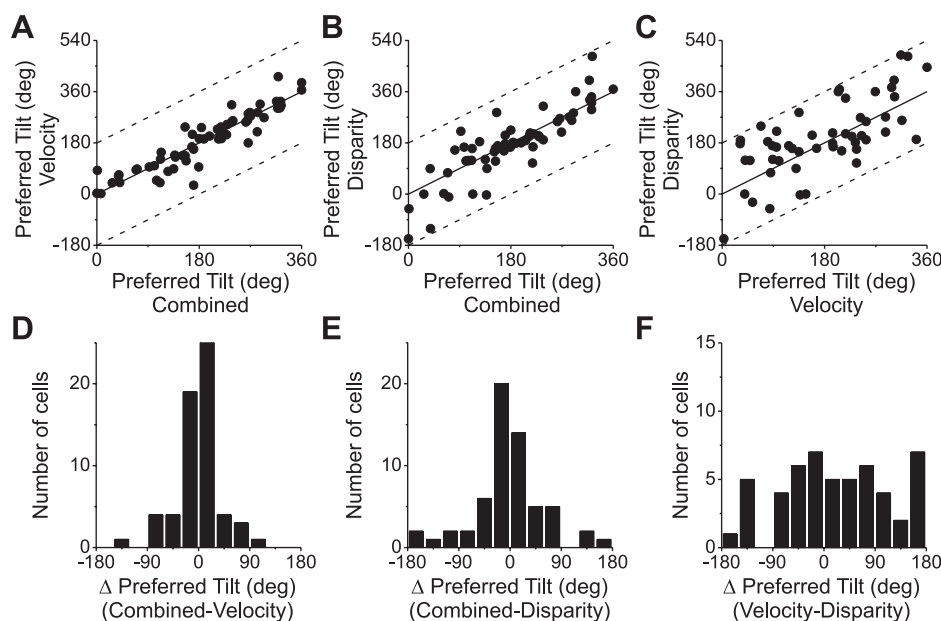


Fig. 5. Comparison of tilt preferences across stimulus conditions. A: tilt preferences for the velocity condition are plotted as a function of those for the combined condition. B: tilt preferences for the disparity condition are plotted against those for the combined condition. C: tilt preferences for the velocity condition are plotted against those measured in the disparity condition. Solid line indicates the unity-slope diagonal, whereas the 2 dashed diagonal lines represent boundaries corresponding to 180° differences between tilt preferences. D–F: distributions of differences (Δ) in tilt preference between the corresponding cue conditions in A–C are shown as histograms.

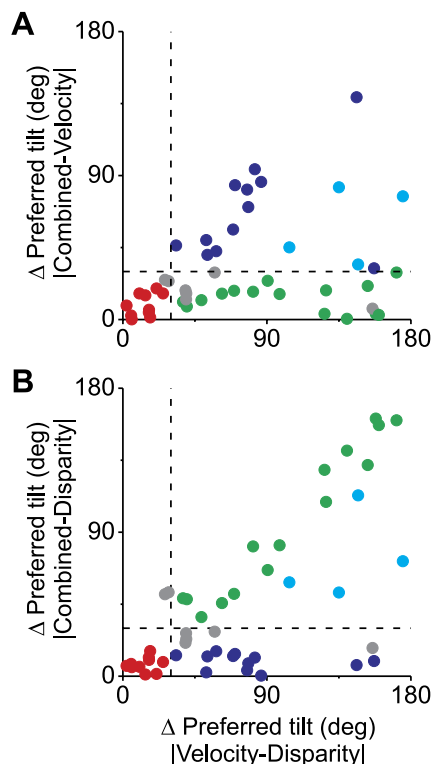


Fig. 6. Tilt preferences in the combined condition tend to be dominated by disparity or velocity. *A*: the absolute value of the difference in tilt preference between the combined and velocity conditions is plotted against the absolute difference in tilt preference between velocity and disparity conditions. Red symbols represent cells with similar tilt preferences (within 30°) across all 3 conditions. Green symbols denote cells with similar tilt preferences in the combined and velocity conditions. Blue symbols denote cells with similar tilt preferences in the combined and disparity conditions. Cyan symbols represent cells for which the tilt preferences in the velocity, disparity, and combined conditions all differ from each other by at least 30° . Gray symbols denote a handful of neurons that did not fit into any of these groups. *B*: analogous data are shown except that the vertical axis now represents the absolute value of the difference in tilt preference between the combined and disparity conditions.

nated by disparity (blue symbols). A complementary pattern of results is seen when the difference in tilt preference between combined and disparity conditions is plotted in a similar fashion (Fig. 6*B*). These results suggest that when disparity and velocity preferences for tilt are discrepant, many cells are dominated by one cue or the other, whereas a minority of neurons have combined preferences that are truly intermediate (cyan symbols).

Linear integration of tilt cues. Thus far, we have characterized tilt selectivity across cue conditions, but what is the mathematical rule by which MT neurons integrate multiple visual cues to represent tilt? Recently, Morgan et al. (2008) reported that visual and vestibular inputs related to self-motion are integrated in area MST by weighted linear summation. Moreover, theoretical studies have suggested that optimal cue integration can be achieved by linear weighting of inputs by single neurons (Fetsch et al. 2011; Ma et al. 2006). Thus we examined whether weighted linear summation could account for the combined responses to tilt cues. We fitted responses in the combined condition with a weighted linear sum of responses in the three single-cue conditions (Eq. 4, see MATERIALS AND METHODS). Data for all three mean depths were fitted simultaneously using a single set of parameters. Figure 7

shows fitting results for the same three example neurons that were shown in Fig. 2. For the cell in Fig. 7*A*, weights of the individual cue responses w_{velocity} , $w_{\text{disparity}}$, and w_{texture} were 0.80, 0.74, and -0.98 , respectively, and the predicted tuning curves matched the data well for all three mean depths ($R^2 = 0.95$). Note that the magnitude of the texture weight was large even though this neuron had little tilt selectivity in the texture condition. Because of this lack of texture selectivity, there is little to constrain the texture weight during the fit, and the resulting texture weight was only marginally different from zero ($P = 0.03$). This was the case for many neurons, as discussed further below. By comparison, w_{velocity} and $w_{\text{disparity}}$ were both significantly greater than zero ($P < 0.001$) for this neuron. For the neuron in Fig. 7*B*, which had rather discrepant tilt preferences for disparity and velocity cues (Fig. 2*B*), combined responses are also predicted well by the linear model ($R^2 = 0.90$), with cue weights w_{velocity} , $w_{\text{disparity}}$, and w_{texture} of 1.46, 0.8, and 0.59, respectively. Finally, for the neuron in Fig. 7*C*, which had clear tilt tuning in the velocity and texture conditions, predicted tilt tuning also matched the data well ($R^2 = 0.96$), with cue weights w_{velocity} , $w_{\text{disparity}}$, and w_{texture} of 0.79, 1.04, and 0.72, respectively. Although the texture and disparity weights were marginally significant for this neuron, the confidence intervals on the weights were an order of magnitude larger for disparity and texture than for velocity.

To quantify the results of the linear model fits, we compared the predicted tilt preference to the measured tilt preference by fitting wrapped Gaussian functions to both predicted and measured responses from the combined condition. Predicted and measured tilt preferences were strongly correlated (Fig. 8*A*; circular correlation coefficient, $r = 0.97$, $P < 0.001$, $N = 68$), and 98.5% of neurons showed measured and predicted tilt preferences that differed by $<30^\circ$ (Fig. 8*B*). Fits of the linear model were generally quite good, with 69% of cells having R^2 values >0.7 (median $R^2 = 0.80$; Fig. 8*C*).

Although linear model fits were generally quite good across the population, model fits are expected to be good when tilt preferences are similar in the disparity and velocity conditions such that all of the tilt tuning curves are fairly similar. A much more critical test of the model is to examine fits for neurons with discrepant tuning for disparity and velocity gradients. Figure 8*D* plots the difference in tilt preference between data and model fits against the difference in tilt preference between disparity and velocity conditions for the subset of neurons with significant tuning in both single-cue conditions. We found only a weak, marginally significant correlation between these variables (Spearman rank correlation, $r = 0.29$, $P = 0.05$, $N = 46$), and many cells with large discrepancies between disparity and velocity tilt preferences show little error in model predictions. There is a similarly weak correlation between the R^2 value of the model fits and difference in tilt preference between disparity and velocity conditions ($r = -0.29$, $P = 0.06$, $N = 46$; data not shown). Thus the linear model is quite successful in predicting tilt tuning in the combined condition, even when tuning differs greatly between the disparity and velocity conditions.

One might argue that any linear model with broad tuning curves could fit the combined responses well. For example, if velocity and disparity tuning curves were sinusoidal and out of phase by 90° , then a weighted sum might fit almost any combined tuning curve. To assess this possibility, we shuffled

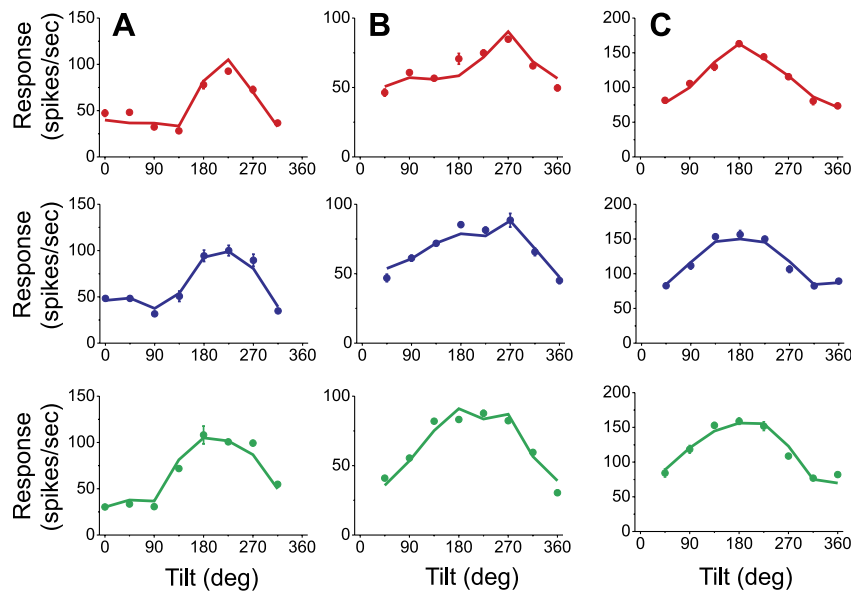


Fig. 7. Predictions of the linear weighted summation model are shown for the same 3 example neurons illustrated in Fig. 2. Each row represents 1 of the 3 mean depths tested for each neuron. Weights of the individual cues (w_{velocity} , $w_{\text{disparity}}$, w_{texture}) are 0.80, 0.74, and -0.98 , respectively, for the cell in A, 1.46, 0.80, and 0.59 for the cell in B, and 0.79, 1.04, and 0.72 for the cell in C. Curves represent model fits.

data across neurons and refit each cell's combined responses with a weighted sum of velocity and disparity tuning curves that were randomly picked from among all cells in the population. The goodness of fit obtained with the shuffled data sets (mean $R^2 = 0.35$) was significantly lower than the goodness of fit (mean $R^2 = 0.75$) obtained by the original analysis (paired t -test, $P < 0.0001$). Thus the structure of the single-cue tuning curves for a given neuron is an important factor for predicting combined responses.

Cue weights. Having established that the linear model provides an adequate description of combined responses to tilt cues, we can now use the weights of the linear fits to characterize the relative contributions of disparity, velocity, and

texture cues across our population of neurons. Histograms in Fig. 9, A–C, show the distributions of velocity, disparity, and texture weights for 87 neurons. Most neurons showed significant weights for velocity (68%; Fig. 9A) and disparity (70%; Fig. 9B), and virtually all of these weights were positive, as

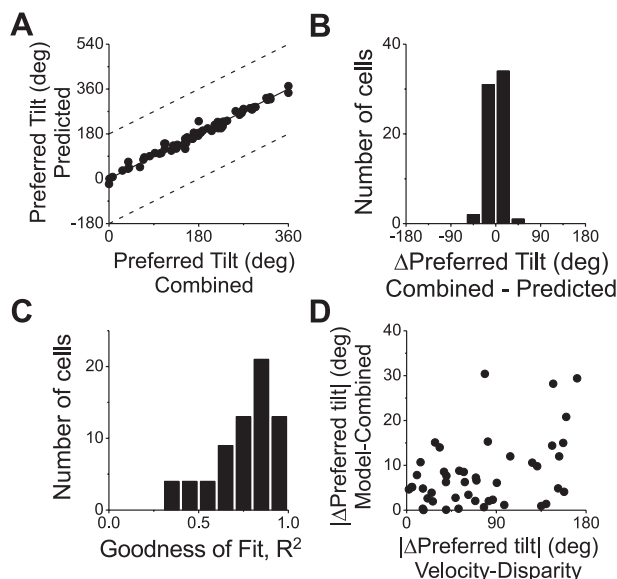


Fig. 8. Summary of linear model fits. A: comparison of predicted and measured tilt preferences in the combined condition. Format is the same as in Fig. 5, A–C. B: distribution of the difference in tilt preferences between measured and predicted responses for the combined condition. C: histogram of goodness of fit (R^2), with a median value of 0.80. D: the absolute value of the difference in tilt preference between measured and predicted responses is plotted as a function of the absolute difference in tilt preference between the velocity and disparity conditions.

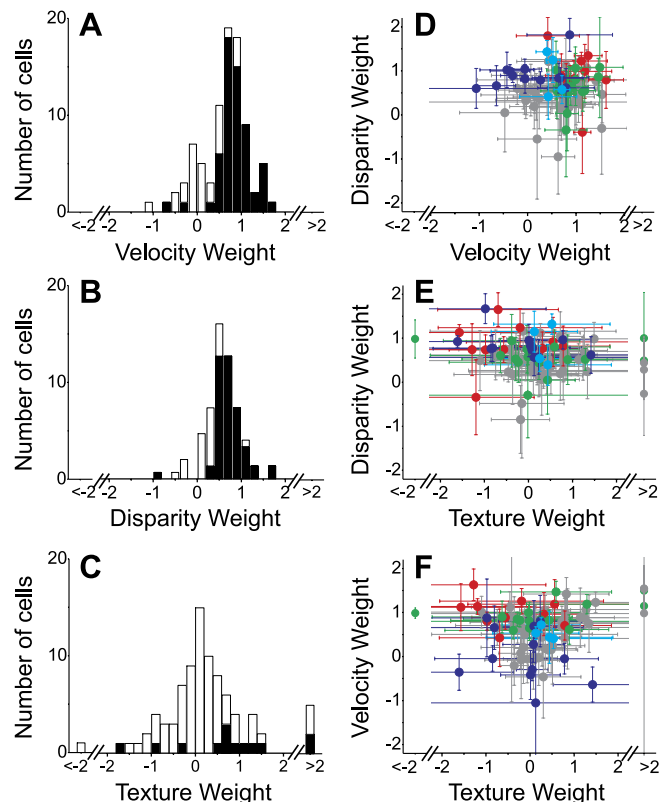


Fig. 9. Summary of cue weights from fits of the linear model. A–C: distributions of weights are shown for each individual cue. Filled bars indicate weights that are significantly different from zero. D: disparity weight is plotted against velocity weight for each neuron. Error bars represent 95% confidence intervals. Cells are color coded according to the similarity of tilt preferences across conditions, as defined in Fig. 6. E: disparity vs. texture weights. F: texture vs. velocity weights.

expected (filled bars). In contrast, only 15% of cells showed a weight for the texture cue that was significantly different from zero (Fig. 9C), and these weights were sometimes negative. Note that even large texture weights (near ± 1) were often not significant (open bars), indicating that the weak texture selectivity did not contribute much to combined responses for most neurons.

Relationships of weights across cue conditions are summarized in Fig. 9, D–F. There was no significant correlation between velocity and disparity weights (Fig. 9D; Spearman's rank correlation, $r = -0.09$, $P = 0.93$, $N = 87$). Notably, neurons with a combined tilt preference that is similar to the velocity preference (green symbols) tended to have velocity weights closer to 1 and disparity weights shifted toward zero. Similarly, cells with a combined tilt preference similar to that in the disparity condition (blue symbols) tended to have velocity weights closer to zero and disparity weights closer to 1. This pattern is consistent with the idea that combined responses of some MT neurons are dominated by the velocity gradient cue, whereas others are dominated by the disparity gradient cue (see also Fig. 6). No significant correlations were observed between disparity weights and texture weights (Fig. 9E; $r = -0.19$, $P = 0.08$, $N = 87$) or between velocity weights and texture weights (Fig. 9F; $r = 0.14$, $P = 0.18$, $N = 87$). Note that the 95% confidence intervals (error bars) on texture weights are generally much larger than for the other two cues. These results suggest that either velocity or disparity cues tend to dominate tilt selectivity in MT, as addressed further below.

Model-based summary of cue contributions. If the texture cue does not contribute strongly to responses in the combined condition, then eliminating the texture cue from linear model fits should have little or no effect. To examine this, we fitted the data with another weighted linear model in which the texture weight was eliminated. As shown in Fig. 10A, the goodness of fit (R^2) of the velocity + disparity model was very similar overall to that of the full model. Only 18% of neurons showed a significantly worse fit when texture was removed from the model (solid symbols, sequential F -test, $P < 0.05$), and the difference in R^2 was relatively small in almost all cases. Mean values of R^2 were 0.76 and 0.72 for the velocity + disparity + texture and velocity + disparity models, respectively (Fig. 10D, left 2 bars), and this difference was significant (paired t -test, $P < 0.001$, $N = 74$).

By comparison, removing either the disparity or velocity term from the linear model impairs the predictions to a much greater extent. When the velocity cue was omitted from the model (disparity + texture model, Fig. 10B), the mean R^2 value fell from 0.76 to 0.42, and this difference was highly significant (paired t -test, $P < 0.001$, $N = 74$). Similarly, removing the disparity cue (velocity + texture model, Fig. 10C) reduced the mean R^2 value to 0.53, an effect that was also highly significant (paired t -test, $P < 0.001$, $N = 74$). As summarized in Fig. 10D, this analysis demonstrates that linear combinations of velocity and disparity responses generally provide a good description of tilt tuning in the combined condition, with relatively little explanatory power gained by including the texture cue. Overall, the velocity + texture model produced marginally better fits than the disparity + texture model (paired t -test, $P = 0.04$, $N = 74$), suggesting that velocity contributes slightly more strongly than disparity overall.

Although responses of many cells are well described by the simple linear model, those of a minority were not. To explore whether nonlinear interactions are needed to predict combined

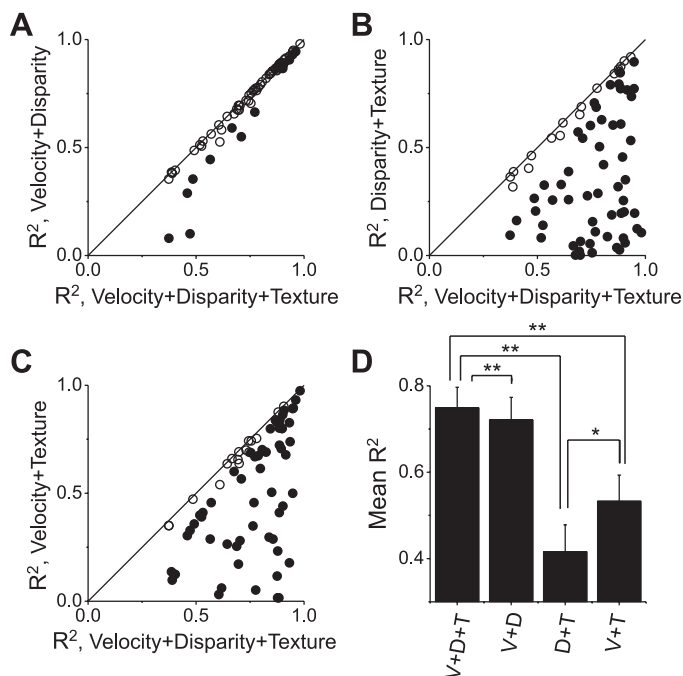


Fig. 10. Comparison of model fits that incorporate 2 vs. 3 cue weights. A: goodness of fit (R^2) for a 2-cue model including disparity and velocity is plotted against R^2 for the 3-cue model (velocity, disparity, and texture). Filled symbols indicate neurons for which the 3-cue model fits significantly better than the 2-cue model (sequential F -test, $P < 0.05$). B: goodness of fit for a 2-cue model including disparity and texture vs. the 3-cue model. C: goodness of fit for a 2-cue model including velocity and texture vs. the 3-cue model. D: mean R^2 values for the 3-cue model and each of the 2-cue models. V, velocity; D, disparity; T, texture.

responses for these cells, we compared the goodness of fit between the linear model and a nonlinear model that included additional terms consisting of products of velocity, disparity, and texture responses (see MATERIALS AND METHODS). Although the nonlinear model has twice as many parameters as the linear model, it provided significantly better fits for only 14% of neurons (solid symbols in Fig. 11; sequential F -test, $P < 0.05$). Overall, the mean R^2 value increased from 0.76 to 0.82 when the nonlinear terms were added, a difference that was significant (paired t -test, $P < 0.001$). The difference in R^2 between the linear and nonlinear models was not significantly different (Wilcoxon rank sum test, $P = 0.1$) for neurons with similar tilt preferences in the velocity and disparity conditions (red symbols) and neurons with discrepant preferences (blue symbols). Thus the nonlinear model does not appear to provide a greater benefit for neurons with mismatched tuning in the velocity and disparity conditions.

We also considered a nonlinear power law model that does not involve interactions among responses to the different stimulus conditions (Britten and Heuer 1999). To directly compare the power law model with the other models, we fitted all models to the tuning curves without subtracting the mean firing rate (since the power law model cannot operate on negative firing rates). In addition, texture responses were excluded from all models in this comparison, given that texture contributes little to combined responses overall. Although the power law model has one more parameter than the linear model, the average R^2 was not significantly different between these two models (paired t -test, $P = 0.11$). In contrast, the nonlinear

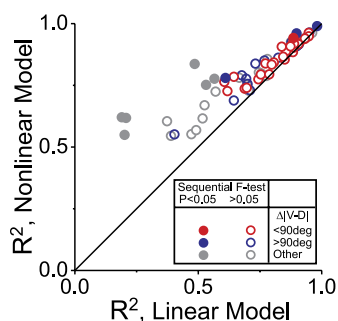


Fig. 11. Comparison of goodness of fit (R^2) values between the linear model and a nonlinear model. Filled symbols indicate neurons for which the nonlinear model fits significantly better than the linear model (sequential F -test, $P < 0.05$). Blue symbols indicate cells that have discrepant tilt preferences between velocity and disparity ($\Delta IV-DI > 90^\circ$), whereas red symbols denote cells that have similar preferences ($\Delta IV-DI < 90^\circ$). Gray symbols represent cells that lack significant tilt tuning in either the velocity or disparity condition (ANOVA, $P > 0.05$), or cells for which the model fit was not good ($R^2 < 0.6$).

model involving interaction terms (Eq. 5) produced an average R^2 value significantly greater than that of the power law model (paired t -test, $P < 0.001$).

These analyses indicate that a combination of 3-D surface orientation cues by MT neurons is reasonably well described by weighted linear summation.

Slant dependency of tilt tuning. Thus far, we have examined tilt tuning for a fixed slant (65°). We previously reported that tilt selectivity in response to disparity gradients was weak for small slants and grew monotonically with increasing slant, whereas tilt preferences remained similar as a function of slant (Nguyenkim and DeAngelis 2003). Here, we examined the effect of slant on tilt selectivity defined by other gradient cues, as summarized in Fig. 12A. In all cue conditions, mean TDI increased with slant ($P < 0.001$, main effect of slant, ANCOVA). There was also a significant interaction between cue type and the effect of slant ($P < 0.01$, ANCOVA), consistent with different slopes of the data in Fig. 12A across cue conditions.

We also examined the effect of slant on tilt preference, combining data across cue conditions to gain statistical power. Comparing slants of 25° and 45° , we found that tilt preferences were strongly correlated (circular-circular correlation, $r = 0.93$, $P < 0.001$, $N = 20$) and seldom differed by more than 45° from each other (Fig. 12B; Rayleigh test, $P < 0.001$). Similarly, when comparing slants of 45° and 65° , we again found a strong correlation ($r = 0.94$, $P < 0.001$, $N = 41$) with close agreement between tilt preferences across slants (Fig. 12C; Rayleigh test, $P < 0.001$). Thus, in all cue conditions, MT neurons show the most robust tilt tuning for large slants.

Comparison of multiunit and single-unit selectivity. Although 3-D surface orientation selectivity has been described in multiple visual areas (Hegde and Van Essen 2005; Nguyenkim and DeAngelis 2003; Srivastava et al. 2009; Sugihara et al. 2002; Taira et al. 2000; Treue and Andersen 1996; Tsutsui et al. 2001, 2002; Xiao et al. 1997), little is known about whether neurons are organized into functional clusters (e.g., columns) according to their tilt or slant tuning. Here we analyzed the tilt selectivity of MU activity and compared it with selectivity of SU activity from the same recording sites. Figure 13, A–D, compares TDI values for MU and SU responses across the four cue conditions. Significant correlations

between MU and SU TDIs were observed for the combined condition ($r = 0.39$, $P < 0.001$, $N = 92$), velocity condition ($r = 0.42$, $P < 0.001$, $N = 92$), and disparity condition ($r = 0.37$, $P < 0.001$, $N = 92$). In contrast, there was no significant correlation between MU and SU TDIs in the texture condition ($r = 0.11$, $P = 0.31$, $N = 92$), which may simply reflect the weak selectivity of MT neurons for texture gradients.

If tilt tuning is clustered in MT, then we would also expect to see similar tilt preferences for MU and SU activity. Indeed, the distribution of differences in tilt preference between MU and SU responses showed a clear peak around zero for all four cue conditions (Fig. 13, E–H). These distributions were significantly nonuniform in all four cases (Rayleigh test; combined: $P < 0.001$, $N = 59$; velocity: $P < 0.001$, $N = 51$; disparity: $P < 0.01$, $N = 31$; texture: $P < 0.001$, $N = 23$). Together, these data suggest that surface orientation-selective MT neurons are clustered together based on their tilt selectivity. Although this analysis does not establish a columnar architecture, it does suggest that MT neurons are organized systematically based on selectivity for 3-D orientation.

DISCUSSION

This study demonstrates that single MT neurons signal 3-D surface orientation (tilt and slant) via selectivity for velocity and disparity gradients, and to a lesser degree, for texture gradients. Selectivity is similar in multiunit activity recorded simultaneously, suggesting that 3-D surface orientation tuning

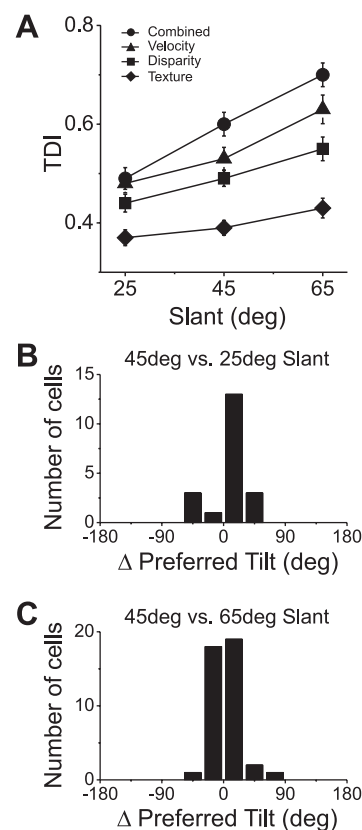
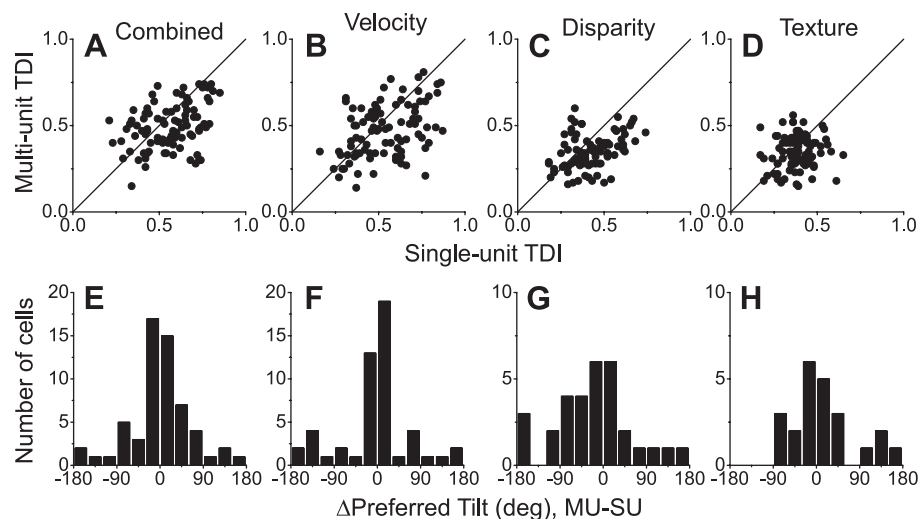


Fig. 12. Slant dependence of tilt tuning. A: mean TDI values are plotted as a function of slant for each of the 4 cue conditions. B: distribution of the difference in tilt preferences between slant values of 25° and 45° . The average difference is 5.7° . C: distribution of the difference in tilt preference between slant values of 45° and 65° , with a mean value of 3.5° .

Fig. 13. Comparison of tilt selectivity between multiunit (MU) and single-unit (SU) responses. *A–D*: TDI values for MU responses are plotted against the corresponding values for SU responses for each of the 4 cue conditions. *E–H*: distributions of the difference in tilt preference between MU and SU responses for each of the 4 cue conditions: combined, velocity, disparity, and texture (left to right).



is clustered in MT. Tilt selectivity is generally enhanced when multiple gradient cues are presented together, indicating that MT may integrate cues to represent surface structure with greater fidelity. This occurs despite the fact that tilt preferences for disparity and velocity gradients are poorly correlated overall. In addition, we found that responses to the combined stimulus, which contains all three gradient cues, are well approximated by a linear weighted sum of responses to the individual cues, with disparity and velocity gradients weighted most heavily. This study provides the first systematic examination of neural integration of multiple gradient cues to surface orientation and suggests that area MT contains an early multicue representation of 3-D surface structure.

Coding of surface orientation based on multiple cues. The ability of humans to integrate multiple visual cues to improve performance in surface orientation or shape discrimination tasks has been well studied psychophysically (Cumming et al. 1993; Cutting and Millard 1984; Hillis et al. 2004; Jacobs 1999; Johnston et al. 1994; Johnston et al. 1993; Knill 2007; Knill and Saunders 2003; Rogers and Graham 1982; Young et al. 1993), but the neural basis for these effects has remained unclear. Although several studies have established that neurons in both the dorsal and ventral streams are selective for the 3-D orientation of planar surfaces (Hegde and Van Essen 2005; Janssen et al. 2000; Nguyenkim and DeAngelis 2003; Sugihara et al. 2002; Taira et al. 2000; Treue and Andersen 1996; Tsutsui et al. 2002; Xiao et al. 1997), most of these studies have examined selectivity to only a single cue type. One important issue is whether tuning for 3-D surface orientation is consistent across cues, and only a few studies, including ours, have addressed this issue. In area CIP, Tsutsui et al. (2002) measured tilt tuning in response to texture gradients and disparity gradients, and they found that the majority of cells had tilt preferences for texture and disparity that were matched to within 45°. Although some neurons showed large discrepancies, disparity and texture preferences for tilt were strongly correlated. In IT cortex, Liu et al. (2004) also examined tilt tuning based on disparity and texture gradients and found a very similar result. Disparity and texture preferences were strongly correlated and generally matched within 45°, with occasional large discrepancies. Thus, at advanced stages along the dorsal and ventral processing streams, neurons have consistent selectivity for

multiple gradient cues, although it should be noted that neither Tsutsui et al. (2002) or Liu et al. (2004) measured responses to combinations of disparity and texture gradients.

In area MT, we found that tilt preferences for disparity and velocity are frequently mismatched, with no significant overall correlation between preferences for the two cues. The potential functional role of neurons with mismatched tilt preferences for disparity and velocity cues is currently not clear. Thus our results may imply that area MT is an early stage in the processing of 3-D orientation cues and that further processing is needed to achieve greater cue invariance. It should be noted, however, that Tsutsui et al. (2001) did not find a significant correlation between tilt preferences in CIP when they compared tilt tuning based on disparity gradients and figural perspective cues. Thus, even in CIP, tilt selectivity does not appear to be aligned for all cues.

A second important issue involves responses to combinations of 3-D orientation cues. Only one previous study, in area CIP (Tsutsui et al. 2001), has compared responses to individual gradient cues with responses to combinations of those cues. Tsutsui et al. (2001) reported that most neurons with tilt tuning in the combined condition also showed selectivity for disparity and perspective cues alone. Combined responses were generally enhanced relative to the single-cue conditions, but an analysis of signal to noise across conditions was not undertaken. Our TDI data show that MT neurons generally have a greater capacity to discriminate between different tilts when gradient cues are combined, relative to single-cue conditions. Thus, despite the fact that tilt preferences for disparity and velocity are frequently misaligned, our data suggest that a population of MT neurons may allow greater discriminability of tilt during cue combination.

Our results suggest that area MT makes a modest contribution to representations of 3-D surface orientation based on texture gradient cues. Although roughly one-half of MT neurons showed significant tilt selectivity, the tuning was often quite weak. As a result, texture contributed little to model fits of the combined responses (Fig. 10). Could the weak tilt tuning that we observed in the texture condition be explained simply because texture gradients are a weaker cue to surface orientation than disparity or velocity gradients? Although we did not train our animals to discriminate surface orientation in this

study, the available human psychophysics literature suggests strongly that texture is not simply a weak cue in our stimuli. Two psychophysical studies of slant discrimination have shown that slant sensitivity to texture gradients was two- to threefold greater than that for disparity gradients when stimulus parameters were comparable to ours (base slants of 60–70° and a viewing distance near 57 cm) (Hillis et al. 2004; Knill and Saunders 2003). In these studies, texture gradients were based on Voronoi patterns, whereas our texture stimulus was a random element stimulus with square elements. However, other psychophysical studies have indicated that sensitivity to slant in texture gradient stimuli is generally similar across texture types, especially when the base slant is large (Rosas et al. 2004; Saunders and Backus 2006). Slant discrimination thresholds in response to random dot stimuli were approximately the same as those for Voronoi patterns when discrimination was performed around a base slant comparable to that of our stimuli (Rosas et al. 2004). Therefore, we conclude that it is very unlikely that the weak tilt tuning that we observed in the texture condition for MT neurons was a reflection of the stimulus.

Rather, it seems very likely that neurons in other visual areas carry substantially greater information about surface orientation based on texture gradients. For example, CIP neurons appear to show tilt tuning for texture gradients that is quite robust, compared with disparity gradients (Tsutsui et al. 2002). This suggests that texture responses in area CIP do not arise through inputs from MT, but rather depend on inputs from other areas such as V3A (Katsuyama et al. 2010; Nakamura et al. 2001). Thus conducting similar experiments in area V3A is likely to be of considerable interest.

Linear weighting of surface orientation cues. Our study is the first to consider the mathematical rule by which neurons integrate different cues to 3-D surface orientation, and little is known generally about how neurons combine multiple cues to represent sensory variables. Morgan et al. (2008) described the first direct measurements of the neural cue integration rule, in the context of multisensory (visual-vestibular) integration in area MSTd. They reported that multisensory responses were well approximated by a weighted linear sum of single-cue responses (see also Fetsch et al. 2011). In addition, Ma et al. (2006, 2008) demonstrated that populations of neurons with Poisson-like spiking statistics can achieve optimal (Bayesian) cue integration by simply summing responses to the individual cue inputs.

Consistent with the results of Morgan et al. and the theory of Ma et al., we found that weighted linear summation provided a good description of combined responses to disparity, velocity, and texture gradients (Fig. 8). Moreover, little predictive power was gained by incorporating nonlinear response terms into the model (Fig. 11). One caveat to this finding is that cue-conflict stimuli were not employed in these experiments, unlike the study of Morgan et al. (2008). Thus, for neurons with congruent tilt preferences in the disparity and velocity conditions, it may be trivial that combined responses are well predicted by a linear weighted sum, as long as the tuning width of responses to disparity and velocity cues are similar. Thus the critical test of the linear model in this study involves neurons with discrepant tilt preferences for disparity and velocity gradients. Crucially, the linear model provided good predictions of combined responses for these neurons as well. There was little dependence of goodness of fit on the difference in tilt preference between disparity and velocity conditions, and model

predictions were generally good even when the tilt preference in the combined condition was intermediate.

Note that we subtracted the mean response (across tilts) from the tuning curve in each cue condition before performing the model fits. Hence, our linear model was required to fit the response modulation in the combined condition with a linear function of the response modulations in the single-cue conditions, but the model was not required to account for the mean response in the combined condition. In general, we observed (e.g., Fig. 2) that the mean response across tilts in the combined condition was not much greater than the mean response in the single-cue conditions. Indeed, the average sum of velocity and disparity weights is 1.38, indicating that MT neurons integrate velocity and disparity cues in a subadditive manner. This might result from the operation of some form of response normalization (Britten and Heuer 1999; Busse et al. 2009; Carandini et al. 1997; Heeger 1992; Ohshiro et al. 2011). Because our model did not incorporate a normalization operation, we did not require the model to fit the mean responses across cue conditions.

Although our results are broadly consistent with the theory of Ma et al. (2006, 2008), most MT neurons did not apply equal weights to their disparity and velocity inputs. Rather, the combined response of many neurons was dominated by either disparity or velocity when tilt preferences for the two cues were discrepant. This dominance was not simply determined by the relative strength of tuning in the disparity and velocity conditions, because there was no significant correlation between the ratio of disparity to velocity weights and the ratio of disparity to velocity TDI values from the single-cue responses ($r = -0.02$, $P = 0.88$, $N = 41$). Thus the factors that determine the relative dominance of disparity and velocity cues in the combined response are not clear.

In closing, our findings suggest that area MT plays a role in combining multiple gradient cues to surface orientation, but subsequent processing in downstream areas may be needed to achieve cue invariance. Although MT neurons appear to perform a simple weighted sum of inputs from different cues, the receptive field mechanisms that give rise to gradient selectivity are less clear. Selectivity for planar surface orientation could arise from many complex receptive field organizations, and the nature of the underlying basis functions is unknown. Thus future studies that directly map receptive field structures underlying 3-D orientation selectivity may be of considerable value.

ACKNOWLEDGMENTS

We thank Akiyuki Anzai for many helpful comments on the manuscript. Present address of J. D. Nguyenkim: Edge Business Innovations, 3303 South County Trail, East Greenwich, RI 02818.

GRANTS

This work was supported by National Eye Institute (NEI) Grant EY013644 and NEI Core Grant EY001319.

DISCLOSURES

No conflicts of interest, financial or otherwise, are declared by the author(s).

AUTHOR CONTRIBUTIONS

Author contributions: T.M.S., J.D.N., and G.C.D. conception and design of research; T.M.S. and J.D.N. analyzed data; T.M.S., J.D.N., and G.C.D.

interpreted results of experiments; T.M.S., J.D.N., and G.C.D. prepared figures; T.M.S. drafted manuscript; T.M.S., J.D.N., and G.C.D. edited and revised manuscript; T.M.S., J.D.N., and G.C.D. approved final version of manuscript; J.D.N. performed experiments.

REFERENCES

- Bradley DC, Chang GC, Andersen RA.** Encoding of three-dimensional structure-from-motion by primate area MT neurons. *Nature* 392: 714–717, 1998.
- Bridge H, Cumming BG.** Representation of binocular surfaces by cortical neurons. *Curr Opin Neurobiol* 18: 425–430, 2008.
- Britten KH, Heuer HW.** Spatial summation in the receptive fields of MT neurons. *J Neurosci* 19: 5074–5084, 1999.
- Busse L, Wade AR, Carandini M.** Representation of concurrent stimuli by population activity in visual cortex. *Neuron* 64: 931–942, 2009.
- Carandini M, Heeger DJ, Movshon JA.** Linearity and normalization in simple cells of the macaque primary visual cortex. *J Neurosci* 17: 8621–8644, 1997.
- Chen A, Gu Y, Takahashi K, Angelaki DE, DeAngelis GC.** Clustering of self-motion selectivity and visual response properties in macaque area MSTd. *J Neurophysiol* 100: 2669–2683, 2008.
- Chowdhury SA, DeAngelis GC.** Fine discrimination training alters the causal contribution of macaque area MT to depth perception. *Neuron* 60: 367–377, 2008.
- Cumming BG, Johnston EB, Parker AJ.** Effects of different texture cues on curved surfaces viewed stereoscopically. *Vision Res* 33: 827–838, 1993.
- Cutting JE, Millard RT.** Three gradients and the perception of flat and curved surfaces. *J Exp Psychol Gen* 113: 198–216, 1984.
- DeAngelis GC, Cumming BG, Newsome WT.** Cortical area MT and the perception of stereoscopic depth. *Nature* 394: 677–680, 1998.
- DeAngelis GC, Newsome WT.** Organization of disparity-selective neurons in macaque area MT. *J Neurosci* 19: 1398–1415, 1999.
- DeAngelis GC, Newsome WT.** Perceptual “read-out” of conjoined direction and disparity maps in extrastriate area MT. *PLoS Biol* 2: e77, 2004.
- DeAngelis GC, Uka T.** Coding of horizontal disparity and velocity by MT neurons in the alert macaque. *J Neurophysiol* 89: 1094–1111, 2003.
- Dodd JV, Krug K, Cumming BG, Parker AJ.** Perceptually bistable three-dimensional figures evoke high choice probabilities in cortical area MT. *J Neurosci* 21: 4809–4821, 2001.
- Fetsch CR, Pouget A, DeAngelis GC, Angelaki DE.** Neural correlates of reliability-based cue weighting during multisensory integration. *Nat Neurosci* 15: 146–154, 2011.
- Girshick AR, Banks MS.** Probabilistic combination of slant information: weighted averaging and robustness as optimal percepts. *J Vis* 9: 8 1–20, 2009.
- Heeger DJ.** Normalization of cell responses in cat striate cortex. *Vis Neurosci* 9: 181–197, 1992.
- Hegde J, Van Essen DC.** Role of primate visual area V4 in the processing of 3-D shape characteristics defined by disparity. *J Neurophysiol* 94: 2856–2866, 2005.
- Hillis JM, Watt SJ, Landy MS, Banks MS.** Slant from texture and disparity cues: optimal cue combination. *J Vis* 4: 967–992, 2004.
- Jacobs RA.** Optimal integration of texture and motion cues to depth. *Vision Res* 39: 3621–3629, 1999.
- Janssen P, Vogels R, Orban GA.** Macaque inferior temporal neurons are selective for disparity-defined three-dimensional shapes. *Proc Natl Acad Sci USA* 96: 8217–8222, 1999.
- Janssen P, Vogels R, Orban GA.** Three-dimensional shape coding in inferior temporal cortex. *Neuron* 27: 385–397, 2000.
- Johnston EB, Cumming BG, Landy MS.** Integration of stereopsis and motion shape cues. *Vision Res* 34: 2259–2275, 1994.
- Johnston EB, Cumming BG, Parker AJ.** Integration of depth modules: stereopsis and texture. *Vision Res* 33: 813–826, 1993.
- Katsuyama N, Yamashita A, Sawada K, Naganuma T, Sakata H, Taira M.** Functional and histological properties of caudal intraparietal area of macaque monkey. *Neuroscience* 167: 1–10, 2010.
- Knill DC.** Robust cue integration: a Bayesian model and evidence from cue-conflict studies with stereoscopic and figure cues to slant. *J Vis* 7: 5, 1–24, 2007.
- Knill DC, Saunders JA.** Do humans optimally integrate stereo and texture information for judgments of surface slant? *Vision Res* 43: 2539–2558, 2003.
- Krug K, Cumming BG, Parker AJ.** Comparing perceptual signals of single V5/MT neurons in two binocular depth tasks. *J Neurophysiol* 92: 1586–1596, 2004.
- Liu Y, Vogels R, Orban GA.** Convergence of depth from texture and depth from disparity in macaque inferior temporal cortex. *J Neurosci* 24: 3795–3800, 2004.
- Ma WJ, Beck JM, Latham PE, Pouget A.** Bayesian inference with probabilistic population codes. *Nat Neurosci* 9: 1432–1438, 2006.
- Ma WJ, Pouget A.** Linking neurons to behavior in multisensory perception: a computational review. *Brain Res* 1242: 4–12, 2008.
- Morgan ML, DeAngelis GC, Angelaki DE.** Multisensory integration in macaque visual cortex depends on cue reliability. *Neuron* 59: 662–673, 2008.
- Nadler JW, Angelaki DE, DeAngelis GC.** A neural representation of depth from motion parallax in macaque visual cortex. *Nature* 452: 642–645, 2008.
- Nadler JW, Nawrot M, Angelaki DE, DeAngelis GC.** MT neurons combine visual motion with a smooth eye movement signal to code depth-sign from motion parallax. *Neuron* 63: 523–532, 2009.
- Nakamura H, Kuroda T, Wakita M, Kusunoki M, Kato A, Mikami A, Sakata H, Itoh K.** From three-dimensional space vision to prehensile hand movements: the lateral intraparietal area links the area V3A and the anterior intraparietal area in macaques. *J Neurosci* 21: 8174–8187, 2001.
- Nguyenkim JD, DeAngelis GC.** Disparity-based coding of three-dimensional surface orientation by macaque middle temporal neurons. *J Neurosci* 23: 7117–7128, 2003.
- Nover H, Anderson CH, DeAngelis GC.** A logarithmic, scale-invariant representation of speed in macaque middle temporal area accounts for speed discrimination performance. *J Neurosci* 25: 10049–10060, 2005.
- Ohshiro T, Angelaki DE, DeAngelis GC.** A normalization model of multisensory integration. *Nat Neurosci* 14: 775–782, 2011.
- Palanca BJ, DeAngelis GC.** Macaque middle temporal neurons signal depth in the absence of motion. *J Neurosci* 23: 7647–7658, 2003.
- Rogers B, Graham M.** Similarities between motion parallax and stereopsis in human depth perception. *Vision Res* 22: 261–270, 1982.
- Rosas P, Wichmann FA, Wagemans J.** Some observations on the effects of slant and texture type on slant-from-texture. *Vision Res* 44: 1511–1535, 2004.
- Saunders JA, Backus BT.** Perception of surface slant from oriented textures. *J Vis* 6: 882–897, 2006.
- Srivastava S, Orban GA, De Maziere PA, Janssen P.** A distinct representation of three-dimensional shape in macaque anterior intraparietal area: fast, metric, and coarse. *J Neurosci* 29: 10613–10626, 2009.
- Sugihara H, Murakami I, Shenoy KV, Andersen RA, Komatsu H.** Response of MSTd neurons to simulated 3D orientation of rotating planes. *J Neurophysiol* 87: 273–285, 2002.
- Taira M, Tsutsui KI, Jiang M, Yara K, Sakata H.** Parietal neurons represent surface orientation from the gradient of binocular disparity. *J Neurophysiol* 83: 3140–3146, 2000.
- Treue S, Andersen RA.** Neural responses to velocity gradients in macaque cortical area MT. *Vis Neurosci* 13: 797–804, 1996.
- Tsutsui K, Jiang M, Yara K, Sakata H, Taira M.** Integration of perspective and disparity cues in surface-orientation-selective neurons of area CIP. *J Neurophysiol* 86: 2856–2867, 2001.
- Tsutsui K, Sakata H, Naganuma T, Taira M.** Neural correlates for perception of 3D surface orientation from texture gradient. *Science* 298: 409–412, 2002.
- Uka T, DeAngelis GC.** Contribution of area MT to stereoscopic depth perception: choice-related response modulations reflect task strategy. *Neuron* 42: 297–310, 2004.
- Uka T, DeAngelis GC.** Contribution of middle temporal area to coarse depth discrimination: comparison of neuronal and psychophysical sensitivity. *J Neurosci* 23: 3515–3530, 2003.
- Uka T, DeAngelis GC.** Linking neural representation to function in stereoscopic depth perception: roles of the middle temporal area in coarse versus fine disparity discrimination. *J Neurosci* 26: 6791–6802, 2006.
- Verhoef BE, Vogels R, Janssen P.** Contribution of inferior temporal and posterior parietal activity to three-dimensional shape perception. *Curr Biol* 20: 909–913, 2010.
- Xiao DK, Marcar VL, Raiguel SE, Orban GA.** Selectivity of macaque MT/V5 neurons for surface orientation in depth specified by motion. *Eur J Neurosci* 9: 956–964, 1997.
- Young MJ, Landy MS, Maloney LT.** A perturbation analysis of depth perception from combinations of texture and motion cues. *Vision Res* 33: 2685–2696, 1993.