

# Population Anisotropy in Area MT Explains a Perceptual Difference Between Near and Far Disparity Motion Segmentation

Finnegan J. Calabro and Lucia M. Vaina

*J Neurophysiol* 105:200-208, 2011. First published 10 November 2010; doi:10.1152/jn.00725.2009

**You might find this additional info useful...**

---

This article cites 38 articles, 11 of which can be accessed free at:

<http://jn.physiology.org/content/105/1/200.full.html#ref-list-1>

Updated information and services including high resolution figures, can be found at:

<http://jn.physiology.org/content/105/1/200.full.html>

Additional material and information about *Journal of Neurophysiology* can be found at:

<http://www.the-aps.org/publications/jn>

---

This information is current as of May 16, 2011.

# Population Anisotropy in Area MT Explains a Perceptual Difference Between Near and Far Disparity Motion Segmentation

Finnegan J. Calabro<sup>1</sup> and Lucia M. Vaina<sup>1,2</sup>

<sup>1</sup>Department of Biomedical Engineering, Boston University; and <sup>2</sup>Department of Neurology and Radiology, Harvard Medical School, Boston, Massachusetts

Submitted 7 August 2009; accepted in final form 3 November 2010

**Calabro FJ, Vaina LM.** Population anisotropy in area MT explains a perceptual difference between near and far disparity motion segmentation. *J Neurophysiol* 105: 200–208, 2011. First published November 10, 2010; doi:10.1152/jn.00725.2009. Segmentation of the visual scene into relevant object components is a fundamental process for successfully interacting with our surroundings. Many visual cues, including motion and binocular disparity, support segmentation, yet the mechanisms using these cues are unclear. We used a psychophysical motion discrimination task in which noise dots were displaced in depth to investigate the role of segmentation through disparity cues in visual motion stimuli (*experiment 1*). We found a subtle, but significant, bias indicating that near disparity noise disrupted the segmentation of motion more than equidistant far disparity noise. A control experiment showed that the near-far difference could not be attributed to attention (*experiment 2*). To account for the near-far bias, we constructed a biologically constrained model using recordings from neurons in the middle temporal area (MT) to simulate human observers' performance on *experiment 1*. Performance of the model of MT neurons showed a near-disparity skew similar to that shown by human observers. To isolate the cause of the skew, we simulated performance of a model containing units derived from properties of MT neurons, using phase-modulated Gabor disparity tuning. Using a skewed-normal population distribution of preferred disparities, the model reproduced the elevated motion discrimination thresholds for near-disparity noise, whereas a skewed-normal population of phases (creating individually asymmetric units) did not lead to any performance skew. Results from the model suggest that the properties of neurons in area MT are computationally sufficient to perform disparity segmentation during motion processing and produce similar disparity biases as those produced by human observers.

## INTRODUCTION

The everyday world is filled with objects that move at all speeds, directions, and depths. To build an accurate representation of the dynamic world around us, the visual system must carve the scene into meaningful components that belong to different objects or surfaces. In addition to scene segmentation, an almost simultaneous process of integration must reassemble similar components into objects and surfaces (Braddick 1993). The twin mechanisms of segmentation and integration are essential for an animal's survival in a dynamic visual world. This is especially true in regard to object motion: grouping two distinct objects and estimating their motion as if they were a single object would add significant error to the estimation of speed and direction, while segmenting too finely would reduce the available information, both preventing observers from

reacting appropriately and efficiently to their surroundings. Either of these would be detrimental to various aspects of motion perception, including the detection of whether and when an observer will collide with an oncoming object. To accurately perceive potential obstacles and threats, the visual system must reach a balance between grouping like features into common percepts (Koffka 1935; Wertheimer 1938), and segmenting objects along visual discontinuities. Although the tasks of integrating, or grouping, and segmenting seem to occur effortlessly (Julesz 1971), they require complex computations the underlying neural circuitry of which is only partially understood.

Among the many cues that the visual system can use for segmentation, motion and binocular disparity are particularly useful in allowing a quick and reliable segmentation of moving objects from their surroundings. Both cues are likely to change as a direct result of an object's movement, and whereas many animals have evolved coloring and texturing to help camouflage them in their habitat, such concealment is considerably more difficult for motion and depth cues. Furthermore, motion and depth cues tend to change together, suggesting an environmental link between them that contributes to their role in scene segmentation. In spite of the evidence that static scenes can be processed in relatively narrow bands of binocular disparity (Stevenson et al. 1992) and that disparity can be used as a cue for segmentation by the motion system (Lappe 1996; Snowden and Rossiter 1999), it is not clear whether a single disparity segmentation mechanism is responsible for both results or whether segmentation is accomplished across a distributed network of cortical areas, each implicitly segmenting the scene as needed.

The depth of a visual element (e.g., a moving object) affects the role it plays in visual scene processing. For example, objects in the background (far disparities) may be useful in computing self-motion (Ito and Shibata 2005), whereas objects in front of the observer's fixation plane pose a threat of colliding with the observer. Because foreground and background objects have different ecological roles, we hypothesize that they are segmented from the visual scene differently based on their depth. We further suggest that this is a reflection of the properties of neurons that perform disparity segmentation for motion and thus may inform neural models of segmentation.

In this study, we used a psychophysical experiment to quantify the effectiveness of disparity segmentation during motion perception (*experiment 1*). Our data show a subtle, but significant, difference in observers' performance on a motion discrimination task in the presence of near- and far-disparity noise. In a control experiment (*experiment 2*), we show that

Address for reprint requests and other correspondence: L. M. Vaina, Dept. of Biomedical Engineering, 44 Cummington St., Boston, MA 02215 (E-mail: vaina@bu.edu).

attention did not contribute to the asymmetric segmentation results of *experiment 1*. We suggest that the properties of neurons in the middle temporal area (MT) are computationally sufficient to explain the depth-dependent segmentation observed in *experiment 1*. To test this hypothesis, we propose a model containing units the properties of which are drawn from neural recordings in area MT (courtesy of Dr. Greg DeAngelis) showing that the near-disparity bias is predicted by properties of MT neurons. To isolate the cause of this effect, we developed an explanatory, physiologically constrained model of joint motion and disparity processing using units with Gabor disparity tuning curves. The results of this model show that the perceptual near-far bias is predicted when preferred disparities are chosen from an anisotropic (skewed-normal) distribution (Bradley and Andersen 1998; DeAngelis and Uka 2003; Maunsell and Van Essen 1983b) but is not found when using a biased sample of individually asymmetric units (anisotropic distribution of phases). The results of the model provide evidence that the joint motion and depth tuning of neurons in area MT are sufficient to account for the elevated near-disparity noise thresholds found in *experiment 1*. This suggests that area MT is a likely candidate for the neural substrate of disparity-based segmentation without relying on other visual areas to first perform scene segmentation. Thus we conjecture that neurons in area MT mediate an efficient segmentation of moving objects.

## METHODS

### Experiment 1

The stimulus was a random dot stereogram (RDS) presented within a circular aperture with a  $10^\circ$  radius (background luminance,  $25.2 \text{ cd/m}^2$ ) and consisting of 314 dots ( $1 \text{ dot}^{\circ 2}$ ) of  $7.5 \text{ arcmin}$  width. The dots were anti-aliased and moved and redrawn on every frame (at  $60 \text{ Hz}$ , or every  $16.7 \text{ ms}$ ) for  $500 \text{ ms}$  with a speed of  $5.0^\circ/\text{s}$ . Throughout the experiments, a chin rest was used to hold viewing distance constant at  $60 \text{ cm}$  and to reduce head movement. Subjects fixated on a white cross, placed at the center of the aperture at  $0 \text{ arcmin}$  disparity. Disparities were measured relative to the fixation cross, with negative values referring to crossed (near) disparities and positive values to uncrossed (far) disparities.

Dots were assigned to be either signal (moving left or right) or noise (repositioned randomly within the stimulus aperture between frames). The signal dots were displayed only in the fixation plane ( $0 \text{ arcmin}$  disparity), and the depth of the noise plane was one of  $0, \pm 2, \pm 5, \text{ or } \pm 12 \text{ arcmin}$ . This arrangement resulted in two transparent planes: one containing only signal dots, the other containing only noise dots. At  $0 \text{ arcmin}$  disparity separation, both planes were presented at the same depth. To prevent subjects from simply locating a

dot in the signal plane and tracking it for the duration of the trial, coherence in the signal plane was set at  $25\%$  (and  $0\%$  in the noise plane). The coherence values reported in the following text refer to the proportion of dots in the signal plane, although for comparison to previous results, it is important to note that only  $25\%$  of these were moving coherently on a given frame, reducing the overall coherence of the stimulus.

The task was a single-interval, two alternative forced choice (2AFC) direction of motion discrimination task (left- or rightward motion). The threshold (proportion coherence) was measured by an adaptive staircase varying coherence (Vaina et al. 2003). The staircase started at  $100\%$  coherence with all dots moving either left or right, in the fixation plane. As coherence dropped, the proportion of dots in the signal (fixation) plane ( $0 \text{ arcmin}$  disparity) decreased, and the noise plane became populated with an increasing proportion of dots randomly repositioned between frames, thus providing masking motion noise. Subjects were asked to fixate on the cross at the center of the display and to report the direction of the stimulus motion (left or right). Staircases consisted of four reversals of an adaptive staircase, following by eight reversals of a classical three-down, one-up staircase (Levitt 1970; Vaina et al. 2003). Thresholds were estimated for each staircase as the average of the last six reversals with overall means computed as the average of four to eight staircases per subject. Error bars for group results indicate SE and for individual subject data as the SD of curve fits to a skewed-normal function, estimated via a bootstrapping procedure (Fig. 1).

### Experiment 2

To determine whether the elevated thresholds for near-disparity noise reported in *experiment 1* are due to an unbalanced spread of attention to the noise plane, we created a version of the stimulus in which attention could not be allocated a priori only to the signal plane. In *experiment 1*, the signal and noise were presented with a  $5.0 \text{ arcmin}$  disparity separation (in which asymmetries were most visible in data from *experiment 1*) with neither present at fixation depth ( $+2.5$  and  $-2.5 \text{ arcmin}$  of disparity). Two conditions were tested. In the *noninterleaved* condition, the motion was positioned in either the near or far depth plane for the entire staircase. Observers were told whether the front or the back plane contained signal dots, allowing them to attend only to the signal plane (they still fixated on the  $0 \text{ arcmin}$  disparity fixation mark). In the *interleaved* condition, the signal was chosen to be in the near or far plane randomly on each trial, and subjects were told that they had to monitor both planes to determine the direction of motion, and the staircases for near- and far-disparity motion were interleaved. Thus in the interleaved condition, observers were unaware of which of the depth plane contained the signal dots in each trial. In both conditions, as in *experiment 1*, observers reported whether the signal dots moved left- or rightward. To reduce the ability of subjects to determine the signal plane and attend to it within the course of a trial, stimulus duration was reduced to  $250 \text{ ms}$ . All other parameters were chosen to match those described for *experiment 1*.

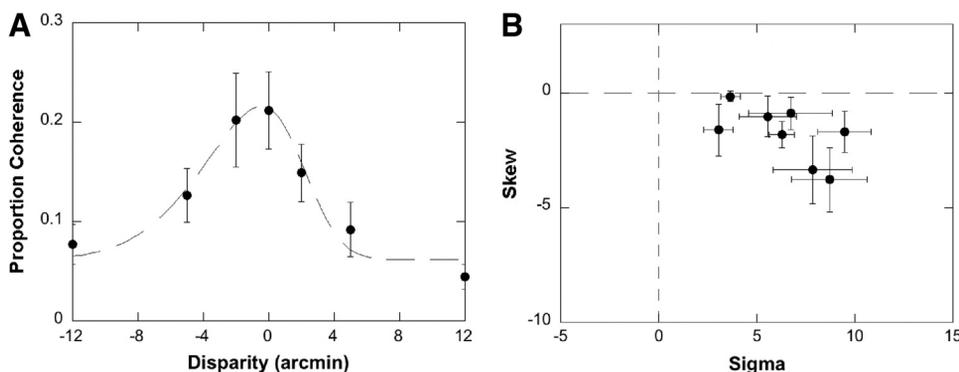


FIG. 1. A: mean thresholds for the motion discrimination task with signal dots presented at  $0 \text{ arcmin}$  of disparity (*experiment 1*) across observers ( $n = 8$ ) as a function of the disparity of the noise dots. ---, a curve fit to a skewed-normal function. Error bars are SE across observers. B: summary of individual results and curve fits. Error bars are SDs of fit parameters estimated by a bootstrapping method.

## Stereo

The stimulus was presented by simultaneously displaying red and green dots viewed through corresponding red and green filters (Berezin Stereo Photography Products, berezin.com). The effective luminance measured through the filters was  $3.5 \text{ cd/m}^2$  against a background of  $0.98 \text{ cd/m}^2$ , for a Michelson contrast of 55%. Dot luminances viewed through the wrong filter were  $1.1 \text{ cd/m}^2$ , for a contrast of 0.05% relative to the background, leading to minimal cross-talk.

## Subjects

Eight observers (mean age,  $24 \pm 4.4$ ) participated in *experiment 1*, and three in the control experiment (*experiment 2*). All observers had normal or corrected to normal vision. Stereoacuity was assessed with a computerized stereo acuity task that measured subjects' abilities to perform a two-dimensional (2-D) object depth discrimination (triangle vs. square subtending the same area). All subjects had disparity thresholds  $<3$  arcmin. Author FC was an experienced observer while all other observers were naive to the purpose of the experiments.

## RESULTS

### *Experiment 1: disparity segmentation for motion*

In *experiment 1*, subjects identified the direction of dots' motion at 0 arcmin disparity in the presence of noise at a variable disparity. Each subject's performance was fit to a skewed normal curve (Azzalini 1985; Azzalini and Capitanio 1999) with free parameters for tuning width (sigma), skew, amplitude and y-offset. A bootstrapping procedure based on the resampling of residuals (Davison and Hinkley 1997; Efron and Tibshirani 1993) was performed estimating the curve fit parameters for 500 repetitions. Mean parameters across these repetitions were taken as the best-fit values, and SDs across repetitions were used to assess the confidence interval of each parameter in the fit.

All observers showed a significant reduction in thresholds as the disparity difference between the noise and signal planes increased [amplitude parameter from a Gaussian curve fit was significantly  $>0$ ,  $t(7) = 3.75$ ,  $P = 0.007$ ]. Figure 1A shows mean thresholds across subjects as a function the disparity of the noise dots (with the signal plane fixed at 0), and Fig. 1B shows the summary of the individual curve fit parameters per subject. The mean tuning width (Gaussian sigma) across subjects was 6.41 arcmin and the y-offset was 5.6%, indicating that for large disparity separations, subjects' thresholds reached a plateau at this level, about one-fourth of their thresholds for no disparity separation (mean, 21.2%).

We found that all subjects had best-fit skew values that were  $<0$ , ranging from  $-0.14$  to  $-3.7$ . Across the group, the mean skew was  $-1.78$  (SD across subjects of 1.22), and this group effect was statistically significant [ $t(7) = -4.11$ ,  $P = 0.004$ ], indicating that our subjects showed a significant threshold elevation in the presence of near-disparity noise dots. This result indicates that the detection of motion was more difficult when observers were presented with near-disparity noise than with far-disparity noise of equal disparity difference.

### *Experiment 2: role of attention*

A possible explanation for near-far asymmetries in disparity segmentation during motion perception is that the disparity-specific bias arises from the involvement of an attention mech-

anism rather than from the motion mechanisms themselves. It has been previously demonstrated that attention can spread along a surface (Egly et al. 1994; He and Nakayama 1995), including occluded parts of the surface (Moore and Fulton 2005; Pratt and Sekuler 2001). However, how disparity would spread between two objects at different depths has not been addressed. In this context, in *experiment 1*, attention may spread from the attended (signal) plane to the noise plane asymmetrically (i.e., attention may be more likely to spread from fixation to a near-disparity plane than to a far-disparity plane). This is plausible because objects at near-disparity are more likely to require an observer action, and thus attention may be disproportionately allocated to object at near depths. Asymmetric segmentation, with higher thresholds for near-disparity noise, could then be attributed to properties of visual attention, and not motion, mechanisms.

This hypothesis stems from the fact that in *experiment 1*, signal dots were always presented in the fixation plane (0 arcmin), and thus observers knew a priori where to attend to optimally determine the direction of stimulus motion. To determine whether disparity-dependent properties of attention mechanisms contributed to the results of *experiment 1*, observers' detected near and far disparity motion signals in blocks of trials for which the two conditions were interleaved or were presented individually. If the cause of the difference in observers' performance on near- and far-disparity noise conditions was the imbalanced spread of attention to either the near- or far-disparity noise planes, then the interleaved condition should not exhibit the same behavior because subjects had to split their attention between the two planes.

Figure 2 shows the data from three observers. In the noninterleaved condition when subjects attended to the motion plane, there was a significant threshold elevation for near-disparity noise (mean elevation of 3.96% coherence, 2-way ANOVA controlling for subject showed a significant main effect of disparity,  $F = 5.05$ ,  $P = 0.03$ ), mirroring the threshold elevations found in *experiment 1*. Similarly, in the interleaved (nonattended) condition, thresholds were also elevated relative to the noninterleaved condition (mean effect size of 5.64% coherence,  $F = 2.93$ ,  $P = 0.09$ ), although this difference did not reach statistical significance. Importantly, the asymmetry between near- and far-disparity segmentation, evidenced by better motion detection for near-disparity signal dots (i.e., far-disparity noise) relative to far-disparity signal, persisted in the interleaved condition where observers could not attend to the signal plane a priori. That the near-far bias persisted when attention was split between the two planes implies that the elevated near-disparity noise thresholds reported in *experiment 1* were not a result of selectively attending to the signal plane.

## Model

The psychophysical results presented in *experiment 1* demonstrate a consistent difference in observers' performance on the segmentation of near- and far-disparity noise in a motion coherence task with thresholds higher when noise dots were in a near-disparity plane than a far-disparity plane. The results of *experiment 2* suggest that this result cannot be explained by selective attention to the signal plane. Because the outcome of these experiments do not directly provide an explanation for

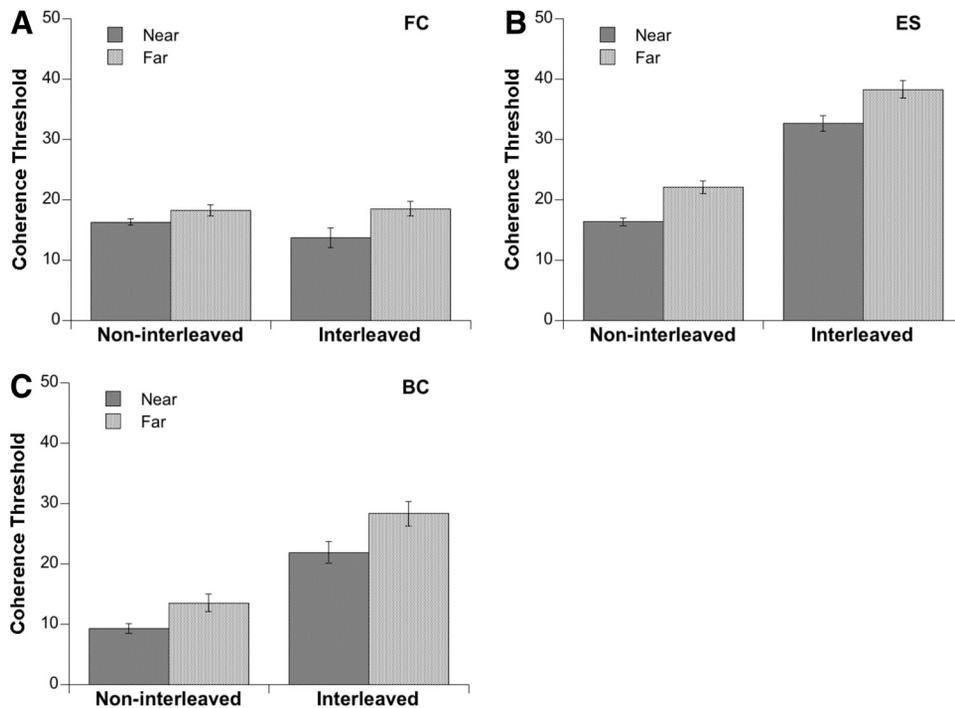


FIG. 2. Results of *experiment 2* for 3 subjects on the interleaved (dark gray) and noninterleaved (light gray) conditions. Data are shown for cases in which the signal was present in the near or far depth plane (with noise occupying the other).

disparity-based segmentation that would account for elevated near-disparity thresholds, we investigated whether the segmentation bias could arise from the disparity sensitivity of the motion processing mechanisms. We developed a physiologically constrained model of direction and disparity tuning in MT to address the sufficiency of area MT to explain the results of *experiment 1* (Fig. 3). Specifically, we tested whether properties of neurons in area MT are computationally sufficient to perform the disparity segmentation we observed and whether the near-far disparity bias is consistent with the disparity tuning properties of these neurons.

The near-far disparity asymmetry we reported could potentially arise due to one of several different physiological processes. Two potential explanations are proposed based on

results from physiological studies of area MT: asymmetries at the population or individual unit level. Large recording samples from MT have shown that more MT neurons have preferred disparities in the near visual field than the far field (Bradley and Andersen 1998; DeAngelis and Uka 2003; Maunsell and Van Essen 1983b), suggesting that an over-representation of near disparities across the population of MT neurons may result in the elevated near-disparities thresholds found in *experiment 1*. In characterizing the tuning curves of MT neurons, DeAngelis and Uka (2003) found that a Gabor model provided better fits than Gaussian tuning curves. Consequently, an alternative explanation for the near-far bias we report emerges from the offset of a cosinusoidal component relative to the Gaussian center location in the Gabor tuning curve,

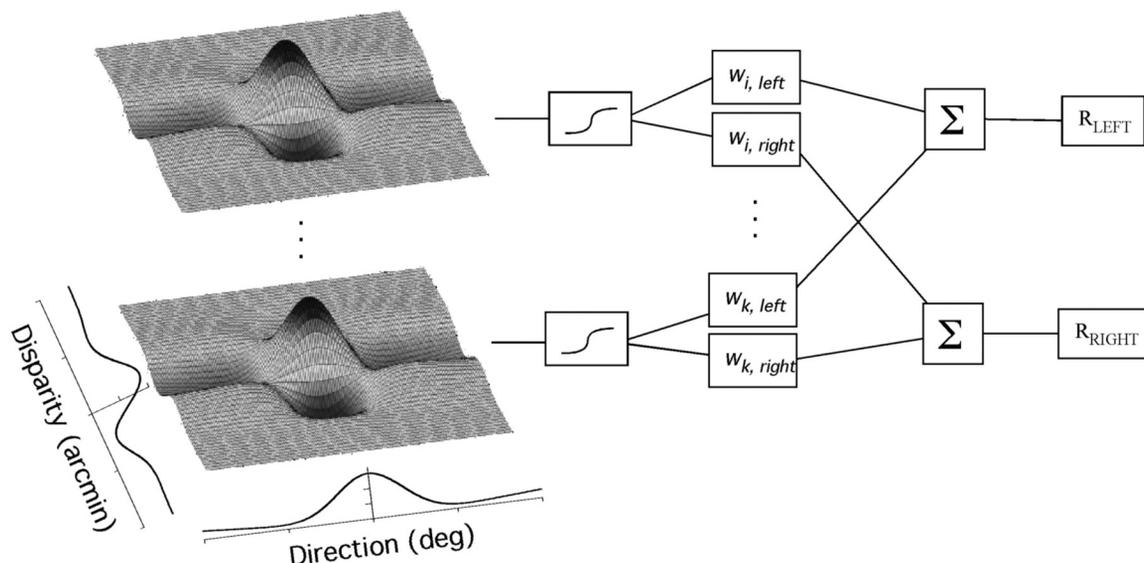


FIG. 3. Schematic view of the model. Each input unit had tuning to both direction and disparity. The response of each unit was passed through a nonlinearity, and then weighted sums were computed to determine an overall leftward and rightward model response.

creating individual unit asymmetry. We will first show that the behavioral asymmetry is predicted by tuning properties of MT neurons using a model based on neural recordings and then use an artificial model based on these neurons to determine whether population and/or individual anisotropies account for the elevated near-disparity thresholds we have reported.

We modeled the disparity sensitivity of MT-like units by a phase modulated Gabor function—the product of a Gaussian curve and a sine wave. Thus the response of unit  $i$  is given by the product of difference of Gaussian direction tuning and Gabor disparity tuning, as

$$r_{i,\delta} = e^{(\delta - \delta_i)^2 / 2\sigma_{i,\delta}^2} \cdot \cos[2\pi f(\delta - \delta_i) + \phi_i]$$

where  $r_{i,\delta}$  is the component disparity response of unit  $i$ ,  $\sigma$  is the tuning width,  $\delta_i$  is the preferred disparity defined as the Gaussian center position,  $f$  is the frequency of the cosinusoidal component (uniformly distributed between 11 and 17° based on data from DeAngelis and Uka 2003), and  $\phi_i$  is the cosinusoidal phase.

Two  $6 \times 6^\circ$  motion vector fields were used as inputs to the model and were presented at different disparities (signal dots at 0 arcmin disparity and noise dots between  $\pm 12$  arcmin), and sampled by  $0.25 \times 0.25^\circ$  input elements (the input space was  $25 \times 25$  pixels, every pixel could contain a motion vector). The response of each unit to this input was calculated as the product of the direction ( $r_\theta$ , modeled as a difference of Gaussian function with a negative Gaussian tuning width 4 times the main tuning width to contribute opponent motion responses) and disparity ( $r_\delta$ ) response, summed over the input space. To normalize across trials in which the disparity of the noise plane varied, the responses of each unit were normalized by the maximum disparity response of that unit. Neurons had a baseline firing rate of 30 Hz with noise represented as an additive term ( $\rho$ ) drawn from a normal distribution with sigma equal to half of the maximum firing rate (50 Hz), chosen as a balance between low relative noise values, in which the steepest part of the tuning curve conveys the most information, and high relative noise values, in which only the peak of the tuning curve is able to discriminate stimuli (Butts and Goldman 2006).

$$r_i = \frac{\sum_x \sum_y r_{i,\theta} r_{i,\delta} + \rho}{N \cdot \sum r_{i,\delta}}$$

The decision stage of the network performed a weighted sum of unit responses after passing the responses through a nonlinearity (sigmoid) to allow response saturation. The two parameters for the nonlinearity were determined from a Monte Carlo simulation: for each combination of slope and bias, the model performance was evaluated and compared with psychophysical performance for a threshold at 0 arcmin disparity separation, and disparity tuning width (i.e., the drop in coherence thresholds as a function of disparity separation). The slope and bias had the effect of scaling the model performance both in terms of raw thresholds and tuning width but did not fundamentally change the behavior of the model. Values were chosen to optimally match model and human thresholds for 0 and 10 arcmin disparity separations for comparison purposes. Near-far asymmetry was not included as part of the criteria. Two output weights were associated with each unit to provide the model

with direction specificity (e.g., units with leftward tuning would have a larger weight associated with the “left” output unit than the “right” output unit). Weighted sums were calculated for both the left- and rightward sets of weights, and the larger response was chosen as the response (left or right) to prevent a directional response bias.

$$r_d = \sum_i w_{i,d} \cdot \frac{1}{1 + e^{-s(r_i - b)}}$$

To determine the weights projecting to left- and rightward output units, the network was trained on sample stimuli of known direction. An equal number of left- and rightward stimuli were presented to the model, and weights were adjusted using the exposure based learning rule proposed by Vaina et al. (1995). In our model, all units the responses of which were among the top 10% had their weights increased by 0.005 (relative to an initial weight of 1). The weights were normalized across the population such that the mean weight was 1 throughout training to keep the total left- and rightward weights balanced.

$$\text{if } r_i \geq r_t: w_{i,d} = \frac{w_{i,d} + \eta}{\sum w_d}, \quad \text{if } r_i < r_t: w_{i,d} = \frac{w_{i,d}}{\sum w_d}$$

To measure the overall model performance on the psychophysical task, the trained networks were run with stimuli at 10 coherence levels ranging from 2 to 40% for 20 noise disparity conditions between  $-30$  and  $+30$  arcmin (most densely sampled near 0 arcmin). Psychometric functions of percent correct versus coherence were fit by sigmoid functions using a bootstrapping procedure based on the resampling of residuals (500 iterations with SDs across all iterations used to measure error in each of the estimated parameters). Thresholds were estimated from the sigmoid fit to compare with the psychophysical thresholds obtained in *experiment 1*. For each condition, to estimate variability due to random effects, thresholds are reported as the mean of the results from five “unique” networks each generated with a new set of units randomly chosen from the specified distributions, and independently trained.

### MT Model

We first used the curve fits of neuronal recordings from 501 MT neurons (courtesy of Dr. Greg DeAngelis, data published in DeAngelis and Uka 2003) in a model of direction and disparity processing by neurons in area MT. These units were modeled using a Gabor disparity tuning function with parameters for preferred disparity, phase, frequency, amplitude, and baseline firing rate.

Because we were interested in the disparity tuning properties of these neurons and because the psychophysical stimulus involved only left- and rightward motion, we did not assign other preferred directions to any of the units because this would give them a negligible contribution to the model output. We therefore limited the units’ preferred directions to 0 and 180° (left or right). To ensure that both directions were equally represented within our model, the full network consisted of 1,002 units—two copies of each of the 501 neuronal recordings—with one copy tuned to leftward motion and its pair tuned to rightward motion. All units were assigned equal weight when projecting to the correct output unit (leftward

units to the left output unit, rightward units to the right output unit), and a weight of 0 to the opposite direction.

The model performance is shown in Fig. 4. A bootstrapping procedure was applied to the data by performing repeated curve fits on subsampled data sets and found a mean skew of  $-1.70 \pm 0.06$  (SD). This shows that the properties of the MT neuron sample lead to elevated motion detection thresholds for near-disparity noise (relative to equal magnitude far-disparity noise), similar to that observed behaviorally in *experiment 1*.

#### Artificial model results

To investigate the cause of the disparity skew seen both in human observers and in a model of MT neurons, we constructed a model in which we manipulated the population statistics of the disparity sensitivities of MT-like units. To compare the effects of anisotropic population distributions and individually asymmetric units, we manipulated the tuning curve parameters of the Gabor disparity tuning model (the center of the Gaussian envelope,  $\delta_i$ , and the phase of the sine wave,  $\phi_i$ ). Manipulating the distribution of Gaussian positions created a population of symmetric (on average) units but with an anisotropic population (more near-tuned neurons than far-tuned units). Manipulating the distribution of phases created a population in which individual units tended to have skewed tuning curves preferring near disparities than far disparities. This allowed us to determine whether the population statistics or individual unit asymmetry (or both) could explain the near-far disparity bias observed in our observers. We applied population distributions (sigma = 55.65 arcmin, skew = 0,  $-2.79$ ) (the sigma and skew were estimated from the distributions reported by DeAngelis and Uka 2003) to each of the two parameters in separate conditions. In each case, the other parameter was modeled as a uniform distribution ( $-60$  through  $+60$  arcmin for  $\delta_i$  and  $-\pi$  through  $+\pi$  for  $\phi_i$ ). This resulted in uncorrelated position parameters. Direction tuning was modeled as a difference of Gaussians function (Gaussian of sigma  $36^\circ$  minus a Gaussian with sigma  $144^\circ$ ) with preferred directions distributed uniformly among all directions.

The analysis of the effect of population distributions revealed that skewed populations led to skewed performance when the population distributions were applied to the center

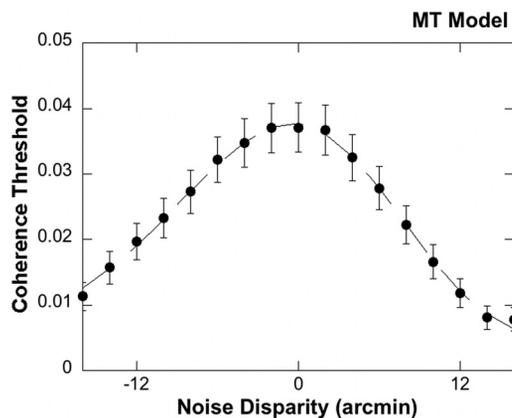


FIG. 4. Results of the model using units the parameters of which matched those of real neurons reported by DeAngelis and Uka (2003). A disparity-skew similar to that shown by our psychophysical subjects is found in the performance of the middle temporal area (MT) model, resulting in elevated thresholds for near disparities (negative disparity values).

position of the Gaussian envelope. Performance curves were fit to the results of each unique model (typical network performance and skewed-normal curve fits shown in Fig. 5, skews summarized in Fig. 6). When a symmetric population was used for the distribution of preferred disparities (Fig. 5A), the performance curve was well fit by a skewed-normal curve with sigma  $6.24 \pm 0.13$  arcmin and skew of  $0.03 \pm 0.02$ . The anisotropic distribution resulted in an increased sigma of  $11.2 \pm 1.47$  arcmin, and a significant performance skew of  $-1.75$  (SD from bootstrap fit of 0.40,  $z = -4.37$ ,  $P < 0.001$  using a 2-tailed  $z$ -test). This skew was similar to that shown by human subjects in *experiment 1* (mean skew of  $-1.78$ , SD across subjects of 1.23) and by the model of actual MT neuron curve fits ( $-1.7 \pm 0.06$ ).

These results were not preserved when the population distributions were applied to the phase parameter (the peak of the sinusoidal component of the Gabor tuning function). The sigma of the performance curve for the model using an anisotropic distribution of phases was again slightly elevated compared with the isotropic model (from 5.94 to 7.75 arcmin), but there was no significant skew in the performance curve resulting from an anisotropic phase population (Fig. 5D) related to changing the population distribution as applied to the phase parameter (mean skew:  $0.02 \pm 0.03$ ,  $z = 0.67$ ,  $P > 0.5$ ).

Thus although the Gabor fits result in a biased response function when the population distribution was applied to the center of the Gaussian envelope, there appears to be no effect on the model performance when the anisotropic distribution of preferred disparities is applied to the phase parameter. Because applying the skewed distribution to the phase parameter did not produce a skewed performance distributions, we suggest that the center and phase parameter have computationally distinct roles in determining the overall model performance.

#### DISCUSSION

The results of *experiment 1* extend previous studies (Snowden and Rossiter 1999) that have reported that binocular disparity can be used as a segmentation cue during motion perception. In a group of eight subjects, we showed that this effect occurs gradually with the increase of disparity separation between signal and noise, with a mean bandwidth (Gaussian sigma) of 6.4 arcmin. Interestingly, we found a consistent disparity-dependent bias among observer performance, showing elevated thresholds in the presence of near-disparity noise (mean skew of  $-1.78$ ). Results from a control test (*experiment 2*) suggested that even when unable to selectively attend to the signal plane, subjects exhibited a significant threshold elevation when presented with near-disparity noise. This indicated that the near-far bias was not likely a property of the attention system.

Based on our results, we hypothesized that the near-disparity bias may be a property of the motion mechanism being used in this task and that it may be used as a criterion in linking potential neural mechanisms to the perception of motion during disparity segmentation.

We investigated whether the results of *experiment 1* could be explained by properties of the neuronal population in area MT, a cortical area the neurons of which are known to be selective to both direction of motion and binocular disparity (Maunsell and Van Essen 1983a,b); using models first based on actual

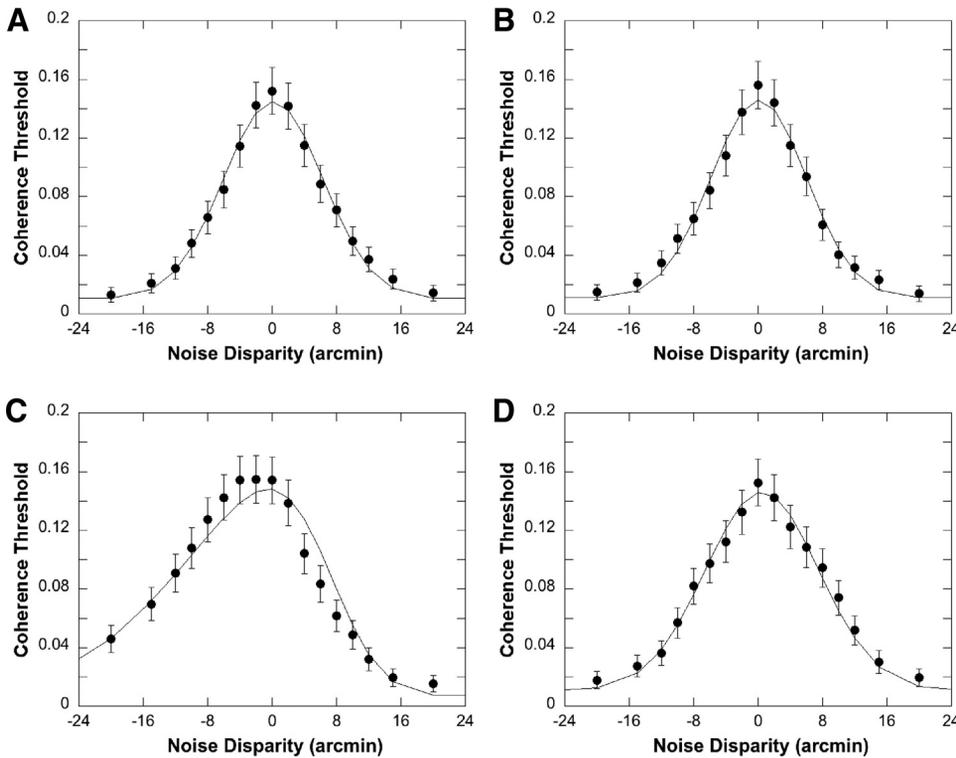


FIG. 5. Typical results from single unique networks. *Top*: symmetric distributions applied to the Gaussian position (A) and phase (B) parameters. *Bottom*: anisotropic distributions applied to Gaussian position (C) and phase (D) parameters. Only the skewed-Gaussian position distribution (C) exhibited significant skew similar to that found by human subjects in *experiment 1*. —, curve fits to a skewed-normal function.

neural recordings in area MT, and then by a simulated set of MT-like units the properties were drawn from the known population statistics of MT. Computational models have of which demonstrated that stimulus discontinuities can be detected from motion cues (Hildreth 1983; Spoerri 1991; Thompson 1980; Vaina et al. 1994, 1998) or from binocular disparity (Julesz 1971; Marr 1982) alone. In this study, we were interested in determining whether the neural processing implemented by MT neurons is computationally sufficient to explain the segmentation employed by subjects in our psychophysical task (*experiment 1*). Neuronal models provide an important tool for exploring and explaining the link between neurophysiology and perceptual phenomena (Baloch et al. 1999; Beardley and Vaina 1998, 2001, 2004; Berns et al. 1993; Cadieu et al. 2007; Giese and Poggio 2003; Pack et al. 2001; Schultz et al. 1997).

We simulated performance of a model of neurons recorded from area MT (DeAngelis and Uka 2003) that accurately

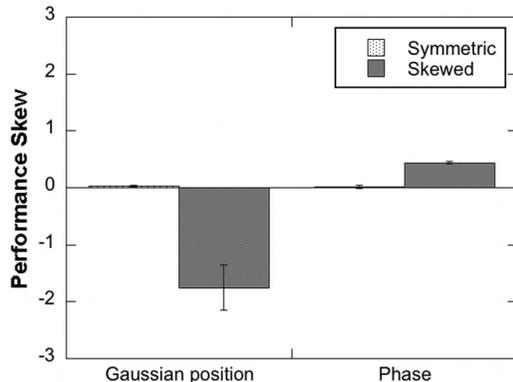


FIG. 6. Summary of performance skew for symmetric and asymmetric distributions applied to the Gaussian center location (*left*) and phase (*right*) of model disparity tuning functions.

replicates the human behavioral results reported in *experiment 1*. Three elements of the model could account for near-far disparity differences: the distribution of preferred disparities (Gaussian center position), the distribution of phases, and the weighting rule. Because the MT neuron simulations used uniform weights, the weighting and decision rules were unlikely to contribute to the near-far difference.

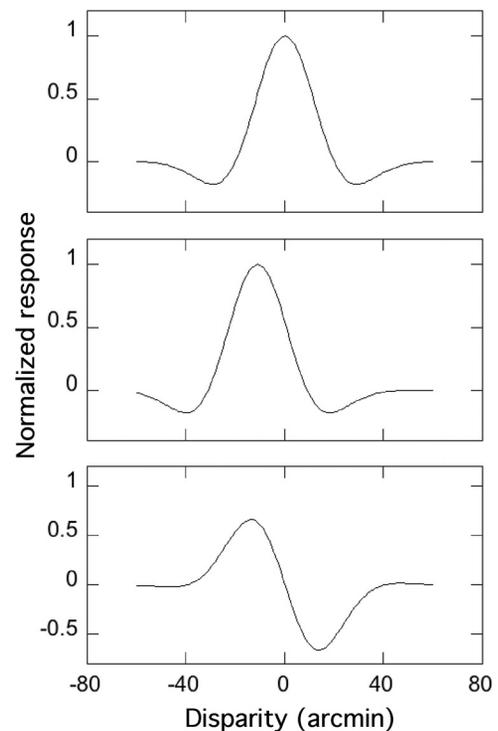


FIG. 7. Comparison of a sample unit tuning curve with a preferred disparity and phase of 0 (A), a shifted preferred disparity (B), and a shifted phase (C).

To test whether the distribution of preferred disparities or phases was responsible for the performance skew, we compared the performance of a model with asymmetries built into the model either at the population (distribution of Gaussian centers) or individual unit (distribution of phases) level. By applying the skewed normal population anisotropies to each parameter independently, we measured the relative contribution of each parameter to the performance skew in our model. The results (Fig. 5) showed that the distribution of Gaussian centers provided similar near disparity bias as that found for our human subjects and the model of MT units, while the phase parameter led to negligible differences between near and far disparity segmentation.

The different effects of manipulating preferred disparity and phase are illustrated by a sample unit shown in Fig. 7 and may be used to explain the difference in results for two anisotropic population models. When the preferred disparity is shifted, the unit continues to have a large response to 0 arcmin disparity, that is, to the disparity containing signal dots. Thus the unit shown in Fig. 7 became more responsive to the noise while maintaining its response to the motion signal. On the other hand, for the phase shift, the unit's response to the signal disparity drops nearly to zero (indicating baseline firing rate). Thus even though this unit is now tuned to the noise disparity, its overall response is quite low because it no longer responds to the motion disparity, causing any modulatory effect of noise dots to have a minimal overall effect. Another way to consider this is that by having an offset phase, the tuning response of the unit as a function of disparity becomes narrower, so that the unit is less likely to respond to both signal and noise. This suggests that preferred disparity (by the Gaussian center position) and the phase may have different computational properties: preferred disparity controls the overall response and performance, including generating the near-disparity bias observed in the model, while an offset phase creates narrower tuning widths allowing more specific disparity responses.

A near-far disparity difference similar to the results of *experiment 1* has been previously reported for a shift in the perceived center of expansion when viewed in the presence of unidirectional motion (Duffy and Wurtz 1993; Grigo and Lappe 1998). The size of this illusion decreases as the expanding and translating motion components are separated in depth, and the decrease is much more pronounced for far disparities (Grigo and Lappe 1998). Lappe and Grigo (1999) hypothesized this effect could be explained by the disparity processing of MT neurons and suggested that a disparity-based weighting is used to selectively process far-tuned MT neurons. The results of our model further suggest that population anisotropies within MT allow asymmetric disparity-dependent processing even though individual neurons may have symmetric tuning.

Our model proposes a physiological implementation of disparity segmentation for motion perception that relies solely on direction and disparity properties of neurons in area MT. The use of physiologically recorded neural properties resulted in the reproduction of a near-far disparity bias similar to that observed psychophysically with human subjects, suggesting a perceptual link between the disparity tuning of MT neurons and the segmentation of motion present at different depths within the visual scene. Although these results suggest that properties of MT neurons are computationally sufficient to segment moving surfaces based on disparity, it is also possible

that interactions among several cortical areas may play a role. A possible candidate, is cortical area V2, whose neurons respond to illusory surfaces (Bakin et al. 2000; Peterhans and von der Heydt 1989; von der Heydt et al. 1984). If these neurons inhibit responses to occluded surfaces, this could contribute a near-far bias similar to what we have observed (*experiment 1*). This possibility has been suggested by studies of the role of disparity on motion-induced blindness, in which static dots disappear less frequently when they are displayed in front of moving dots (Graf et al. 2002). This effect has been linked to attention (Bonneh et al. 2001; Driver and Vuilleumier 2001). Because in our study the near-far disparity bias remained when selective attention was removed (*experiment 2*), a possible contribution of surface completion would have to occur inattentively (i.e., by inhibiting the perception of the far disparity plane).

#### ACKNOWLEDGMENTS

We thank Dr. Greg DeAngelis for generously providing the curve fit parameters for the disparity tuning of MT neurons recorded in his lab.

#### GRANTS

This work was supported by National Institute of Neurological Disorders and Stroke Grant R01NS-064100 to L. M. Vaina.

#### DISCLOSURES

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

#### REFERENCES

- Azzalini A.** A class of distributions which includes the normal ones. *Scand J Stat* 12: 171–178, 1985.
- Azzalini A, Capitanio A.** Statistical applications of the multivariate skew normal distribution. *J R Stat Soc B Statist Methodol* 61: 579–602, 1999.
- Bakin JS, Nakayama K, Gilbert CD.** Visual responses in monkey areas V1 and V2 to three-dimensional surface configurations. *J Neurosci* 20: 8188–8198, 2000.
- Baloch AA, Grossberg S, Mingolla E, Nogueira CA.** Neural model of first-order and second-order motion perception and magnocellular dynamics. *J Opt Soc Am A Opt Image Sci Vis* 16: 953–978, 1999.
- Beardsley SA, Vaina LM.** Computational modelling of optic flow selectivity in MSTd neurons. *Network Comput Neural Sys* 9: 467–493, 1998.
- Beardsley SA, Vaina LM.** A laterally interconnected neural architecture in MST accounts for psychophysical discrimination of complex motion patterns. *J Comput Neurosci* 10: 255–280, 2001.
- Beardsley SA, Vaina LM.** Linking perception and neurophysiology for motion pattern processing: the computational power of inhibitory connections in cortex. In: *Optic Flow and Beyond*, edited by Vaina LM, Beardsley SA, Rushton SK. Dordrech, The Netherlands: Kluwer Academic, 2004, p. 183–222.
- Berns GS, Dayan P, Sejnowski TJ.** A correlational model for the development of disparity selectivity in visual cortex that depends on prenatal and postnatal phases. *Proc Natl Acad Sci USA* 90: 8277–8281, 1993.
- Bonneh YS, Cooperman A, Sagi D.** Motion-induced blindness in normal observers. *Nature* 411: 798–801, 2001.
- Braddick O.** Segmentation Versus Integration in visual motion processing. *Trends Neurosci* 16: 263–268, 1993.
- Bradley DC, Andersen RA.** Center-surround antagonism based on disparity in primate area MT. *J Neurosci* 18: 7552–7565, 1998.
- Butts DA, Goldman MS.** Tuning curves, neuronal variability, and sensory coding. *PLoS Biol* 4: e92, 2006.
- Cadiou C, Kouh M, Pasupathy A, Connor CE, Riesenhuber M, Poggio T.** A model of V4 shape selectivity and invariance. *J Neurophysiol* 98: 1733–1750, 2007.
- Davison AC, Hinkley DV.** Bootstrap Methods and Their Application. Cambridge: New York: Cambridge University Press, 1997, p. x, 582 p.

- DeAngelis GC, Uka T.** Coding of horizontal disparity and velocity by MT neurons in the alert macaque. *J Neurophysiol* 89: 1094–1111, 2003.
- Driver J, Vuilleumier P.** Perceptual awareness and its loss in unilateral neglect and extinction. *Cognition* 79: 39–88, 2001.
- Duffy CJ, Wurtz RH.** An illusory transformation of optic flow fields. *Vision Res* 33: 1481–1490, 1993.
- Efron B, Tibshirani R.** An Introduction to the Bootstrap. New York: Chapman & Hall, 1993 p. xvi, 436 p.
- Egly R, Driver J, Rafal RD.** Shifting visual attention between objects and locations: evidence from normal and parietal lesion subjects. *J Exp Psychol Gen* 123: 161–177, 1994.
- Giese MA, Poggio T.** Neural mechanisms for the recognition of biological movements. *Nat Rev Neurosci* 4: 179–192, 2003.
- Graf EW, Adams WJ, Lages M.** Modulating motion-induced blindness with depth ordering and surface completion. *Vision Res* 42: 2731–2735, 2002.
- Grigo A, Lappe M.** Interaction of stereo vision and optic flow processing revealed by an illusory stimulus. *Vision Res* 38: 281–290, 1998.
- He ZJ, Nakayama K.** Visual attention to surfaces in three-dimensional space. *Proc Natl Acad Sci USA* 92: 11155–11159, 1995.
- Hildreth EC.** *The Measurement of Visual Motion*. Cambridge, MA: MIT Press, 1983.
- Ito H, Shibata I.** Self-motion perception from expanding and contracting optical flows overlapped with binocular disparity. *Vision Res* 45: 397–402, 2005.
- Julesz B.** *Foundations of Cyclopedian perception*. Chicago, IL: University of Chicago Press, 1971.
- Koffka K.** *Principles of Gestalt Psychology*. New York: Harcourt Brace, 1935.
- Lappe M.** Functional consequences of an integration of motion and stereopsis in area MT of monkey extrastriate visual cortex. *Neural Comp* 8: 1449–1461, 1996.
- Lappe M, Grigo A.** How stereovision interacts with optic flow perception: neural mechanisms. *Neural Netw* 12: 1325–1329, 1999.
- Levitt H.** Transformed up-down methods in psychoacoustics. *J Acoust Soc Am* 49: 467477, 1970.
- Marr D.** *Vision*. New York: Freeman, 1982.
- Maunsell JH, Van Essen DC.** Functional properties of neurons in middle temporal visual area of the macaque monkey. I. Selectivity for stimulus direction, speed, and orientation. *J Neurophysiol* 49: 1127–1147, 1983a.
- Maunsell JH, Van Essen DC.** 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, 1983b.
- Moore CM, Fulton C.** The spread of attention to hidden portions of occluded surfaces. *Psychon Bull Rev* 12: 301–306, 2005.
- Pack C, Grossberg S, Mingolla E.** A neural model of smooth pursuit control and motion perception by cortical area MST. *J Cognit Neurosci* 13: 102–120, 2001.
- Peterhans E, Von der Heydt R.** Mechanisms of contour perception in monkey visual cortex. II. Contours bridging Gaps. *J Neurosci* 9: 1749–1763, 1989.
- Pratt J, Sekuler AB.** The effects of occlusion and past experience on the allocation of object-based attention. *Psychon Bull Rev* 8: 721–727, 2001.
- Schultz W, Dayan P, Montague PR.** A neural substrate of prediction and reward. *Science* 275: 1593–1599, 1997.
- Snowden RJ, Rossiter MC.** Stereoscopic depth cues can segment motion information. *Perception* 28: 193–201, 1999.
- Spoerri A.** *The Early Detection of Motion Boundaries*. Cambridge, MA: MIT, 1991 (Technical report).
- Stevenson SB, Cormack LK, Schor CM, Tyler CW.** Disparity tuning in mechanisms of human stereopsis. *Vision Res* 32: 1685–1694, 1992.
- Thompson WB.** Combining motion and contrast for segmentation. *IEEE Trans PAMI* 2: 543–549, 1980.
- Vaina LM, Grzywacz NM, Saiviroonporn P, LeMay M, Bienfang DC, Cowey A.** Can spatial and temporal motion integration compensate for deficits in local motion mechanisms? *Neuropsychologia* 41: 1817–1836, 2003.
- Vaina LM, Grzywacz NM, Kikinis R.** Segregation of computation underlying perception of motion discontinuity and coherence. *Neuroreport* 5: 2289–2294, 1994.
- Vaina LM, Grzywacz NM, LeMay M, Bienfang D, Wolpaw E.** Perception of motion discontinuities in patients with selective motion deficits. In: *High Level Visual Motion*, edited by Watanabe T. Cambridge, MA: MII Press, 1998, p. 213–247.
- Vaina LM, Sundaeswaran V, Harris JG.** Learning to ignore: psychophysics and computational modeling of fast learning of direction in noisy motion stimuli. *Brain Res Cogn Brain Res* 2: 155–163, 1995.
- von der Heydt R, Peterhans E, Baumgartner G.** Illusory contours and cortical neuron responses. *Science* 224: 1260–1262, 1984.
- Wertheimer M.** Laws of organization in perceptual forms. In: *A Sourcebook of Gestalt Psychology*, edited by Ellis WD. London: K. Paul, Trench, Trubner, 1938, p. 71–88.