

# A Neural Representation of Depth from Motion Parallax in Macaque Visual Cortex

JACOB W. NADLER<sup>1</sup>, DORA E. ANGELAKI<sup>1</sup> AND GREGORY C. DEANGELIS<sup>1,2</sup>

<sup>1</sup> *Dept. of Anatomy & Neurobiology, Washington University School of Medicine, St. Louis, MO 63110*

<sup>2</sup> *Dept. of Brain and Cognitive Sciences, University of Rochester, NY 14627*

Address for correspondence:  
Dr. Gregory C. DeAngelis  
Professor of Brain and Cognitive Sciences  
Center for Visual Science  
University of Rochester  
245 Meliora Hall  
Rochester, NY 14627  
Office Phone: (585)275-8677  
FAX: (585)271-3043  
Email: gdeangelis@cvs.rochester.edu

**Acknowledgements:** We are grateful to Lawrence Snyder, Akiyuki Anzai, Takahisa Sanada, Yong Gu, and Christopher Fetsch for helpful comments. We thank Christopher Broussard for technical development, and Amanda Turner, Erin White and Kim Kocher for monkey care and training. This work was supported by NEI Institutional NRSA 5-T32-EY13360-07 (to JWN) and NEI grants EY013644 (to GCD) and EY017866 (to DEA).

**Perception of depth is a fundamental challenge for the visual system, particularly for observers moving through their environment. The brain makes use of multiple visual cues to reconstruct the three-dimensional structure of a scene. One potent cue, motion parallax, frequently arises during translation of the observer because the images of objects at different distances move across the retina with different velocities. Human psychophysical studies have demonstrated that motion parallax can be a powerful depth cue<sup>1-5</sup>, and motion parallax appears to be heavily exploited by animal species that lack highly developed binocular vision<sup>6-8</sup>. However, little is known about the neural mechanisms that underlie this capacity. We used a virtual-reality system to translate macaque monkeys while they viewed motion parallax displays that simulated objects at different depths. We show that many neurons in the middle temporal (MT) area signal the sign of depth (i.e., near vs. far) from motion parallax in the absence of other depth cues. To achieve this, neurons must combine visual motion with extra-retinal (non-visual) signals related to the animal's movement. Our findings suggest a new neural substrate for depth perception, and demonstrate a robust interaction of visual and non-visual cues in area MT. Combined with previous studies that implicate area MT in depth perception based on binocular disparities<sup>9-12</sup>, our results suggest that MT contains a more general representation of three dimensional space that leverages multiple cues.**

Humans can make precise judgments of depth based on motion parallax, the relative retinal image motion between objects at different distances<sup>1-5</sup>. However, motion parallax alone is not sufficient to specify the sign of depth, that is whether an object is near or far relative to the plane of fixation<sup>13-15</sup>. Rather, the direction of image motion relative to observer motion is crucial to specifying depth-sign (Fig. 1). Objects located nearer than the plane of fixation sweep across the visual field in a direction *opposite* to head translation (Fig. 1, black arrows). In contrast, objects located farther than the plane of fixation move in the *same* direction as the head (Fig 1, white arrows). Thus, image motion in a particular direction (e.g. leftward) can be associated with either a near (Fig. 1a) or far (Fig. 1b) object. Under some conditions, the brain could interpret such ambiguous visual motion by using other cues such as occlusion, size, or perspective. However, to compute depth-sign in the absence of these pictorial cues, the visual system

needs access to extra-retinal signals related to observer movement. We exploited this fact to probe for a neural correlate of depth from motion parallax.

We performed extracellular microelectrode recordings in area MT of two macaque monkeys that were trained to maintain visual fixation on a world-fixed target while being translated by a motion platform. While tracking the target, the animal viewed a display in which a circular patch of random dots was placed over the neuron's visual receptive field (RF). In each trial, motion of the dots was computed to accurately simulate a surface placed at one of nine depths, which are expressed in terms of their equivalent binocular disparities (Fig. 1c, Supplementary Fig. S1). Stimuli were viewed monocularly, and the random-dot stimulus was scaled such that the retinal image maintained a constant size, retinotopic location, and dot density, independent of simulated depth (Fig. 1c). By eliminating all pictorial depth cues, we forced our visual stimulus to be depth-sign ambiguous.

Our experiment consisted of two main stimulus conditions that were randomly interleaved. In the *Motion Parallax (MP)* condition, the animal was translated sinusoidally at 0.5 Hz (Fig. 1d) while a 3D graphics engine performed the necessary projections to render the fixation point and random-dot surface as stationary world-fixed objects. During translation, the animal was required to make compensatory smooth eye movements to track the fixation point (Fig. 1d, Supplementary Fig. S2). In the *Retinal Motion (RM)* condition, we replicated the visual image seen during the MP condition. However, the animal remained stationary and, thus, did not have to make any eye movements. Because the retinal stimulation was the same in both conditions, any differences in neural response should be due to the action of extra-retinal signals.

Figure 2 illustrates results obtained from a single neuron. Peri-stimulus time histograms (PSTHs) are shown for the RM and MP conditions (Fig. 2a, b), with responses grouped (into columns) according to the phase of the real or simulated movement of the observer. Note that near and far simulated depths have opposite retinal velocity profiles (grey traces). In the RM condition (Fig. 2a), responses of this directionally-selective MT neuron follow the retinal velocity of the stimulus such that near and far simulated depths of the same magnitude (e.g., -2° vs. +2°) elicit very similar responses. Given the amplitude and speed of the head movements, resulting retinal

velocities were slow (ranging from  $\sim 0^\circ/\text{s}$  for a simulated depth of  $0^\circ$  up to  $\pm 5^\circ/\text{s}$  for the most extreme simulated depths of  $\pm 2^\circ$ , Supplementary Fig. S1). Because this neuron preferred fast speeds, simulated depths closer to the plane of fixation evoked less activity. When average firing rate is plotted as a function of simulated depth, the neuron shows a depth-tuning curve that is symmetric around  $0^\circ$  equivalent disparity (Fig. 2c, open symbols). A strikingly different pattern of response occurred in the MP condition (Fig. 2b). Responses to near stimuli (e.g.,  $-2^\circ$ ) were enhanced while responses to far stimuli (e.g.,  $+2^\circ$ ) were suppressed relative to the RM condition. As a result, the depth-tuning curve is strongly asymmetric in the MP condition, with response declining monotonically from near to far (Fig. 2c, filled symbols). Since the retinal motion stimulus is the same for the two conditions, this selectivity for depth sign must arise from extra-retinal signals.

Depth-tuning curves for six additional cells are shown in Fig. 3. Three of these example neurons prefer near depths (left) in the MP condition, while the other three prefer far depths (right). Note that the tuning curves are generally monotonic within the range tested. Indeed, we found that only 28% of MT neurons with significant depth-sign selectivity had tuning curves that deviated significantly from monotonicity in the MP condition (see Methods), including the example cell in Fig. 3a. This tendency toward monotonic tuning lies in contrast to the peaked tuning curves that many MT neurons exhibit in response to binocular disparities<sup>16,17</sup>, but this does not preclude these cells from participating in a representation of depth based on multiple cues.

Depth-sign selectivity was quantified using a Depth-sign Discrimination Index (DSDI, see Methods). Neurons with strong preferences for near and far depths will have DSDI values approaching -1 and +1, respectively. All neurons with well-isolated action potentials were studied ( $n = 144$ ), although some responded poorly at all depths because the range of stimulus speeds was not well matched to their speed preference. In the MP condition (Fig. 4a, top histogram), DSDIs are spread across a wide range and roughly two thirds (100/144) of the neurons show significant selectivity ( $p < 0.05$ , permutation test). Neurons that prefer near depths (66/100) were significantly more common than neurons that prefer far depths (34/100) ( $p = 0.0014$ , chi-square test). This held true for both animals individually ( $p < 0.04$ , chi-square tests). By contrast, the range of DSDIs in the RM condition is much narrower ( $p < 0.0001$ , Levene's F-Test), and far fewer cells are

significantly tuned (17/144). While this fraction (12%) is larger than expected by chance (5%), this can be explained by a motion parallax equivalent of vertical disparity (Supplementary Fig. S3). Overall, the average |DSDI| in the MP condition ( $0.41 \pm 0.02$  SE) was significantly larger ( $p << 0.001$ , paired t-test) than that in the RM condition ( $0.15 \pm 0.01$  SE). This difference indicates that many MT neurons are modulated by extra-retinal signals to generate selectivity for depth sign.

Retinal image motion is identical in the MP and RM conditions provided that the monkey accurately tracks the fixation target during platform motion. Thus, we must exclude the possibility that differences in selectivity between the MP and RM conditions result from inaccurate eye pursuit. Pursuit was generally good and saccades were rare (Fig. 1d, Supplementary Fig. S2). However, both animals tended to slightly under-pursue the fixation target by ~4% (average pursuit gain = 0.963). A careful analysis revealed that there was no significant correlation between the accuracy of pursuit and DSDI values measured in the MP condition (Supplementary Fig. S4). In addition, we corrected DSDI values for imperfect pursuit and found only small changes that would not alter our conclusions (Supplementary Fig. S5). Thus, depth-sign selectivity in MT cannot be explained by inaccuracies in eye movements.

Figure 4b reveals a significant negative correlation ( $r = -0.57$ ,  $p < 0.0001$ , Spearman rank correlation) between depth-sign preference in the MP condition and preferred retinal speed. MT neurons that prefer far depths tended to prefer slow speeds, whereas near-preferring neurons were tuned to a broad range of speeds (except for speeds very close to zero). This correlation may reflect an adaptation to the ecological constraint that the range of retinal image speeds (due to observer motion) is larger for near objects than far objects. Under the conditions of our experiment, any object nearer than one-half the viewing distance will have a retinal speed greater than all possible far objects. Alternatively, this correlation may simply be a by-product of the mechanism by which extra-retinal signals interact with visual motion in MT. In either case, the correlation between DSDI and speed preference likely explains the higher proportion of near-preferring neurons in our sample. Neurons with speed preferences above  $5^\circ/\text{s}$  tended to be near-preferring, and the majority of neurons in MT have speed preferences that exceed this value<sup>18,19</sup>. A similar predominance of near-preferring neurons has been reported for

binocular disparity tuning in MT<sup>17</sup>, though disparity preferences were not found to be strongly correlated with speed preferences<sup>16</sup>.

It is sensible that depth from motion parallax would be represented in neural circuits that are sensitive to visual motion and that play a role in coding three-dimensional scene structure. Area MT is well known as a motion processing area<sup>20</sup>. In addition, most MT neurons are tuned for depth defined by binocular disparities<sup>16,17</sup>, and MT has been causally linked to some forms of depth perception<sup>9,11</sup>. Notably, motion and disparity cues are known to interact in MT to disambiguate relative image motion that results from 3D object structure<sup>21-23</sup>. Our findings show that, even in the absence of disparity cues, MT neurons can use extra-retinal signals to disambiguate retinal image motion and compute depth. This substantially enhances the flexibility of depth signals in MT. Neurons that are selective for both disparity and motion parallax should allow for a more robust representation of depth, and may mediate improvements in depth perception seen when disparity and motion parallax cues are presented together<sup>3,5</sup>.

It is unclear whether selectivity for depth from motion parallax emerges in area MT, or whether it arises earlier in the visual pathways. Neurons that are sensitive to relative image motion have been described in primary visual cortex of monkeys<sup>24</sup> but this does not necessarily imply a role in computing depth from motion parallax. We have demonstrated selectivity for depth sign by using ambiguous retinal image motion and showing that an extra-retinal signal overcomes this ambiguity. The strength of the extra-retinal influences in our data may appear surprising based on previous work. MT neurons are not thought to be strongly modulated by pursuit eye movements<sup>25</sup>, and we are not aware of any published evidence of substantial vestibular signals in MT. Responses of MT neurons have been shown to be modulated by eye position<sup>26</sup>, but such ‘gain fields’ cannot explain our results because the same variation in eye position has opposite modulatory effects on responses to near and far stimuli (Fig. 2b).

What might be the mechanism that generates depth-sign selectivity in MT? One possibility is that an extra-retinal signal related to head or eye movement simply sums with responses to visual motion, thus enhancing responses to one depth sign (e.g., near) and suppressing responses to the opposite depth sign<sup>27</sup>. Our data suggest that the mechanism is more complicated. In ‘null’ trials containing no visual stimulus in the

receptive field, 75% of MT neurons still show significant response modulations during eye/head motion (e.g., Fig. 2b, top). However, these modulations generally fail to predict, in both sign and magnitude, the difference in activity between the MP and RM conditions (Supplementary Fig. S6). For the neuron of Fig. 2, if the response modulations seen in the null trials (Fig. 2b, top) simply added to responses in the RM condition, the neuron should prefer far depths. Instead, it prefers near stimuli (see also Fig. S6a). Thus, the interaction between visual motion and extra-retinal signals appears to be complex and nonlinear, and requires further study.

In summary, we have demonstrated that single neurons in area MT carry reliable information about depth sign from motion parallax. While the existence of motion parallax as a potent cue for depth perception has long been established<sup>1,4</sup>, our findings provide the first evidence of a neural substrate for this perceptual capacity. We cannot yet prove that these signals in area MT are used by the monkey to perceive depth (they could reflect extra-retinal signals used for another purpose), but our findings enable a direct causal test in trained animals. Such proof notwithstanding, our results establish a new potential neural mechanism for processing depth information, and suggest that area MT may be involved in integrating multiple cues to depth.

## Methods Summary

We recorded extracellular single-unit activity from area MT using tungsten microelectrodes (FHC, Bowdoinham, ME) in two adult rhesus monkeys (*Macaca mulatta*). A custom-made virtual-reality system<sup>28</sup> was used to provide stimuli consisting of sinusoidal translation and/or visual motion. Animals were trained to maintain fixation on a visual target during translation of the motion platform. Custom-written OpenGL software was used to generate visual stimuli that depicted a random-dot surface at one of several possible depths in a virtual environment. The visual stimulus was viewed monocularly by the animal and all pictorial depth cues were removed from the stimulus to render the visual motion ambiguous with respect to depth sign. Thus, to compute depth sign based on motion parallax, neurons must combine visual motion with extra-retinal signals generated by physical translation of the animal (e.g., vestibular or eye movement signals).

Our experimental design compared neural responses in two conditions: a Motion Parallax (MP) condition in which the combination of physical translation and visual motion specified depth unambiguously, and a Retinal Motion (RM) control condition in which the same visual motion stimulus was presented in the absence of extra-retinal signals, thus rendering it depth-sign ambiguous. Neuronal responses were measured as mean firing rates and the significance of depth-sign selectivity was assessed using permutation tests. Eye position data were filtered (200 Hz lowpass) and analyzed to quantify the accuracy of pursuit eye movements. All procedures were approved by the Institutional Animal Care and Use Committee at Washington University and were in accordance with National Institutes of Health guidelines.

## Methods

### Subjects and task

Two awake male monkeys (*Macaca mulatta*) were prepared for neurophysiological experiments as described in detail elsewhere<sup>28</sup>. Monkeys were head restrained and seated in front of a 60x60 cm tangent screen that subtended roughly 90°x90° of visual angle at the viewing distance of 32 cm. The display screen and stereoscopic projector (Christie Digital Mirage 2000) were mounted on a six degree-of-freedom motion platform (MOOG 6DOF2000E) that allowed us to translate the animal along any direction in the frontoparallel plane (fore-aft movements were not used here). Platform movements and visual stimuli were controlled by computer at 60 Hz, and the measured transfer function of the system (verified by accelerometer measurements) allowed us to accurately record the platform's position at all times<sup>28,29</sup>. A room-mounted laser was used to ensure that platform motion was precisely synchronized with the video display (to within ~1 ms). Additional details are available in a previous publication<sup>28</sup>.

For each isolated neuron, we first obtained quantitative measurements of direction tuning and receptive field location, as described previously<sup>30</sup>, and we used those measurements to specify the location, size, and direction axis of the random-dot stimulus. To measure speed tuning, each neuron was tested with random-dot patterns that had speeds of 0, 0.5, 1, 2, 4, 8, 16 and 32 degrees/s. Preferred speeds were obtained from the peak of a fitted Gamma function<sup>19</sup>. For the main experimental conditions, the OpenGL

graphics library was used to render a fixation target and a random-dot surface in a world-fixed virtual environment (Fig. 1c). The random-dot surface was a portion of a cylinder (analogous to the geometric horopter) that was oriented perpendicular to the axis of translation (Supplementary Fig. S1). Five hundred ms after the monkey achieved fixation, the random dot stimulus was presented in the neuron's receptive field. In each trial, one of nine simulated depths was chosen pseudo-randomly from the following set of equivalent disparities:  $0.0^\circ$ ,  $\pm 0.5^\circ$ ,  $\pm 1.0^\circ$ ,  $\pm 1.5^\circ$ , and  $\pm 2.0^\circ$  (depth from motion parallax is commonly expressed in units of the equivalent binocular disparity<sup>4</sup>). ‘Null’ trials, in which no stimulus was presented over the receptive field, were also interleaved. Stimuli were viewed monocularly and the random-dot patch was positioned and scaled so as to eliminate all other cues to depth. The random-dot surface was transparent when it overlapped the fixation target such that no occlusion cues were present (see Supplementary Fig. S1 for additional details regarding stimulus generation). For almost all cells, visual stimuli were presented to the eye contralateral to the recording hemisphere.

Our experiment consisted of two stimulus conditions. In the Motion Parallax (MP) condition, the animal was translated through one cycle of a 0.5 Hz sinusoid having a total displacement of 4 cm, which is slightly more than one interocular separation (Fig. 1d). The 2 s movement trajectory was windowed with a high powered Gaussian to smooth out the beginning and end of the movement. The axis of platform motion was chosen according to the visual direction tuning of each MT neuron, such that random-dot motion oscillated along the neuron's preferred-null axis. Thus, all platform movements were along an axis in the fronto-parallel plane. Motion of the random-dot patch was consistent with that produced by a stationary surface. Thus, the boundaries of the random-dot patch moved relative to the receptive field (typically by <25% of the receptive field diameter) during each trial. Importantly, the excursion of the patch relative to the receptive field was matched for near and far simulated depths, such that it could not generate a depth-sign preference.

Two opposite phases of platform motion were used (Fig. 1d, thick curves), such that retinal image motion began in the neuron's preferred direction on one-half of the trials, and began in the null direction on the remaining half. Throughout the movement,

the monkey's only task was to fixate the world-fixed target, and successful completion of the trial required that his gaze remained within an electronic fixation window. To allow the monkey an opportunity to make an initial catch-up saccade (if necessary) at the onset of pursuit, the fixation window was initially 3.0-4.0 degrees square, and shrunk to 1.5-2.0 degrees after 250 ms. Horizontal and vertical eye position was monitored with a scleral search coil and captured at a sampling rate of 250 Hz. For illustrative purposes, eye traces in Figure 1d and Supplementary Figure S2 were smoothed with a boxcar filter (position: 50ms wide; velocity: 200ms wide).

In the Retinal Motion (RM) condition, the monkey remained stationary and we replicated the visual image seen during the MP condition by translating the OpenGL camera along the same trajectory that the monkey followed in the MP condition. As the OpenGL camera translated, it also rotated to maintain 'aim' at the fixation target. This simulates the smooth tracking eye movements of the animal, and generates visual stimuli that match those in the MP condition. The monkey performed 6-10 repetitions of each stimulus condition at each simulated depth, randomly interleaved.

### Analysis

Single-unit data were analyzed using custom software written in Matlab (Mathworks Inc.). For generation of peri-stimulus time histograms (PSTHs, Fig. 2a,b), firing rate was computed in 50 ms bins. To quantify selectivity for depth sign, we combined data across the two phases of platform motion and computed a mean firing rate across the total duration of each trial. Spikes were counted within a temporal window that began 80ms after stimulus onset and ended 80ms after stimulus offset (to compensate for response latency). For each neuron we computed a Depth-Sign Discrimination Index (DSDI):

$$DSDI = \frac{1}{4} \sum_{i=1}^4 \frac{R_{far(i)} - R_{near(i)}}{|R_{far(i)} - R_{near(i)}| + \sigma_{avg}}.$$

For each pair of depths symmetric around zero (e.g. +/-1 degree), we calculated the difference in response between far ( $R_{far}$ ) and near ( $R_{near}$ ), relative to response variability ( $\sigma_{avg}$ , the average standard deviation of the two responses). We then averaged across the four matched pairs of depth to obtain the DSDI which ranges from -1 to 1. Neurons that respond more strongly to near stimuli will have negative values, while neurons that prefer far stimuli will have positive values. A DSDI was calculated separately for the MP and RM conditions. DSDIs were classified as significantly different from zero (or not) by permutation test (1000 permutations,  $p < 0.05$ ).

The above metric has the advantage of taking into account trial-to-trial variations in response while quantifying the magnitude of response differences between near and far. However, we also analyzed our data using a standard contrast measure and obtained very similar results (Supplementary Fig. S7).

Tuning curves in the MP condition were classified as monotonic if: i) the mean firing rates were a strictly monotonic function of equivalent disparity (i.e., steadily increasing or decreasing), or ii) a second-order polynomial ( $y = a_1 x + a_2 x^2 + a_3$ ) did not provide a significantly better fit than a linear function ( $y = a_1 x + a_2$ ) by sequential F-test ( $p > 0.05$ ), or iii) the best-fitting second-order polynomial was itself strictly monotonic within the range of the data. Tuning curves not meeting any of these conditions were classified as non-monotonic (e.g., Fig. 3a).

### **Supplementary information line**

Supplementary Information accompanies the paper on [www.nature.com/nature](http://www.nature.com/nature).

### **Author information**

Reprints and permissions information is available at [npg.nature.com/reprintsandpermissions](http://npg.nature.com/reprintsandpermissions). The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to [gdeangelis@cvs.rochester.edu](mailto:gdeangelis@cvs.rochester.edu).

## Figure Legends

**Figure 1.** Schematic illustration of motion parallax and stimulus design. **a**, As the head moves to the right, the image of a near object moves leftward, while the image of a far object moves rightward. **b**, The opposite occurs during leftward head movement. Without pictorial depth cues, an extra-retinal signal is needed to determine depth sign. **c**, Random-dot stimuli were scaled so that size and density were identical across simulated depths. Three depths—far (+1°), near (-1°), and zero—are illustrated. **d**, Thick black and gray curves represent the motion trajectory for two possible starting phases. Thin curves represent average eye position and velocity traces for a single session, in equivalent stimulus units.

**Figure 2.** A neuron selective for depth from motion parallax. **a**, Responses in the Retinal Motion (RM) condition for five of nine simulated depths tested. One column of PSTHs is shown for each starting phase of motion. Grey traces represent retinal image velocity, with peaks representing motion in the neuron's preferred direction. 'Null' responses were obtained when no random-dots were presented. **b**, The Motion Parallax (MP) condition. Responses to near stimuli are accentuated, while responses to far stimuli are suppressed relative to the RM condition. **c**, Depth-tuning curves for the RM (open symbols, DSDI = -0.73) and MP (filled symbols, DSDI = 0.15) conditions. Error bars represent SEs, and the dashed horizontal line indicates average spontaneous activity.

**Figure 3.** Tuning curves from six additional MT cells. Filled and open symbols show firing rates ( $\pm$  SE) in the MP and RM conditions, respectively. Asterisks denote significant DSDI values ( $p < 0.05$ ). **a, c, e**) Neurons that prefer near stimuli in the MP condition. Speed preferences are 16, 27, and 17 deg/sec, respectively. **b, d, f**) Neurons that prefer far stimuli. Cells in b and d prefer slow speeds (1.0 and 0.0 deg/sec, respectively) and thus have RM tuning curves that are symmetrically peaked. The neuron in f has a moderate speed preference (5.2 deg/sec).

**Figure 4.** Many MT neurons are selective for depth defined by motion parallax. **a**, DSDI values for the RM condition are plotted against those for the MP condition ( $N=144$  neurons from two monkeys; M1: circles, M2: triangles). In the marginal histograms, solid bars represent DSDI values significantly different from zero ( $p < 0.05$ ). **b**, DSDI values in the MP condition are correlated with speed preferences ( $r = -0.57$ ,  $p < 0.0001$ , Spearman rank correlation). Individual cells are coded for significance of the DSDI in the MP condition (filled:  $p < 0.05$ ; open:  $p > 0.05$ ).

## References

1. Rogers, B. & Graham, M. Motion parallax as an independent cue for depth perception. *Perception* 8, 125-34 (1979).
2. Rogers, B. J. Motion parallax and other dynamic cues for depth in humans. *Rev Oculomot Res* 5, 119-37 (1993).
3. Rogers, B. J. & Collett, T. S. The appearance of surfaces specified by motion parallax and binocular disparity. *Q J Exp Psychol A* 41, 697-717 (1989).
4. Rogers, B. J. & Graham, M. E. Similarities between motion parallax and stereopsis in human depth perception. *Vision Res* 22, 261-70 (1982).
5. Bradshaw, M. F. & Rogers, B. J. The interaction of binocular disparity and motion parallax in the computation of depth. *Vision Res* 36, 3457-68 (1996).
6. Ellard, C. G., Goodale, M. A. & Timney, B. Distance estimation in the Mongolian gerbil: the role of dynamic depth cues. *Behav Brain Res* 14, 29-39 (1984).
7. van der Willigen, R. F., Frost, B. J. & Wagner, H. Depth generalization from stereo to motion parallax in the owl. *J Comp Physiol A Neuroethol Sens Neural Behav Physiol* 187, 997-1007 (2002).
8. Kral, K. Behavioural-analytical studies of the role of head movements in depth perception in insects, birds and mammals. *Behav Processes* 64, 1-12 (2003).
9. DeAngelis, G. C., Cumming, B. G. & Newsome, W. T. Cortical area MT and the perception of stereoscopic depth. *Nature* 394, 677-80 (1998).
10. Uka, T. & DeAngelis, G. C. Contribution of middle temporal area to coarse depth discrimination: comparison of neuronal and psychophysical sensitivity. *J Neurosci* 23, 3515-30 (2003).
11. Uka, T. & DeAngelis, G. C. Linking neural representation to function in stereoscopic depth perception: roles of the middle temporal area in coarse versus fine disparity discrimination. *J Neurosci* 26, 6791-802 (2006).
12. Uka, T. & DeAngelis, G. C. Contribution of area MT to stereoscopic depth perception: choice-related response modulations reflect task strategy. *Neuron* 42, 297-310 (2004).
13. Farber, J. M. & McConkie, A. B. Optical motions as information for unsigned depth. *J Exp Psychol Hum Percept Perform* 5, 494-500 (1979).
14. Rogers, S. & Rogers, B. J. Visual and nonvisual information disambiguate surfaces specified by motion parallax. *Percept Psychophys* 52, 446-52 (1992).
15. Nawrot, M. Eye movements provide the extra-retinal signal required for the perception of depth from motion parallax. *Vision Res* 43, 1553-62 (2003a).
16. DeAngelis, G. C. & Uka, T. Coding of horizontal disparity and velocity by MT neurons in the alert macaque. *J Neurophysiol* 89, 1094-111 (2003).
17. Maunsell, J. H. & Van Essen, D. C. Functional properties of neurons in middle temporal visual area of the macaque monkey. II. Binocular interactions and sensitivity to binocular disparity. *J Neurophysiol* 49, 1148-67 (1983b).
18. Maunsell, J. H. & Van Essen, D. C. Functional properties of neurons in middle temporal visual area of the macaque monkey. I. Selectivity for stimulus direction, speed, and orientation. *J Neurophysiol* 49, 1127-47 (1983a).
19. Nover, H., Anderson, C. H. & DeAngelis, G. C. A logarithmic, scale-invariant representation of speed in macaque middle temporal area accounts for speed discrimination performance. *J Neurosci* 25, 10049-60 (2005).

20. Born, R. T. & Bradley, D. C. Structure and function of visual area MT. *Annu Rev Neurosci* 28, 157-89 (2005).
21. Bradley, D. C., Qian, N. & Andersen, R. A. Integration of motion and stereopsis in middle temporal cortical area of macaques. *Nature* 373, 609-11 (1995).
22. Bradley, D. C., Chang, G. C. & Andersen, R. A. Encoding of three-dimensional structure-from-motion by primate area MT neurons. *Nature* 392, 714-7 (1998).
23. Dodd, J. V., Krug, K., Cumming, B. G. & Parker, A. J. Perceptually bistable three-dimensional figures evoke high choice probabilities in cortical area MT. *J Neurosci* 21, 4809-21 (2001).
24. Cao, A. & Schiller, P. H. Neural responses to relative speed in the primary visual cortex of rhesus monkey. *Vis Neurosci* 20, 77-84 (2003).
25. Newsome, W. T., Wurtz, R. H. & Komatsu, H. Relation of cortical areas MT and MST to pursuit eye movements. II. Differentiation of retinal from extraretinal inputs. *J Neurophysiol* 60, 604-20 (1988).
26. Bremmer, F., Ilg, U. J., Thiele, A., Distler, C. & Hoffmann, K. P. Eye position effects in monkey cortex. I. Visual and pursuit-related activity in extrastriate areas MT and MST. *J Neurophysiol* 77, 944-61 (1997).
27. Nawrot, M. & Joyce, L. The pursuit theory of motion parallax. *Vision Res* 46, 4709-25 (2006).
28. Gu, Y., Watkins, P. V., Angelaki, D. E. & DeAngelis, G. C. Visual and nonvisual contributions to three-dimensional heading selectivity in the medial superior temporal area. *J Neurosci* 26, 73-85 (2006).
29. Gu, Y., DeAngelis, G. C. & Angelaki, D. E. A functional link between area MSTd and heading perception based on vestibular signals. *Nat Neurosci* 10, 1038-47 (2007).
30. Nguyenkim, J. D. & DeAngelis, G. C. Disparity-based coding of three-dimensional surface orientation by macaque middle temporal neurons. *J Neurosci* 23, 7117-28 (2003).

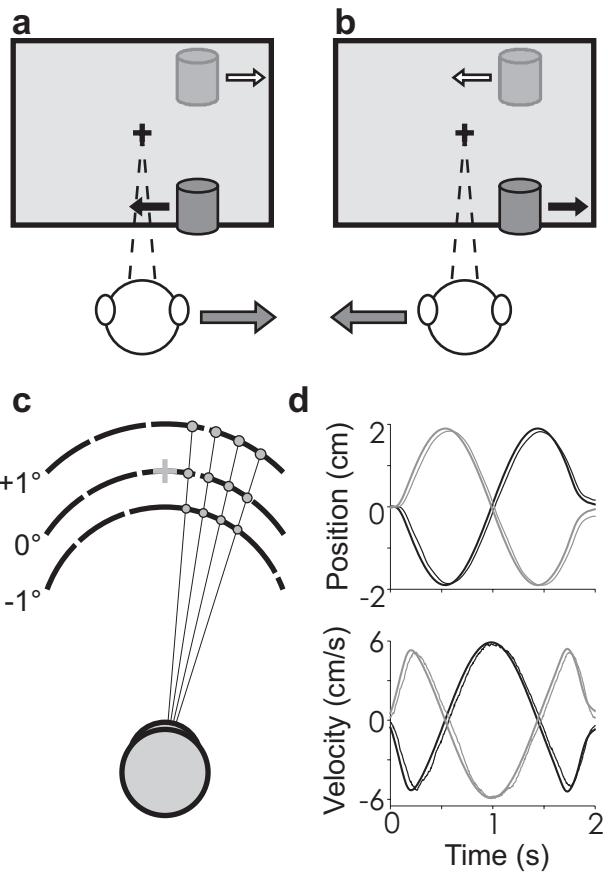


FIGURE 1

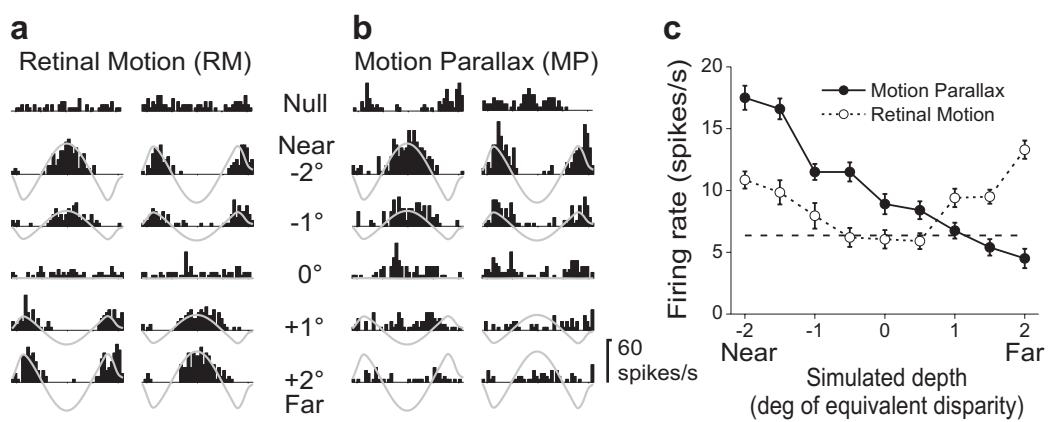
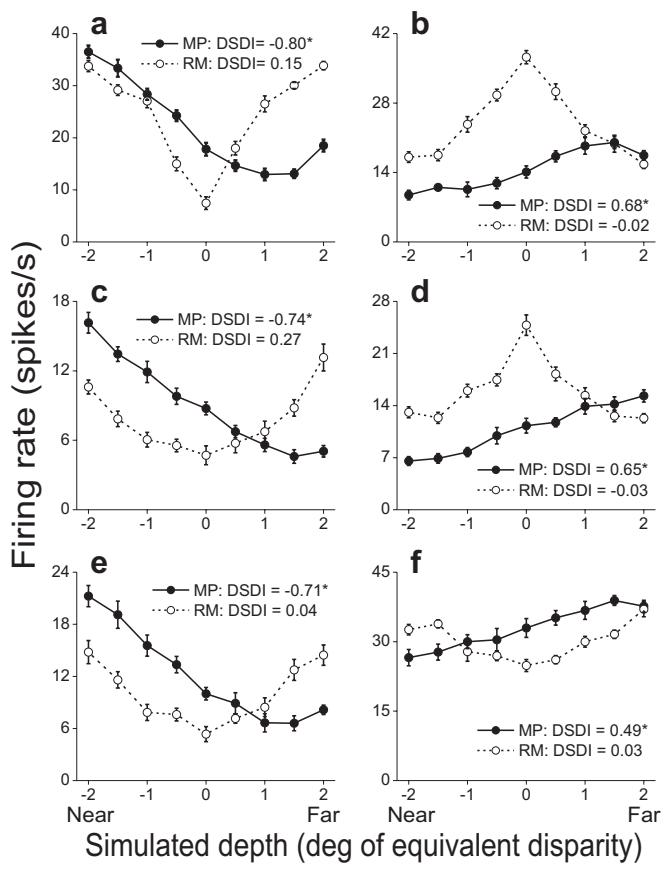


FIGURE 2



**FIGURE 3**

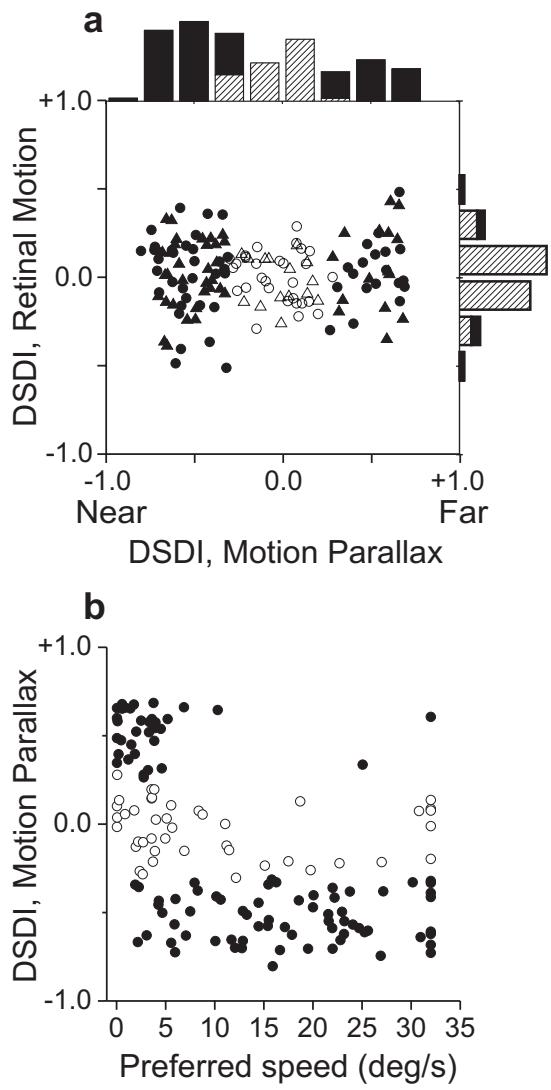


FIGURE 4