Supplementary Figure 1

Illustration of two additional viewing geometries.

(a) Illustration of the case of pure translation of the eye relative to the scene. An observer's head translates from left to right while the eye remains stationary relative to the head. This produces no perspective distortion under planar image projection. For a dynamic version, see Supplementary Movie 2. (b) Illustration of the case of a pure eye rotation, with no eye or head translation (e.g., smooth pursuit of a target). This produces dynamic perspective distortion in the image plane but not in spherical (retinal) coordinates (see Supplementary Movie 3).
Supplementary Figure 2

Summary of stimulus and mask dimensions.

(a) Each red circle represents the size and location of the random-dot patch that was placed over the receptive field of a single MT neuron. Each blue circle (centered on a red circle) indicates the size of the mask region that was used to prevent background dots from entering the receptive field. (b) The sizes of the random-dot patch (red) and the mask (blue) are plotted against receptive field eccentricity. Each neuron is represented by a pair of red and blue data points that are vertically aligned. Masks were generally 2-3 times larger than the stimulus patch (geometric mean of the ratio of diameters = 2.79), and the mask was large enough to encompass the fixation target for 85/96 neurons (mask radius was not saved for the initial 5 neurons tested).
Quantification of depth-sign discrimination capacity of single MT neurons using ROC analysis.

For each depth magnitude, the ability of each MT neuron to discriminate between near and far stimuli was quantified by applying ROC analysis to distributions of responses corresponding to the neuron’s preferred and non-preferred depth signs (as defined by the sign of DSDI). The area under the ROC curve represents the ability of an ideal observer to discriminate between the preferred and non-preferred depth signs—a value of 0.5 corresponds to chance performance. (a) Distribution of ROC areas for each depth magnitude tested in the Retinal Motion condition; arrowheads show the median values. Filled bars indicate neurons with ROC values that are significantly different from 0.5 (permutation test, P < 0.05). The overall median value across all depth magnitudes was 0.63. (b) Distributions of ROC areas for the Dynamic Perspective condition. Median values are 0.78, 0.88, 0.88, and 0.80 for depth magnitudes of 0.5, 1.0, 1.5, and 2.0 deg, respectively. The overall median across depth magnitudes is significantly greater than that for the Retinal Motion condition (n = 412, P = 6.7 × 10^{-23}, Wilcoxon signed rank test). (c) Distributions of ROC areas for the Motion Parallax condition. Median ROC areas for all depth magnitudes are 1, and the overall median is significantly greater than that for the Dynamic Perspective condition (n = 412, P = 8.3 × 10^{-16}, Wilcoxon signed rank test). (d) ROC areas for each neuron were averaged across the four depth magnitudes and the average ROC area was plotted against the absolute DSDI value for each neuron: Retinal Motion (black), Dynamic Perspective (magenta), and Motion Parallax (blue). The two metrics are strongly correlated (n = 309, ρ = 0.98, P = 6.9 × 10^{-223}, Spearman rank correlation), indicating that DSDI is an effective measure of how well neurons discriminate depth sign.
Supplementary Figure 4

Depth-sign selectivity in the retinal motion condition is correlated with dynamic perspective information (DPI) in the stimulus.

For each point in the image, in stimulus coordinates (x, y), eye rotation relative to the scene induces a component of velocity orthogonal to the axis of translation (Eqn. 4), and this component is proportional to the product of the location coordinates of that image point, xy. Thus, we can approximate dynamic perspective information within a region of interest as $\sum |xy|$ over that region (Eqn. 5). For each neuron, the absolute value of DSDI in the Retinal Motion condition is plotted as a function of DPI computed over the stimulus region overlying the receptive field. The two variables are significantly correlated ($\rho = 0.24$, $P = 0.02$, Spearman rank correlation), indicating that significant depth-sign selectivity in the Retinal Motion condition generally arises when receptive fields are large and located away from the visual field meridia, such that DPI is larger within the region of the stimulus.
Supplementary Figure 5

Eye movements do not drive depth-sign tuning in the dynamic perspective condition.

Eye movements were quantified by computing pursuit gain, which is the amplitude of the 0.5Hz component of eye velocity divided by the corresponding frequency component of target velocity. (a) Pursuit gain in the Dynamic Perspective condition is plotted against that in the Retinal Motion condition. Filled circles and open triangles represent data from monkeys M1 and M2, respectively. Green symbols denote cases in which pursuit gain is significantly different between the Dynamic Perspective and Retinal Motion conditions, whereas red symbols denote cases with no significant difference. (b) Absolute value of DSDI is plotted against pursuit gain for the Dynamic Perspective condition. Data are from monkey M1, and show no significant correlation ($\rho = -0.18$, $P = 0.22$, Spearman rank correlation). (c) Corresponding data from monkey M2, format as in b. Again, the correlation is not significant ($\rho = -0.25$; $P = 0.07$).
Supplementary Figure 6

Depth-sign selectivity in the dynamic perspective condition is not correlated with surround suppression.

We examined whether the modulatory effects of background motion in the Dynamic perspective condition are correlated with surround suppression in MT neurons. Surround suppression was quantified by analyzing size tuning curves and computing the percentage of surround suppression (see Methods). We found no significant correlation between the magnitude of depth-sign selectivity in the Dynamic Perspective condition and the percentage of surround suppression (n = 102; \( \rho = -0.19; P = 0.06 \), Spearman rank correlation). In fact, the correlation is slightly negative, indicating that cells with strong surround suppression tend to have slightly weaker depth-sign selectivity. This result is consistent with the finding that MT neurons still showed depth-sign selectivity in the DPbalanced condition, in which the velocity distributions of near and far dots moving in opposite directions are matched. Effects of surround suppression were likely minimized because we masked a fairly large region around the classical receptive field, thus removing most visual stimulation from the suppressive surround (Supplementary Fig. 2).