

Mechanisms of Heading Perception in Primate Visual Cortex

David C. Bradley, Marsha Maxwell, Richard A. Andersen,*
Martin S. Banks, Krishna V. Shenoy

When we move forward while walking or driving, what we see appears to expand. The center or focus of this expansion tells us our direction of self-motion, or heading, as long as our eyes are still. However, if our eyes move, as when tracking a nearby object on the ground, the retinal image is disrupted and the focus is shifted away from the heading. Neurons in primate dorso-medial superior temporal area responded selectively to an expansion focus in a certain part of the visual field, and this selective region shifted during tracking eye movements in a way that compensated for the retinal focus shift. Therefore, these neurons account for the effect of eye movements on what we see as we travel forward through the world.

When we move along a straight path, the retinal image appears to expand, creating a vector field on the retina in which all vector directions point away from the focus of expansion. When the eyes are still, this focus corresponds to the direction of self-motion (the heading), which humans can easily identify (1, 2). When the eyes move, the problem becomes more complicated. If we turn our eyes leftward to maintain fixation on an object to the left of our motion path, a rightward laminar flow component is added to the expanding retinal field. This combined field is similar to the original expansion field (when the scene contains little depth variation), but the apparent

focus position becomes shifted to the left of the heading (Fig. 1). Humans can still estimate heading during pursuit eye movements (2, 3), showing that they can correct for this shift. Recent psychophysical experiments have shown that this correction requires eye-movement information (3, 4). One might therefore expect the brain site or sites responsible for heading computation to process smooth-pursuit eye-movement signals as well as optic flow.

We examined the dorso-medial superior temporal (MSTd) area because its neurons are responsive to optic-flow stimuli (for example, expansion) (5-7) as well as to smooth-pursuit eye movements (7-9). Many MSTd cells are also selective for the position of the focus of expansion (10). Our goal was to determine how optic flow and eye velocity signals might be used in area MSTd to compute the direction of self-motion.

We trained a rhesus monkey to fixate on a point on a computer screen (11). The point was either stationary or was moved at a constant velocity of 15.7°/s (12). While the monkey fixated, we showed a large, expanding random-dot pattern for 1 s. The position of the focus in this expansion was varied along an axis to simulate different directions of self-motion, or headings. This axis was made parallel with each neuron's preferred pursuit direction (that eliciting the strongest response), which was determined in preliminary tests (13).

During the stimulus presentation we recorded single-neuron activity with a microelectrode inserted in MSTd cortex (14). The optic flow preference (that is, expansion, contraction, or rotation) of each MSTd cell was first determined. If the cell responded best to expansion, we carried out the following tests (cells preferring contraction and rotation were also studied and are discussed below). First, while the monkey fixated a stationary point, we recorded the cell's firing rate as a function of the position of the flow focus on the screen (the simulated heading) (15). This allowed us to measure the neu-

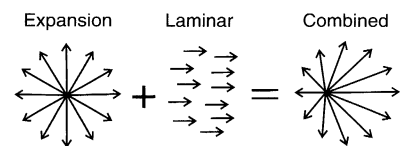


Fig. 1. When we move forward, the visual world appears to expand. If the eyes are still, the focus position tells us our direction of heading. However, leftward eye movement adds rightward laminar flow motion to the retinal image, which shifts the focus. To recover the heading direction, we must correct for this focus shift.

D. C. Bradley, M. Maxwell, R. A. Andersen, K. V. Shenoy, Division of Biology, California Institute of Technology, Pasadena, CA 91125, USA.
M. S. Banks, School of Optometry, University of California, Berkeley, CA 94720-2020, USA.

*To whom correspondence should be addressed.

ron's focus tuning, that is, its response profile as a function of heading. These tuning curves were usually well defined and roughly sigmoidal or gaussian in shape. The tuning curve peaks were fairly evenly distributed from -40° to 40° , suggesting a continuum of preferred headings.

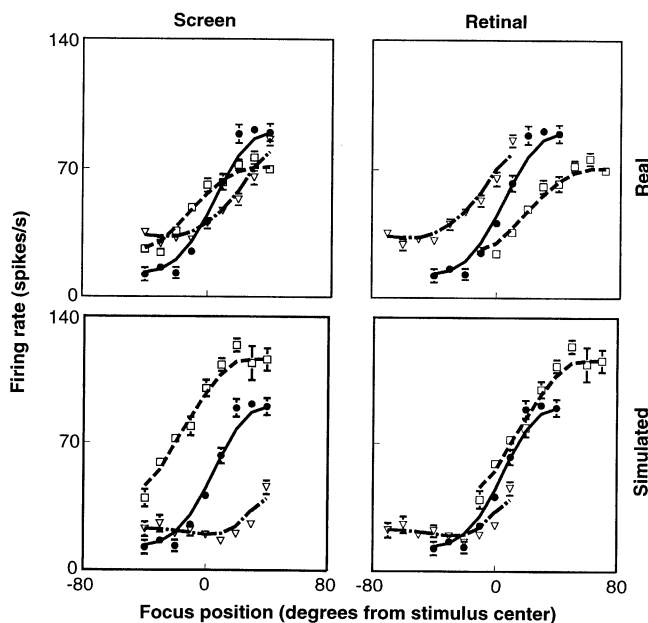
When the eyes are fixed, the heading is given by the retinal focus position. Therefore, we cannot tell whether the fixed-eye focus tuning curves represent heading or simply the position of the focus on the retina. To make this distinction, we showed the same expanding patterns during a smooth-pursuit eye movement of 15.7° per second. The eye movement was made in the neuron's preferred direction or the antipreferred (opposite) direction to elicit the largest effect. Data were recorded over a 500-ms interval during which the eye moved 8° . The average eye position during the recording period was the same as in the fixed-eye experiment. However, the eye's instantaneous velocity displaced the apparent focus by 30° in the direction of the eye movement. The simulated heading in screen coordinates remained unchanged.

If MSTd activities reflect the heading, we expect their fixed- and moving-eye focus tuning to be the same when expressed in screen coordinates (16). On the other hand, if MSTd response is determined solely by the pattern of retinal image motions, we expect the tuning curves to be the same when expressed in retinal coordinates. We found both types of cell, as well as intermediate types (17). A neuron for which responses were similar in screen coordinates is shown in Fig. 2. In this cell, an eye movement in the preferred or

antipreferred direction caused the focus tuning to shift by $\sim 30^\circ$ in retinal coordinates (top right panel) such that the two moving-eye fields and the fixed-eye field were nearly superimposed when plotted in screen coordinates (top left). Thus, activity in this neuron reflected the position of the screen focus rather than the retinal focus.

The above experiment suggests that an eye-movement signal is used to shift the retinal focus tuning. This is because the expansion stimulus is two-dimensional and thus devoid of retinal cues, such as motion parallax, that might give information about the rate of eye movement (18, 19). To prove that eye-movement signals are used, we presented a retinal stimulus virtually identical to that in the eye-movement experiment above, except that the eyes were held stationary. This stimulus was achieved by having the animal fixate a stationary point while the expansion stimulus, including its border, was moved in the opposite direction of the simulated eye movement. It is important to note that the retinal image over time is the same in this and the real eye-movement condition; the only difference is that the eye moves in the real movement condition and does not move in the simulated movement condition (20). In most instances the simulated pursuit experiment caused no shift in retinal focus tuning (Fig. 2, bottom right). Therefore, the shifts in focus tuning observed in our experiments can be attributed only to the influence of an extraretinal eye-movement signal. The effect of this signal in bringing focus tuning curves into screen coordinates is seen by

Fig. 2. An MSTd heading cell. In all panels, the solid lines and solid circles represent fixed-eye focus tuning (identical in all four graphs), the dashed lines and open squares are preferred-direction eye movements (real or simulated), and the dot-and-dashed-lines and open triangles are antipreferred-direction eye movements (real or simulated). Data in the left and right columns are identical, except that pursuit curves in the right column were shifted by 30° relative to screen coordinates (thus giving retinal coordinates). The moving-eye focus tuning curves align in screen coordinates (top left panel) and thus encode the direction of heading. However, for simulated eye movements, the fields align in retinal coordinates. Smooth curves are five-point moving averages of the data. Data points are shown as the mean \pm SEM for four replicates, where each replicate is the mean firing during the middle 500 ms of the stimulus-presentation interval.



comparing pursuit curves between the lower and upper left panels: in the simulated condition, the curves are not aligned in screen coordinates (lower left) whereas they are in the real pursuit condition (upper left).

To summarize results for the entire neuron sample, we measured the cross-correlation coefficient (R_c) between fixed- and moving-eye focus tuning curves for a variety of relative shifts between the two curves (21). A high R_c value indicates that the two curves are well aligned at a given relative shift. When pursuit eye movements were simulated, the optimal shift (that maximizing R_c) was near zero with respect to retinal coordinates (Fig. 3A). In contrast, during real eye movements, the relative shift was $\sim 15^\circ$ away from retinal coordinates and toward screen coordinates. This does not necessarily imply that all neurons shifted by 15° ; instead, optimal shifts were relatively

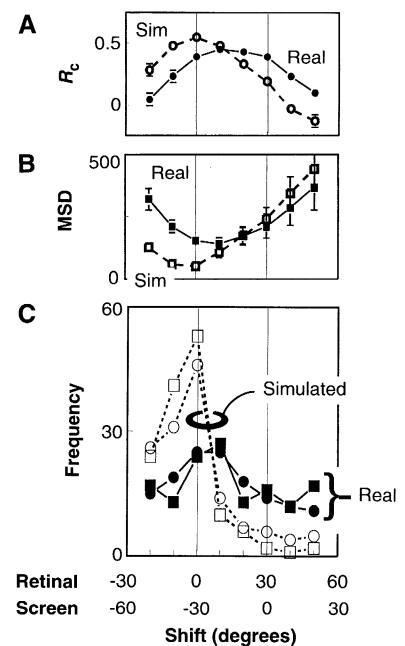


Fig. 3. Population data representing the entire MSTd sample ($n = 139$). x axis represents relative shift between fixed- and moving-eye focus tuning curves. Each panel compares results for real and simulated (Sim) pursuits. **(A)** For each relative shift, a correlation was done between the overlapping regions of the two curves. High R_c values imply good alignment on the x axis. One cross-correlogram (R_c versus shift) was defined for each neuron and for each pursuit direction; cross-correlograms were then averaged over cells to form the data of this panel. **(B)** MSD between focus tuning curves. Averages computed as for R_c (above). **(C)** Frequency distribution for "optimal" relative shift, defined in terms of maximal R_c (circles) or minimal MSD (squares). One optimal shift value was defined for each cell (for the preferred pursuit direction); the histogram reflects the frequency of each optimal shift. In **(A)** and **(B)** data are means \pm SEMs ($n = 139$).

evenly distributed between -20° and 50° (Fig. 3C). By comparison, this distribution for simulated eye movements was heavily stacked in the region of zero shift. Thus, the effect of actual pursuit was to cause many neurons to shift their focus tuning toward a screen coordinate frame. Results were similar when the optimal shift was calculated by minimizing the mean squared difference (MSD) between fixed- and moving-eye tuning curves (Fig. 3, B and C). Also, results were similar for preferred- and antipreferred-direction eye movements; the shifts were opposite in the two cases (Fig. 2).

Although many MSTd focus tuning curves shift toward screen coordinates during pursuit, others do not (Figs. 3C and 4A). Additional experiments will be needed to understand this variability and how it is related to the perception of heading. Perhaps only part of the MSTd population is involved in heading computation, whereas the rest perform functions that do not require pursuit compensation. Another explanation, however, is that the distribution of shifts reflects the transformation from retinal to screen coordinates. In this instance, one would expect to find neurons representing various stages of the coordinate transformation, that is, nonshifting cells (retinal coordinates), shifting cells (screen coordinates), and possibly cells corresponding to intermediate stages of the transformation. Presumably, heading judgments would be based on the activity of heading cells rather than on those taking part in the transformation. However, it is possible that all MSTd activities are sensed at once, convey-

ing a mean shift of about 15° (Fig. 3C); such a result would be consistent with psychophysical experiments showing that subjects may only partly correct for pursuit eye movements when judging stimulus motion in screen coordinates (22, 23).

That area MSTd may convert retinal to screen coordinates is supported by the finding that in most nonshifting neurons, response amplitude was modulated by pursuit (24) (average amplitude change, $25 \pm 3\%$) (Fig. 4A). If the output of two such neurons were summed, the resulting focus tuning could shift if the two input tuning curves were offset relative to each other. To test this idea, we modeled focus tuning curves for heading cells as the sum of two mutually offset sine functions (representing the nonshifting cells) (25). By changing the amplitude of the sines, it was possible to re-create the measured heading cell tuning curves consistently (mean $r^2 = 0.72 \pm 0.03$ when all heading cell tuning curves were fit with this model). An example is shown in Fig. 4B. The shift of this cell's tuning curve occurs because its two input functions are offset relative to each other and the gain changes on these functions alter their relative inputs to the heading cell. Thus, our modeling suggests that heading (shifting) cells could be built up from nonshifting, retinal cells in area MSTd.

The data in Fig. 3 were pooled from neurons tuned not only for expansion (41%), but also for contraction (33%) and rotation (27%). The data were combined because the results from the three cell types were nearly identical (26). Thus, for expansion, contraction, and rotation cells, focus tuning shifted

during eye movement (compared with simulated eye movement) by $18^\circ \pm 3^\circ$, $15^\circ \pm 3^\circ$, and $17^\circ \pm 4^\circ$, respectively, in all instances toward a screen coordinate frame (27). This is remarkable because expansion, contraction, and rotation patterns require focus tuning shifts in different directions to compensate for focus shifts during eye movements. For example, an eye movement to the left shifts the retinal focus of an expansion to the left (Fig. 1), but the retinal focus of a clockwise rotation is shifted downward (Fig. 5). The expansion cell would thus require a leftward tuning shift and the rotation cell a downward shift to compensate for the retinal focus shift. This is indeed what MSTd neurons appear to do, suggesting a general mechanism that compensates for the effect of pursuit eye movements on various types of retinal flow, not just expansion.

In conclusion, many MSTd neurons are tuned to the position of the focus in an optic flow pattern (for example, expansion, contraction, and rotation). In addition, many neurons shift their retinal focus tuning during eye movement in the same direction as the shift in the retinal focus induced by that movement (28). This focus tuning shift can be attributed only to an eye-movement signal to area MSTd because an identical retinal stimulus in the absence of eye movement produces no such shift. In the case of an expansion stimulus, focus tuning shifts are in the appropriate direction for encoding the heading. However, the shifts for rotation and contraction cells may indicate a more general phenomenon of perceptual stability in the face of retinal motion due to smooth-pursuit eye movements. Therefore, for instance, when the eyes track across a rotating wheel, the wheel does not appear to move up or down (29). Thus, area MSTd may play a general role in using eye-movement signals to compensate for self-induced image motion—that is, the motion we see because of the eye movements themselves. One important outcome of this compensation is the ability to compute the direction of self-motion.

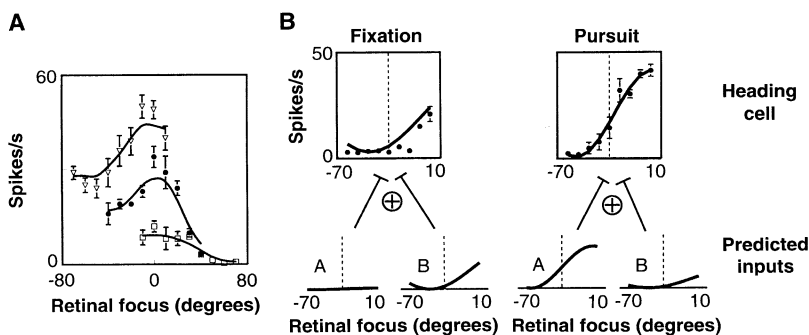


Fig. 4. (A) Example of a retinal (nonshifting) neuron in which the response amplitude is modulated by pursuit. The three curves, corresponding to stationary fixation and opposite directions of pursuit, all peak at $\sim 0^\circ$, but their amplitudes vary substantially. This type of cell could serve as input to the heading (shifting) cells. (B) Example of a heading neuron and how it might be constructed from retinal neurons. The upper graphs show the neuron's measured focus tuning during fixation (left) and preferred-direction pursuit (right). The neuron shifted its retinal focus tuning during pursuit in such a way as to compensate for the retinal focus shift induced by that pursuit. Circles: mean response; curves: model fit. (Lower panels) Predicted input functions. Each function is characterized by three sine-wave parameters and multiplied by a gain; two functions are summed to make the focus tuning curve of a heading cell. All parameters were adjusted by nonlinear regression to fit the data (upper panels). The focus tuning shift during pursuit was achieved by increasing the gain on function A while decreasing the gain on function B. The sine-wave parameters (other than gain) were identical for the fixation and pursuit conditions; only the gains were adjusted to simulate the focus tuning shift. The gain-modulated sine functions resemble neurons in the sample that have gain-modulated (nonshifting) focus tuning.

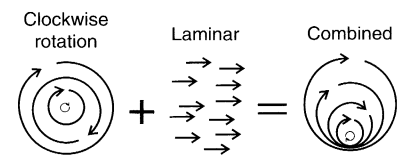


Fig. 5. The focus in a clockwise rotating pattern shifts downward during a leftward pursuit eye movement, which produces rightward laminar flow. To correct for this shift, clockwise-rotation cells must shift their retinal focus tuning curves downward. This is equivalent to the focus tuning shifts of expansion cells, which go in the direction of eye movement, because in both instances the neurons are changing their flow preference to accommodate more rightward linear motion.

REFERENCES AND NOTES

- J. J. Gibson, *Perception of the Visual World* (Houghton Mifflin, Boston, 1950).
- W. H. Warren and D. J. Hannon, *Nature* **336**, 162 (1988).
- C. S. Royden, M. S. Banks, J. A. Crowell, *ibid.* **360**, 583 (1992).
- The eye-movement signal could be a copy of the smooth-pursuit command sent to the heading computation site, or a proprioceptive signal from the eye muscles and the periorbital tissue.
- H. Saito *et al.*, *J. Neurosci.* **6**, 145 (1986).
- M. S. Graziano, R. A. Andersen, R. J. Snowden, *ibid.* **14**, 54 (1994).
- H. Sakata, H. Shibutani, K. Kawano, *J. Neurophysiol.* **49**, 1364 (1983).
- H. Komatsu and R. H. Wurtz, *ibid.* **60**, 580 (1988).
- W. T. Newsome, R. H. Wurtz, H. Komatsu, *ibid.*, p. 604.
- C. J. Duffy and R. H. Wurtz, *J. Neurosci.* **15**, 5192 (1995).
- The monkey was seated in a chair with its head fixed, facing a 100° by 100° projection screen. After maintaining fixation within a 5° by 5° window for the required time (3 s), the monkey was rewarded with a drop of water. Eye position was monitored with a scleral search coil [S. J. Judge *et al.*, *Vision Res.* **20**, 535 (1980)]. All procedures with animals were approved by the Caltech Institutional Animal Care and Use Committee.
- Expanding random dot fields were generated by simulating the approach to a wall at 38 cm/s at a distance of 57 cm. This condition generated a radial speed of 15.7° per second at an eccentricity of 30° from the focus. Dot lifetimes were 300 ms (to limit acceleration cues), after which they were renewed at random locations. The stimulus was 50° by 50° square and contained 77 dots. The focus position was varied in 10° steps from -40° to 40° along an axis parallel to the neuron's preferred pursuit direction (distances are relative to the center of the stimulus, which was approximately centered in the neuron's spatial receptive field). For the ±30° and ±40° focus positions, the focus was situated outside the visible stimulus. We generated the stimulus for the simulated eye movements by moving the entire stimulus on the screen (by computer computation of the dot positions); this is equivalent to adding a constant-velocity component to each point in the flow field, which shifts the focus and causes the stimulus borders to move across the screen. The room was completely dark except for the stimulus.
- Before studying a neuron's focus tuning, we determined its preferred pursuit direction by having the monkey follow a point moving at 0°, 45°, 90°, and so on, up to 315°. The direction eliciting the strongest response was taken as the preferred direction; the opposite direction was designated antipreferred.
- Single neurons were monitored extracellularly with varnish-coated tungsten microelectrodes advanced dorsoventrally through the dura and into area MSTd, which was identified on the basis of receptive field size (most were >50° wide and crossed over the vertical meridian), flow preference (for example, expansion and rotation), position invariance, and modulation during smooth pursuit. Spike times were stored on a personal computer for subsequent analysis. All data are expressed as the mean firing rate during the middle 500 ms of a 1-s stimulus-presentation period. Results are based on a sample of 139 neurons from one hemisphere. These results are corroborated by recent data from a second animal ($n = 18$; K. S. Shenoy, D. C. Bradley, R. A. Andersen, *Soc. Neurosci. Abstr.*, in press).
- In practice, all stimuli were presented in pseudorandom order.
- "Screen coordinates" is an operational term. Because the monkey's eyes and head were in a constant position (except for the small eye displacement during pursuit), we cannot tell whether MSTd cells code for heading in eye, head, body, or world coordinates. Our results imply simply that these neurons account for eye movement, the first step necessary for computing the heading.
- If we arbitrarily define retinal cells as shifting by -20°, -10°, or 0°, intermediate cells by 10° or 20°, and heading cells by 30°, 40°, or 50°, the percentages of the different cell types are about 42%, 31%, and 27%, respectively, where the shift estimate was based on maximum cross-correlation between the fixed- and moving-eye curves.
- H. C. Longuet-Higgins and K. Prazdny, *Proc. R. Soc. London Ser. B* **208**, 385 (1980).
- W. H. Warren, in *Perception of Space and Motion* (Academic Press, New York, 1995), chap. 3, pp. 263-325.
- Because we did not correct for tangent error on the screen, dot speeds (for the pursuit target as well as the stimulus dots) were slightly underestimated at eccentric positions. However, this error was only ~8% at 12.5° eccentricity, ~20% in the corner (35° eccentricity), and 0 in the center. Therefore, the real and simulated pursuit conditions differed slightly in terms of the distribution of speeds across the retina.
- To compute R_c for a given relative shift between curves, we correlated the responses on one curve with the responses on the other curve, pairing by x -axis value. Only overlapping regions of the curves could be cross-correlated. MSD was computed between corresponding data points on the two curves, again, only for overlapping regions.
- A. Mack and E. Herman, *Vision Res.* **18**, 55 (1978).
- T. C. A. Freeman, J. A. Crowell, M. S. Banks, *Invest. Ophthalmol. Vis. Sci.* **37**, S454 (1996).
- Heading cell receptive fields also showed gain changes. In fitting our model (Fig. 4) to the heading cell data, these gain changes were taken into account, along with the receptive field shifts. It is uncertain whether these gain changes encode information about the heading; it is possible that they simply average to zero.
- Sine functions were used because different parts of a sine function can approximate either a gaussian or a sigmoid function. Each sine function was characterized by amplitude, frequency, and phase, as well as by two "gain" parameters that were applied in the pursuit condition (one gain for each pursuit direction). All parameters were adjusted simultaneously, fitting three receptive fields (fixed-eye and two pursuit directions) concurrently, with the use of nonlinear least-squares regression. Analysis was done on a subset of 36 neurons, the receptive fields of which shifted during eye movements.
- The speed of rotating patterns was scaled so that dot speed was identical to the expanding patterns (15.7°/s) at 30° eccentricity. Eye movements cause the focus in a rotating pattern to shift orthogonally to the eye movement. Therefore, for rotating patterns the focus was also varied orthogonally—otherwise it would not be possible to measure relative shifts between fixed- and moving-eye receptive fields.
- Shifts were calculated on the basis of MSD.
- Because the retinal focus shift depends on the pursuit speed and the rate of image expansion, receptive field shifts of different sizes are required to compute heading under different conditions. Because we tested only one pursuit and one expansion rate, we do not know whether individual receptive fields shift by varying amounts or whether a population code is used to read out the heading from different neurons depending on the pursuit and expansion rates.
- Preliminary psychophysical experiments in our lab suggest that humans compensate at least partly for eye movements while pursuing across rotating stimuli.
- We are grateful to D. Ward and B. Gillikin for technical assistance and to W. Warren and J. Crowell for helpful comments. This work was funded by the National Eye Institute, the Sloan Foundation for Theoretical Neurobiology at Caltech, the Office of Naval Research, and the Air Force Office of Scientific Research.

19 April 1996; accepted 19 July 1996