NEURAL MECHANISMS FOR SELF-MOTION PERCEPTION IN AREA MST

Richard A. Andersen, Krishna V. Shenoy, James A. Crowell, and David C. Bradley
Division of Biology, California Institute of Technology, Pasadena, California

I. Introduction
II. Area MST—Optic Flow Selectivity
   A. Spiral Space
   B. Position and Form/Cue Invariance
   C. Anatomical Organization
III. Area MST—Shifting Receptive Fields
   A. Speed Tuning
   B. Gaze Rotations with Head Movements
   C. Gain Field Model for Compensation
   D. Psychophysics
IV. Conclusion
    References

I. Introduction

Research on the neural circuitry responsible for perception of self-motion has focused on the medial superior temporal area, particularly the dorsal division (MSTd). Cells in this area are selective for the location of the focus of expansion and to pursuit eye movements, two signals necessary for recovering the direction of self-motion (Gibson, 1950). Research reviewed here shows many interesting correlates between the perception of self-motion and the activity of MST neurons. In particular, the focus tuning curves of these cells adjust to take into account motions during eye movements using extra-retinal signals, similar to the results of human perceptual experiments. Eye rotations due to head movements are also compensated for perceptually, and the focus tuning of MST neurons are also compensated for during head-generated eye rotations. Finally, the focus tuning curves compensate for both the direction and the speed of eye rotations, similar to that found in psychophysical studies. However, there are also several aspects of MSTd activity that do not completely mesh with the perception of self-motion; these differences suggest that area MSTd is not the final stage or the only locus of brain activity which accounts for this percept. Finally we offer a “gain field” model, which explains how area MSTd neurons can compensate for gaze rotations.
II. Area MST—Optic Flow Selectivity

In the middle 1980s, two groups discovered cells in the medial superior temporal (MST) area that were sensitive to complex visual-motion patterns similar to those encountered during self-motion, termed optic flow (Sakata et al., 1985, 1986, 1994; Saito et al., 1986; Tanaka et al., 1986, 1989; Tanaka and Saito, 1989; Gibson, 1950). A number of types of motion patterns were used in these studies (e.g., expansion/contraction, rotation, and laminar motion); cells tended to respond selectively to particular types of motion patterns. Duffy and Wurtz (1991a, b) made the important observation that even though some MSTd neurons were selective for a single type of pattern (single-component cells in their terminology), many others had sensitivity for two (double-component) or even three (triple-component) types of motion patterns. For instance, a triple-component cell might respond to expansion, clockwise rotation, and leftward laminar motion.

A. Spiral Space

One powerful class of computational models for recovering the direction of self-motion is based on linear analyses of local regions of the flow field (Longuet-Higgins and Prazdny, 1980; Koenderink and Van Doorn, 1981; Rieger and Lawton, 1985; Hildreth, 1992). These techniques are used to recover the expansion component of flow, due to observer translation, from the complex visual motions produced by eye rotations. The finding of neurons in MST that were sensitive for expansion/contraction, curl, or laminar motion led to the idea that the brain might in fact be analyzing complex flow fields in terms of these particular components. In other words, the brain might use these simple types of motion patterns as a basis set for describing the more complex patterns observed during self-motion.

The general idea that the brain represents features of the environment using basis sets has been quite fruitful; for instance, color vision is based on three types of photoreceptors with different spectral sensitivities whose relative activations can represent many colors. However, many central locations in the brain tend to use a continuum of selectivities rather than a few basis functions. For instance, the direction of motion in the frontal parallel plane could, in principle, be represented by the relative activities of up/down and left/right detectors (a basis-set description); however, we know that neurons in central motion pathways have cells tuned to a continuous range of motions that completely tile the set of possible motion directions. In fact, the finding by Duffy and Wurtz of double- and triple-component cells suggested that the brain
was not using a limited set of basis functions to represent optic flow. However, it was still quite possible that the set of basis functions might also include those sensitive to combinations of the three basic types.

In order to test this idea of basis versus continuous representation of optic flow, Graziano et al. (1994) examined the tuning of MSTd neurons in a spiral space. Figure 1A depicts this space; expansion/contraction is represented on the vertical axis, and clockwise/counterclockwise rotation on the horizontal axis. Noncardinal directions in this space represent spirals, and the distance from the origin indicates the magnitude of neural activity. The decomposition hypothesis would predict that MSTd neurons sensitive to rotations or expansions/contractions should have tuning curves which peak along the cardinal axes in this space (e.g., cells sensitive to expansion should also respond selectively to expansion). However, if MSTd is representing optic flow with a continuous array of detectors, then one would predict that there would be many cells which preferred different types of spirals. The data in Fig. 1 shows one example of a cell that is not tuned along one of the cardinal axes, and this result was quite common indicating that the continuous-representation hypothesis is correct. Not only were cells selective for expansions and rotations, but a large number of cells also preferred clockwise and counterclockwise expanding spirals and clockwise and counterclockwise contracting spirals.

In another study, Orban and colleagues (1992) examined the responses of expansion cells when rotary or curl motions were added to the expansion stimulus. They found that the addition of curl reduced neural activity when compared with the response to the expansion pattern alone. This result indicates again that the MSTd cells are not extracting the expansion component from these compound stimuli; if they were, the response of the cells should not have been affected by the introduction of additional motions.

B. Position and Form/Cue Invariance

The MSTd neurons prefer the same patterns of motion when the stimuli are displaced, even if the same part of the retina receives different motion directions depending on the position of the pattern within the receptive field (Graziano et al., 1994). This finding indicates that MSTd neurons extract the overall pattern of motion, even though locally the motion stimulus may be completely different. The cells also demonstrated scale invariance, giving similar responses for large or small flow fields of the same type. The preceding study explored invariance locally and was designed to examine the effect of opposing motions at the same retinal location. Thus typically the invariance tests were made over a 20° diameter area of the MSTd receptive fields. Although the tuning was
Fig. 1. A cell tuned to clockwise contracting spiral. (A) In this polar plot, the angle represents the type of stimulus, and the radius represents the response magnitude. The line directed at $-47^\circ$ indicates the cell’s best direction as determined by a Gaussian curve fit. The error bar shows a standard error of 2.5° for determining the best direction. The response is also shown in histograms at the perimeter, summed over 10 trials. CW, clockwise; CCW, counterclockwise. (B) The tuning curve from (A) plotted in Cartesian coordinates and fitted with a Gaussian function. The icons on the x-axis indicate the type of stimulus. The error bars show standard error across ten trials. Reprinted with permission from Graziano et. al. (1994), J. Neurosci. 14, 54-67.
amazingly constant in terms of pattern selectivity (i.e., tuning in spiral space), the magnitude of the response sometimes decreased as the stimuli were placed farther away from the center of the receptive field. This effect may explain why some studies reported less invariance, since they examined the magnitude of response to a single pattern of motion as the criterion for position invariance, rather than the invariance in selectivity for different patterns of motion (Duffy and Wurtz, 1995). In a few cells, Geesaman and Andersen (1996) did examine spiral tuning over a 50° diameter area and still found remarkably good invariance. However, since models for self-motion perception which predict a lack of invariance for large displacements in the receptive field have been proposed, it would be important to study more thoroughly pattern selectivity over the entire receptive field of MSTd neurons (Lappe et al., 1996).

If area MST is important for processing optic flow for navigation, it is crucial that the cells exhibit form/cue invariance (i.e., that the exact features providing the motion signals are not important to the overall selectivity of the cells). Geesaman and Andersen (1996) found that MSTd neurons do, in fact, exhibit form/cue invariance; regardless of whether the motion pattern was provided by a single object, a more typical pattern of independent features (dots), or even by non-Fourier motion cues, the cells exhibited the same pattern selectivity. This finding also implies that MSTd cells may have a dual role, processing not only the perception of self motion but also the perception of object motion.

C. ANATOMICAL ORGANIZATION

Several groups have noted that there appears to be a local clustering of cells with similar preferred stimulus pattern selectivity within MSTd. The topography of MST was examined directly by Geesaman et al. (1997). They used a double-label deoxyglucose technique in which the animals viewed one motion pattern in the presence of one metabolic tracer and a second pattern during the administration of a second tracer. Different patterns of patchy cortical columns were found within the superior temporal sulcus, including MSTd, for the different stimulus patterns. Interestingly, when expansion and contraction columns were labeled with the two tracers, they were found to be more widely separated than when expansion and rotation columns were labeled; in other words, cells selective for patterns that were more widely separated in spiral space tended to be farther apart in cortex. Britten (1998) has recently performed electrophysiological mapping experiments and found a columnar organization for different motion patterns within MST.
III. Area MST—Shifting Receptive Fields

As mentioned earlier, one daunting problem for navigation using optic flow is to separate the translational from the rotational component of optic flow (see also van den Berg, this volume, and Lappe and Hoffman, this volume). With the fall of models based on local linear operators, part of the remaining models use templates. According to this class of model, MST neurons are proposed to contain templates for a variety of optic flow conditions, including those which take into account different rates of translation and eye rotation (Perrone and Stone, 1994). The one drawback of this class of model is that it requires a large number of templates, possibly many more than there are neurons in MSTd. An attractive means for reducing the required number of templates is to adjust them dynamically using an extraretinal signal. Psychophysical experiments have shown that an extraretinal signal of eye pursuit speed and direction is used to compensate perceptually for eye rotation (Royden et al., 1992, 1994). Thus the possibility exists that a smaller set of templates could be dynamically shifted to account for eye rotation, rather than a larger number of templates being used to represent all possible eye rotation speeds and directions (Andersen et al., 1996).

This possibility was directly tested in experiments by Bradley et al. (1996). It has been known for some years that cells in MSTd are not only active for visual motion stimuli, but also for the direction and speed of eye pursuits (Newsome et al., 1988; Kawano et al., 1984, 1994). This “pure” pursuit neural activity is weaker than is commonly seen for moving visual stimuli. Recently Duffy and Wurtz (1995) showed that MSTd neurons are spatially tuned for the focus of an expansion stimulus. Bradley and colleagues reasoned that this pursuit-related signal may be used to adjust the focus tuning of MST neurons during eye rotation.

In these experiments, the spatial focus tuning of each MSTd neuron was determined with the eyes stationary, as shown in Fig. 2. Next the tuning curve was remapped but with the animal pursuing in the preferred or opposite pursuit direction of the neuron (also determined in earlier tests). The activity of the MSTd neuron was measured when the eye was in approximately the same orbital position as in the previous mapping test, the only difference being that in one case the eye is moving and in the other it is not. The laminar motion caused by the eye rotation combines with the expansion and shifts the focus on the retina in the direction of the eye movement (see Fig. 1, van den Berg, this volume). It was found that the focus tuning curves of many MSTd cells shifted during pursuit, and if they shifted, they were much more likely to shift in the di-
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 the 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 the 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 ± SEM for four replicates, where each replicate is the mean firing during the middle 500 ms of the stimulus-presentation interval. Reprinted with permission from Bradley, D. C., Maxwell, M. A., Andersen, R. A., Banks, M. S., and Shenoy, K. V. (1996). Neural mechanisms for heading perception in primate visual cortex. Science 273, 1544–1547. Copyright 1996 American Association for the Advancement of Science.

rection of the eye movement. These shifts produced activity that was constant with respect to the location of the focus in the world. In other words, these cells code the same focus location regardless of whether the eyes are stationary or moving. In a final experiment, we had the animal hold the eye stationary and simulated the same retinal image as was created when the eyes were moving. This was achieved by moving the display on the screen in the opposite direction to that in which the eye moved in the pursuit condition. We routinely found that the focus tuning curves did not compensate under these conditions, indicating that the
compensation observed during real pursuit must be due to an extraretinal pursuit signal. In these experiments not only expansion-selective cells but also contraction- and rotation-selective neurons compensated for pursuit. When a rotary motion is combined with an eye movement, the laminar motion caused by the eye movement shifts the retinal focus of rotation orthogonally to the direction of the eye movement. An eye movement combined with contraction will produce a shift in the focus in the opposite direction of eye pursuit. Interestingly, the focus tuning of curl- and contraction-selective cells compensated in the correct direction—orthogonal for rotation and opposite for contraction. These results suggest that all templates in MST compensate, not just those for expansion. This general compensation can be useful for self-motion perception. For instance, pursing an object on a ground plane to the side while moving forward will produce a spiral or curl-like motion pattern on the retina (see Fig. 5e, Lappe and Hoffman, this volume). Moving backward in the world produces contraction. Thus, templates tuned to these flow patterns can correctly compensate for pursuit eye movements. In these cases, however, unlike expansion, the direction of translation cannot be recovered directly from the focus location, and an additional mapping step is required to recover the direction of self-motion. These results raise a possibility similar to that raised by the cue invariance experiments, namely that MSTd has a more general role in motion perception than just the computation of self-motion. For instance, a horizontal pursuit eye movement over a rotating umbrella does not lead to the perception that the center of rotation displaces up or down, even though the focus of the rotation on the retina does move up or down.

Finally, Britten and van Wezel (1998) have found that microstimulation of MST can affect the perception of heading. They found larger effects during eye pursuit, consistent with the idea that area MST is the site for combining extraretinal pursuit and visual motion signals for self-motion perception.

A. Speed Tuning

Compensation for pursuit eye movements must take into account not only the direction of eye movement but also the speed of pursuit. Experiments currently underway in our laboratory indicate that MSTd neurons' focus tuning curves shift during pursuit in a monotonically increasing fashion with increases in pursuit speed (Shenoy et al., 1998). Thus, the direction and speed of pursuit are both taken into account. Other important variables are the speed of observer translation and distance of objects in the environment; these two variables determine the
rate of expansion on the retina during forward self-motion. Duffy and Wurtz (1997) have recently reported that most MSTd neurons' responses are modulated by the expansion rate.

B. Gaze Rotations with Head Movements

Large gaze rotations involve head as well as eye movements and are quite common during locomotion. If area MSTd is responsible for self-motion perception, then it would be expected that its neurons also show compensation during pursuit movements of the head. We have recently tested this idea by requiring monkeys to suppress their VOR to maintain fixation on a target during whole body rotations (Shenoy et al., 1996, 1999). In this paradigm, the head and body are rotated together, and the animal follows a spot of light that moves with them. Compensatory focus tuning shifts were found that were very similar to the compensation observed during pursuit eye movements. In fact, MSTd neurons that compensated during pursuit eye movements generally also compensated during VOR cancellation (Shenoy et al., 1997, 1999). The source of this compensation may be vestibular canal signals. When monkeys are rotated in the dark, many MST neurons are modulated by vestibular stimulation (Thier and Ericksen, 1992). This modulation is larger if the animal suppresses its VOR by tracking a fixation point, suggesting the possibility that an eye pursuit signal directing the eye in the direction of gaze rotation, and thereby canceling the VOR, may also be a factor.

C. Gain Field Model for Compensation

Many, but not all, MSTd neurons show compensatory changes in their focus tuning during the pursuit conditions described earlier. These changes are not necessarily smooth translations of the focus tuning curves on the retina, and many appear as distortions in which one part of the focus tuning curve is suppressed or enhanced. A different, but also very major effect of pursuit on MSTd neurons is to modulate the magnitude of the response (Bradley et al., 1996; Shenoy et al., 1999). Generally the gain increases when the pursuit is in the neuron’s preferred pursuit direction.

Figure 3A shows an example of an MSTd neuron which does not show shift compensation, but does exhibit gain modulation by pursuit. We have proposed a simple three-neuron model to explain how the “shifted” tuning curves are produced (Bradley et al., 1996). Figure 3B shows two model input cells that have different focus tuning and are differentially gain-modulated by pursuit. The outputs of these cells are the inputs to a compensating neuron that sums their responses. When
Fig. 3. (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 $-0^\circ$, 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. Reprinted with permission from Bradley D. C., Maxwell, M. A., Andersen, R. A., Banks, M. S., and Shenoy, K. V. (1996). Neural mechanisms for heading perception in primate visual cortex. Science 273, 1544–1547. Copyright 1996 American Association for the Advancement of Science.
the eyes pursue in one direction, one input cell is modulated upward and the other downward. This produces a compensatory shift in the tuning curve of the output cell. The change in gain modulation of the input cells reverses for pursuit in the opposite direction, producing a compensatory shift in the opposite direction in the output cell. Using this very simple model with few parameters, we were able to recreate accurately the focus tuning compensation found in MSTd neurons (Bradley et al., 1996).

Similar gain modulation effects were found for gaze rotations during VOR cancellation (Shenoy et al., 1999). Thus, this gain model could account for compensation for gaze rotations due to eye or head movements. van den Berg and Beintema (1997, 1998) have recently proposed a model for self-motion perception which uses a similar gain modulation via an extraretinal pursuit signal. Lappe (1998) has recently proposed a model which uses the distribution of compensatory shifts to arrive at self-motion perception performance similar to that found in humans.

D. PSYCHOPHYSICS

As mentioned earlier, we have found that MST neurons compensate for eye rotations due to passive head movements. However, it has never been determined if humans correctly perceive the direction of self-motion during head-generated gaze rotations. We have recently examined this issue and have found that subjects have the same complete compensation for gaze rotation with head movements as is found with eye movements (Crowell et al., 1997, 1998a). It is difficult to isolate experimentally the source of the extraretinal signal for pursuit eye movement compensation because it could be due to efference copy or muscle and orbital proprioceptive signals. Such experiments would require, among other things, moving the eyes passively. On the other hand, the head is much more accessible for examining the source of compensation signals. There are three obvious sources for head movement compensation—vestibular, neck-proprioceptive, and efference copy signals.

We examined the relative importance of these three extraretinal signals for self-motion perception by having observers make judgments about displays simulating linear forward motion across a ground plane. In conditions in which compensation is incomplete or nonexistent, our subjects inaccurately perceived motion along a path that curved in the direction of the gaze movement. We found that self-motion perception is most accurate when all three extraretinal signals are present, less accurate when only two (efference copy and neck-proprioception or vestibular and neck-proprioception) are present, and very inaccurate
when only one signal (vestibular or neck-proprioception) is present (Crowell \textit{et al.}, 1997, 1998a). This above finding is very interesting in the context of our physiological findings. When vestibular stimulation was presented alone in the psychophysical experiments, the human subjects performed the same VOR cancellation task as the monkeys in the physiology experiments. However, in humans there was no perceptual compensation. This perceptual finding makes good ecological sense. There are an infinite number of self-motion paths that all create the same retinal velocity field; for example, a linear translation with a pursuit gaze shift and a curvilinear translation can give rise to the same instantaneous retinal velocity pattern, although over time the two would differ. Thus, it is plausible that the visual system uses extraretinal signals to distinguish between these possibilities. However, a vestibular canal signal considered in isolation is consistent either with a head rotation during linear self-motion or with self-motion moving on a curved path. In fact, because head turns also generate efference copy and neck-proprioceptive signals, a canal signal alone is actually more consistent with curvilinear self-motion. The fact that MST neurons demonstrate compensation that is not observed perceptually during VOR cancellation suggests that area MSTd may not be the final stage in the computation of self-motion. In MSTd, there are populations both of compensating and non-compensating cells (Shenoy \textit{et al.}, 1997, 1999); it is thus possible that different populations of MSTd neurons are read out at a downstream sight depending on the presence or absence of other cues such as neck-proprioception and efference copy.

As mentioned earlier, not only do MSTd expansion neurons compensate during pursuit, but so do curl-selective cells. We have recently asked whether humans compensate equally well for the shift of the focus of rotation of curl stimuli and the focus of expansion of expanding stimuli during pursuit (Maxwell \textit{et al.}, 1997; Crowell \textit{et al.}, 1998b). We find that human subjects do compensate with these types of stimuli, but only partially—approximately 45\% of the amount required using expansions and only 25\% using curls. This is another inconsistency between psychophysics and physiology, since MST curl cells compensate by approximately the same amount as expansion cells. However, there are many fewer curl than expansion cells in MSTd, and this difference may contribute to the smaller degree of compensation. Consistent with this population code idea, Geesaman and Qian (1996) have found that dots moving at the same velocity are perceived as moving much faster if they are part of a global expansion compared to a global curl. They also propose that the predominance of expansion- over curl-selective cells may contribute to this perceptual difference.
To summarize, comparison of perceptual and physiological responses to simulated self-motion has uncovered many parallels and some inconsistencies. The many parallels suggest that cortical area MSTd plays an important role in self-motion perception; the inconsistencies suggest that it is not the only or final cortical site involved.

IV. Conclusion

Physiological experiments which demonstrate a possible neurophysiological foundation for self-motion perception are described. The particularly challenging problem of correctly estimating observer translation during eye rotations appears to be accomplished by a convergence and interaction of pursuit and visual signals in area MSTd. Specifically, the focus tuning of MSTd neurons is shifted by the extraretinal signal to compensate for eye rotation. This shift appears to be accomplished by a gain modulation mechanism. Such a gain mechanism has previously been shown to provide a possible basis for spatial constancy (Andersen and Mountcastle 1983; Andersen 1997). Eye position and vestibular signals have been found to gain-modulate retinal visual signals in the posterior parietal cortex (Brotchie et al., 1995). These modulations allow representation of objects in head, body, or world coordinates. The results from MSTd could be considered a velocity analog of this mechanism, in which eye velocity signals gain-modulate visual motion signals. These results suggest that the gain modulation mechanism is a very general method for performing computations in the brain, especially those computations related to spatial constancy and spatial perception.

References


