

A Neural Model of Motion Processing and Visual Navigation by Cortical Area MST

Stephen Grossberg, Ennio Mingolla and Christopher Pack

Department of Cognitive and Neural Systems and Center for Adaptive Systems, Boston University, 677 Beacon Street, Boston, MA 02215, USA

Cells in the dorsal medial superior temporal cortex (MSTd) process optic flow generated by self-motion during visually guided navigation. A neural model shows how interactions between well-known neural mechanisms (log polar cortical magnification, Gaussian motion-sensitive receptive fields, spatial pooling of motion-sensitive signals and subtractive extraretinal eye movement signals) lead to emergent properties that quantitatively simulate neurophysiological data about MSTd cell properties and psychophysical data about human navigation. Model cells match MSTd neuron responses to optic flow stimuli placed in different parts of the visual field, including position invariance, tuning curves, preferred spiral directions, direction reversals, average response curves and preferred locations for stimulus motion centers. The model shows how the preferred motion direction of the most active MSTd cells can explain human judgments of self-motion direction (heading), without using complex heading templates. The model explains when extraretinal eye movement signals are needed for accurate heading perception, and when retinal input is sufficient, and how heading judgments depend on scene layouts and rotation rates.

Introduction

A neural model has been developed to provide a functional explanation and quantitative simulations of experimental data concerning cells in the dorsal medial superior temporal area (MSTd) of monkey primate cortex. The model also explains and qualitatively simulates psychophysical data about human navigation, notably about the computation of heading. The model explains these data using a small number of mechanisms that are individually well-known to exist in cortex. It shows how interactions among these mechanisms lead to emergent properties that behave like the data. These mechanisms are the cortical magnification factor, Gaussian motion-sensitive receptive fields, spatial pooling of motion-sensitive signals and subtractive extraretinal signals from eye movement commands. By combining these elements in previously unexpected ways, the present paper derives unified quantitative explanations of data that have previously been difficult to interpret, and which, on the surface, do not even seem to be related.

Area MSTd has attracted a great deal of experimental interest because of its role in processing complex visual motion patterns. Cells in this area have large receptive fields that respond selectively to the expansion, rotation and spiral motion stimuli that are generated during observer motion (Saito *et al.*, 1986; Duffy and Wurtz, 1991a; Graziano *et al.*, 1994). This type of stimulation is called optic flow, and it can be used to guide observer navigation through the world (Gibson, 1950). In particular, optic flow can be used to compute useful quantities such as heading, which specifies the direction of self-motion relative to the direction of gaze, and is therefore useful for pursuing objects and navigating around them.

MSTd receives its primary input from the medial temporal (MT) area, which calculates motion direction and speed in

relatively small regions of the visual field. A fundamental question concerns how local MT motion estimates can be organized into the global selectivity for optic flow that is found in MSTd (Duffy and Wurtz, 1991a). Saito *et al.* (Saito *et al.*, 1986) suggested that a simple template model would suffice, in which optic flow selectivity is derived by integrating over MT cells that are selective for a preferred local direction of optic flow at each point in the receptive field (Fig. 1) (Koenderink and Van Doorn, 1977). For example, an MSTd cell selective for expansion integrates the responses of MT cells with direction preferences pointing away from a particular point.

This simple template model has been tested in studies which examine the spatial properties of MSTd cell receptive fields. Several studies have investigated whether these receptive fields are positionally variant or invariant, i.e. do their response selectivities or amplitudes change as a function of stimulus position within their receptive fields? Duffy and Wurtz found that many MSTd cells responded to a single preferred optic flow stimulus irrespective of its position within the receptive field (Duffy and Wurtz, 1991a). Graziano *et al.* (1994) also found that MSTd cell selectivity was position-invariant in this sense (Graziano *et al.*, 1994). Duffy and Wurtz (Duffy and Wurtz, 1991a) suggested that the property of position invariance contradicted Saito *et al.*'s template model (Saito *et al.*, 1986).

In contrast, Lappe *et al.* found that nearly all the MSTd cells from which they recorded responded to one type of optic flow stimulus (e.g. expansion) in one part of the visual field, and the opposite type (e.g. contraction) in a different part of the visual field, suggesting that MSTd response selectivity is position-varying (Lappe *et al.*, 1996). A model which could explain both types of results would be helpful in interpreting the functional role of MSTd cells. For example, a position-invariant, expansion-selective cell could signal the approach of an object, irrespective of its position in the visual field. Such a cell could not, however, be used to compute self-motion direction, since the retinal position of the center of an expansion stimulus corresponds to the direction of heading (Gibson, 1950).

The Saito *et al.* model is too simple to explain the above types of neurophysiological data (Saito *et al.*, 1986). It also does not address how MSTd cells may facilitate navigation by helping to compute estimates of heading. More elaborate template models (Perrone and Stone, 1994) have proposed mechanisms for computing heading, but they do not address the anatomical mechanisms by which their templates could self-organize during brain development (see Discussion). The present paper shows how these data can be explained without assuming complex templates, instead suggesting a possible explanation of MSTd receptive field properties and their role in navigation based on known properties of primate visual cortex. In particular, the model suggests how the preferred motion direction of the most active MSTd cells can explain human psychophysical data about

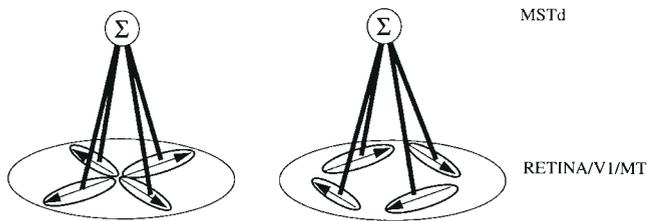


Figure 1. Simple template model of optic flow processing. An expansion cell (left) integrates the responses of local motion detectors arranged radially about a specific point. Similarly a cell selective for circular motion integrates the responses of cells tuned to local motion in a circular pattern.

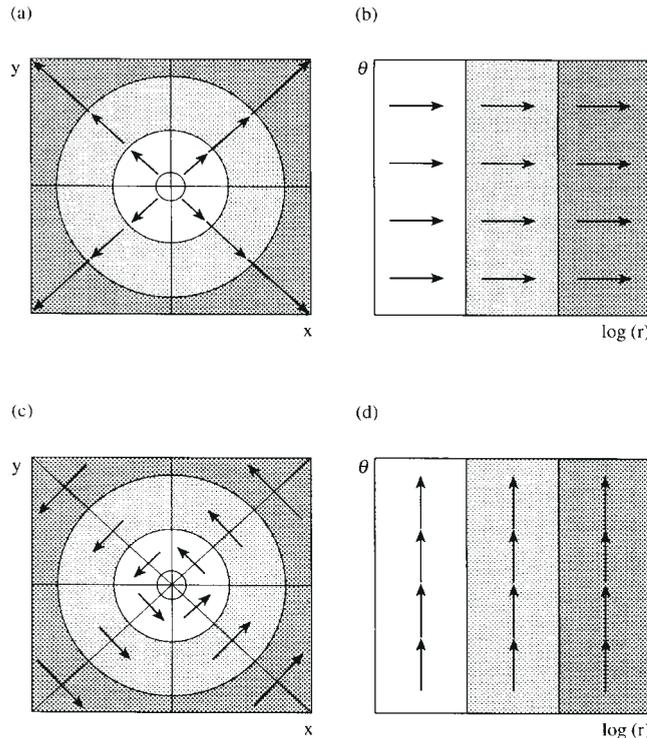


Figure 2. Log polar mapping of optic flow. (a) An expansion stimulus, in which points move outwards along a single axial direction. This results in motion along a single axis of the cortical surface (b). (c) A circular motion stimulus, in which points rotate around the central point, changing their angular position, but not their radial distance from the center. This results in motion along the orthogonal axis of the cortical surface (d).

perceived heading direction. In order to arrive at these hypotheses, the model exploits the fact that the mapping of visual information from retina to cortex obeys a cortical magnification factor, whereby foveal information has a higher cortical resolution than extrafoveal information (Daniel and Whitteridge, 1961; Fischer, 1973; Tootell *et al.*, 1982; van Essen *et al.*, 1984). This property can be well-approximated mathematically by a log polar transformation, or map, of retinal signals into cortical activations (Schwartz, 1977). The log map has the pleasing property that it transforms expansion, rotation and spiral motions around the fovea into linear motions, in different directions, on the cortex.

Figure 2 illustrates the mapping of expansion and circular motions from Cartesian (x,y) coordinates onto the log polar radial coordinate ($\log r$) and angular coordinate (θ) of primary visual cortex. The expansion stimulus consists of motion of

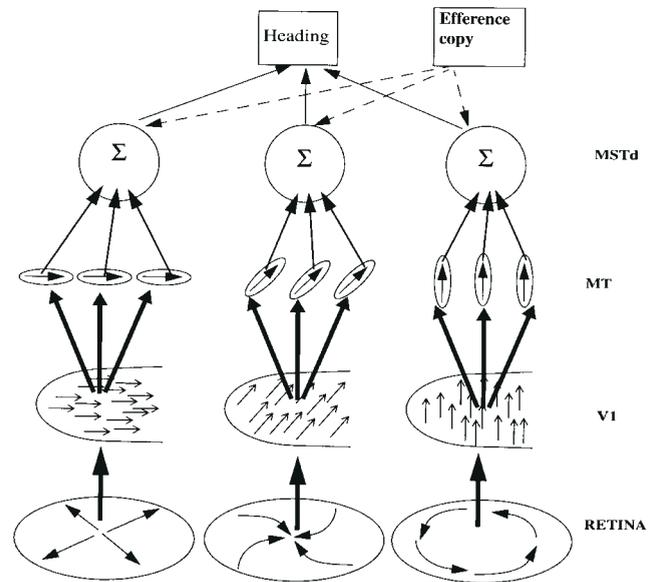


Figure 3. Model schematic. Retinal representation of the optic flow field is mapped onto primary visual cortex (V1) using the log polar cortical magnification factor. Model MT cells exhibit Gaussian motion direction tuning in this coordinate system. MSTd cells sum over MT cells selective for similar log polar motion directions. The output of these MSTd cells is combined with eye movement information to interpret global characteristics of the flow field. See text for details.

individual points along lines at a constant angle with increasing radius (Fig. 2a). Figure 2b indicates that the resulting log polar vectors are comprised of motion along the radial axis (horizontal), with no motion along the angular axis (vertical). For a circular stimulus (Fig. 2c), moving points increase their angular coordinates with no change in radius. Figure 2d shows that, in log polar coordinates, there is motion along the angular axis (vertical), but no motion along the radial axis (horizontal). Thus, expansion and circular motions in Cartesian coordinates define horizontal and vertical motions, respectively, in log polar coordinates. A similar analysis can be carried out for any spiral combination of expansion and circular motion. Such spiral motions in Cartesian coordinates are transformed into linear motion in oblique directions in log polar coordinates. Thus the log polar map defines a natural coordinate system within which each of these motions defines a distinct and statistically coherent motion direction in the cortex. These log polar motion directions are proposed to be spatially integrated by MSTd cells in a manner similar to that envisioned by the template model (Fig. 3). One key difference between these formulations is that optic flow selectivity in a log polar coordinate system is defined with respect to the fovea, while each template in the Saito *et al.* model was defined with respect to the cell's receptive field. Another crucial difference is that MT and MSTd receptive fields in the present model integrate signals that code similar motion directions in log polar space, whereas the Saito *et al.* model integrated over widely different motion directions in Cartesian space (Saito *et al.*, 1986). In like manner, the present model suggests how MSTd cell selectivity builds upon the local receptive field properties of the cortical magnification factor, rather than on complex and specialized interactions that define an explicit heading algorithm, as is often assumed (Lappe and Rauschecker, 1993; Perrone and Stone, 1994).

Because the cortical magnification factor is computed no later than cortical area V1, model MT cells were assumed to compute

their preferred local directions of optic flow in a log polar coordinate system. In particular, at each position, a model MT cell has a Gaussian receptive field that is tuned around a preferred motion direction (Albright, 1984). This Gaussian receptive field renders each MT cell decreasingly sensitive to motion directions that are progressively further from its preferred motion direction. Model MSTd receptive fields sum inputs from MT cells with the same preferred log polar motion direction over a spatial region around their receptive field center (Fig. 3). The model also characterized the probability with which a prescribed motion direction is represented in MSTd, with expansion motion being most probable, as would be expected based on its frequency when approaching objects during navigation.

Remarkably, these elementary assumptions are sufficient to quantitatively simulate many neurophysiologically recorded properties of MSTd cells. It is shown below that model optic flow selectivity matches that of MSTd cells, even for physiological studies in which optic flow stimuli were not centered on the fovea (Graziano *et al.*, 1994; Duffy and Wurtz, 1995; Lappe *et al.*, 1996). With regard to position-varying responses, the properties of the model are quantitatively similar to those found in MSTd (Graziano *et al.*, 1994; Duffy and Wurtz, 1995; Lappe *et al.*, 1996). Furthermore, the model predicts the surprising result that MSTd cells seem to optimize the size of their directionally selective tuning curves to maximize the amount of position invariance that can be achieved within a positionally variant coordinate system like the foveally centered log polar map. This prediction warrants further experimental investigation.

Finally, we test the model on stimuli from psychophysical experiments. These simulations show how MSTd cell responses can be used to qualitatively simulate human psychophysical data about heading, using the linking hypotheses described above, under a wide variety of experimental conditions. These results are consistent with data suggesting that single cells in MSTd are sufficient to support psychophysical judgments for a range of motion perception tasks, including heading perception (Celebrini and Newsome, 1994; Britten and Van Wezen, 1998). The model sheds light on a long-standing controversy in the heading perception field by quantifying the circumstances under which extraretinal eye movement signals improve heading perception and those under which they do not. MSTd cells are sensitive to such extraretinal eye movement signals (Komatsu and Wurtz, 1988; Erickson and Thier, 1991; Bradley *et al.*, 1996). They are relevant to an understanding of navigation using optic flow because eye rotations can distort the optic flow motion patterns that would otherwise be caused by object or observer motion. Extraretinal signals that are caused by these eye rotations can, in principle, be subtracted from the total optic flow pattern, and thereby greatly simplify the computation of heading, both in theory (Cameron *et al.*, 1998) and in experiments (Royden *et al.*, 1994). However, a number of studies, which are discussed below, have suggested that these extraretinal signals are not always needed to explain heading performance. Model MSTd cell properties provide a natural explanation of why this is so, and when it is so. These results were briefly reported previously (Pack *et al.*, 1997, 1998a).

Methods

This section defines the equations of the mathematical model. The Results section describes the neurophysiological and psychophysical data that are simulated, and how the model simulates them. The Results can be read independently of the mathematical equations.

Log Polar Mapping

Each retinal position can be transformed from two-dimensional Cartesian

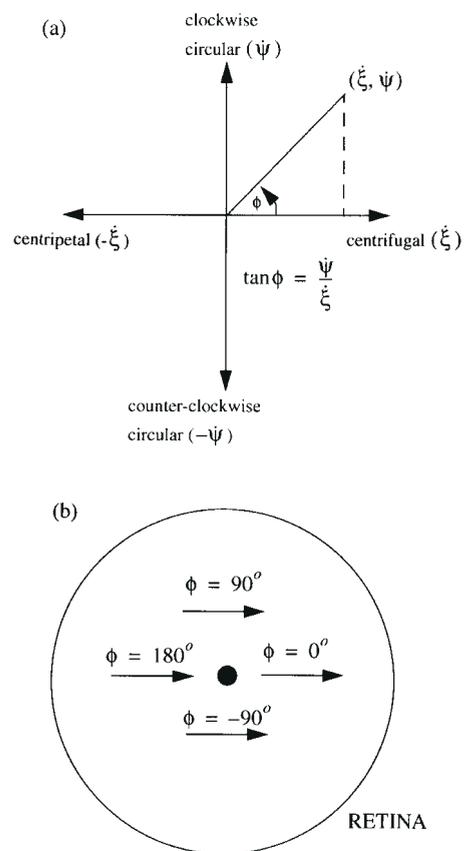


Figure 4. (a) Geometric interpretation of motion direction in log polar space. Any motion direction in this coordinate system can be represented in terms of the ratio of circular to radial motion ($\dot{\psi}/\dot{\xi}$). (b) Space-variant processing of motion. Each arrow represents a stimulus moving across the retina. For each arrow, the log polar motion direction (ϕ) depends on the retinal stimulus location relative to the fovea (indicated by the black circle).

coordinates (x,y) into polar coordinates (ρ,η) that describe the radial ρ and angular η position of the point with respect to the fovea. The log polar cortical mapping is defined in terms of these radial and angular quantities by:

$$\xi = \log(\rho + a) \quad (1)$$

$$\psi = \eta \quad (2)$$

Parameter a was set equal to 0.3° to approximate the foveal extent of the cortical map that is defined by the cortical magnification factor (Schwartz, 1994).

Model Direction Selectivity

Motions on the retina are transformed by the log polar map before they can activate model MT and MSTd cells. These cells have receptive fields that are tuned to a preferred motion direction in log polar coordinates (Fig. 3). To describe motion directions in log polar motion coordinates, we first define the speeds, or time derivatives, of these coordinates; namely $(\dot{\xi},\dot{\psi})$. These quantities define the directions of pure circular motion ($\dot{\psi}$) and radial motion ($\dot{\xi}$). Their ratio ($\dot{\psi}/\dot{\xi}$) can be used to define an arbitrary direction of motion, including expansion, circular and spiral motion. Using trigonometry, we may also define an angle ϕ such that the tangent of this angle, namely $\tan(\phi)$ equals the ratio ($\dot{\psi}/\dot{\xi}$) (Fig. 4a). Equivalently, ϕ equals the arctangent of the ratio, namely

$$\phi = \arctan(\dot{\psi}/\dot{\xi}) \quad (3)$$

For example, $\phi = 0^\circ$ occurs when $\dot{\psi} = 0$ and $\dot{\xi} > 0$. Then motion is radial away from the fovea (or centrifugal). When $\phi = 180^\circ$, then $\dot{\psi} = 0$, but $\dot{\xi} < 0$, so motion is radial towards the fovea (or centripetal). When $\phi = 90^\circ$, $\dot{\xi} = 0$ and $\dot{\psi} > 0$, so motion is clockwise around the fovea. Finally, when $\phi = -90^\circ$, $\dot{\xi} = 0$ but $\dot{\psi} < 0$, so motion is counterclockwise around the fovea. When both $\dot{\xi}$ and $\dot{\psi}$ are nonzero, then more complex motions may be defined. In particular, then the motion is a mixture of radial and circular, namely it spirals away from ($\dot{\xi} > 0$) or towards ($\dot{\xi} < 0$) the fovea. A key model hypothesis is that MSTd cells are sensitive to motion in these log polar directions.

In particular, each model MT cell has a Gaussian receptive field that is tuned around a particular preferred direction. The choice of a Gaussian tuning profile is motivated by the finding that a Gaussian function provides an excellent fit to MT cell selectivity (Albright, 1984). Because of log polar preprocessing and motion selectivity, the tuning function is defined over a range of log polar motion directions. Letting p_i be a prescribed preferred direction, the response Ω of a model MT cell at position (x, y) in Cartesian visual space with this directional preference is given by the tuning equation:

$$\Omega_i(x, y) = \left(\frac{1}{\sigma\sqrt{2\pi}} \right) \exp \left\{ -0.5 \left(\frac{p_i - \phi(x, y)}{\sigma} \right)^2 \right\} \quad (4)$$

where σ is the standard deviation of the Gaussian tuning curve. In all simulations, the value of σ was set to 38° , which is the average standard deviation of direction tuning in MT cells (Albright, 1984). It should be noted that MT direction tuning, measured physiologically in retinal coordinates, can be compared directly to local direction tuning in log polar coordinates, since changes in the two coordinate systems are equivalent. That is, an angular change in motion direction on the retina is equivalent to the same change in ϕ in equation (3).

The model assumes that an MSTd cell with preferred motion direction p_i sums inputs from a spatial neighborhood of MT cells with the same direction preference. This model MSTd cell response, χ_i is defined by:

$$\chi_i = \sum_x \sum_y \Omega_i(x, y) \quad (5)$$

The summation of model MT cell outputs at MSTd cells is consistent with data showing that MSTd cells respond more strongly to larger inputs (Duffy and Wurtz, 1991b; Lagae *et al.*, 1994). The values of x and y , which represent locations for input vectors in Cartesian visual space (see Appendix 1), are constrained to be in the circular receptive field region defined for each cell i by:

$$R_i = \left\{ (x, y) : \sqrt{(x - X_i)^2 + (y - Y_i)^2} \leq r_i \right\} \quad (6)$$

where (X_i, Y_i) defines the center of the receptive field, and r_i is the receptive field radius. The model simulates 196 direction-selective cells ($i = 1, 2, \dots, 196$) with receptive field centers forming a 14×14 grid extending 30° into the visual periphery, which is consistent with the findings of Tanaka and Saito (Tanaka and Saito, 1989). We set $r_i = 26^\circ$ for all cells. This yielded a square root of receptive field area of 46° , which is consistent with Tanaka and Saito's finding that the mean square root of MSTd receptive field area was 45° , and that receptive field size is essentially independent of the eccentricity of the center (Tanaka and Saito, 1989). Inputs were calculated within each receptive field at 1° intervals, unless otherwise specified. The magnitudes of local vectors were not considered in the definition of the cell selectivities, since these quantities have little effect on MSTd cell selectivity (Tanaka *et al.*, 1989).

The model next specified the probability with which prescribed preferred motion directions p_i in equation (4) occur. It was assumed that the most probable motion direction is centrifugal expansion motion, which is activated whenever objects approach an observer or an observer approaches an object. Other motion directions were assumed to be chosen with a random Gaussian distribution centered around this most frequent direction. Stated mathematically, for each cell i , a direction preference Γ was chosen from the Gaussian distribution defined by:

$$\exp(-B\Gamma^2) \quad (7)$$

where Γ represents a uniform distribution of log polar motion angles in degrees, and equation (7) defines the probability of a particular log polar motion preference p_i being assigned to cell i . Because $\Gamma = 0^\circ$ defines the maximal possible value of the exponential function, equation (7) generates a distribution of motion preferences centered on 0° (centrifugal expansion motion), with a standard deviation determined by the value of B . It was determined empirically that this type of directional distribution, with a value of $B = 0.00009$, simulates the directional preferences that have been found experimentally (Graziano *et al.*, 1994) (see Results). A slightly different distribution of preferences was found by Geesaman *et al.* (Geesaman *et al.*, 1997), but we did not investigate the effect of using this distribution. Once B was fixed in this way, it was used to simulate many other types of data that, on the surface, have no obvious connection with Graziano *et al.*'s data (Graziano *et al.*, 1994).

Extraretinal Input

MSTd cells that are sensitive to optic flow receive an extraretinal input which subtracts the part of the flow field that is caused by eye rotation (Komatsu and Wurtz, 1988; Erickson and Thier, 1991; Bradley *et al.*, 1996). To simulate the psychophysical effects of eye movement corollary discharges, the presence of a real eye rotation was assumed to subtract the rotational component of the flow field. Cameron *et al.* have modeled how such a subtraction can be calibrated through learning (Cameron *et al.*, 1998). In the present model, this was accomplished by simply removing the part of the flow field due to eye rotation from the input equations (see Appendix 1). Although this assumption is clearly a simplification, it provided a straightforward method of testing how the same model MSTd cells process optic flow with or without eye movement signals. How visual and extraretinal information can be combined in area MSTd is a complex question, which we have begun to address in other modeling (Pack *et al.*, 1998b), and psychophysical studies (Pack and Mingolla, 1998).

Model Heading Computation

MSTd is generally assumed to be involved in computations of self-motion. To determine if model MSTd cells could play a role in estimating heading, the log polar direction preference of the most active model MSTd cell was used to represent the activity of MSTd in response to an optic flow stimulus. The model response to a heading stimulus is therefore:

$$\Theta = p_I \quad \text{where } \chi_I = \max_i[\chi_i] \quad (8)$$

To simplify the simulations of heading perception, speed sensitivity was not incorporated into the response profiles of the model MSTd. Various data indicate that heading perception shows very little dependence on image velocity (Warren *et al.*, 1991) [but Dyre and Andersen have given exceptions (Dyre and Andersen, 1997)].

Results

Neurophysiology

Saito *et al.* showed that cells in MSTd respond selectively to large optic flow stimuli defined by expansion, contraction or circular motion (Saito *et al.*, 1986). It was subsequently found that these stimulus selectivities form part of a continuum of selectivity to spiral stimuli, which are linear combinations of radial and circular motion (Graziano *et al.*, 1994). In recent years a number of neurophysiological studies have been aimed at uncovering the mechanisms by which MSTd cell properties are derived. The typical paradigm is to isolate a single MSTd cell and to measure its response amplitude to different spiral stimuli. Measurements are also made of responses to the cell's preferred spiral stimulus centered at different locations within the receptive field, where the center of an optic flow stimulus is simply the point relative to which stimulus motion is defined. In an expansion stimulus, all motion trajectories (which begin as dots placed in random locations) point away from a center point, and this center point can be placed anywhere in the visual field. For circular

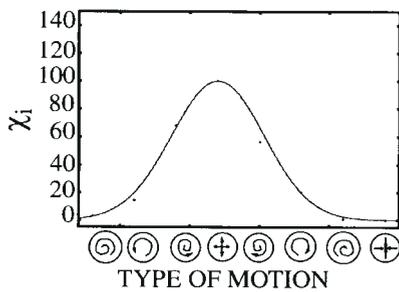


Figure 5. Tuning curve for a spiral-selective model cell fit to a Gaussian with standard deviation 50° , $r = 0.99$ (bottom). Spiral stimuli similar to those used by Graziano *et al.* (Graziano *et al.*, 1994) were placed in the center of each cell's receptive field to obtain the selectivity. This cell is typical of the model cell population, which had an average r value of 0.98, and standard deviation of 59.8° .

motion, all trajectories rotate around the center point, whereas spiral stimuli are linear combinations of expansion and circular motion. We defined these stimuli mathematically (Appendix 1), and used them to simulate key neurophysiological studies. The model parameters σ and B in equations (4) and (7) were constrained by physiological results. Other MSTd cell properties are shown to be emergent properties of these constraints.

Spiral Tuning

Graziano *et al.* reported that MSTd cell sensitivity to expansion and circular motion reflects a continuum of Gaussian response selectivity to spiral motion stimuli (Graziano *et al.*, 1994). An important methodological consideration is that Graziano *et al.* used spiral motion stimuli that were *centered on the receptive field of each cell*. From this result, it is natural to assume that MSTd spiral tuning is defined with respect to the receptive field center (Saito *et al.*, 1986; Tanaka *et al.*, 1989; Duffy and Wurtz, 1991b), but this proposition has never been tested. We therefore simulated the experiment of Graziano *et al.*, to discover if model cells tuned to log polar motion directions that are defined *with respect to the fovea* could produce tuning curves for stimuli centered on cell receptive fields at non-foveal positions. The answer is 'yes'. An example of model spiral tuning is shown in Figure 5.

In order to demonstrate this finding, spiral stimuli, as defined in Appendix 1, were centered on a cell's receptive field and the cell's responses were calculated. For each cell i the spiral stimuli were centered at the receptive field center (X_i, Y_i) in equation (6), and limited to 20° in diameter, as in the experiments of Graziano *et al.* (Graziano *et al.*, 1994). This was achieved by simply limiting the values of x and y in the input equations, as defined in Appendix 1, to be within 10° of the stimulus center. Each response profile was then fit to Gaussians of varying mean and standard deviation and the best Gaussian fit was determined by minimization of least squared error.

As illustrated by Figure 5, model MSTd cells exhibited Gaussian tuning to spiral stimuli centered on their receptive fields. The mean standard deviation of the Gaussian for the entire model cell population was found to be 59.8° , and the mean goodness of fit was $r = 0.98$. For comparison, the Gaussian fit of Graziano *et al.* (Graziano *et al.*, 1994) to their data had a mean standard deviation of 61° , and a mean $r = 0.97$. Thus the model's assumptions of Gaussian tuning to log polar motion direction imply more than Gaussian spiral tuning with respect to the fovea, as summarized in Figure 2. Surprisingly, this hypothesis also implies spiral tuning for stimuli centered on cells with non-

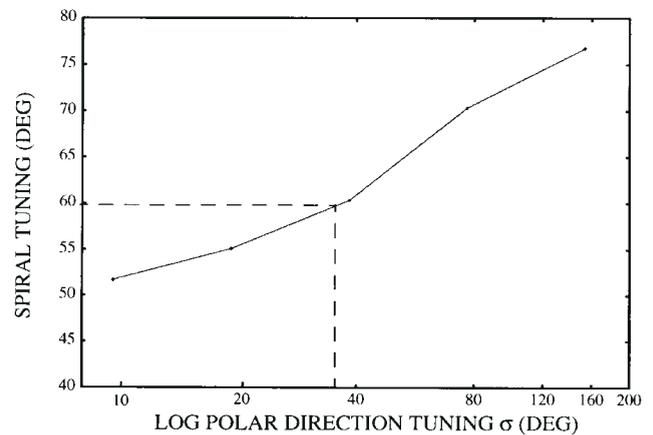


Figure 6. Average standard deviation of spiral tuning curves over all model cells for different values of the parameter σ , which controls the standard deviation of direction tuning curves in log polar space (equation 4). The horizontal dotted line indicates the width of the average spiral tuning curve for a model cell (59.8°). The average spiral tuning found by Graziano *et al.* was 61° (Graziano *et al.*, 1994). The vertical dotted line shows that this spiral tuning emerges as a result of the average tuning width (38°) of MT cells found by Albright *et al.* (Albright *et al.*, 1984).

foveally centered receptive fields, and this model emergent property quantitatively matches data from MSTd cell recordings.

We studied this relationship further by plotting the average standard deviation of spiral tuning, as a function of the tuning width σ of model MT cells, as defined in equation (4). Figure 6 shows that the average standard deviation of spiral tuning ($\sim 61^\circ$) found by Graziano *et al.* in MSTd emerges from the model's use of the average standard deviation of direction tuning (38°) that was found in MT by Albright (Albright, 1984). The key hypothesis that makes this predictive linkage work is that spiral tuning is defined with respect to the fovea, in log polar coordinates.

Spiral Preferences

The next simulation examined the distribution of model cells that prefer each type of spiral stimulus. This was determined by selecting the spiral stimulus that yielded the best response for each cell. Figure 7a summarizes the data of Graziano *et al.* (Graziano *et al.*, 1994), showing that the distribution was biased heavily toward cells that prefer expansion, with very few cells responding best to contraction. In the model, this distribution was controlled by parameter B in equation (7), and the value of B was set to provide a good visual fit to Graziano *et al.*'s data. As such, this result is not an emergent property of the model, but was used to constrain the distribution of cell types, which plays an important role in other simulated emergent properties, such as those that are described below.

Spiral Position Invariance

The log polar coordinate system defines a space-variant representation (Fig. 4b), meaning that the interpretation of motion direction depends on the location of the stimulus in the visual field. As a result, it is expected that moving the center of an optic flow stimulus may change the way in which neurons encode the stimulus. The degree to which the neuronal response changes with displacements of stimulus location can be used to quantify deviations from position invariance.

Graziano *et al.* found that their spiral-tuned cells exhibited some degree of position invariance (Graziano *et al.*, 1994). This was measured by presenting the full set of spiral stimuli at two

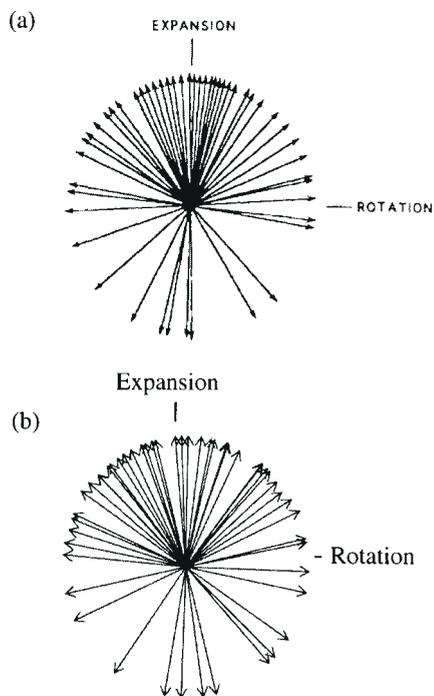


Figure 7. Preferred spiral direction for (a) 57 MSTd cells and (b) 49 model cells. Each arrow represents a cell. For each cell, the preferred direction was found by fitting the tuning curve with a Gaussian function. [Part (a) reprinted with permission from Graziano *et al.* (1994).]

different locations in each cell's receptive field. The stimuli used were 16.5° in diameter, and the two locations were separated by a vertical displacement of 8.5° . Graziano *et al.* categorized optic flow stimuli using a method similar to that defined in Figure 3a (Graziano *et al.*, 1994). In their 'spiral space', clockwise motion corresponded to 0° , expansion motion to 90° , counterclockwise motion to 180° and contraction to 270° . A spiral stimulus consisting of expansion combined with clockwise circular motion corresponded to 45° , and other spirals were defined analogously. The angular difference in spiral space between the stimuli that evoked the strongest response at each position was then used as a measure of the cell's position invariance. A difference of 0° would indicate complete invariance, while no invariance would be indicated by an average difference of 90° . Graziano *et al.* found that the mean difference was 10.7° for all spiral tuned cells, including those tuned to the cardinal directions of expansion, contraction and rotation (Graziano *et al.*, 1994).

Model MSTd cells display a similar type of position invariance (Fig. 7b). We presented each model cell with a set of spiral stimuli, as defined in Appendix 1. The stimuli were presented at an upper and lower position vertically displaced by 8.5° with respect to the center of the receptive field and the best response was calculated in each case. The mean difference in response selectivity for all spiral-tuned cells in the model was 4.3° . Thus, model MSTd cells are strongly position-invariant for small displacements of the stimulus, despite the fact that the underlying log polar representation is space-variant. However, this position invariance is not absolute, but rather is dependent on the size of the stimulus displacement. As described in subsequent simulations, the model predicts position-varying responses for larger displacements of the stimulus, and also for variations in the model parameter σ (in equation 4) that controls direction selectivity in model MT cells.

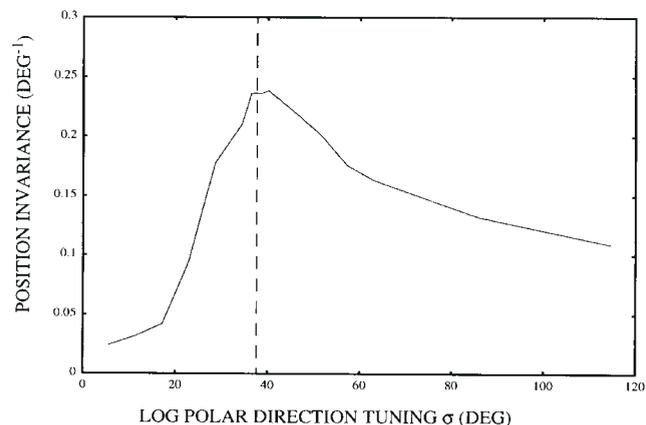


Figure 8. Position invariance in model cells as a function of the standard deviation of the Gaussian selectivity for log polar motion direction. This value is given by σ in equation (4), and controls the model's tuning to spiral stimuli (Fig. 6). Position invariance peaks near the value observed physiologically (61°), indicated by the dotted line.

Dependence of Position Invariance on Spiral Tuning

As mentioned in the Methods section, we set the model direction tuning parameter to $\sigma = 38^\circ$ to match the direction tuning for MT cells. During our simulations, it was observed that this parameter had a strong influence on the position invariance of model MSTd cells. To examine this effect quantitatively, average position invariance was calculated across the model MSTd cell population as a function of the parameter σ , using the same method as Graziano *et al.* (Graziano *et al.*, 1994), as described for the previous simulation. This measure tests the average change in spiral stimulus preference due to 8.5° vertical displacements of the center of optic flow stimulation. Position invariance was quantified as the reciprocal of this value, so that large changes in stimulus preference implied little position invariance, and conversely.

Remarkably, model position invariance peaks very near the point at which biologically observed spiral tuning curves emerge (Fig. 8). In other words, direction tuning in areas MT and MSTd of the primate visual system seems to be optimized for realizing the maximally position-invariant computation of optic flow that is possible in space-variant log polar coordinates. This result is crucial to understanding how MSTd processes optic flow, since it can be argued that a position-invariant system cannot be specialized for guiding self-motion (Geesaman and Andersen, 1996), given that effective computation of heading depends upon spatial localization of the center of expansion motion. However, it should be noted that the observed position invariance in the model and in MSTd is far from absolute. In subsequent simulations, it is demonstrated that the model's ability to compute self-motion is largely unaffected by its limited measure of position invariance. The functional implications of this result are examined further in the Discussion.

Reversal of Selectivity

The position invariance found by Graziano *et al.* appears to depend on how much the center of the test stimulus is displaced (Graziano *et al.*, 1994). Small displacements ($<10^\circ$) yielded strongly position-invariant responses, but other studies have shown less position invariance for larger displacements of the motion stimulus (Duffy and Wurtz, 1991b; Orban *et al.*, 1992; Lagae *et al.*, 1994; Lappe *et al.*, 1996). These latter studies quantified the failure of position invariance in terms of a reversal

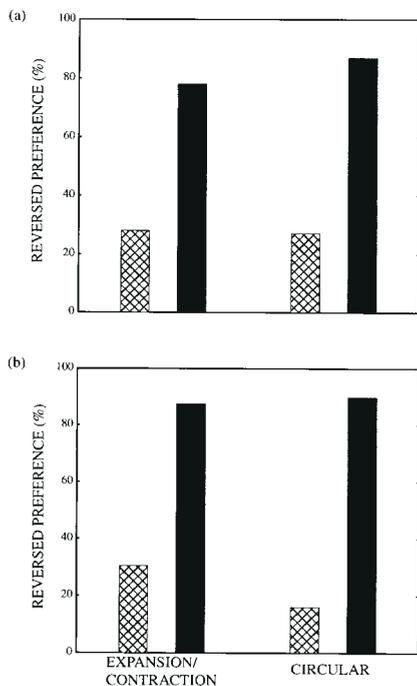


Figure 9. Percentage of cells showing a reversal of direction selectivity for MSTd (a) and for the model (b). A reversal of direction selectivity was defined as a preference (direction index > 0.5) for one type of optic flow (e.g. expansion) in one part of the visual field, and a preference for the opposite type of optic flow (e.g. contraction) in another part of the visual field. Textured bars indicate testing at 15° eccentricity. Dark bars indicate testing at 40° eccentricity. [Part (a) adapted from Lappe *et al.* (1996)].

in direction selectivity when stimulus position was changed. Many MSTd cells respond to one type of motion (e.g. expansion) for a stimulus at one position in the visual field, and the opposite type of motion (e.g. contraction) for a stimulus in a different portion of the visual field. The displacement necessary to cause a reversal in selectivity is generally between 15° and 80°.

Lappe *et al.* tested the reversal of selectivity in MSTd cells by presenting full-field optic flow stimuli centered at various locations in the visual field (Lappe *et al.*, 1996). The 17 stimuli were centered at different locations on a ring around the fixation point. One ring had a radius of 15° eccentricity, and the other 40° eccentricity. Each ring contained eight stimulus centers, and the remaining stimulus was centered on the fixation point. A cell was considered to have reversed its selectivity if it was found to be selective (direction index > 0.5) for one direction of motion at one location, and selective for the opposite direction of motion at another stimulus location. Comparisons were made within each ring, and the central stimulation point was included in both rings, so that each test for reversal consisted of placing the stimulus at nine different points in the visual field. The results indicate that rotation cells reversed selectivity in 27% and 87% of the cases for the 15° and 40° rings respectively. Expansion cells reversed selectivity for 28% and 78% of the cells for the inner and outer rings, respectively (Fig. 9a).

The model was tested using the same stimulus conditions. Since Lappe *et al.* used full-field stimulation, each model cell received input across its entire receptive field (Lappe *et al.*, 1996). A cell was identified as direction-selective for a stimulus centered on a given point if its response to that stimulus was more than twice the response to the opposite stimulus. Rotation cells reversed their selectivity in 16% and 88% of the cases for the inner and outer rings respectively. Expansion cells showed a

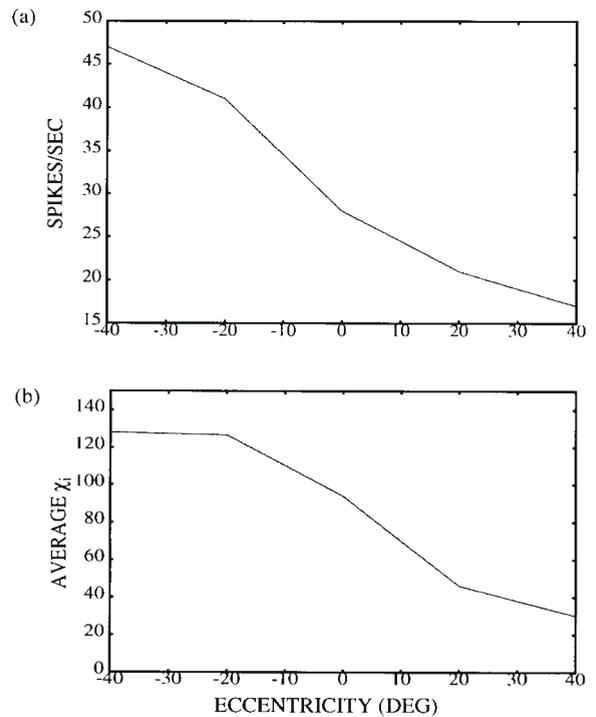


Figure 10. Average response curves for cells in MSTd (a) and the model (b). Average responses are computed by measuring each cell's activity along its preferred axis of stimulation in the visual field. These responses were then averaged across the cell population. [Part (a) adapted from Lappe *et al.* (1996)].

reversal for 29% and 90% of the cells for the inner and outer rings respectively. The model simulations are depicted in Figure 9b. The model hereby reconciles the results of Lappe *et al.* (Lappe *et al.*, 1996) showing position-varying responses with those of Graziano *et al.* (Graziano *et al.*, 1994) on approximate position invariance. The log polar space-variance of the model is crucial to understanding this result, since the degree of position invariance depends on the size of the stimulus displacement.

How does a model cell reverse its selectivity for optic flow stimulation? As mentioned previously, the reversal of selectivity is an extreme type of position-dependent response in which the spiral preference changes by 180°. This is in large part due to the fact that opposite types of stimuli can contain similar motion types in local regions. For instance, to a cell with a receptive field centered near the fovea, an expansion stimulus centered at the far right of the visual field appears similar to a contraction stimulus centered at the far left of the visual field. Both types of stimuli contain primarily leftward motion across the fovea. Since most cells are centered within the central 30° of the visual field (Tanaka *et al.*, 1989), changes in stimulus position across large regions of space increase the chance of a reversal in selectivity. These results therefore support the model approach of basing cell selectivity on local motion directions. Models which consist of templates for global motion patterns tend to exhibit greater position invariance for larger stimulus displacements (Perrone and Stone, 1998), and therefore could not explain this result.

Average Response Curve

Lappe *et al.* also measured the average response curve for the MSTd cell population (Lappe *et al.*, 1996) (Fig. 10a). Using the ring configuration described above, they found that MSTd cells exhibited a monotonic (sigmoidal) change in activity as the center point of a preferred optic flow stimulus was moved across

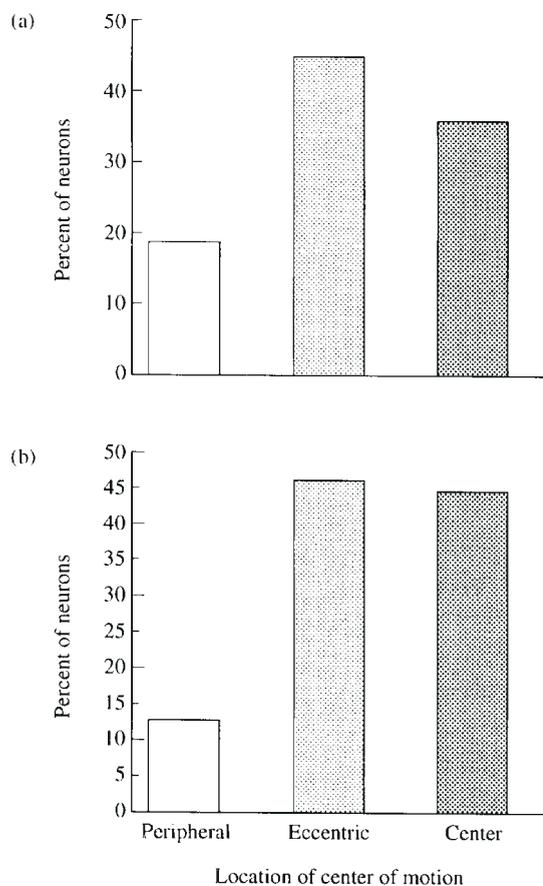


Figure 11. Preferred location for center of expansion motion for MSTd cells (a) and model cells (b). The preferred location is simply the position in the visual field yielding the strongest response. Peripheral and Eccentric locations are 90° and 45° from the fovea respectively. [Part (a) adapted with permission from Duffy and Wurtz (1995)].

the visual field in a particular direction. This was quantified as a response gradient for optic flow stimulation centered at points along a line that connected the fovea to the point of maximal response. For example, if a cell responded best to stimulation in the left part of the visual field, then the gradient was measured from left to right.

This effect was simulated using the same stimuli as in the previous simulation. The best response location was calculated, and the response gradient was calculated along stimulus locations that were colinear with the fovea and the best response center. For stimuli with peaks at the fovea, the gradient was calculated along a line connecting the fovea and the second best response. The results were then averaged for all cells in the population. The results for the expansion stimuli are shown in Figure 10b.

Preference for Center of Motion

Lappe *et al.*'s results indicate that, on average, MSTd neurons respond more strongly as the center of an optic flow stimulus is moved farther into the retinal periphery (Lappe *et al.*, 1996). However, Duffy and Wurtz, using a similar experimental paradigm, found that many MSTd cells responded more strongly to motion centered on the fovea than to any other motion stimulus (Duffy and Wurtz, 1995). They also found that some MSTd cells showed a decrease in response as the center of the optic flow stimulus was moved beyond a given point in the

periphery. These results are not necessarily inconsistent, since the peripheral stimuli in Lappe *et al.*'s study were presented in eight different locations, and were therefore weighted more than the central stimuli in the computation of the average response (Lappe *et al.*, 1996). Also, Lappe *et al.* only tested their stimuli out to 40° eccentricity, whereas Duffy and Wurtz (1995) moved their stimuli as far as 90° (Duffy and Wurtz, 1995). Thus the monotonically increasing response found by Lappe *et al.* (Lappe *et al.*, 1996) could be a result of their not having tested a large enough range of optic flow positions. We tested the model against the stimuli used by Duffy and Wurtz (Duffy and Wurtz, 1995) to see if we could reconcile their results with the apparently contradictory findings of Lappe *et al.*

In the Duffy and Wurtz experiments (Duffy and Wurtz, 1995), optic flow stimuli were presented at different locations forming two concentric rings around the fovea. Each ring consisted of eight stimulus centers, and the rings were located at 45° and 90° eccentricity. Each cell was identified as preferring the central (0°) location, one of the eight eccentric (45°) locations or one of the eight peripheral (90°) locations. The results (Fig. 11a) indicate that, for expansion stimuli, a preference for one of the eight eccentric positions was most common, followed by the central position and the eight peripheral positions. However, the single stimulus site preferred by most cells was the central stimulus.

The experiment of Duffy and Wurtz was simulated by presenting optic flow stimuli at the visual field locations used in their experiment (Duffy and Wurtz, 1995), and calculating the stimulus that gave the best response for each cell. For model cells that responded to expansion, eccentric preferences were most common, as shown in Figure 11b. Duffy and Wurtz reported increased peripheral preferences for rotation cells, but in their sample, eccentric preferences were still the most common (Duffy and Wurtz, 1995). For model rotation cells (not shown), peripheral preferences were also most common. Duffy and Wurtz found that the preferred centers of optic flow stimulation for a given cell were spatially contiguous in the visual field (Duffy and Wurtz, 1995). This was also verified in model MSTd cells.

Duffy and Wurtz also observed that cells which preferred motion fields centered on the fovea were more selective in their responses than cells which preferred motion centered in the periphery (Duffy and Wurtz, 1995). That is, cells with center preferences showed drastically decreased responses when the center of motion was moved off the fovea, but cells with peripheral preferences could tolerate larger displacements. This finding is just the type of property that one expects from log polar cortical magnification, because small retinal displacements near the fovea are magnified on the cortex, while small retinal displacements in the visual periphery are compressed on the cortex. The results from the studies of Duffy and Wurtz (Duffy and Wurtz, 1995) and Lappe *et al.* (Lappe *et al.*, 1996) on preferences for optic flow stimulus locations are of particular relevance to the model hypothesis of log polar motion tuning. Model cells exhibit the unimodal distribution of log polar preferred directions that is defined by equation (7). This distribution favors expansion motion with respect to the fovea (i.e. $\phi = 0^\circ$). Similarly, Duffy and Wurtz reported that the single most commonly preferred center for expansion motion was the fovea (Duffy and Wurtz, 1995). On the other hand, as in the data, there are more model cells that respond to expansion motion at one of the eccentric positions than at the fovea. One reason for this is simply that there are many eccentric positions, and only one central position. This is relevant because circular motion

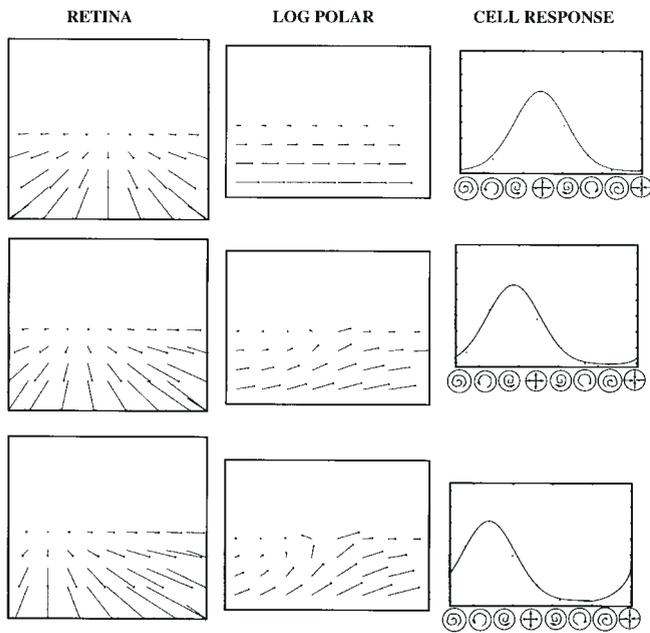


Figure 12. Log polar mapping of heading stimuli. The left column shows the optic flow stimulation that would appear on the retina for forward observer motion over a ground plane with no eye movements for various heading eccentricities. The center column shows the log polar representation of each of these optic flow stimuli. The right column shows the response of model MSTd cells tuned to various log polar motion directions. By analyzing motion in a log polar coordinate system, these cells are capable of indicating the magnitude of the heading angle.

around the fovea can be generated by centering an expansion stimulus at an eccentric location (see Fig. 12), and most model cells prefer a component of circular motion about the fovea (specified by $|\phi| > 0$ in equation 3) when both polarities of circular motion are considered.

The only constraint placed on the distribution of direction preferences in the model was that it approximate the distribution of spiral preferences found by Graziano *et al.* (Graziano *et al.*, 1994). In Graziano *et al.*'s study (Graziano *et al.*, 1994), stimulus position was held constant at the cell receptive field, while different types of spiral optic flow stimuli were tried. In the studies of Duffy and Wurtz (Duffy and Wurtz, 1995) and Lappe *et al.* (Lappe *et al.*, 1996), stimulus type was held constant, while stimulus position was varied across the entire visual field. The current model was able to simulate the findings of all three studies with a simple unimodal distribution of cell types (equation 7). This suggests that an important factor in understanding MSTd cell properties is the space-variant manner in which they relate stimulus position to motion direction, and that this relationship can be captured by a log polar coordinate system. Thus, while previous models have focused on explaining the representation of optic flow in MSTd as a consequence of the need to process heading information, the current model suggests that the optic flow in MSTd may be a natural consequence of the basic anatomy of cortical magnification. The next section considers how navigationally useful quantities such as heading can be obtained from this representation.

Heading Simulations

A number of psychophysical experiments have examined human heading perception in response to different types of optic flow stimulation. Using computer-generated stimuli, experimenters

typically show subjects a simulated visual self-motion trajectory and ask them to indicate their perceived heading direction. One consistent result has been that heading perception depends on the simulated structure of the environment. Observers often perform differently if the simulated self-motion consists of walking along a ground plane, as opposed to moving through a formless cloud of points. The addition of depth cues seems to improve heading perception (Van den Berg and Brenner, 1994a,b), as does the addition of texture and occlusion cues (Cutting *et al.*, 1997).

A controversial question regards the extent to which heading can be determined on the basis of optic flow alone. This is measured psychophysically by presenting observers with a motion sequence depicting the changing flow field that would occur if an eye rotation were combined with forward motion in some direction. Eye rotations are simulated while the observer fixates a stationary point. After viewing each optic flow stimulus, the observer indicates the perceived heading. Some studies indicate that heading can be perceived accurately under simulated eye movement conditions (Van den Berg and Brenner, 1994a). However, other studies report that heading perception is highly inaccurate unless a real eye movement is made (Royden *et al.*, 1994). In a real eye movement condition, subjects pursue a moving fixation point, while only the component of the flow field due to forward observer motion is displayed. Although the retinal stimulation is the same in both cases, the presence of an eye movement signal appears to improve heading accuracy. It has been suggested (Royden *et al.*, 1994) that the eye movement signal causes the rotational component to be removed from the brain's representation of the optic flow field. The simulations summarized below show that model MSTd cells can predict human heading judgments for different environmental layouts and eye movement conditions, thereby clarifying when eye movements can improve accuracy and when they are unnecessary.

For heading simulations, the visual field was limited to a diameter of 35° to approximate a typical experimental configuration (Royden *et al.*, 1994; Van den Berg and Brenner, 1994a). The input equations are described in Appendix 1. If the rotation rate changed over time, the flow field was calculated to correspond to the mean eye rotation (real or simulated) during the trial. The output Θ of the model was the preferred log polar direction of the most active cell in response to the optic flow stimulus, as specified by equation (8).

Figure 12 illustrates how model MSTd cells can be related to heading azimuth for the case of observer motion over a ground plane. The left column of Figure 12 shows typical optic flow stimuli corresponding to various heading directions. The middle column shows that the dominant direction of motion in log polar coordinates changes systematically with heading direction. The right column shows how cells tuned to particular directions of log polar motion can be used to estimate the heading angles shown in the left column. In particular, the dominant directions of log polar motion in the middle column are transformed in the right column into Gaussian profiles that peak at different spiral preferences. The progression down the rows of the left column from a centered heading angle to a progressively eccentric heading angle is transformed in the rows of the right column into a progression from expansion to spiral to circular motion. Figure 4a and equation (3) show mathematically how this progression from expansion to spiral to circular motion corresponds to increasing magnitudes of log polar motion direction $|\phi|$. Taken together, the three columns in Figure 12 illustrate how in-

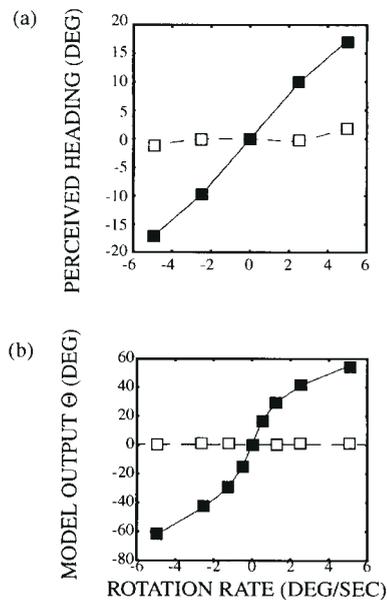


Figure 13. Heading perception for straight-ahead movement across a ground plane while fixating a moving object. (a) Royden *et al.* found large differences between actual heading (0°) and perceived heading for simulated (closed squares) but not for real (open squares) eye movements (Royden *et al.*, 1994). (b) The model also shows illusory dependence on rotation rate for simulated but not for real eye rotations. [Part (a) adapted with permission from Royden *et al.* (1994)].

creasingly eccentric heading angles correspond to increasing magnitudes of the maximally activated log polar motion direction $|\phi|$, as in equation (8). For all simulations, the output was compared to the azimuth of the actual heading direction, since this was the relevant quantity in the psychophysical experiments. The Discussion suggests how the model could be expanded to encode both azimuth and elevation of heading.

Moving Object, Ground Plane

Royden *et al.* tested heading perception for the situation where an observer moves forward while fixating a moving object (Royden *et al.*, 1994). If a real eye movement tracked the object, then subjects accurately perceived their heading as straight ahead. However, if eye movements were simulated, then heading judgments were strongly biased, as in Figure 13a. To simulate these data, the model input was the flow field generated from observer motion across a ground plane at 1.9 m/s at an eye height of 1.6 m [the same values used in Royden *et al.*'s Experiment 4 (Royden *et al.*, 1994)]. For the simulated eye movement case, the model input also contained a rotational component, depicting the eye rotation necessary to track a moving object at rates of 0–5°/s. Figure 13b shows how the model fits these data, relating the log polar motion direction ϕ to heading angle. In particular, when actual eye rotations occur (open symbols), heading estimates are accurate in all cases, since the eye movement signal subtracts the rotational flow, leaving pure expansion along the line of sight, to which the model is inherently sensitive. In the case of simulated eye movements, heading estimates progressively deteriorate with increasing rate of rotation.

Stationary Object, Ground Plane, Limited Depth Range

Van den Berg and Brenner reported that observers could accurately perceive their heading if the fixation point was rigidly attached to the simulated ground plane (Van den Berg and

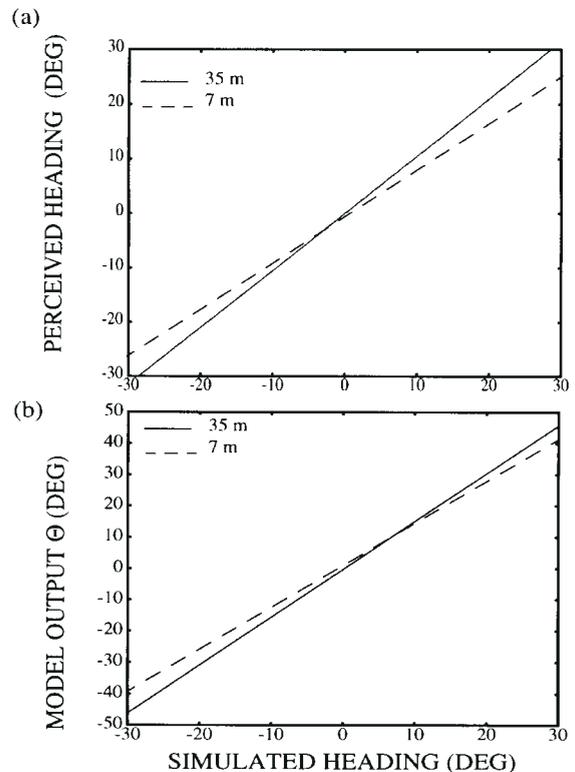


Figure 14. Heading perception for movement across a ground plane of limited depth while fixating a stationary point. (a) Linear fits to the results for one subject from Van den Berg and Brenner's Experiment 1 (Van den Berg and Brenner, 1994a) and (b) model simulation. The dotted line indicates depth visible up to 7 m beyond fixation. The solid line indicates a depth range of 35 m. [Part (a) adapted with permission from Van den Berg and Brenner (1994a)].

Brenner, 1994a) (Fig. 14a). Using a range of eye rotations similar to that of Royden *et al.* (Royden *et al.*, 1994), they found that heading accuracy was similar whether a real or a simulated eye movement was made. The model input simulated the conditions used by Van den Berg and Brenner (Van den Berg and Brenner, 1994a) for testing heading perception over a ground plane that extended 40 m in depth. The fixation point was chosen on the ground plane at a distance of 5 m from the observer. Thus the ground plane was truncated at 35 m beyond the fixation point (Fig. 14a, solid line). Forward observer motion was simulated at 3 m/s at an eye height of 1.3 m for 16 heading angles between approximately -20° and 20° . Figure 14b (solid line) shows the model output fit to a line of slope 1.54 by minimization of squared error ($r = 0.96$). The results show that sensitivity was maintained in log polar space for this stimulus configuration, across a range of rotation rates similar to that used by Royden *et al.* (Royden *et al.*, 1994). The model suggests that a real eye movement is not necessary for computation of heading over a ground plane, because the retinal stimulation approximates a spiral, for which model cells are inherently selective.

For the real eye movement case, the rotational portion of the flow field is subtracted off, leaving an expansion stimulus centered on the direction of heading, rather than a spiral centered on the fovea. We tested the model with this stimulus in order to examine to what extent the simulated and real eye movement conditions could be equated by processing in log polar coordinates. Self-motion over a ground plane with rotational flow subtracted does indeed generate a reasonably coherent direction in log polar space (Fig. 12). Simulation results

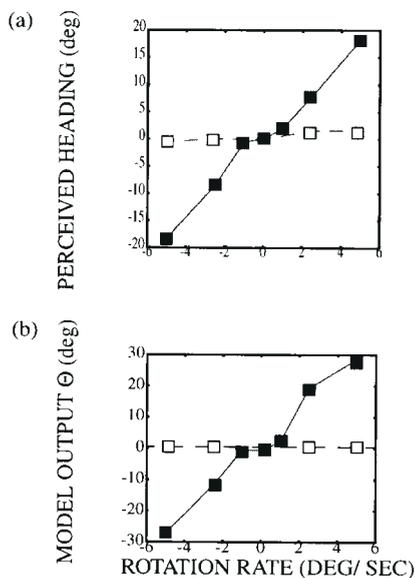


Figure 15. Heading perception for straight-ahead movement through a dot cloud while fixating a stationary point. (a) Results from Royden *et al.*'s Experiment 7 (Royden *et al.*, 1994) showing large differences between actual heading (0°) and perceived heading for simulated (closed squares) versus real (open squares) eye movements faster than $1^\circ/\text{s}$. (b) The model can compute heading for rotations slower than $1^\circ/\text{s}$, but shows increasing errors beyond that point, when heading judgments are based on motion in the lower visual field. Heading is accurate for real eye rotations. [Part (a) adapted with permission from Royden *et al.* (1994)].

indicated that the best-fitting line relating log polar motion direction to heading angle had a slope of 1.75 ($r = 0.93$), indicating a bias away from the fixation point relative to the simulated eye movement condition. A similar bias was found in real eye movement trials (Van den Berg and Brenner, 1994a; Van den Berg, 1996). Taken together, these results indicate that a single set of spiral-tuned cells can compute heading in a consistent fashion in the presence or absence of a subtractive efference copy signal. These results also suggest that such flow fields contain sufficient visual information for the brain to extract heading direction without the need for the type of high-level reasoning posited by Royden *et al.* (Royden *et al.*, 1994).

In the previous condition, the ground plane was truncated at 35 m beyond the fixation point, which itself was at 5 m from the observer. Van den Berg and Brenner also tested heading perception for observer translation over ground planes which terminated at 7 m beyond the fixation point (Van den Berg and Brenner, 1994a) (Fig. 14a, dashed line). Changing the angle between heading and gaze generated rotation rates between 0 and $6^\circ/\text{s}$. The model cell responses to these stimuli were fit to a line by minimization of least squared error for comparison with the data (Fig. 14b, dashed line). Reducing the range of visible points from 35 m to 7 m beyond the fixation point lowers the slope of both data and simulation, and thus causes the model to bias its heading judgments toward the fixation point.

The reason for this is that forward movements generate radial optic flow vectors that decrease in magnitude with their distance from the observer. In contrast, eye rotations cause movements whose effects on the optic flow pattern are constant across all depths. Therefore the furthest locations are dominated by the flow that is caused by eye movements. Removing these locations, by restricting the range of points that are visible beyond the fixation point, decreases the relative influence of eye rotations on the optic flow pattern that is caused by forward

movement. The result is an optic flow pattern that is closer to one caused by forward movement without eye rotation. Since purely forward movement implies centrifugal optic flow, and thus a zero heading angle (see Fig. 4a), the estimate of heading is biased toward the fixation point. The model hereby clarifies how manipulations of scene geometry can indirectly affect heading judgments.

Stationary Object, Dot Cloud

Royden *et al.* tested heading perception for the case where an observer moves through a dot cloud which is devoid of structure, while fixating a single point in the cloud (Royden *et al.*, 1994) (Fig. 15a). The simulated depth of the fixation point along the line of sight determined the speed of eye rotation. As before, Royden *et al.* found that heading perception was accurate for real eye rotations (Royden *et al.*, 1994) (open symbols in Fig. 15a). When the eyes did not rotate to fixate a single point, but the optic flow incorporated simulated eye rotations, then heading was highly inaccurate (closed symbols in Fig. 15a).

The model input replicated the conditions used by Royden *et al.* in their Experiment 7 (Royden *et al.*, 1994). The speed of observer motion was set at 2.5 m/s through a cloud of 615 dots randomly positioned at distances of 0–37.3 m. Fixation of dots at different depths yielded rotation rates between 0 and $5^\circ/\text{s}$. In all cases, the actual heading was straight ahead (0°). Initial simulations indicated that the model calculated heading as being nearly straight ahead for all rotation rates in both the real and simulated eye movement conditions. In the real eye rotation condition, the rotational part of the flow field is subtracted off, leaving an expansion flow around the fovea (0° heading). In the simulated eye movement condition, the flow field is dominated by planar (leftward or rightward) motion, for which the model has no selectivity. In this case, the model output is not consistent with the psychophysical data.

However, with further testing we noticed another factor: if stimulation was limited to the lower visual hemifield, the pattern of model outputs was quite similar to that of Royden *et al.*'s subjects (Royden *et al.*, 1994). Interestingly, the model exhibited accurate heading judgments in the simulated eye movement case for rotation rates $<1^\circ/\text{s}$, and increasingly inaccurate heading judgments thereafter, as found for some observers in the studies of Warren and Hannon (Warren and Hannon, 1990) and Royden *et al.* (Royden *et al.*, 1994). Figure 15b shows the model output in the case where input was limited to the lower visual field. Of course, this result should be interpreted with caution, since there is no inherent reason to suspect that observers in the experiment of Royden *et al.* ignored the upper part of the display. However, it further demonstrates the importance of the geometric structure of the stimulus in heading experiments, and suggests a possible explanation for the accurate judgments seen psychophysically at low rotation rates. In general, the model performs worse on an environment defined by random dots than on one defined by a ground plane, because the former contains more planar motion, which does not necessarily contain a coherent direction of log polar motion. The Discussion suggests ways in which planar motion sensitivity could be incorporated into the existing model to improve heading estimates.

Stationary Object, Approach to a Wall

When the simulated environment contains no depth, as in an approach to a wall, observers tend to perceive themselves as heading in the direction of gaze unless a real eye movement is made. Warren and Hannon tested observers in this situation by

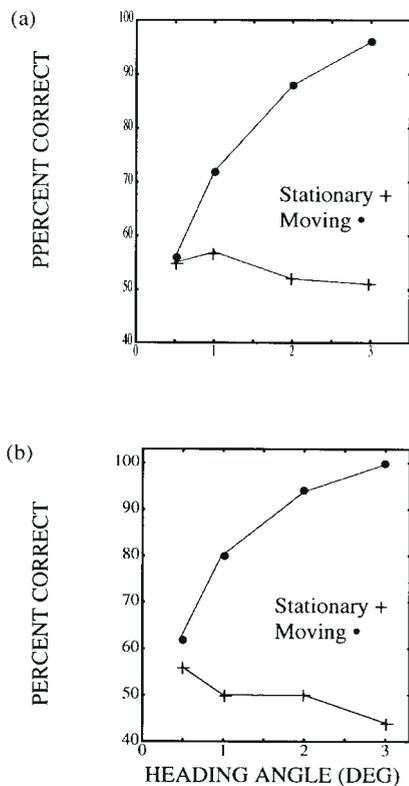


Figure 16. Heading perception for movement toward a wall while fixating a stationary object. (a) Results from Warren and Hannon's Experiment 4 (Warren and Hannon, 1990) and (b) the model showing high heading accuracy for real (circles) eye movements and low accuracy for simulated (+) eye movements. [Part (a) adapted with permission from Warren and Hannon (1990)].

having them judge their heading as being to the left or right of fixation (Warren and Hannon, 1990). When a real eye movement was made, observers could make correct judgments when heading and gaze deviated by only a few degrees (Fig. 16a, upper curve); however, when eye movements were simulated, observers performed near chance (Fig. 16a, lower curve).

The model was next tested for the same stimuli used in Warren and Hannon's Experiment 7 (Warren and Hannon, 1990), under the assumption that actual eye movements cause subtraction of the resulting full-field motion (see Mathematical Model section – Extraretinal input). Warren and Hannon asked observers to estimate their heading as being to the left or to the right of the fixation target, and measured the percentage of correct responses as a function of heading angle (Warren and Hannon, 1990). In order to generate similar probabilistic perceptual judgments in the model, random noise (see Appendix 2) was added to the cell responses from equation (5). The log polar motion preference of the most active cell was then chosen, and the presentation was repeated 50 times for each simulated angle of approach. For the simulated eye movement case, there was no subtraction of the rotational flow. The results are shown in Figure 16b. Without extraretinal information, there is no way for the model to determine heading, and performance is essentially at chance in this condition (Fig. 16b, lower curve). With the aid of an extraretinal signal, performance is substantially better, and nears 100% for heading angles as small as 3° (Fig. 16b, upper curve).

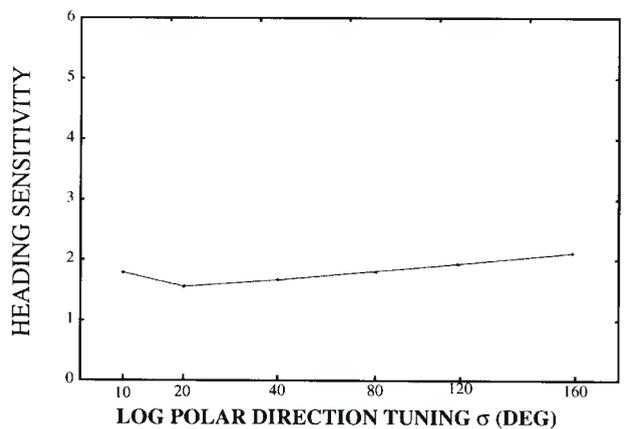


Figure 17. Heading sensitivity is unaffected by changes in the standard deviation of the Gaussian selectivity for log polar motion direction. This is specified by the value of σ in equation (4).

Dependence of Heading Sensitivity on Spiral Tuning

Figure 8 shows the surprising model prediction that the spiral tuning of MSTd cells, whose selectivity, or sharpness, is scaled by the parameter σ in equation (4), is optimized to provide maximal position invariance. However, position invariance is not desirable for computing self-motion, since heading computation depends on the ability to locate the position of the focus of expansion, not merely its presence. It has been suggested that heading can be extracted at the population level (Geesaman and Andersen, 1996). We examined this possibility by measuring heading sensitivity as a function of σ . As in Figure 14, heading sensitivity was quantified by measuring the model's heading estimate Θ (see equation 8) in response to different heading angles. In particular, for each value of σ , the inputs were chosen to depict observer motion over a ground plane at 2.5 m/s, with rotational flow removed, for 10 heading angles between 0° and 20°. For each σ , model heading sensitivity was quantified as the slope of the best fitting line to the model output Θ as a function of these heading angles. In all cases, the fit to this line was excellent with $r > 0.95$. Figure 17 shows that heading sensitivity (i.e. the slope of each line) is largely unaffected by changes in σ , which is calibrated in Figure 17 both in terms of spiral tuning and log polar direction tuning (see Fig. 6 for the conversion factor). This shows that σ , and thus the model's position-invariance, could be optimized, as in Figure 8, without altering the model's ability to compute estimates of self-motion like heading. Properties like those summarized in Figures 8 and 17 illustrate that emergent properties of simple neural mechanisms can be quite unexpected when they act together in a neural system.

Discussion

The current study models how patterns of optic flow on the retina give rise to observed properties of receptive fields in cortical area MSTd. In particular, the model suggests how conversion of the retinal image into log polar coordinates by the cortical magnification factor in V1, Gaussian tuning of log-polar motion direction in MT, and spatial pooling of MT outputs in MSTd can explain many challenging properties of MSTd cells, such as spiral tuning curves and position invariance. Emergent properties of the model include quantitative accounts of neurophysiological observations, such as spiral tuning with respect to receptive field centers and changes in cell response

properties as a function of stimulus position. Moreover, the model demonstrates that motion processing in MT and MSTd seems to be optimized for calculating position-invariant motion patterns without sacrificing the ability to compute self-motion information. Psychophysical simulations indicate how the neurophysiologically observed spiral tuning properties can serve as a basis for human perception of heading through interactions with eye movement signals.

Organization of Visual Pathways

Our work indicates that the global topography of the primate visual cortex plays an important role in shaping the responses of cells at subsequent levels of cortical visual processing. Evidence for a global structure has been seen in cortical motion processing region MT of the primate (Albright, 1989) and homologous region LS in the cat (Rauschecker *et al.*, 1987). The distributions of directional preferences among cells in monkey and cat extrastriate cortex are biased such that these cells are more likely to prefer motion away from the fovea than towards the fovea. Similarly, cells in parietal cortex, to which MSTd projects, seem to compute motion direction relative to the fovea (Steinmetz *et al.*, 1987). The current work demonstrates that this foveocentric structure plays a key role in generating the visual response properties of many cells in MSTd. While the primary projection to MSTd comes from MT (Maunsell and Van Essen, 1983), a second route by which optic flow information may reach MSTd is the tectopulvinar pathway (Ballard, 1987). This pathway bypasses the primary visual cortex altogether, passing motion signals from the superior colliculus through the pulvinar to MT (Standage and Benevento, 1983). These connections may serve as a more primitive pathway for navigation or for 'defensive' responses to approaching stimuli (Dean *et al.*, 1989). The pulvinar appears to have a topographic organization that is similar to that of V1 and MT (Standage and Benevento, 1983), although this has not been examined quantitatively to our knowledge.

MSTd Position Invariance

The log polar mapping defines a space-variant system, meaning that the interpretation of a motion direction depends on where in the visual field the stimulus occurs. The model demonstrates how this key assumption of space variance can explain a number of paradoxical findings relating to optic flow sensitivity in MSTd. Most of these findings involve the degree to which MSTd cells exhibit position invariance for stimuli placed within their large receptive fields.

Area MSTd is generally assumed to process self-motion by locating the focus of expansion. However, Graziano *et al.* have reported position invariance of MSTd cells (Graziano *et al.*, 1994), which might have prevented them from calculating self-motion. This is because the heading angle is determined by the position of the focus of expansion, so changes in heading would need to be registered as changes in the position of an expansion stimulus. A position-invariant cell would be incapable of registering these changes, as it would respond similarly to the presence of an expansion motion regardless of the locus of stimulation. Geesaman and Andersen suggested that, just as inferotemporal neurons may extract form information in a position-invariant way, MSTd neurons may be used to compute the trajectories of moving objects, since this computation should not depend on the position of the object in the visual field (Geesaman and Andersen, 1996). Position invariance therefore

suggests an alternative role for MSTd cells in extracting motion patterns derived from object trajectories in space.

The current model illustrates how MSTd cells can process self-motion information while maintaining some degree of position invariance. The position-variant log polar mapping that occurs in V1 provides useful image compression properties (Schwartz, 1994). The model goes beyond this important insight to show how these properties can coexist with a limited degree of position invariance. In particular, the model suggests that this is accomplished by the receptive field structure of the motion-direction cells in MT and MSTd which derive their inputs from the V1 log polar map. The degree of direction selectivity of these cells is controlled by the parameter σ in equation (4) (see also Figs 6 and 8). Figure 8 shows the surprising model prediction that the spiral tuning of 61° found by Graziano *et al.* actually maximizes the position invariance of MSTd receptive field cells (Graziano *et al.*, 1994), and Figure 17 shows that this occurs without impairing heading sensitivity. Figures 8 and 17, taken together, suggest that the cortex somehow optimizes its motion selectivity, possibly in response to early visual experience, to obtain the most spatially invariant directional estimates for the purpose of tracking object motion.

Earlier modeling work on visual motion perception suggests how a long-range motion filter that is predicted to occur between V1 and MT can be used for motion perception and for tracking moving objects; this work has been reviewed by Grossberg (Grossberg, 1998). This long-range motion filter has been used to simulate many properties of long-range apparent motion (Grossberg and Rudd, 1992; Francis and Grossberg, 1996; Baloch and Grossberg, 1997) that are proposed to be useful for motion tracking, including properties of MT cells (Baloch *et al.*, 1999). It remains to be tested if this long-range motion filter is related to the MT values of 38° (Albright *et al.*, 1984) and the resultant MSTd values of 61° (Graziano *et al.*, 1994) that the present model uses for computing heading. If so, then these results, taken together, clarify how MST can accomplish object motion tracking while also computing motion properties that are used in visual navigation.

Visual Navigation

To test the model's navigational properties, the model was presented with stimuli typical of heading experiments. The motion preference of the most active model MSTd cell was compared to perceived heading direction. The model provides good fits to human heading performance, including biases toward fixation for limited depth (Van den Berg and Brenner, 1994a), inability to distinguish between heading and fixation for approach to a wall (Warren and Hannon, 1990), and variations in perceived heading for simulated eye rotations when fixating a moving object (Royden *et al.*, 1994).

The model heading computation illustrates how a subtractive efference copy can be used by the same cells that analyze visual information. For the case of observer motion over a ground plane, the model computes heading similarly whether an efference copy is available or not (see Results). This is particularly important in explaining heading computation in MSTd, because there is substantial variability in the extent to which MSTd cells compensate for eye rotation (Bradley *et al.*, 1996).

Although the heading model is sufficient to explain the magnitude of key biases in human heading perception, in its present form it is limited in its ability to compute heading direction in general. This is because visual navigation requires the computation of a three-dimensional observer trajectory,

whereas the model presently transforms two-dimensional inputs into two-dimensional outputs. Thus the model can distinguish backward self-motion from forward self-motion, and estimate the angle between heading and gaze. However, it cannot distinguish self-motion in the horizontal plane (azimuth) from self-motion in the vertical plane (elevation). Model outputs correlate well with the azimuth of perceived heading in psychophysical experiments, but a complete model must be able to compute both azimuth and elevation. A simple way to remedy this problem would be to use separate groups of model cells to signal azimuth and elevation, consistent with the finding that these quantities are encoded independently (D'Avossa and Kersten, 1996).

An alternative possibility is that a separate population of MSTd cells is selective to visual motion resulting from movement of the observer in the frontoparallel plane, i.e. a plane in front of an observer that is perpendicular to the ground plane. In this regard, Duffy and Wurtz have reported the existence of MSTd cells which are selective to frontoparallel motion stimuli moving horizontally, vertically or obliquely (Duffy and Wurtz, 1997). These cells could provide estimates of vertical and horizontal observer motion direction, since these types of self-motion generate unidirectional flow across the visual field. However, these cells are incapable of signaling forward or backward observer motion, and cannot indicate the magnitude of heading angle in any straightforward way. Thus, information computed by MSTd cells that are selective for frontoparallel motion is complementary to information computed by spiral-tuned cells. The eye movement signals (Komatsu and Wurtz, 1988), disparity sensitivity (Roy *et al.*, 1992) and vestibular inputs (Thier and Erickson, 1992) to MSTd cells could be used to help distinguish among rotations of the eye, head and body.

Another limitation of the current model is the method by which visual and extraretinal information sources of information are combined. While it is clear that MSTd compensates at least partially for optic flow induced by eye rotations (Erickson and Thier, 1991; Bradley *et al.*, 1996), the mechanism by which this occurs is not well understood. A computationally simple method of achieving this compensation is to remove the rotational flow from the input representation in MT, which would leave the optic flow representation in MSTd immune to the effect of eye rotations. However, this solution is contradicted by physiological evidence (Erickson and Thier, 1991) showing that MT cells do not compensate for retinal motion caused by eye movements.

Lappe suggested a plausible approach, in which an extraretinal pursuit signal is used to counteract the precise amount of excitation or inhibition generated by rotational optic flow at the level of individual MST cells (Lappe, 1998). A similar approach could be adopted in the context of the current model. However, it remains to be seen whether MST cells exhibit the quantity and specificity of connections required to implement this scheme.

A fourth dimension in visual navigation is the computation of time-to-contact (TTC), which is a measure of the time until a moving observer reaches a point in the environment. It has been shown in the field of computer vision that the log polar transformation greatly simplifies the computation of TTC (Tistarelli and Sandini, 1993), and our modeling work suggests how this property is used by MSTd to derive TTC estimates (Pack *et al.*, 1998a) (C. Pack, S. Grossberg and E. Mingolla, in preparation).

Comparison with Other Models

A number of previous models have been suggested to explain the

role of MST in heading computations. The model of Perrone and Stone hypothesized that MSTd cells serve as heading templates against which optic flow patterns can be matched (Perrone and Stone, 1994). Each model template is constructed by hardwiring connections from MT cells that encode particular locations, directions, speeds and depths to a model MSTd cell which represents a particular heading direction. By sampling a subset of the possible flow fields, the model is able to match human psychophysical data on heading perception, and properties of many of the model heading templates match properties of MSTd cells (Perrone and Stone, 1998). One problem with the Perrone and Stone model is that each template covers the entire visual field, which is inconsistent with physiological measurements (Tanaka *et al.*, 1989; Duffy and Wurtz, 1991a,b), and prevents the model from explaining the reversal of direction selectivity observed in nearly every MSTd cell tested by Lappe *et al.* (Lappe *et al.*, 1996). The model also seems to be inconsistent with perceptual data under conditions wherein eye rotation is not linked to self-motion (Crowell, 1997). A more general problem is the complexity and specificity of the model's templates.

The model proposed by Lappe and Rauschecker matches optic flow patterns to populations of cells that encode particular heading directions (Lappe and Rauschecker, 1993). This model also relies on a complex hardwiring of inputs; in particular, a modified version of the subspace algorithm of Heeger and Jepson (Heeger and Jepson, 1992) is built into the synaptic weights. As a result, the model provides a mathematically efficient solution to the extraction of heading from optic flow, and has successfully predicted some properties of MSTd cells (Lappe *et al.*, 1996). On the other hand, the use of hard-wired templates to compute heading in the models of Lappe and Rauschecker (Lappe and Rauschecker, 1993) and Perrone and Stone (Perrone and Stone, 1994) is hard to reconcile with the strong dependence of heading perception on the cognitive state of the observer (Van den Berg, 1996).

Another issue related to heading concerns how extraretinal inputs should be integrated into the processing of optic flow. The present model introduces such inputs in the simplest possible way and shows how, in the absence of structured flow fields, extraretinal signals can interact with visual selectivity in MSTd to improve heading judgments. Lappe (Lappe, 1998) has also demonstrated how this can be accomplished within the existing framework of the Lappe and Rauschecker model (Lappe and Rauschecker, 1993), but this model lacks such features as spiral tuning and position invariance, and restricts its discussion to how extraretinal input may influence heading estimates.

One similarity between these models and the current model is the specialization for centrifugal flow. The current model incorporates this bias into the distribution of cell types found in equation (7). The model of Lappe and Rauschecker (Lappe and Rauschecker, 1993) incorporates this specialization explicitly, and the Perrone and Stone model (Perrone and Stone, 1994) incorporates it through its finer sampling of heading directions close to the line of sight. Thus, all the models have taken into account that the primate visual system contains a preponderance of cells preferring centrifugal flow (Albright, 1989). It has been argued that this statistical constraint may explain some aspects of heading perception. In particular, Lappe and Rauschecker (Lappe and Rauschecker, 1994) have suggested that the failure of subjects to achieve accurate heading perception in the study of Royden *et al.* (Royden *et al.*, 1994) is due to the lack of centrifugal flow in their displays. The current model also

performs better in the presence of centrifugal flow (see Heading Simulations – stationary fixation, ground plane).

The current model connectivity may also be viewed as a type of template model, albeit one whose ‘templates’ are defined in terms of known properties of cortical organization. Using just the local direction preferences of spatially pooled cell receptive fields that are superimposed on a log polar map, the model can simulate a wide range of MSTd data. The elaborate connectivity suggested by Lappe and Rauschecker (Lappe and Rauschecker, 1993) and Perrone and Stone (Perrone and Stone, 1994) is hereby shown to be logically unnecessary for understanding MSTd cell properties. The current model can also simulate observer percepts for a number of heading experiments, although it is incomplete in this regard for reasons specified above. In light of the model’s explanatory successes, it may be argued that human navigational properties emerged from rather general anatomical properties such as cortical magnification, rather than from specialized mathematical heading algorithms.

In the same vein, it is also important to consider how MSTd cell selectivity can develop. The Lappe and Rauschecker (Lappe and Rauschecker, 1993) and Perrone and Stone (Perrone and Stone, 1994) models require hardwired connection patterns to compute a complex heading algorithm. It is difficult to imagine how these connections could self-organize during cortical development. A recent model by Zemel and Sejnowski suggests that MSTd cell selectivity could develop through a specific type of optimization of synaptic weights, subject to the constraint that MSTd cells attempt to encode faithfully the pattern of inputs (Zemel and Sejnowski, 1998). This would require a complex type of MSTd learning law, for which there is as yet no biological evidence.

In contrast, the current model suggests how MSTd cells can develop based on selectivity to a single stimulus dimension (motion direction), which is consistent with known properties of cortical self-organization. In particular, cells selective to a particular stimulus dimension are often clustered in visual cortex, including area MT, which contains columns of direction selectivity (Albright *et al.*, 1984), and recent models of cortical self-organization illustrate how maps on this level of complexity could self-organize (Obermayer *et al.*, 1992; Swindale, 1992; Miller, 1994; Olson and Grossberg, 1998). The current model suggests a functional role for columns of direction selectivity in log polar space that self-organize within V1, MT and MSTd, and several studies (Lgae *et al.*, 1994; Geesaman *et al.*, 1997; Britten, 1998) confirm that the expected results of such self-organization – namely cells tuned to expansion, rotation and contraction – are indeed clustered into columns in MSTd.

Notes

This work was carried out as part of C.P.’s dissertation research. Authors of this article are listed alphabetically. S.G. supported in part by the Defense Research Projects Agency and the Office of Naval Research (ONR N00014-92-J-1309, ONR N00014-95-1-0409, and ONR N00014-95-1-0657). E.M. supported in part by the Defense Research Projects Agency and the Office of Naval Research (ONR N00014-94-1-0597, ONR N00014-95-1-0409, and ONR N00014-95-1-0657). C.P. supported in part by the Air Force Office of Scientific Research (AFOSR F49620-92-J-0334) the Defense Advanced Research Projects Agency (ONR N00014-92-J-4015), and the Office of Naval Research (ONR N00014-91-J-4100, ONR N00014-92-J-1309, ONR N00014-94-1-0597, ONR N00014-95-1-0409, and ONR N00014-95-1-0657).

Address correspondence to Professor Stephen Grossberg, Department of Cognitive and Neural Systems, Boston University, 677 Beacon Street, Boston, MA 02215, USA. Email: steve@cns.bu.edu.

References

- Albright TD (1984) Direction and orientation selectivity of neurons in visual area MT of the macaque. *J Neurophys* 52:1106–1130.
- Albright TD (1989) Centrifugal directionality bias in the middle temporal visual area (MT) of the macaque. *Vis Neurosci* 2:177–188.
- Albright TD, Desimone R, Gross CG (1984) Columnar organization of directionally selective cells in visual area MT of the macaque. *J Neurophysiol* 51:16–31.
- Baloch A, Grossberg S (1997) A neural model of high-level motion processing: line motion and formation dynamics. *Vis Res* 37:3037–3059.
- Baloch A, Grossberg S, Mingolla E, Nogueira CAM (1999) A neural model of first-order and second-order motion perception and magnocellular dynamics. *J Opt Soc Am A* 16:953–978.
- Ballard D (1987) Cortical connections and parallel processing: structure and function. In: Vision, brain, and cooperative computation (Arbib M, Hanson A, eds), pp. 563–621. Cambridge, MA: MIT Press.
- Bradley D, Maxwell M, Andersen RA, Banks MS, Shenoy K (1996) Mechanisms of heading perception in primate visual cortex. *Science* 273:1544–1547.
- Britten K, Van Wezen (1998) Clustering of response selectivity in the medial superior temporal area of extrastriate cortex in the macaque monkey. *Vis Neurosci* 15:553–558.
- Cameron S, Grossberg S, Guenther F (1998) A self-organizing neural network architecture for navigation using optic flow. *Neural Comput* 10:313–352.
- Celebrini S, Newsome W (1994) Neuronal and psychophysical sensitivity to motion signals in extrastriate area MST of the macaque monkey. *J Neurosci* 14:4109–4124.
- Crowell JA (1997) Testing the Perrone and Stone (1994) model of heading estimation. *Vis Res* 37:1653–1671.
- Cutting JE, Vishton PM, Fluckiger M, Baumberger B, Gerndt J (1997) Heading and path information from retinal flow in naturalistic environments. *Percept Psychophys* 59:426–441.
- Daniel PM, Whitteridge D (1961) The representation of the visual field on the cerebral cortex in monkeys. *J Physiol* 159:203–221.
- D’Avossa G, Kersten, D (1996) Evidence in human subjects for independent coding of azimuth and elevation for direction of heading from optic flow. *Vis Res* 36:2915–2924.
- Dean P, Redgrave P, Westby G (1989) Event or emergency? Two response systems in the mammalian superior colliculus. *Trends Neurosci* 12:137–147.
- Duffy CJ, Wurtz RH (1991a) Sensitivity of MST neurons to optic flow stimuli. I. A continuum of response selectivity to large-field stimuli. *J Neurophysiol* 65:1329–1345.
- Duffy CJ, Wurtz RH (1991b) Sensitivity of MST neurons to optic flow stimuli. II. Mechanisms of response selectivity revealed by small-field stimuli. *J Neurophysiol* 65:1346–1359.
- Duffy CJ, Wurtz, RH (1995) Response of monkey MST neurons to optic flow stimuli with shifted centers of motion. *J Neurosci* 15:5192–5208.
- Duffy CJ, Wurtz RH (1997) Planar directional contributions to optic flow responses in MST neurons. *J Neurophysiol* 77:782–796.
- Dyre B, Andersen G (1997) Image velocity magnitudes and perception of heading. *J Exp Psychol Hum Percept Perform* 23:546–565.
- Erickson RG, Thier P (1991) A neuronal correlate of spatial stability during periods of self-induced visual motion. *Exp Brain Res* 86: 608–616.
- Fischer B (1973) Overlap of receptive field centers and representation of the visual field in the cat’s optic tract. *Vis Res* 13:2113–2120.
- Francis G, Grossberg S (1996). Cortical dynamics of form and motion integration: persistence, apparent motion, and illusory contours. *Vis Res* 36:149–173.
- Geesaman B, Andersen RA (1996) The analysis of complex motion patterns by form/cue invariant MSTd neurons. *J Neurosci* 16:4716–4732.
- Geesaman B, Born RT, Andersen RA, Tootell RBH (1997) Maps of complex motion selectivity in the superior temporal cortex of the alert macaque monkey: a double-label 2-deoxyglucose study. *Cereb Cortex* 7:749–757.
- Gibson JJ (1950) Perception of the visual world. Boston, MA: Houghton Mifflin.
- Graziano MSA, Andersen RA, Snowden R (1994) Tuning of MST neurons to spiral motions. *J Neurosci* 14:54–67.
- Green D, Swets J (1974) Signal detection theory and psychophysics. New York: Kreiger Press.

- Grossberg S (1998) How is a moving target continuously tracked behind occluding cover? In: High-level motion processing (Watanabe T, ed.). Cambridge, MA: MIT Press.
- Grossberg S, Rudd M (1992) Cortical dynamics of visual motion perception: short-range and long-range apparent motion. *Psychol Rev* 99:78–121.
- Heeger DJ, Jepson A (1992) Subspace methods for recovering rigid motion I: Algorithm and implementation. *Int J Comput Vis* 7:95–117.
- Koenderink JJ, van Doorn AJ (1977) How an ambulant observer can construct a model of the environment from the geometrical structure of the visual inflow. In: Kibernetik (Hauske G, Butendant E, eds). Munchen: Oldenbourg.
- Komatsu H, Wurtz RH (1988) Relation of cortical areas MT and MST to pursuit eye movements. III. Interaction with full-field visual stimulation. *J Neurophysiol* 60:621–644.
- Lagae L, Maes H, Raiguel S, Xaio DK, Orban GA (1994) Responses of macaque STS neurons to optic flow components: a comparison of areas MT and MST. *J Neurophysiol* 71:1597–1626.
- Lappe M (1998) A model of the combination of optic flow and extra-retinal eye movement signals in primate extrastriate visual cortex. *Neural Networks* 11:397–414.
- Lappe M, Rauschecker JP (1993) A neural network for the processing of optic flow from ego-motion in man and higher mammals. *Neural Comput* 5:374–391.
- Lappe M, Rauschecker JP (1994) Heading detection from optic flow. *Nature* 369:712–713.
- Lappe M, Bremmer F, Pökel M, Thiele A, Hoffman KP (1996) Optic flow processing in monkey STS: a theoretical and experimental approach. *J Neurosci* 16: 6265–6285.
- Longuet-Higgins HC, Prazdny K (1980) The interpretation of a moving retinal image. *Proc R Soc Lond* 208:385–397.
- Maunsell JH, Van Essen DC (1983) The connections of the middle temporal area (MT) and their relationship to a cortical hierarchy in the macaque monkey. *J Neurosci* 3:2563–2586.
- Miller KD (1994) A model of the development of simple cell receptive fields and the ordered arrangement of orientation columns through activity-dependent competition between ON- and OFF-center inputs. *J Neurosci* 14:409–441.
- Obermayer K, Blasdel GG, Schulten K (1992) Statistical-mechanical analysis of self-organization and pattern formation during the development of visual maps. *Phys Rev A* 45:7568–7589.
- Olson SJ, Grossberg S (1998) A neural network model for the development of simple and complex cell receptive fields within cortical maps of orientation and ocular dominance. *Neural Networks* 11:189–208.
- Orban GA, Lagae L, Verri A, Raiguel S, Xiao DK, Maes H, Torre V (1992) First-order analysis of optical flow in monkey brain. *Proc Natl Acad Sci USA* 89:2595–2599.
- Pack C, Mingolla E (1998) Global induced motion and visual stability in an optic flow illusion. *Vis Res* 38:3083–3093.
- Pack C, Grossberg S, Mingolla E (1997) How does the visual cortex use optic flow to navigate? *Soc Neurosci Abstr* 23:348.
- Pack C, Grossberg S, Mingolla E (1998a) How does cortical area MST use optic flow to navigate? *ARVO Abstr* 39:1520.
- Pack C, Grossberg S, Mingolla E (1998b) Cortical processing of visual motion for smooth pursuit eye movement. *Soc Neurosci Abstr* 24:3440.
- Pack C, Grossberg S, Mingolla E (1999) A neural model of attentive time-to-contact estimation by cortical area MSTd. In preparation.
- Perrone JA, Stone LS (1994) A model of self-motion estimation within primate extrastriate visual cortex. *Vis Res* 34:2917–2938.
- Perrone JA, Stone LS (1998) Emulating the visual receptive-field properties of MST neurons with a template model of heading estimation. *J Neurosci* 18:5958–5975.
- Rauschecker JP, von Grunau MW, Poulin C (1987) Centrifugal organization of direction preferences in the cat's lateral suprasylvian visual cortex and its relation to flow field processing. *J Neurosci* 7:943–958.
- Roy JP, Komatsu H, Wurtz RH (1992) Disparity sensitivity of neurons in monkey extrastriate area MST. *J Neurosci* 12:2478–2492.
- Royden CS, Banks MS, Crowell JA (1994) Estimating heading during eye movements. *Vis Res* 34:3197–3214.
- Saito H, Yukiie M, Tanaka K, Hikosaka K, Fukada Y, Iwai E (1986) Integration of direction signals of image motion in the superior temporal sulcus of the macaque monkey. *J Neurosci* 6:145–157.
- Schwartz EL (1977) Afferent geometry in the primate visual cortex and the generation of neuronal trigger features. *Biol Cybernet* 28:1–24.
- Schwartz EL (1994) Computational studies of the spatial architecture of primate visual cortex: columns, maps, and protomaps. In: *Cerebral cortex*, vol. 10 (Peters A, ed.), pp. 359–411. New York: Plenum Press.
- Simpson W (1993) Optic flow and depth perception. *Spatial Vis* 7:35–75.
- Standage GP, Benevento LA (1983) The organization of connections between the pulvinar and visual area MT in the macaque monkey. *Brain Res*, 262:288–294.
- Steinmetz MA, Motter BC, Duffy CJ, Mountcastle VB (1987) Functional properties of parietal visual neurons: radial organization of directionalities within the visual field. *J Neurosci* 7:177–191.
- Swindale N (1992) A model for the coordinated development of columnar systems in primate striate cortex. *Biol Cybernet* 66:217–230.
- Tanaka K, Saito H (1989) Analysis of the visual field by direction, expansion/contraction, and rotation cells clustered in the dorsal part of the medial superior temporal area of the macaque monkey. *J Neurophysiol* 62:626–641.
- Tanaka K, Fukada Y, Saito H (1989) Underlying mechanisms of the response specificity of expansion/contraction and rotation cells in the dorsal part of the medial superior temporal area of the macaque monkey. *J Neurophysiol* 62:642–656.
- Thier P, Erickson R (1992) Responses of visual-tracking neurons from cortical area MSTl to visual, eye and head motion. *Eur J Neurosci*, 4:539–553.
- Tistarelli M, Sandini G (1993) On the advantages of polar and log-polar mapping for direct estimation of time-to-impact from optical flow. *IEEE Trans PAMI* 15:401–410.
- Tootell RBH, Silverman MS, Switkes E, DeValois RL (1982) Deoxyglucose analysis of retinotopic organization in primate striate cortex. *Science* 218:902–904.
- Van den Berg AV (1996) Judgments of heading. *Vis Res* 36:2337–2350.
- Van den Berg AV, Brenner E (1994a) Humans combine the optic flow with static depth cues for robust perception of heading. *Vis Res* 34:2153–2167.
- Van den Berg AV, Brenner E (1994b) Why two eyes are better than one for judgments of heading. *Nature* 371:700–702.
- Van Essen DC, Newsome, WT, Maunsell JHR (1984) The visual representation in striate cortex of macaque monkey: asymmetries, anisotropies, and individual variability. *Vis Res* 24:429–448.
- Warren WH, Hannon DJ (1990) Eye movements and optical flow. *J Opt Soc Am* 7:160–169.
- Warren WH, Blackwell AW, Kurtz K, Hatsopoulos N, Kalish M (1991) On the sufficiency of the velocity field for perception of heading. *Biol Cybernet* 65:311–320.
- Zemel RS, Sejnowski TJ (1998) A model for encoding multiple object motions and self-motion in area MST of primate visual cortex. *J Neurosci* 18:531–547.

Appendix 1: Input Equations

Retinal Input: Spiral Stimuli

Spiral stimuli are used in many neurophysiological experiments to probe cell selectivity. In Cartesian coordinates, expansion motion centered at a point (x_0, y_0) to a point (x, y) can be defined by:

$$v_x = \kappa(x - x_0) \quad (9)$$

$$v_y = \kappa(y - y_0) \quad (10)$$

where κ scales the rate of expansion, with negative values indicating contraction. Similarly, rotation about a point (x_0, y_0) to a point (x, y) can be defined by:

$$v_x = \vartheta(y - y_0) \quad (11)$$

$$v_y = -\vartheta(x - x_0) \quad (12)$$

where positive values indicate clockwise motion and negative values indicate counterclockwise motion. To generate a spiral stimulus, equation (9) can be combined with (11), and (10) can be combined with (12):

$$v_x = \kappa(1 - c)(x - x_0) + c\vartheta(y - y_0) \quad (13)$$

$$v_y = \kappa(1 - c)(y - y_0) - c\vartheta(x - x_0) \quad (14)$$

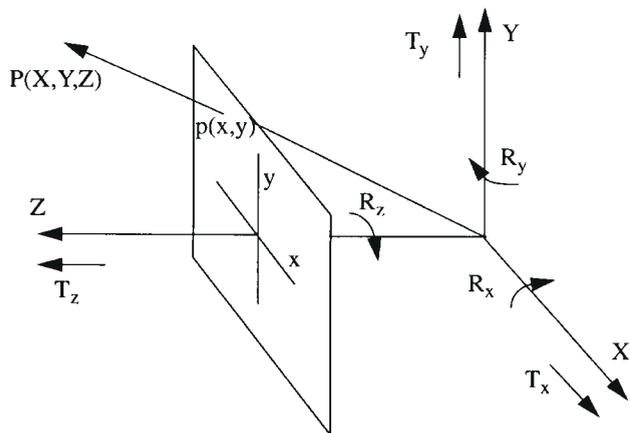


Figure 18. Coordinate frame for optic flow equations. Visual space is depicted in Cartesian coordinates centered on the line of sight. See text for details. [Adapted with permission from Simpson (1993)].

where c defines a continuum of spiral stimuli ranging from pure expansion ($c = 0$) to pure rotation ($c = 1$).

For neurophysiological simulations, the type of spiral was determined by equations (13) and (14). Increasing the value of c in the range (0,1) specified a spiral stimulus with an increasing proportion of circular motion. The radial component of each spiral stimulus was determined by the sign of κ , with positive values indicating expansion and negative values indicating contraction. The circular component was determined by the sign of ϑ , with positive values indicating clockwise rotation and negative values indicating counterclockwise rotation.

Log Polar Mapping of Spiral Stimuli

The above equations can be converted to polar coordinates (ρ, η) by:

$$\dot{\rho} = v_x \cos \eta + v_y \sin \eta \quad (15)$$

$$\dot{\eta} = \frac{v_y \cos \eta - v_x \sin \eta}{\rho} \quad (16)$$

Substituting equations (13) and (14) into (15) and (16) yields:

$$\dot{\rho} = \left[\kappa(1-c)(x-x_0) + c\vartheta(y-y_0) \right] \cos \eta + \left[\kappa(1-c)(y-y_0) - c\vartheta(x-x_0) \right] \sin \eta \quad (17)$$

$$\dot{\eta} = \frac{1}{\rho} \left\{ \left[\kappa(1-c)(y-y_0) - c\vartheta(x-x_0) \right] \cos \eta - \left[\kappa(1-c)(x-x_0) + c\vartheta(y-y_0) \right] \sin \eta \right\} \quad (18)$$

The log polar mapping is given by:

$$\xi = \log(\rho + a) \quad (19)$$

$$\psi = \eta \quad (20)$$

This yields the radial and circular components of speed:

$$\dot{\xi} = \frac{d}{dt} \log(\rho + a) = \frac{\dot{\rho}}{\rho + a} \quad (21)$$

$$\dot{\psi} = \dot{\eta} \quad (22)$$

Retinal Input: the Optic Flow Field

In order to generate a realistic stimulus for heading experiments, the

optic flow field for a moving observer can be characterized mathematically in terms of instantaneous motion vectors (Longuet-Higgins and Prazdny, 1980). Translational movement of the observer along a straight line produces an expanding motion pattern, while rotation of the eye in space generates a streaming pattern which is constant across visual space. Mathematically, the flow field is a projection of these motion vectors in three dimensions onto a flat surface approximating the retina. Thus, a point $P(X, Y, Z)$ has retinal coordinates $(x, y) = (X/Z, Y/Z)$, assuming a projection plane at unit distance from the origin (see Fig. 18). Then for any translational velocity $T = (T_x, T_y, T_z)$ and rotational velocity $R = (R_x, R_y, R_z)$ in three-dimensional space (X, Y, Z) , the resulting motion at retinal point (x, y) is given by

$$v_x = x \frac{T_z}{Z} - \frac{T_x}{Z} + \left[xyR_x - (1+x^2)R_y + yR_z \right] \quad (23)$$

$$v_y = y \frac{T_z}{Z} - \frac{T_y}{Z} + \left[(1+y^2)R_x - xyR_y - xR_z \right] \quad (24)$$

where Z represents the depth of the point $P(X, Y, Z)$ along the optical axis. In this context, the computational challenge to the visual system is to extract the relevant values of T , R and Z from the values of v_x and v_y , distributed across the retina.

Log Polar Mapping of Optic Flow

It is now possible to consider the cortical representation of points in the environment during self-motion. This involves converting the optic flow equations (23) and (24) into log polar coordinates. Following Tistarelli and Sandini (Tistarelli and Sandini, 1993), the retinal velocity of environmental points during observer motion can be described in polar coordinates by radial and circular components. Substituting equations (23) and (24) into (15) and (16) yields

$$\dot{\rho} = \left[\frac{xT_z - T_x}{Z} + xyR_x - (x^2 + 1)R_y + yR_z \right] \cos \eta + \left[\frac{yT_z - T_y}{Z} + (y^2 + 1)R_x - xyR_y - xR_z \right] \sin \eta \quad (25)$$

$$\dot{\eta} = \frac{1}{\rho} \left\{ \left[\frac{yT_z - T_y}{Z} + (y^2 + 1)R_x - xyR_y - xR_z \right] \cos \eta - \left[\frac{xT_z - T_x}{Z} + xyR_x - (x^2 + 1)R_y + yR_z \right] \sin \eta \right\} \quad (26)$$

Replacing the Cartesian points (x, y) with their polar values ($\rho \cos \eta, \rho \sin \eta$) leaves

$$\dot{\rho} = \frac{1}{\rho} (\rho T_z - T_x \cos \eta + T_y \sin \eta) + (\rho^2 + 1) (R_x \sin \eta - R_y \cos \eta) \quad (27)$$

$$\dot{\eta} = \frac{1}{\rho} \left[\left(\frac{T_x}{Z} + R_y \right) \sin \eta + \left(R_x - \frac{T_y}{Z} \right) \cos \eta \right] - R_z \quad (28)$$

Equations (27) and (28) describe the optic flow field in polar coordinates. The logarithmic magnification of the central visual field relative to the periphery requires a further transformation, as specified in equations (19) and (20).

Removal of Rotational Flow

For heading simulations in which a real eye movement was assumed to

remove the rotational part of the flow field, we set $R_x = R_y = R_z = 0$ in equations (27) and (28).

Appendix 2: Simulation Techniques

Noise

For the simulation that required judgments of the percentage of responses that were greater than a threshold, model cell outputs were perturbed with random noise (Green and Swets, 1974). In each case, the random noise was calculated from a Gaussian probability distribution centered around zero. The algorithm for generating the noise N was as follows:

1. Select a random number n_1 from the range $[-1,1]$.
2. Select another random number n_2 from the same range.
3. If

$$n_2 < \exp(-Gn_1^2) \quad (29)$$

then $N = n_1$. Otherwise, return to step 1.

The value of N was then multiplied by a constant (set to 0.5 in the simulations) and added to the cell output. For all simulations, $G = 1$.