A Neural Model of Visually Guided Steering, Obstacle Avoidance, and Route Selection

David M. Elder, Stephen Grossberg, and Ennio Mingolla

Department of Cognitive and Neural Systems
Center for Adaptive Systems
and
Center of Excellence for Learning in Education, Science, and Technology
Boston University
677 Beacon Street, Boston MA, 02215
Phone: 617-353-7858
Fax: 617-353-7755

Journal of Experimental Psychology: Human Perception & Performance, in press
Submitted: April, 2007
Revised: March, 2009

All correspondence should be addressed to

Professor Stephen Grossberg
Department of Cognitive and Neural Systems
Boston University
677 Beacon Street
Boston, MA 02215
Phone: 617-353-7858
Fax: 617-353-7755
Email: steve@cns.bu.edu

Abstract

A neural model is developed to explain how humans can approach a goal object on foot while steering around obstacles to avoid collisions in a cluttered environment. The model uses optic flow from a 3D virtual reality environment to determine the position of objects based on motion discontinuities, and computes heading direction, or the direction of self-motion, from global optic flow. The cortical representation of heading interacts with the representations of a goal and obstacles such that the goal acts as an attractor of heading, while obstacles act as repellers. In addition the model maintains fixation on the goal object by generating smooth pursuit eye movements. Eye rotations can distort the optic flow field, complicating heading perception, and the model uses extraretinal signals to correct for this distortion and accurately represent heading. The model explains how motion processing mechanisms in cortical areas MT, MST, and posterior parietal cortex can be used to guide steering. The model quantitatively simulates human psychophysical data about visually-guided steering, obstacle avoidance, and route selection.

Key Words: Heading Perception, Steering, Optic Flow, Obstacle, Goal, Pursuit Eye Movement, Gain Fields, Peak Shift, V2, MT, MST, PPC, LIP

1. Introduction

Attractors and repellers in visually-guided navigation. Many important steering tasks are guided by visual information, including walking through a cluttered environment (Fajen & Warren, 2003), driving (Land & Horwood, 1995; Hildreth, Beusmans, Boer, & Royden, 2000), vehicle braking (Lee, 1976), and piloting an aircraft (Gibson, Olum, & Rosenblatt, 1955; Beall & Loomis, 1997). Steering through a cluttered environment involves the selection of a path that avoids obstacles and simultaneously approaches the intended goal. Human steering is guided by visual information about the spatial layout of the environment and the direction of self-motion through the environment, as well as proprioceptive feedback and other sensory inputs.

Movement through the world creates a full-field motion pattern on the retina, called *optic flow*, which contains information about the direction of self-motion, or heading (Gibson, 1950). In principle, optic flow can be used to compute heading (Longuet-Higgins & Prazdny, 1980). During translational movement, the optic flow field contains a singularity, called the *focus of expansion*, which specifies the direction of heading in the absence of an eye rotation.

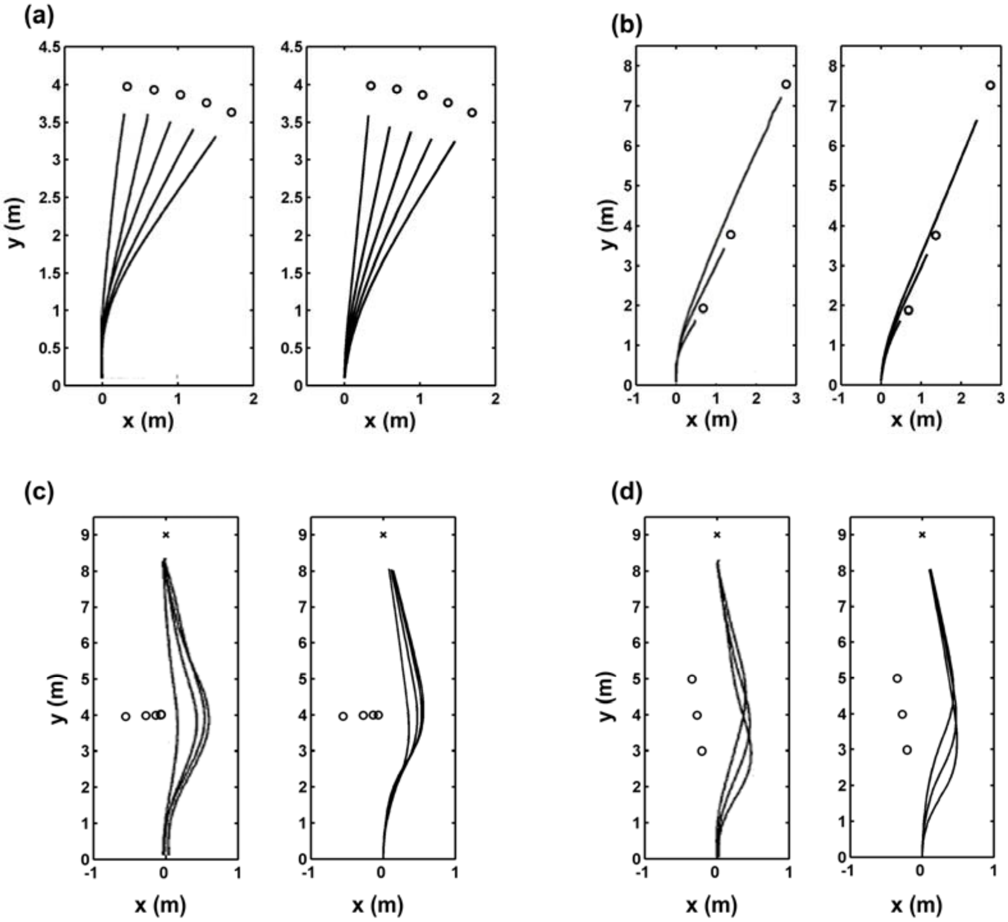


Figure 1. (a), left panel: Mean goal approach paths taken by subjects in Experiment 1a, Fajen & Warren (2003); right panel: Goal approach trajectories steered by the STARS model. Goals are placed at an initial distance of 4 meters and initial angles of 5°, 10°, 15°, 20°, and 25°. (b), left panel: Mean goal approach paths taken by subjects in Experiment 1b of Fajen & Warren (2003); right panel: Goal approach trajectories steered by STARS. Goals are placed at initial distances of 2, 4, and 8 meters and an initial angle

of 20º. (c), left panel: Mean obstacle avoidance paths taken by subjects in Experiment 2a of Fajen & Warren (2003); right panel: Obstacle avoidance trajectories steered by STARS. The goal is placed at an initial distance of 9 meters and obstacles at an initial distance of 4 meters and initial angles of -8º, -4º, -2º, and -1º. (d) left panel: Obstacle avoidance paths taken by subjects in Experiment 2b of Fajen & Warren (2003); right panel: Obstacle avoidance trajectories steered by STARS. The goal is placed at an initial distance of 9 meters and obstacles at initial distances of 3, 4, and 5 meters and an initial angle of -4º. [Left panels in (a) and (b) are reprinted with permission from Figure 4 of Fajen and Warren (2003). Left panels in (c) and (d) are reprinted with permission from Figure 10 of Fajen and Warren (2003).

Whereas optic flow relates observer motion to the available visual stimuli, recent research clarifies how visual information governs the dynamics of human navigational behavior. Fajen and Warren (2003) studied the dynamics of human steering behavior in simple goal approach and obstacle avoidance tasks using an immersive virtual reality system. They found that human performance on these tasks can be described by a simple dynamical control model, which is referred to as the FW model herein. The goal object acts as an *attractor* of heading, while obstacles act as *repellers* of heading. Figures 1a and 1b show goal approach trajectories steered by the FW model compared to the results produced by our model. Figures 1c and 1d show obstacle avoidance trajectories steered by the FW model and our model. Fajen and Warren's analysis identified as relevant control variables the distances to the goal and obstacles and the angles between the objects and the current heading direction, as illustrated in Figure 2. Other studies of locomotion emphasize the importance of the perceived location of the goal object (Rushton, Harris, Lloyd, & Wann, 1998) and gaze direction (Wilkie & Wann, 2002, 2003a, 2003b, 2005).

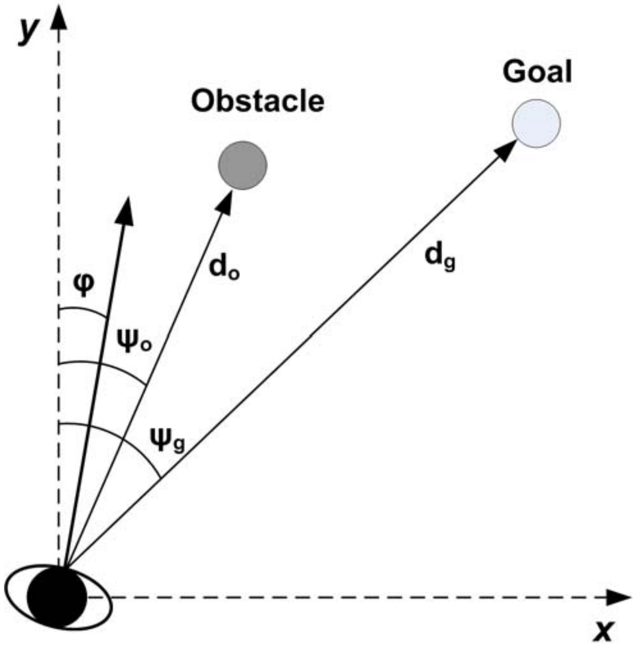


Figure 2. Scene geometry used to control steering in the FW model (Fajen & Warren, 2003). Goal and obstacle distances are denoted d_g and d_o, heading angle is denoted φ, and goal and obstacle angles are denoted ψ_g and ψ_o. Angles are computed in world

coordinates, relative to the *y*-axis. Adapted with permission from Figure 1 in Fajen and Warren (2003).

The FW model computes trajectories from a third-person perspective, utilizing complete geometric information about the environment. Such information may not, however, be directly available to the brain's visual system from a first-person perspective. Behavioral and dynamical systems analyses like the FW model do not attempt to explain how the primate central nervous system produces intelligent steering behavior from visual and motor information, in keeping with its ecological approach (Gibson, 1979). A central goal of our work is to understand how attractor-repeller steering dynamics can emerge from a biological visual system.

In the laboratory, humans can accurately perceive heading direction relative to a stationary object using simulated optic flow (Warren, Morris & Kalish, 1988). The accuracy of heading perception varies across the visual field and is highest when the focus of expansion is near the fovea and when the optic flow stimulus covers the central visual field (Warren & Kurtz, 1992; Crowell & Banks, 1993; Atchley & Andersen, 1999). Real world navigational situations involve the traversal of environments containing stationary and moving objects. Typically an independently moving object in the optic flow field can impair heading judgment when the object obscures the focus of expansion (Warren & Saunders, 1995; Royden & Hildreth, 1996). Subjects can tolerate velocity noise in the optic flow display, suggesting that heading perception primarily relies on the directional pattern of motion in optic flow (Warren, Blackwell, Kurtz, Hatsopoulos, & Kalish, 1991).

While translational heading perception is robust under a variety of conditions, rotational flow can create problems for heading perception. Rotational flow is produced by eye and head rotations as well as movement of the body along a curved path. Heading can be perceived accurately while making a smooth eye movement to track a fixed point in a scene (van den Berg, 1993) and while tracking a point moving independently in a scene (Royden, Crowell, & Banks, 1994). However, when a person views an optic flow field that contains rotational flow due to a *simulated* eye movement without actually moving their eyes, heading perception is impaired for sufficiently rapid eye velocities (Royden *et al.*, 1994). In simulated eye rotation experiments, subjects often report the sensation of traveling along a curved path, rather than traveling along a straight path while rotating the eye (Royden, Banks, & Crowell, 1992). At low simulated velocity (< 1 deg/sec), heading judgments are relatively accurate (Warren & Hannon, 1988), but at higher velocities, they are systematically biased in the direction of the simulated eye rotation (Royden *et al*. 1992, 1994). Even at low velocities, heading perception can be inaccurate if the simulated rotation is designed to mimic tracking of an independently moving object (Cutting, Vishton, & Braren, 1995).

The depth structure of the scene also influences the accuracy of heading perception during simulated eye rotations. Scenes containing with depth information, especially ground planes, offer additional cues about the direction of heading and the type of path traveled, either linear or curved. Several studies (van den Berg 1993, 1996; Cutting *et al*, 1995) have reported that heading perception remains is accurate during a simulated eye rotation in a scene with structured depth information. Although random dot clouds contain depth information, the distribution of depth information in a random cloud is less structured than in a ground plane, and studies have shown that heading perception during a simulated eye rotation using a random dot cloud stimulus is inaccurate (van den Berg, 1992; Royden *et al*., 1994).

Neurophysiological substrates of navigation. We now briefly describe the known neurophysiological mechanisms involved in heading perception and steering. Visual motion is processed in several areas of the primate brain, including cortical area V1 (Hubel & Wiesel, 1962; Pack, Livingstone, Duffy, & Born, 2003). Cortical area MT receives a projection from area V1 (Zeki, 1971; Sincich & Horton, 2003; Born & Bradley, 2005), and is specialized for the processing of visual motion (Zeki, 1974; Albright, 1984). Cells in MT respond to the direction and velocity of motion (Allman, Miezin, & McGuinness, 1985) as well as the orientation of moving contours (Albright, 1984). MT cells can be divided into two broad groups (Allman *et al.*, 1985; Born & Tootell, 1992). One type of cell prefers large motion patterns, and fails to respond to small stimuli moving in the receptive field. These *additive* cells integrate motion within their receptive fields, in a manner consistent with the processing of a large-scale optic flow field. The second MT cell type prefers small motion stimuli, and suppresses firing in the presence of larger motion patterns. These *subtractive* cells typically have antagonistic surround regions, and are maximally excited when surround in the direction opposite to the preferred motion direction in the receptive field center. Subtractive cells are useful for processing object motion, and exhibit different binocular disparity preferences in their center and surround regions (Bradley & Andersen, 1995), suggesting a role for MT in object segmentation based on motion and depth discontinuity.

Area MT projects to several areas, including area MST (Maunsell & Van Essen, 1983a; Boussaoud, Ungerleider, & Desimone, 1990). The populations of additive and subtractive cells in MT connect anatomically to distinct regions in area MST (Berezovskii & Born, 2000). Ventral MST (MSTv) contains cells with relatively small receptive fields that respond to the movement of small objects (Tanaka & Saito, 1993; Eifuku & Wurtz, 1998). MSTv cells are generally unresponsive to full-field motion patterns. Dorsal MST (MSTd) contains cells with large receptive fields and which respond to full-field motion patterns, but not smaller stimuli (Tanaka & Saito, 1989; Duffy & Wurtz, 1991). MSTd cells respond to optic flow patterns such as radial expansion, rotation, and spirals. The response characteristics of MSTd cells make them suitable for the processing of heading from optic flow. Some MSTd cells respond also to the direction of background motion generated by an eye rotation (Komatsu & Wurtz, 1988). Thus, both MT+/MT- and MSTd/MSTv exhibit complementary computational properties in order to carry out navigational and object tracking behaviors (Grossberg, 2000).

Visual areas such as V1, V2, MT, and MST represent space using a retinotopic coordinate system. However, planning of motor actions, particularly steering, is easier in head-centered (craniotopic) or body-centered coordinates. Converting from retinotopic coordinates to craniotopic or body-centered coordinates must take into account the orientation of the eyes within the head as well as the orientation of the head on the body. Several areas in the higher visual system, including MSTd (Squatrito & Maioli, 1997) and area 7a (Andersen, Essick, & Siegel, 1985, 1987) have cells whose visual responses to retinotopic stimuli are modulated by the position of the eyes in the orbit via an eye position *gain field.* Gain fields allow visually responsive cells to integrate extraretinal information about the position of the eyes. Cells exhibiting gain fields have retinotopic receptive fields that elicit responses when the preferred visual stimulus falls within the receptive field on the retina, regardless of the position of the eyes. However, changing the position of the eyes causes the overall firing rate of the cell to increase or decrease, typically as a linear function of eye position. Gain fields are believed to play an intermediate computational role in the transformation from retinotopic to craniotopic coordinates. Several modeling studies propose how gain fields may work (Zipser & Andersen, 1988; Salinas & Abbott, 1995; Pouget & Sejnowski, 1997).

Introduction to STARS. We have developed a neural network model for Steering, Tracking, And Route Selection (the STARS model) that approaches goals, avoids obstacles, and follows realistic trajectories through cluttered environments. The central tasks the STARS model must solve are the following:

1. Extract the goal and obstacle positions from the optic flow field. Model MT^- subtractive cells detect discontinuities in the flow field to segment an object from the background. MSTv cells group MT^- subtractive signals into Gaussian-shaped activity maps that code object position.

2. Compute heading from optic flow. Model MSTd cells use MT^+ cell inputs to detect expanding optic flow fields and compute a map representing heading direction. Extraretinal signals correct for the effects of eye rotation.

3. Compute steering from goal and obstacle position and heading information. The goal, obstacle, and heading maps are converted into craniotopic coordinates and summed to produce a peaked steering map. The Gaussians corresponding to obstacles are subtracted from those representing the goal and the heading directions. Shifts in the peak location to the left or right of the center of the map cause the model to steer left or right.

4. Maintain fixation on the goal object. STARS commands a smooth pursuit eye movement signal by computing the background optic flow due to eye rotation and the retinal slip due to imperfect tracking of the target. These signals allow the model to estimate relative target velocity and maintain fixation on the goal by changing gaze direction.

The STARS model unifies and extends functional mechanisms described in previous models, especially the visual navigation model of Cameron, Grossberg, and Guenther (1998), the MSTd heading model of Grossberg, Mingolla, and Pack (1999), and the smooth pursuit control model of Pack, Grossberg, and Mingolla (2001), clarifies how cortical motion processing mechanisms can produce competent steering performance, and quantitatively simulates recent behavioral data on steering dynamics.

This paper has the following organization. Section 2 describes the model architecture and how model processing stages interact to produce steering behavior. Section 3 presents computer simulations illustrating model steering and trajectories through various environments. Section 4 discusses how the model accounts for behavioral data and discusses the model's computational stages in light of known neurophysiological and anatomical data. Appendix A presents model equations, and Appendix B describes the simulation system used to run the computer experiments.

2. Model Description

Model overview. STARS extracts information about heading and object location from the optic flow field and to uses this information to produce realistic steering behavior, whereby the goal attracts and obstacles repel heading. Because the model contains several interacting systems, we begin by briefly describing their function before presenting a detailed description. Whereas the FW model computes steering trajectories from a third-person, top-down view of the environment, STARS operates from a first-person perspective, using visual information available to a moving observer. STARS computes a steering trajectory by adding together three peaked distributions of activity; namely, spatial maps that encode the heading direction, goal position, and obstacle position in head-centered coordinates, as shown in Figure 3a. The result is another peaked distribution of activity, whose peak is shifted to the left or right as a result of oppositely signed interactions between the goal and obstacle distributions and the heading distribution. The direction and magnitude of this peak shift causes a change in angular steering velocity.

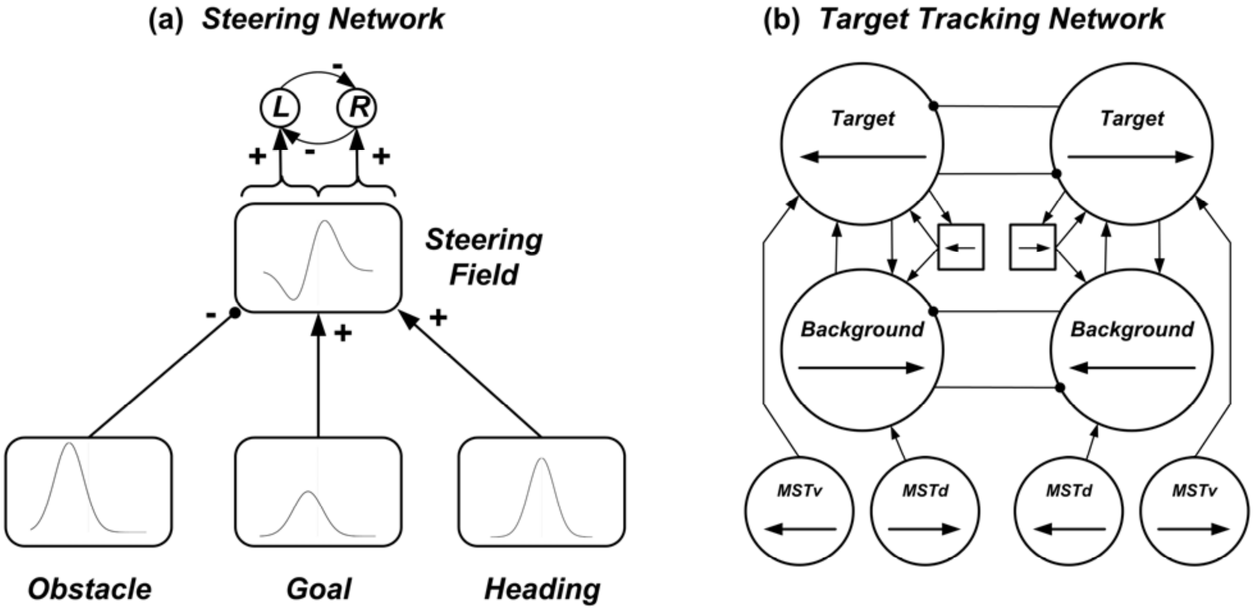


Figure 3. (a) STARS model steering control mechanism. Peaked population responses in a spatial map are used to represent the obstacle, goal, and heading directions. They are summed to produce a shifted steering peak in PPC. A winner-take-all circuit finds the maximally activated cell in the PPC steering field. If the maximum is on the left, it activates the left turning cell *L*, and if it is on the right, it activates the right turning cell *R*. The activity of the turning cells *L* and *R* cause the model to execute a turn to the left or right. (b) STARS model target tracking mechanism. The filled arrows indicate excitatory connections, and the filled circles indicate inhibitory connections. Background motion cells integrate directional signals from the MSTd planar flow cell population, and target tracking cells integrate directional signals from the MSTv fixation depth population. The small boxes represent smooth pursuit cells which generate eye movements to the left and right and provide an efference copy to the target and background motion cells. See text for details. Adapted with permission from Pack, Grossberg, & Mingolla (2001, Figure 3).

STARS computes goal and obstacle positions by partitioning a scene into near, fixation, and far depth regions. This spatial partitioning is computed directly from the scene geometry and supplied to the model. The Discussion section describes other models that can compute depth segmentation using binocular disparity. STARS model maintains fixation on the goal object, the goal remains at the fixation depth. Objects in the near depth region are treated as obstacles. In our simulations, no objects appear more distant than the goal. The smooth pursuit control system that maintains goal fixation is adapted from Pack, Grossberg, and Mingolla (2001). Both the steering and pursuit control systems in STARS are reactive (Figure 3b), requiring only immediately available visual information to intelligently steer while tracking the goal with eye movements. No explicit path planning mechanism is needed.

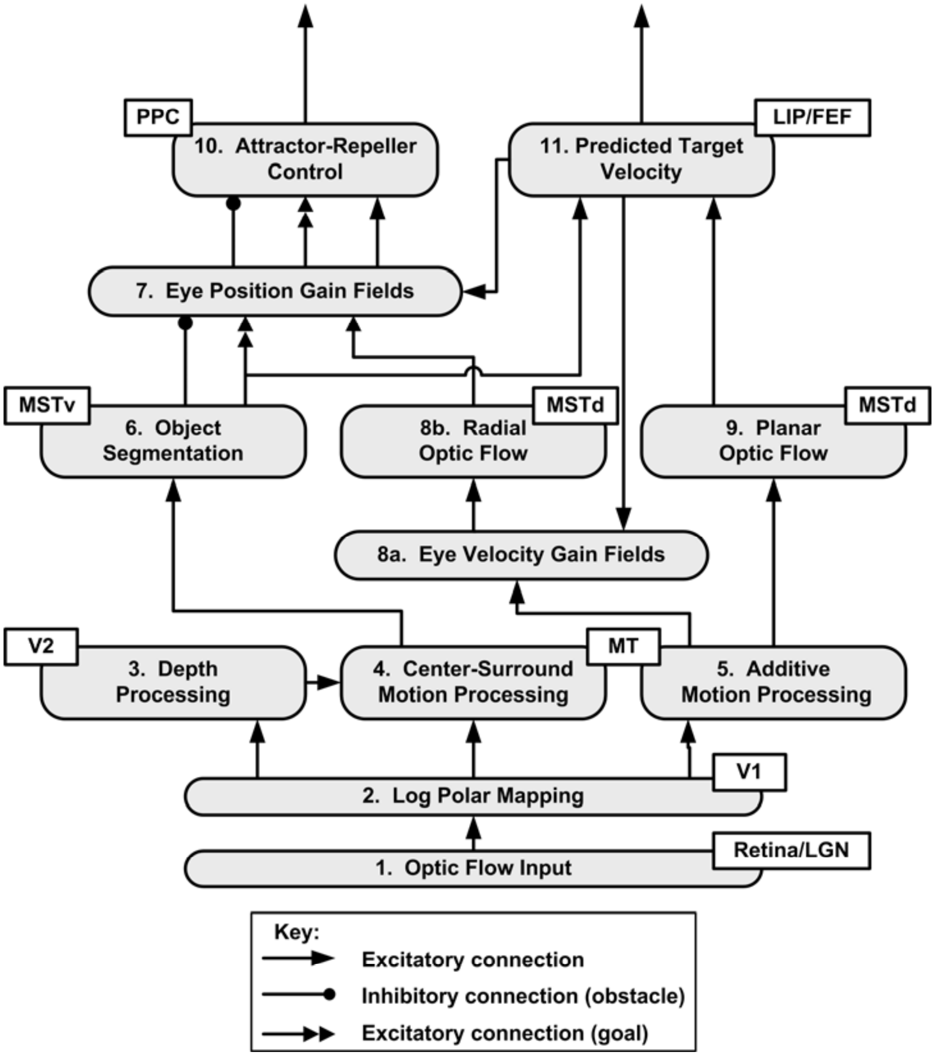


Figure 4. Functional architecture of the STARS model. See text for details.

Figure 4 shows a block diagram of the STARS model architecture. The optic flow field is registered on the retina (Level 1) and is transformed to cortical log-polar coordinates in area V1 (Level 2). Area V2 (Level 3) processes scene depth, and depth-selective MT^- subtractive cells (Level 4) detect discontinuities in the optic flow field. Additive MT^+ cells (Level 5) perform local motion integration, building a distributed representation of the optic flow field. Object motion cells in MSTv (Level 6) encode the obstacles at near depth and the goal at fixation depth. Eye position gain fields (Level 7) help transform the goal, obstacle, and heading directions into craniotopic coordinates. Eye velocity gain fields (Level 8a) correct for the effects of eye rotation, before building a retinotopic representation of heading direction in MSTd (Level 8b). Planar flow cells in MSTd (Level 9) encode the direction of background flow due to eye rotation. The craniotopic representations of goal, obstacle, and heading are combined in posterior parietal cortex (PPC, Level 10) to produce a steering signal. The goal representation in MSTv (Level 6) and the background flow cells in MSTd (Level 9) interact to produce a smooth pursuit eye movement command (Level 11), whereby the model fixates the goal throughout a steering simulation.

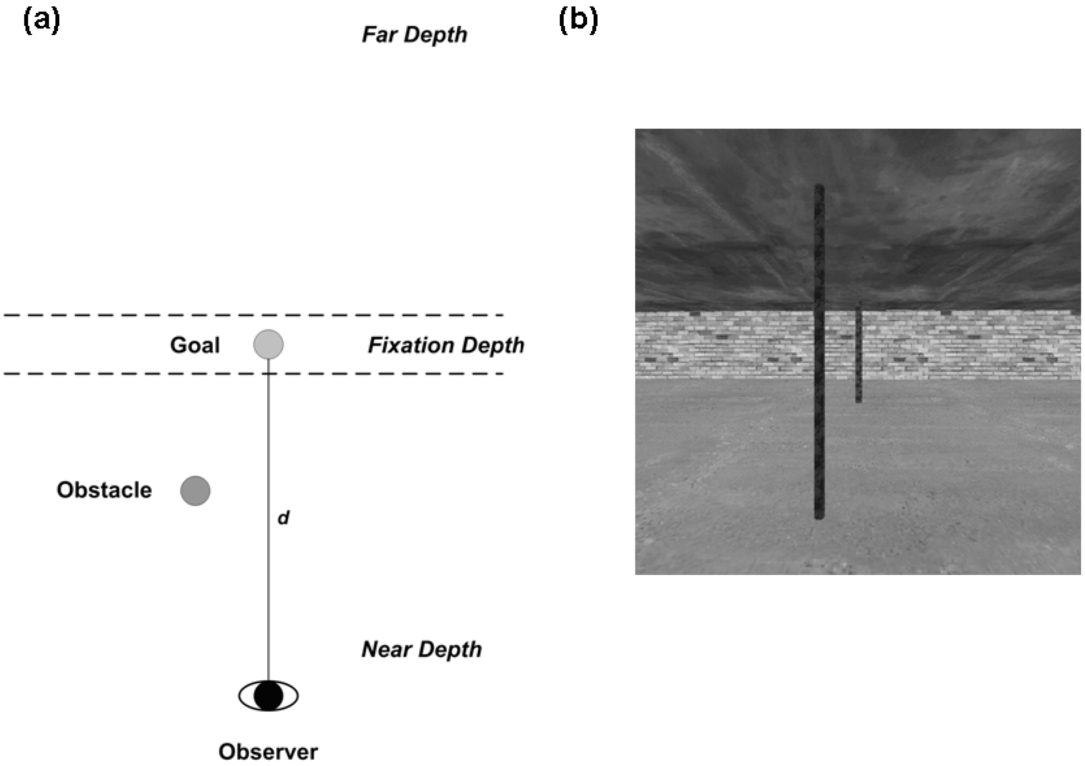


Figure 5. (a) The simulated scene is segmented into 3 depth regions, relative to the fixation point. The observer is at the bottom, at a distance *d* from the goal object. The goal object occupies a narrow band of space at *fixation depth*. Objects nearer than the goal, such as the obstacle, are at *near depth*, and objects further away are at *far depth*. (b) The 3D simulation environment consists of a square room with textured ceiling, floor, and walls. Goal and obstacle objects are represented by vertical cylinders from the floor to the ceiling. The closer pole is the obstacle. The model input is an optic flow field, generated analytically from the geometric specification of the environment and the simulated observer's trajectory.

Simulation Environment. The model was simulated in a 3D computer graphics environment. The environment is a square room with textured walls, floor, and ceiling, and containing a goal and obstacle (Figure 5a). The goal and obstacle are vertical cylinders extending from the floor to the ceiling (Figure 5b). Although our 3D scenes are rendered to look realistic, the neural model does not actually use the color, lighting, or texture information from the frame images. Only the optic flow field, computed from the scene geometry, provides input to the model. We used this simplification to concentrate on steering dynamics. The environments are designed so that the goal object appears 10 simulated meters from the initial viewpoint and any obstacles appear closer than the goal. Our 3D environments are similar to the environments used in the experiments conducted by Fajen and Warren (2003), although our environments are enclosed by textured walls and a ceiling, whereas their environment only consisted of a floor and cylindrical poles.

Level 1: Optic flow input. The input to the STARS model is a dense optic flow field computed analytically from the 3D scene geometry and observer trajectory, using the mathematical formulation of Longuet-Higgins and Prazdny (1980); see Appendix A, Equations (1) and (2). An optic flow vector is computed for each pixel in the retinal optic flow image on each time step.

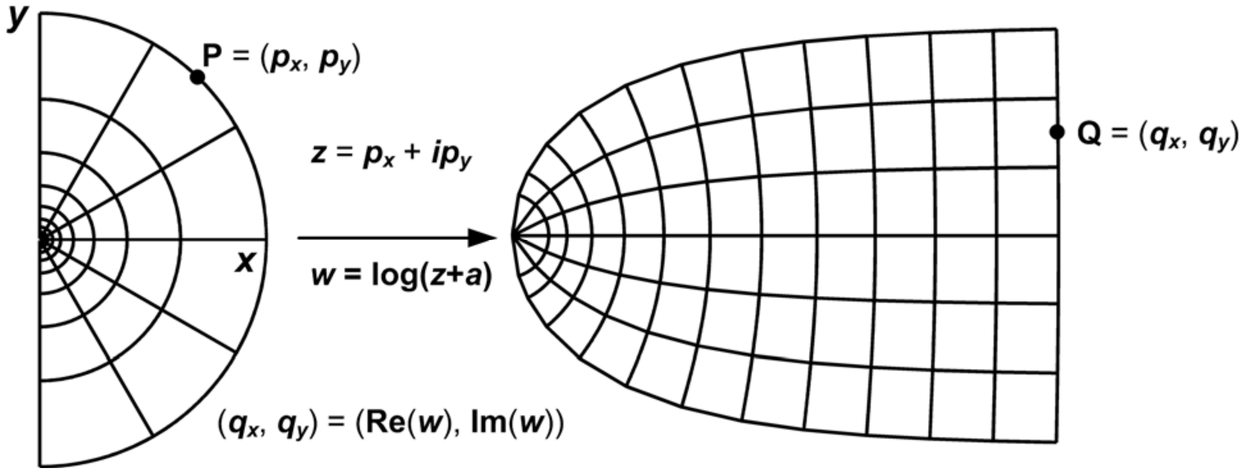
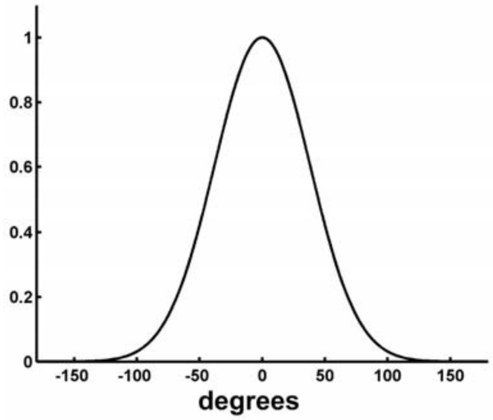


Figure 6. Mapping of the right visual hemifield from retinal to cortical log-polar coordinates.

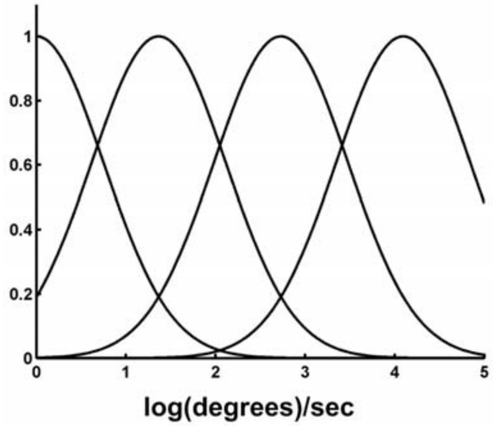
Level 2: Log polar mapping in area V1. Level 2 transforms the Cartesian optic flow field into log-polar cortical coordinates; see Figure 6 and Appendix A, Equations (3) and (4). The log-polar transformation maps points on the retina to points on the surface of primary visual cortex (Daniel & Whitteridge, 1961; Schwartz, 1977, 1980). The log-polar map accounts for several properties of area V1. The fovea is magnified in V1, and the parafoveal and peripheral regions are compressed. Radial lines on the retina project to nearly parallel horizontal lines in V1, while concentric rings on the retina map to nearly parallel vertical lines in V1. The log-polar map allows for high-resolution visual processing of foveated objects, at the cost of lower-resolution processing of objects in the periphery. Optic flow fields are transformed into log-polar coordinates by computing the Jacobian of the log-polar mapping at each point and multiplying the Cartesian flow vectors by the Jacobian.

Level 3: Depth processing in area V2. The STARS model assumes that subtraction cells in MT$^-$ derive their disparity selectivity from the V2-to-MT pathway (Maunsell & Van Essen, 1983a). The neural computation of depth from disparity is a complex problem (see, e.g., Ozawa, DeAngelis, & Freeman, 1990; Howard & Rogers, 1995; Grossberg & Howe, 2003; Grossberg & Swaminathan, 2004; Cao & Grossberg, 2005; Berzhanskaya, Grossberg, & Mingolla, 2007; Fang & Grossberg, 2008; Ponce, Lomber, & Born, 2008), and requires a significant number of mechanisms not directly related to the problem of steering control. For simplicity, we do not simulate binocular disparity in the V1-V2 network. Rather, we assume that such a system exists, and partition visual space into coarse depth regions corresponding to near, fixation, and far depths (or, regions having crossed, zero, and uncrossed disparity); see Appendix A, Equations (5) and (6).

(a) Direction Tuning



(b) Log-Polar Speed Tuning



(c) Cartesian Speed Tuning

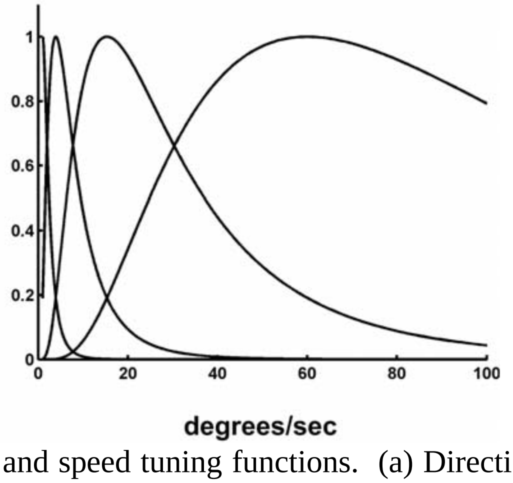


Figure 7. MT cell direction and speed tuning functions. (a) Directional tuning function for a cell preferring motion to the right (0º). Seven other Gaussian tuning curves are used in the simulations, with maxima spaced equally in 45º increments. (b) Four speed tuning functions in log-polar coordinates. (c) speed tuning functions in Cartesian coordinates.

We use planes perpendicular to the line of sight to partition space. The positions of these planes are updated on each simulation time step and are derived from the distance between the eye and the fixation point. Because the model tracks the goal object, the fixation point is always on the goal object, so the goal appears in the fixation depth region, while obstacles appear in near space. No objects appear in the far depth region, although the ceiling, floor, and walls extend into this space.

Level 4: MT$^-$ subtractive cells. STARS extracts the positions of the goal and obstacle from the flow field using motion and depth discontinuities. MT$^-$ cells receive log-polar optic flow input from V1 and depth input from V2; see Appendix A, Equations (7), (8), and (12). These cells are tuned for direction, velocity of motion, and depth (or disparity). Their spatial receptive fields have a circular center preferring motion in a particular direction, surrounded by region preferring motion in the opposite direction. Hence, the cell is maximally activated when the receptive field center receives motion input in the preferred direction and the surround receives motion input in the anti-preferred direction; see Appendix A, Equation (7). When we refer to a cell's preferred direction, we mean the preferred direction of the receptive field center. Directional tuning is modeled using a Gaussian function across direction; see Figure 7a and Appendix A, Equation (9). Such motion interactions are known to exist *in vivo* (Allman, Miezin, & McGuinness, 1985; Born, 2000).

MT$^-$ cells are tuned for stimuli moving at particular log-polar speeds. Due to the spatial anisotropy of the log-polar mapping, a particular log-polar speed corresponds to different Cartesian speeds depending on the eccentricity of the moving stimulus (Figures 7b and 7c). This property allows model MT cells to encode a wide range of Cartesian speeds using a small number of preferred log-polar speeds. The speed tuning of MT cells is modeled by a Gaussian function; see Appendix A, Equation (10).

Finally, model MT$^-$ cells also respond preferentially to stimuli at different depths in the center and surround (Bradley & Andersen, 1998); compare Appendix A, Equations (8) and (12). The cell responds maximally to a central stimulus at the preferred depth and a surround stimulus at a nearer or farther depth. We call MT cells *near-*, *fixation-*, or *far-depth* cells if the receptive field center prefers stimuli moving in the near, fixation, or far depth region, respectively, as computed in area V2 (Figure 8).

To summarize, model MT- cells have tunings for direction, speed, and depth, as well as direction and depth opponencies organized by the spatial center-surround structure of cell receptive fields. At each discrete position in log-polar space, we simulate a cluster of cells having 8 preferred directions, 4 preferred speeds, and two preferred depths, for a total of 64 MT$^-$ cells. This cell cluster comprises a motion hypercolumn in MT$^-$.

Level 5: MT$^+$ additive cells. The additive MT$^+$ cell population builds a distributed representation of the log-polar optic flow field. Unlike the MT$^-$ cells, MT$^+$ cells lack opponent-motion mechanisms. Rather, an additive MT$^+$ cell has a single Gaussian spatial receptive field within which the cell integrates the optic flow input signal; see Appendix A, Equations (14) and (15). MT+ cells are tuned for direction and speed, but not depth. In this way, additive MT$^+$ cells can sample a sufficient distribution of the optic flow field. At each position in the MT$^+$ additive cell map, we simulate a cluster of 32 cells, each having one of eight preferred directions and four preferred speeds. The direction and speed tunings are same as those used by MT$^-$ cells.

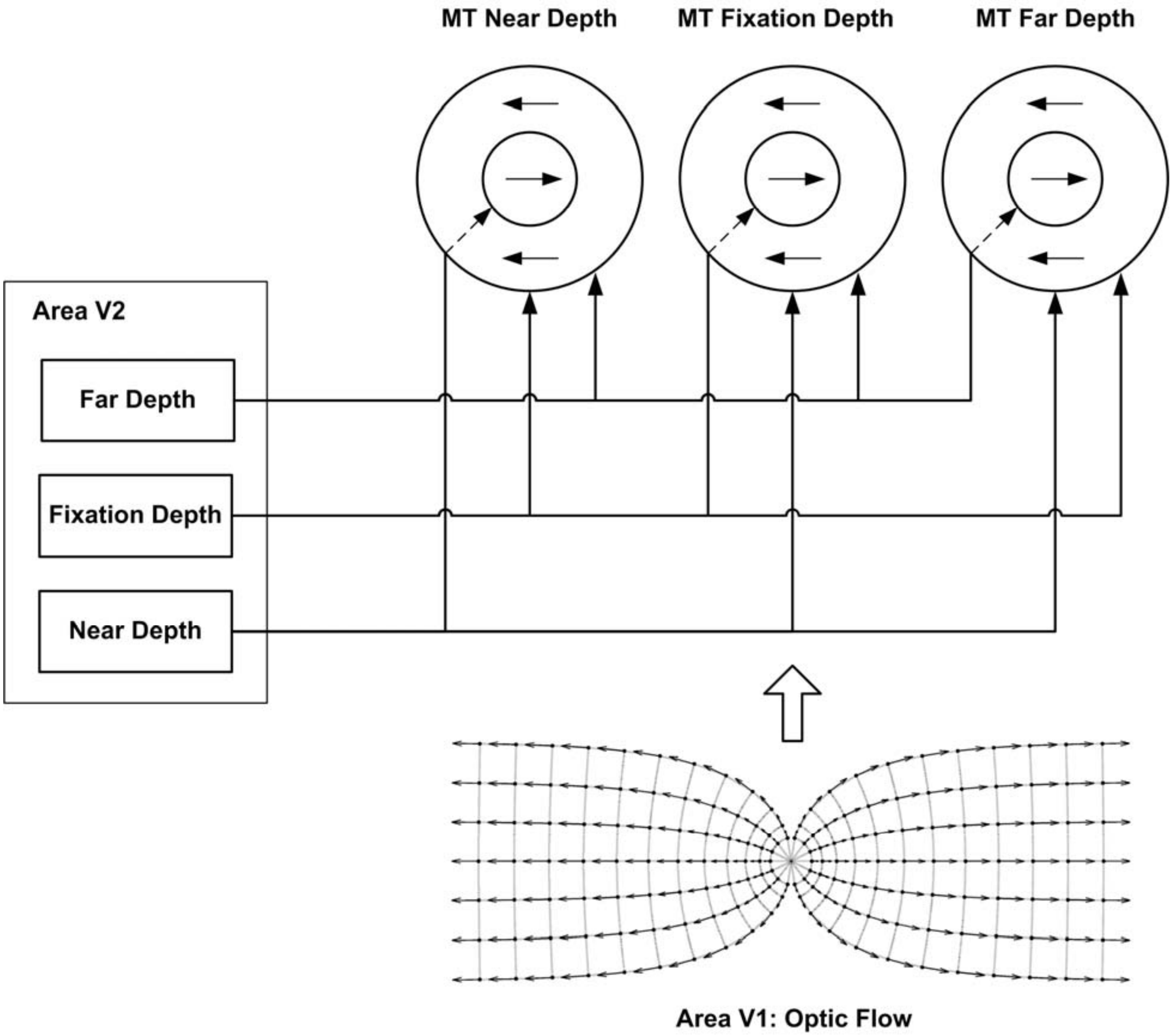


Figure 8. Inputs to MT$^-$ subtractive cells. Three MT$^-$ cells are shown, preferring motion in the receptive field center at near, fixation, and far depths. The receptive field centers prefer motion to the right (solid arrows), while the surrounds prefer motion to the left. The receptive field centers receive depth signals from V2 corresponding to their preferred depth, while the surrounds receive inputs corresponding to the anti-preferred depths. For instance, the fixation depth cell receives fixation depth inputs in the receptive field center, and near and far depth inputs in the surround. MT$^-$ far depth cells are shown for completeness, but are not simulated.

Level 6: MSTv object tracking cells. Cells in MSTv encode the position and direction of motion of objects in the optic flow field. The STARS model contains two depth-selective populations of MSTv cells, one encoding objects at near depth (obstacles) and one encoding objects at fixation depth (the goal). The near-depth MSTv cells receive input from near-depth MT- cells, and likewise for fixation-depth MSTv cells. MSTv cells are directionally selective, summing inputs from the corresponding MT$^-$ cell population with the same preferred direction. MSTv cells are not tuned for specific preferred speeds. Rather, the inputs from MT$^-$ are weighted proportionally to the preferred speed of the MT$^-$ cell, so that MT$^-$ cells preferring faster speeds produce larger responses; see Appendix A, Equation (16). This generates in MSTv responses that scale with the speed of the underlying flow field, consistent with MSTv speed responses observed physiologically (Tanaka *et al.*, 1993). Greater activation in an MSTv cell thus means there is higher motion velocity within the cell's receptive field. In a static environment, this can signal the presence of a nearby object in depth.

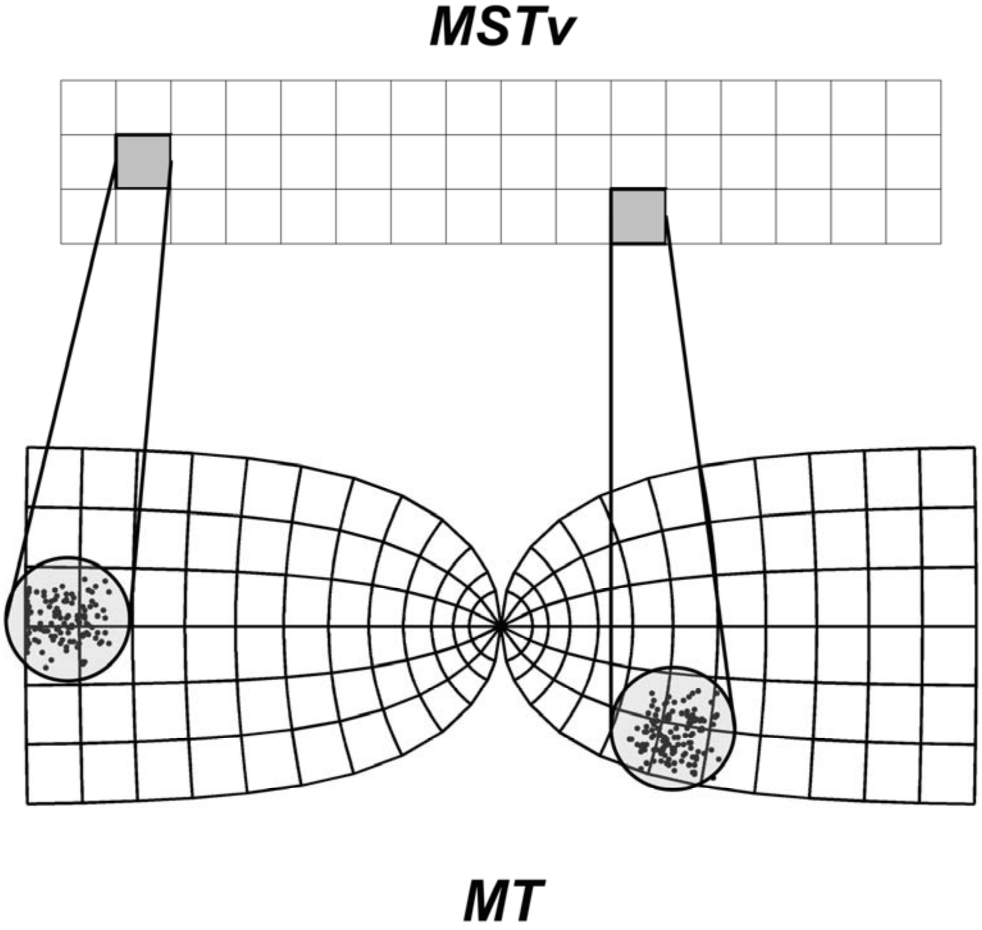


Figure 9. Model MSTv cells have receptive fields arranged in a rectangle covering the width of the visual field and a narrower portion of the height of the visual field. Inputs to each cell are chosen by mapping the cell's receptive field center into log-polar coordinates and choosing 128 MT cells from a Gaussian distribution centered on the log-polar coordinate of the receptive field center. Each MSTv cell receives inputs from approximately 3% of the cells in MT.

MT$^-$ inputs to an individual MSTv cell are organized as follows: A 2D coordinate is assigned to the MSTv cell representing the center of its visual receptive field in Cartesian coordinates. Then 128 additional coordinates are selected randomly from a 2D Gaussian distribution with a mean equal to the coordinates of the receptive field center. The log-polar coordinates corresponding to each randomly chosen Cartesian coordinate are computed, and the nearest MT cell cluster in log-polar space is selected. In this way, a Cartesian spatial map is created in MSTv with inputs from MT$^-$ cells in log-polar space (Figure 9). Constructing MSTv inputs from a subset of the population of MT$^-$ cells keeps the simulations tractable in terms of memory and run-time. Since the model only produces horizontal steering and smooth pursuit movements, model MSTv samples the visual space via MT more broadly in the horizontal than the vertical dimension. A 256 × 8 array of MSTv cells at each depth is simulated with receptive field centers spanning -45º to 45º horizontally and -8º to 8º vertically.

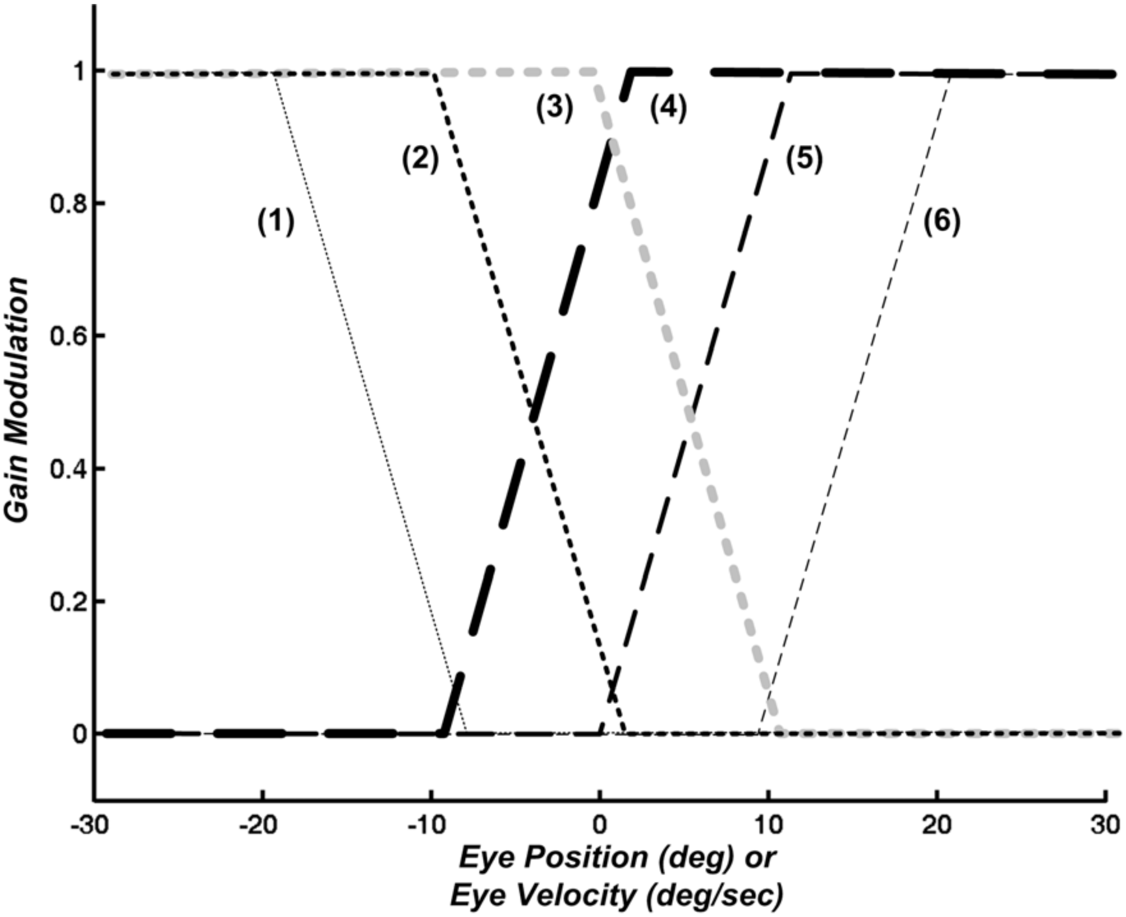


Figure 10. Illustration of the 6 piecewise-linear eye position gain fields. Gain fields (1)-(3) saturate when the eye is in the left half of the orbit, and (4)-(6) saturate when the eye is in the right half of the orbit. These same gain field functions are used for the eye velocity gain fields in Level 7a.

Level 7: Eye position gain fields. Steering commands derive from representations of environmental objects and heading direction in craniotopic coordinates. For simplicity, the model assumes that the head is fixed with respect to the body, and any gaze shifts are the result of eye rotations only. Under this assumption, craniotopic coordinates are aligned with body-centered coordinates. The model converts retinotopic object position and heading direction provided by MSTv and MSTd to craniotopic coordinates using eye position gain fields. The model uses 3 piece-wise linear gain fields for both the left and right eye positions within the orbit. These 6 gain fields are plotted in Figure 10 as a function of eye position. Also see Appendix A, Equations (17) and (18). The retinotopic obstacle representation in near-depth MSTv, goal representation in fixation-depth MSTv, and heading representation in MSTd all project to populations of gain-modulated cells, as shown in Figure 11. Craniotopic representations of obstacle position, goal position, and heading are computed from weighted sums of the outputs of the gain-modulated cell populations; see Appendix A, Equations (17) – (25). The weights themselves are trained using the self-organizing map learning algorithm (Grossberg, 1976; Kohonen, 1982).

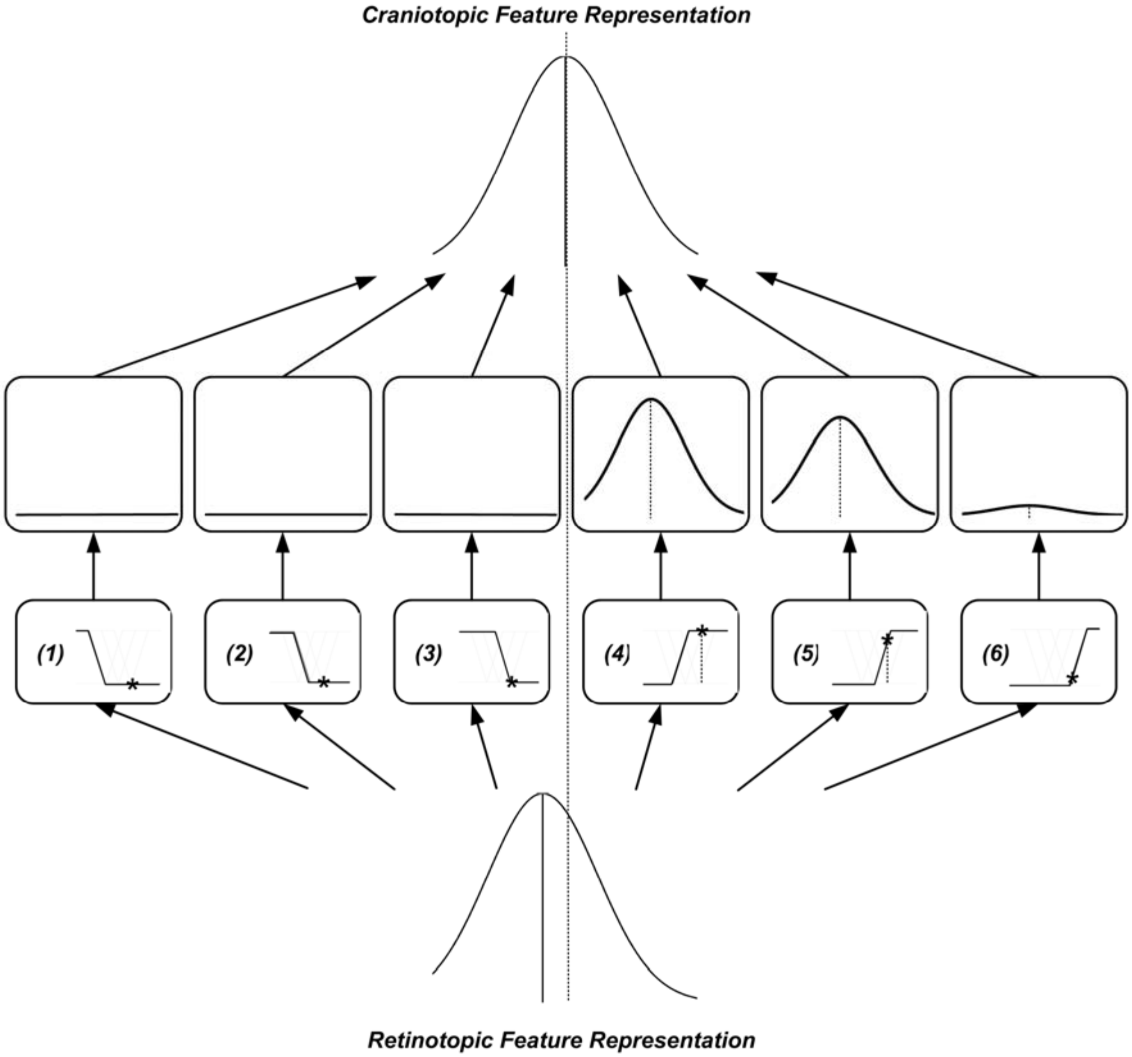


Figure 11. Mapping of retinotopic features to craniotopic coordinates via eye position gain fields. A retinotopic feature, such as a goal and obstacle in MSTv or heading in MSTd, is copied 6 times and each copy is gain-modulated by one of the eye position gain fields. The 6 gain-modulated copies are multiplied by adaptive weights and summed to form a craniotopic representation of the feature in the model PPC.

Levels 8a and 8b: MSTd radial flow cells. The large receptive fields of area MSTd cells receptive fields are selective for different optic flow patterns, including expansion, contraction, circular rotation, spiral motion, and planar flow (Duffy & Wurtz, 1991, 1997). STARS contains two populations of MSTd cells selective for radially expanding flow (Level 8b) and planar flow (Level 9). Model MSTd cells receive weighted inputs from the MT^+ additive cell population (Level 5, but gain-modulated in Level 8a). The radial optic flow cell population in Level 8b of the model encodes the direction of heading from optic flow; see Appendix A, Equations (26) – (32).

To accurately represent the direction of heading from optic flow, MSTd cells must compensate for the effects of eye rotation on the global flow pattern. The bulk of psychophysics research on the matter suggests that extraretinal, oculomotor signals generated during an eye rotation play a key role in compensating for the effects of eye rotation on heading perception. Bradley et al. (1996) found that a significant number of cells in MSTd are gain-modulated in the presence of smooth eye movements, and that some MSTd cells have heading responses that are invariant with respect to eye movements. This study suggests that gain-modulated cells in MSTd comprise an intermediate computational step from which eye rotation invariant responses are formed.

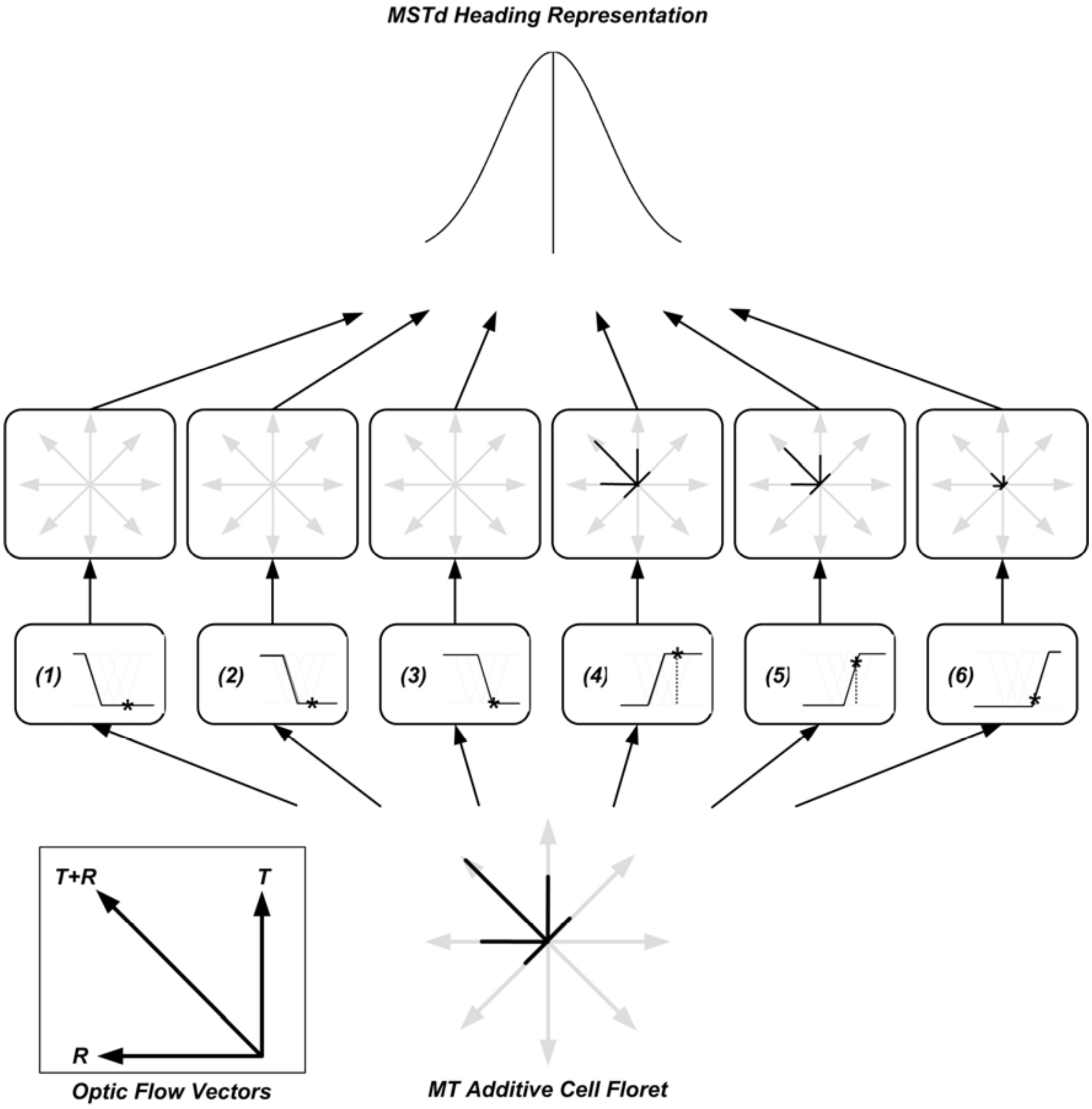


Figure 12. Removal of rotational flow via eye velocity gain fields. The inset figure shows the input optic flow vector *T*+*R* in an MT additive cell floret's receptive field. This flow vector is the sum of a rotational vector *R* generated by an eye rotation to the right and a translational flow vector *T* due to self-motion through the environment.

Recovery of the translational flow vector T is achieved using eye velocity gain fields. The flow vector shown here would be found in the upper half of the visual field near the vertical meridian. In the main figure, each MT additive cell floret is copied 6 times and each copy is gain-modulated by one of the eye velocity gain fields. For clarity, we have only shown here one of the four speeds in the MT additive cell population, but the other three are computed as well in the full model implementation. The rotated and scaled MT additive cell floret activity is shown in the dashed circle, and is computed from a weighted sum of the gain-modulated floret activities. Note that this rotated and scaled MT floret is not explicitly computed in the model. Instead the model computes the MSTd heading signal directly from the gain-modulated copies of the MT additive cell florets.

STARS makes smooth eye rotations to maintain fixation on the goal during navigation, so it can generate extraretinal signals. Eye velocity gain fields (Level 8a) correct for eye rotation in MSTd (Figure 4). Individual MT^+ additive cells project to interneurons in the early layers of MSTd. These interneurons also receive top-down extraretinal inputs from the model's oculomotor system, in Level 11. The extraretinal signals are piecewise-linear functions of eye velocity (Figure 10), and together comprise a set of eye velocity gain fields, similar to those reported by Bradley *et al.* (1996). The radial flow cells in MSTd receive weighted inputs from the gain-modulated interneurons in the early layers of MSTd. The weights are trained using a self-organizing map (SOM), as described in Appendix A, Equations (33) and (34) (Grossberg, 1976; Kohonen, 1982). When the eyes rotate, the gain patterns change and a different subset of interneurons becomes more highly activated. This means that the optic flow patterns generated by heading in the presence of a horizontal eye rotation, and heading along a forward path that has been shifted horizontally in the absence of eye rotation, are encoded by different populations of MT inputs, and can trigger different heading responses in MSTd. Eye velocity gain fields were used in a similar setting in the model proposed by Beintema and van den Berg (1998).

Like the Level 6 MSTv cells, the inputs to MSTd cells are selected from a random set of MT+ additive cell populations. MSTd cells have larger receptive fields, allowing each MSTd cell to respond to optic flow in a wider portion of the visual space.

Level 9: MSTd planar flow cells. The MSTd cells in Level 9 respond to planar background flow during an eye rotation (Figure 4). Because they respond to the component of the flow field caused by eye rotations, they are not subject to the problems encountered by the radial flow cells in Level 7 of the model. These cells have large spatial receptive fields, and they encode the direction and speed of the background flow during an eye rotation; see Appendix A, Equations (35) and (36). The inputs to MSTd planar flow cells are weighted outputs from the MT^+ cell population, without extraretinal gain modulation. These weights are also trained using a self-organizing map.

Level 10: Parietal steering mechanism. The craniotopic representations of goal, obstacle, and heading are combined in model posterior parietal cortex (PPC) to produce a steering signal. The goal and heading representations have excitatory projections to the PPC steering field, but the obstacle representation has an inhibitory projection (Figure 3a); see Appendix A, Equation (37). The three peaked spatial activity maps representing the goal, obstacle, and heading are summed together point-wise to produce activity in the PPC steering field. Two mutually inhibitory steering cells sum the left and right halves of the steering field, as shown in Figure 3b; see Appendix A, Equations (38) – (41). Any imbalance in the steering field causes one of the steering cells to become more activated than the other (cf., Gnadt and Grossberg, 2008). The magnitude of the steering cell's response controls the rate of turning to the left or right. Since the head and body are always aligned with the direction of travel, the craniotopic heading peak is always centered. If a goal object is present in the environment—e.g. to the right of the direction of travel—then the peak in the steering field will be shifted to the right, and the model will steer to the right, approaching the goal. As the model turns toward the goal, the craniotopic heading and goal peaks become closer, causing the model to slow its turn. This is how the model implements the attractor portion of attractor-repeller steering control.

If an obstacle is also present in the environment (see Figure 5), then more complicated interactions occur in the steering field. Because the obstacle projections to the steering field are inhibitory, an obstacle to the left of the direction of heading, for example, shifts the steering peak to the right, away from the obstacle. The goal peak simultaneously pulls the steering peak in the direction of the goal, which prevents the model from continuing to steer away from the obstacle. As the obstacle moves out of the periphery of the retina, the visible goal guides steering. The navigator turns back toward the goal, completing its approach. The excitatory and inhibitory interactions of the goal and obstacle peaks with the heading peak implement attractor-repeller control of steering.

Level 11: Frontal predictive eye movements use complementary MSTv and MSTd signals. STARS computes an eye rotation signal to maintain fixation on the goal object during navigation; see Appendix A, Equation (42) and Level 11 in Figure 4. This mechanism involves parietal areas such as LIP and frontal areas such as the frontal eye fields, including inputs from the fixation-depth object motion cells in MSTv and the background planar flow cells in MSTd. The Level 11 network consists of two target tracking and two background flow cells. The target tracking cells prefer either left or right motion of the pursuit target (i.e., the navigational goal), and sum inputs from the fixation-depth MSTv cells preferring left or right object motion, respectively. The background flow cells encode the direction of motion of the background during a pursuit movement. The left (right) flow cell prefers background motion to the left (right), which occurs during a rightward (leftward) pursuit eye movement. The two background flow cells sum inputs from the planar flow cells in MSTd (Level 9).

In order to achieve predictive tracking, the two target tracking cells inhibit one another, as do the two background flow cells; see Appendix A, Equations (43) – (46). In addition, and essential for accurate prediction, the right-tuned target tracking cell and the left-tuned background flow cell are mutually excitatory, as are the left-tuned target tracking cell and the right-tuned background flow cell. As the smooth pursuit velocity matches the target's velocity, the target becomes stationary on the retina, and leads to no net motion signal at the target tracking cell in MSTv. Smooth pursuit is maintained because the background flow cell is activated in a compensatory way during accurate pursuit and provides excitation from MSTd to the corresponding MSTv target tracking cell as target motion signals approach zero on the retina (Pack, Grossberg, and Mingolla, 2001). The graded speed response of MSTv cells enables the smooth pursuit system to track targets moving at different speeds.

Steering Control Variables. The FW model computes a steering trajectory using third-person information obtained directly from the scene geometry, namely the distance and angles to the goal and obstacle. The FW model describes the motion of a body with inertia. Since inertial bodies cannot make instantaneous changes to angular velocity, the FW model uses a second-order equation which controls angular acceleration, not velocity. STARS steers from a first-person view which cannot directly compute the distance and angle to the goal and obstacle. The peak shift mechanism in the STARS steering control system uses indirect estimates of object distance and angle in a cortical map. In craniotopic coordinates, the peaked activity distribution representing the heading direction is centered, because the model always instantaneously moves in the direction that its "body" is oriented. As noted above, excitatory and inhibitory interactions between the goal and obstacle activity distributions and the heading peak produce a new peak that is shifted to the left or right. As the angle between the goal and the heading direction increases, the steering peak in PPC shifts further to the side. This in turn produces a larger steering movement by the STARS model. The STARS model hereby tracks object angle by adding shifted peaked activity distributions.

Distance is implicitly represented in the STARS model by the overall magnitude of the goal and obstacle activity distributions: the MSTv cell populations have graded speed responses which increase to objects moving at higher velocities. In a static environment, higher velocities often indicate closer objects because velocity scales inversely with distance. Increasing the magnitude of the goal or obstacle representation causes a larger shift in the PPC steering peak. STARS hereby controls steering using indirect representations of the heading direction, the angles between objects, and the distance to objects.

As noted above, to account for inertia, the FW model uses a second-order equation which directly controls angular acceleration, not velocity. The steering output of the STARS model comes from a first-order equation (Appendix A, Equation (41)). Acceleration control comes from the V-shaped weighting function applied prior to determining the angular velocity (Appendix A, Equations (39)-(40)). This function ensures that larger deflections in the combined steering peak produce larger angular acceleration. This also suggests that the STARS model could, in some situations, turn more rapidly than is physically plausible for an inertial body, although this was never observed in the simulations. The output of STARS provides a control signal for a downstream motor control system that is not modeled here, which would be responsible for generating muscular forces and compensating for the inertia of the walking human. For example, the FLETE neural model (Bullock & Grossberg, 1991; Contreras-Vidal, Grossberg, and Bullock, 1997) clarifies how cerebellum and spinal cord interact with movement trajectory commands that are computed through interactions with the neocortex and basal ganglia to ensure that the body accurately follows outflow movement commands.

3. Simulation Results

STARS simulation results show model trajectories in several environments, based on the experiments of Fajen and Warren (2003). We compare these simulated trajectories with the mean paths taken by subjects in Fajen and Warren's experiments. Then we show how the model's internal computations in areas MT, MSTv, MSTd, and PPC collaborate to produce realistic steering behavior. $T = 10(d - 1)$

To evaluate STARS performance during steering, we first compute a baseline trajectory using the FW model. We, then simulate the model in an identical virtual environment for timesteps, where d is the initial distance to the goal in meters. This produces a simulated trajectory consisting of T segments, and we subdivide the baseline trajectory into the same number of segments. We take successive pairs of points along each trajectory, form a quadrilateral, and compute its area. The sum of the areas of all T quadrilaterals gives a measurement of the error. In particular, if the summed area equals zero, the model would perfectly replicate the baseline trajectory. A similar error measure was used by Patla, Tomescu, and Ishac (2004). In addition, we compute the maximum linear distance between the trajectory steered by the model and the baseline, which gives a worst-case measure of how far the model deviates from the correct path.

We fit data produced by the FW model, not the raw data from subjects in Fajen and Warren's study, to which we did not have access. However, the FW model fits the data with high accuracy. The FW model fits the goal angle data with $r^2 = .979$ (see Figure 1a), the goal distance data with $r^2 = .982$ (see Figure 1b), the obstacle avoidance data with $r^2 = .975$ (see Figures 1c and 1d).

Goal Approach. Fajen and Warren (2003) studied how initial angle and distance influence trajectories steered toward a goal object. The first experiment examined goal angle. Subjects walked toward a goal in a virtual environment which appeared at a distance of 4 meters and an angle of 5º, 10º, 15º, 20º, or 25º to the left or right of the initial direction of travel. Subjects turned more rapidly as the initial angle increased, and followed nearly linear paths to the goal after completing the turn. The second part of the experiment varied goal distance. Here, the goal appeared 20º to the left or right of the initial direction of travel and at a distance of 2, 4, or 8 meters. Subjects turned at a higher rate for shorter goal distances.

The trajectories steered by the model in the goal direction simulation are shown in Figure 1a, and the trajectories steered in the goal distance simulation are shown in Figure 1b. The simulations replicated the finding that the turning rate increases with goal angle and decreases with goal distance. The error areas and maximum linear deviations are summarized in Table 1. In the worst case area error, STARS deviated from the baseline trajectory by a maximum of 10.5 cm, much less than the width of an average human body. In this simulation, the model completed its turn toward the goal slightly later than the baseline trajectory did, and thus followed a path to the goal which paralleled the baseline.

Goal angle	Goal distance	Area error	Maximum deviation
5º	4 m	0.0846 m^2	9.16 cm
10º	4 m	0.0760 m^2	9.29 cm
15º	4 m	0.0375 m^2	8.62 cm
20º	4 m	0.0382 m^2	8.00 cm
25º	4 m	0.1632 m^2	10.50 cm
20º	2 m	0.1064 m^2	16.00 cm
20º	4 m	0.0382 m^2	8.00 cm
20º	8 m	0.0401 m^2	8.40 cm

Table 1. Summary of errors for goal approach simulations. See text for details.

Obstacle Avoidance. In the Fajen and Warren (2003) study of how obstacle angle and distance influence steering behavior, subjects walked toward a visible goal object initially located straight ahead at a distance of 9 meters. In the first part of the experiment, single obstacles were placed at an initial distance of 4 meters and an angle of 8º, 4º, 2º, or 1º. Subjects first turned away from the obstacle, and then turned back toward the goal just before passing the obstacle (Figure 1c). Obstacles with smaller initial angles induced faster initial turns, and thus deflected the trajectory more than obstacles with larger angles. In the second part of the experiment, obstacles were placed at a fixed initial angle of 4º and at distances of 3, 4, or 5 meters (Figure 1d). Obstacles with smaller distances induced earlier and larger turns than did more distance obstacles. These results suggest that the repulsive effect of an obstacle is a decreasing function of distance and angle. In other words, obstacles which are nearby and impede on the current trajectory cause larger deflections in steered trajectories.

Obstacle angle	Obstacle distance	Area error	Maximum deviation
-8°	4 m	0.179 m^2	4.90 cm
-4°	4 m	0.137 m^2	3.72 cm
-2°	4 m	0.097 m^2	2.52 cm
-1°	4 m	0.148 m^2	3.70 cm
-4°	3 m	0.093 m^2	2.53 cm
-4°	4 m	0.137 m^2	3.72 cm
-4°	5 m	0.097 m^2	3.45 cm

Table 2. Summary of errors for obstacle avoidance simulations. See text for details.

The trajectories steered by STARS in the obstacle direction and distance simulations are shown in Figures 1c and 1d. As noted above, STARS captured the observed obstacle avoidance behavior, making earlier and larger detours around obstacles with smaller initial angles or shorter initial distances. The simulations are compared with baseline trajectories in Table 2. The worst trajectory was in the 8° conditionbut still had a maximum deviation of less than 5 cm over the entire run.

Route Selection. A third set of Fajen and Warren (2003) experiments examined the conditions producing qualitative changes in the steered trajectory. In the previous obstacle avoidance experiments, the goal was directly in front of the observer's starting point and the obstacle to the left or right of the initial path to the goal. The path to the goal never crossed in front of the obstacle. However, if the initial obstacle angle was between the initial heading direction and goal angle, subjects could take an outside path, passing to the left of the obstacle, or an inside path, passing to the right of the obstacle (as in Figure 13). Taking an inside route to the goal requires the heading direction to cross in front of the obstacle.

Fajen and Warren (2003) set up environments where the goal was located 15° from the initial heading direction and a distance of 5, 7, or 9 meters (Figures 13a-13d). An obstacle was placed at 4 meters and an angle of 7°, 11°, 13°, or 14°. This gave initial goal-obstacle offset angles of 8°, 4°, 2°, and 1°. Fajen and Warren found that subjects took inside and outside trajectories in all environmental configurations, but tended to prefer inside paths for the 4° and 8° offset angle conditions and outside paths for the 1° condition. In the 2° condition, subjects preferred inside paths when the goal was 5 meters away and outside paths when the goal was 7 or 9 meters away. The FW model produced trajectories which switched from inside to outside under the same conditions as the experimental subjects. A parameter was changed to decrease the repulsion of the obstacle as a function of distance.

STARS reproduced this route selection behavior, also requiring a parameter change (see Appendix A, ***Route Selection Parameters***). STARS simulations are shown in Figure 9a-9d and 9f. In some model paths, the obstacle occluded the fixated goal. The model target tracking system (Level 11) cannot track an occluded object, so eye position and rotation were manually controlled in these simulations. Because there was some variability in subjects' route selection choices, STARS was simulated on each of the 12 environments 100 times, each time perturbing the obstacle angle and distance by adding random noise drawn from a uniform distribution with size equal to 12.5% of the initial obstacle angle and distance. The number of inside trajectories was computed for each environment, and the percentage of inside trajectories is shown in Figure 9f. Fajen and Warren (2003) used a similar procedure to evaluate the performance of their model.

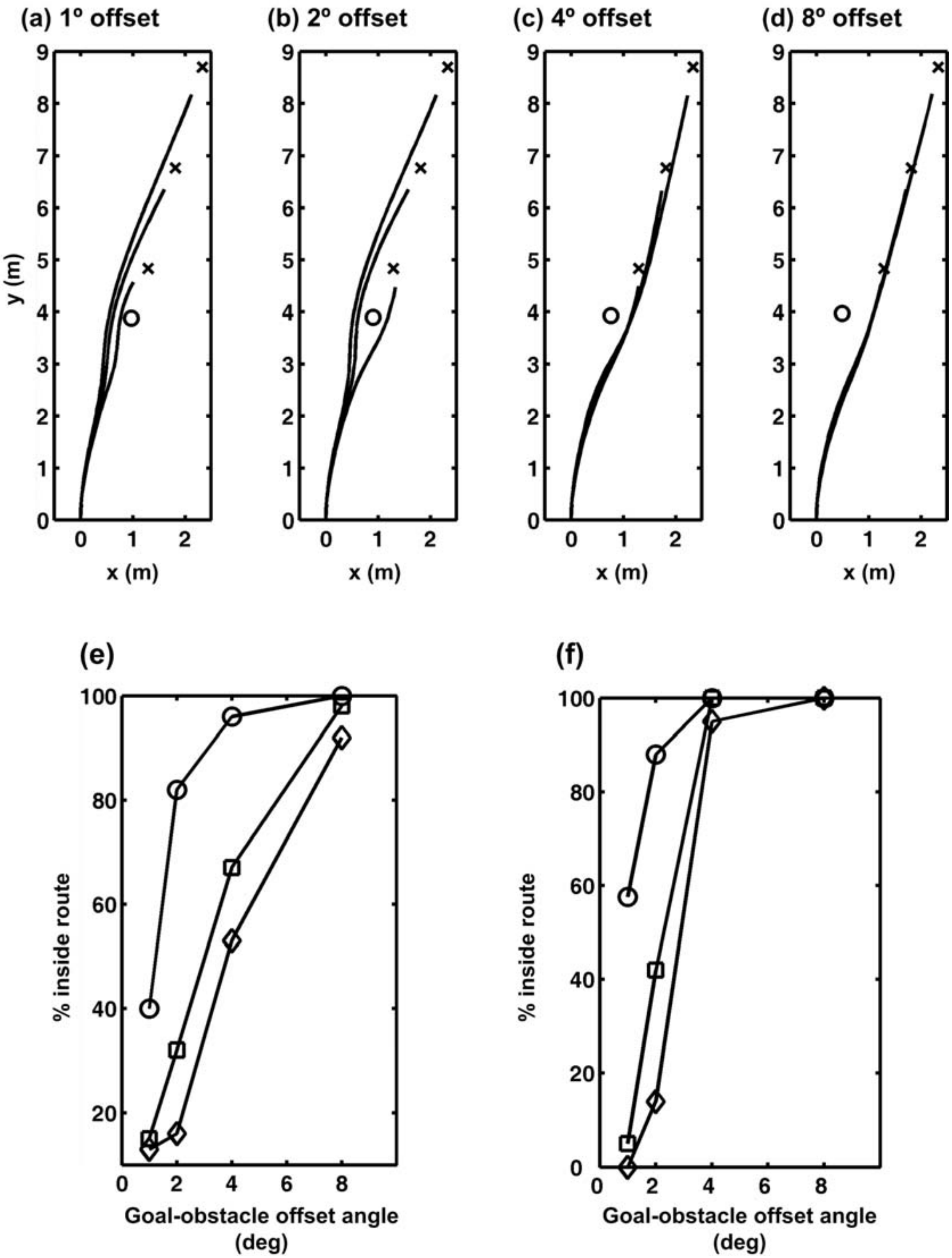


Figure 13. (a) - (d) show paths steered by the STARS model in the 1º, 2º, 4º, and 8º conditions. In parts (e) and (f), circle denotes the 5-meter goal distance, square denotes the 7-meter goal distance, and diamond denotes the 9-meter goal distance. (e) shows the percentage of inside routes taken by the FW model as a function of goal-obstacle offset angle, and (f) shows the percentage of inside route taken by STARS. The percentages in (e) and (f) were obtained by introducing 10% noise in the FW simulations and 12.5% noise in the STARS simulations and simulating each model 100 times on each environment. See text for details.

Heading vs. Perceived Goal Direction. Although heading can be perceived from optic flow, some researchers have argued that perceived heading might not play an important role in steering control (Rushton, Harris, Lloyd, & Wann; 1998; Wann & Land, 2000). The study performed by Rushton and colleagues (1998) asked human subjects to walk toward a visible goal object while wearing prism goggles, which deflected the entire scene by 16º. They found that subjects walked along curved trajectories as if they were guided primarily by the perceived position of the goal object, not the optic flow field. The experimenters suggested that humans rely primarily on perceived goal location, not heading from optic flow, when steering.

A subsequent study by Warren et al. replicated parts of the earlier experiment using a virtual reality environment (Warren, Kay, Zosh, Duchon, & Sahuc, 2001). Their subjects walked toward a goal object viewed through a head-mounted VR display. Because the subjects viewed a 3D virtual environment, the experimenters could vary the amount of optic flow available to subjects as well as simulate optic flow inconsistent with the direction of travel through the (real) environment. The virtual environments were designed so that the optic flow field was shifted 10º to the left or right of the true direction of travel. The experimental conditions were designed so that if a subject pursued an egocentric direction strategy, they would walk in a curved path, but if they followed an optic flow-based strategy, they would walk in a straight path to the target. Warren and colleagues found that subjects pursue an egocentric direction strategy when little flow is available, but follow an optic flow-based strategy when the scene generates dense flow. Subjects tend to “crab-walk” along paths at an angle relative to the orientation of the body. These results suggest that humans tend to use optic flow information when it is available, but can rely on the egocentric direction of the goal when it is not.

STARS replicated the Warren et al. (2001) data; see Figure 10. Since the model cannot “crab-walk” like human subjects, we shifted the direction the model’s viewpoint moves by 10º, which is equivalent to shifting visual space relative to the locomotor axis with a prism. We produced an optic flow field inconsistent with the veridical direction of travel by computing the flow field as if the model’s viewpoint had moved straight ahead, rather than along the shifted trajectory. This produced an optic flow field shifted -10º from the true direction of travel. We varied the amount of optic flow available to the model by removing pieces of the scene geometry, such as the floor, walls, and ceiling.

The model explains these data as follows. Shifting the optic flow field away from the veridical direction of travel causes a shift in the craniotopic heading peak away from the midline. This shift introduces a bias in the direction the model tends to steer by unbalancing the steering peak in Level 10. Because the locomotor axis is shifted relative to the visual space, each attempt by the model to move forward actually takes the model along a slanted trajectory to the right of the goal. STARS compensates for the relative shift of the locomotor axis by turning to the left. The steering bias created by the shifted optic flow field helps the model compensate and reduces steering error. Decreasing the amount of flow in the scene (Figure 14a) causes a decrease in the magnitude of the retinotopic and craniotopic steering peaks, which reduces the amount of steering compensation and increases steering error. The simulation results in Figure 14 are similar to the data obtained by Warren et al. (2001).

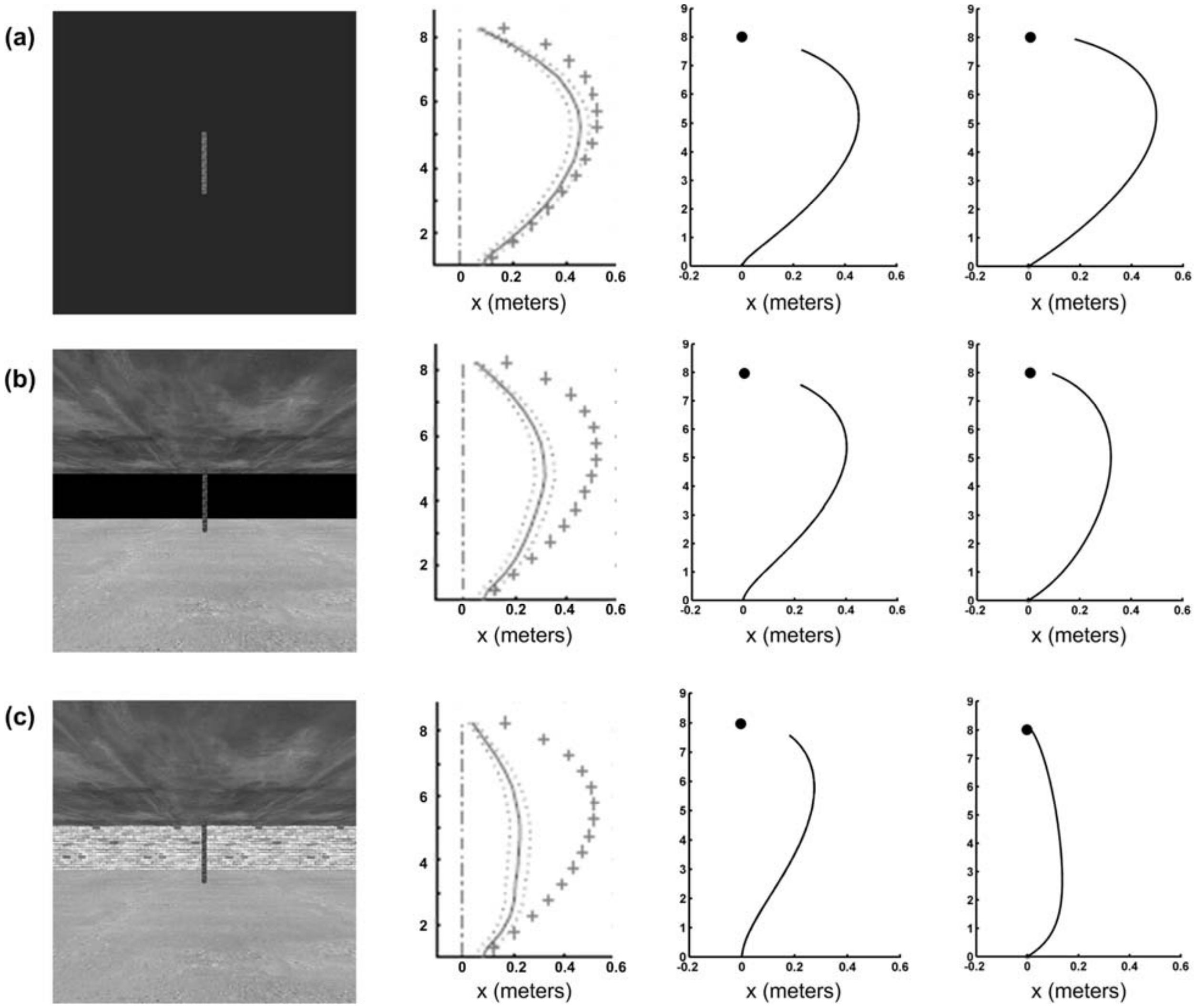


Figure 14. Simulation of Warren et al. (2001) data. (a) shows the scene and trajectory steered for the low optic flow condition. (b) shows the scene and trajectory steered for the medium optic flow condition. (c) shows the scene and trajectory steered for the full optic flow condition. The path straightens as optic flow increases because optic flow helps overcome the misperceived location of the goal. The leftmost column shows a screenshot of the environment at the start of the simulation. The second column shows the experimental data obtained by Warren et al. (2001), and the third column shows results generated using the model proposed in that study. The rightmost column shows simulation results obtained by STARS. See text for details. Warren et al. (2001) data reprinted with permission.

Local Motion Processing – MT. The accurate steering behavior reported above is the product of several interacting systems in the model. The following sections show how the model computations during a typical simulation, and how the final steering behavior arises from the model's representations of the heading direction and goal and obstacle positions. Figure 15a shows the position of the model at time step 20 along the trajectory steered in the 4-meter, 4° obstacle avoidance simulation. This is just before the model passes by the obstacle and begins turning back toward the goal. Figure 15b shows a screenshot of the environment from the viewpoint at time step 20. The following sections describe the internal model state at this timestep.

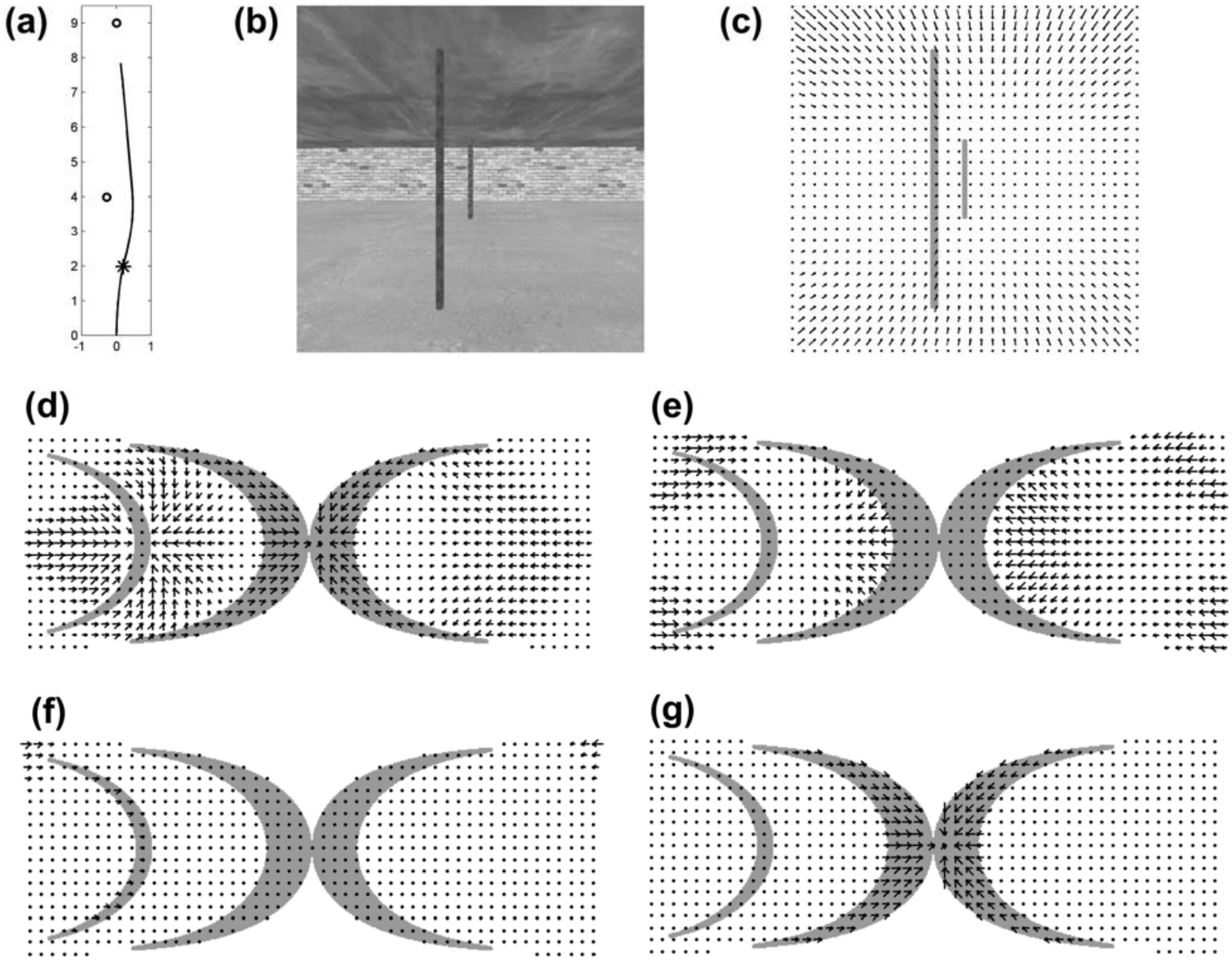


Figure 152. Optic flow and area MT simulation details. (a) shows the trajectory around an obstacle at 4 meters, -4º. The asterisk shows the position of the model at time step 20, after traveling 2 meters. The remaining parts of this Figure and Figure (16) refer to the model's internal state in Levels 4, 5, 6, 8b, and 10 at this time step. (b) shows the scene from the viewpoint of the model at time step 20. (c) shows the optic flow field in Cartesian coordinates. The gray bars highlight the position of the obstacle (left) and goal (center) in the flow field. (d) shows the MT^+ additive cell population tuned to speed 1 (slowest) at time step 20. (e) shows the additive cell population tuned to speed 2. The populations tuned to speeds 3 and 4 are not shown. (f) shows the near-depth MT^- subtractive cell population tuned to speed 3. The most highly active cells are responding to the obstacle in the flow field. Speeds 1, 2, and 4 are not shown. (g) shows the fixation-depth MT^- subtractive cell population tuned to speed 2. The most highly active cells are responding to the goal in the flow field. Speeds 1, 3, and 4 are not shown. (f) and (g) show the speed-tuned population with the highest overall activity level to best illustrate the segmentation of the obstacle and goal.

The input to the model as it moves through the environment is the optic flow field, shown in Cartesian coordinates in Figure 15c. The additive MT^+ cells produce a distributed encoding of the optic flow field, suitable for further processing in MSTd. At each location (i, j) in MT+, we simulate a cluster of 32 individual cells, each tuned to one of eight directions and four speeds. A single spatial receptive field is shared by each cell in a particular cluster, and covers a symmetric 31 × 31 block of optic flow vectors in the input log-polar flow field. A representative subset of the MT^+ responses to the optic flow field at time step 20 are shown in Figures 15d and 15e.

The subtractive MT^- cells detect discontinuities in the optic flow field caused by objects separated in depth from the background. Two populations of cells are simulated, one selective for motion at near depth and one selective for motion at fixation depth. At each position (i, j) in each depth-selective population, we simulate 32 individual cells, tuned to one of eight preferred directions and four speeds. The cell clusters share two common receptive fields defining the response profiles to motion in the center and surround regions. Figure 15f shows the activity of a subset of near-depth cells, which respond primarily to the obstacle, and Figure 15g shows the activity of a subset of fixation-depth cells, which respond primarily to the goal object.

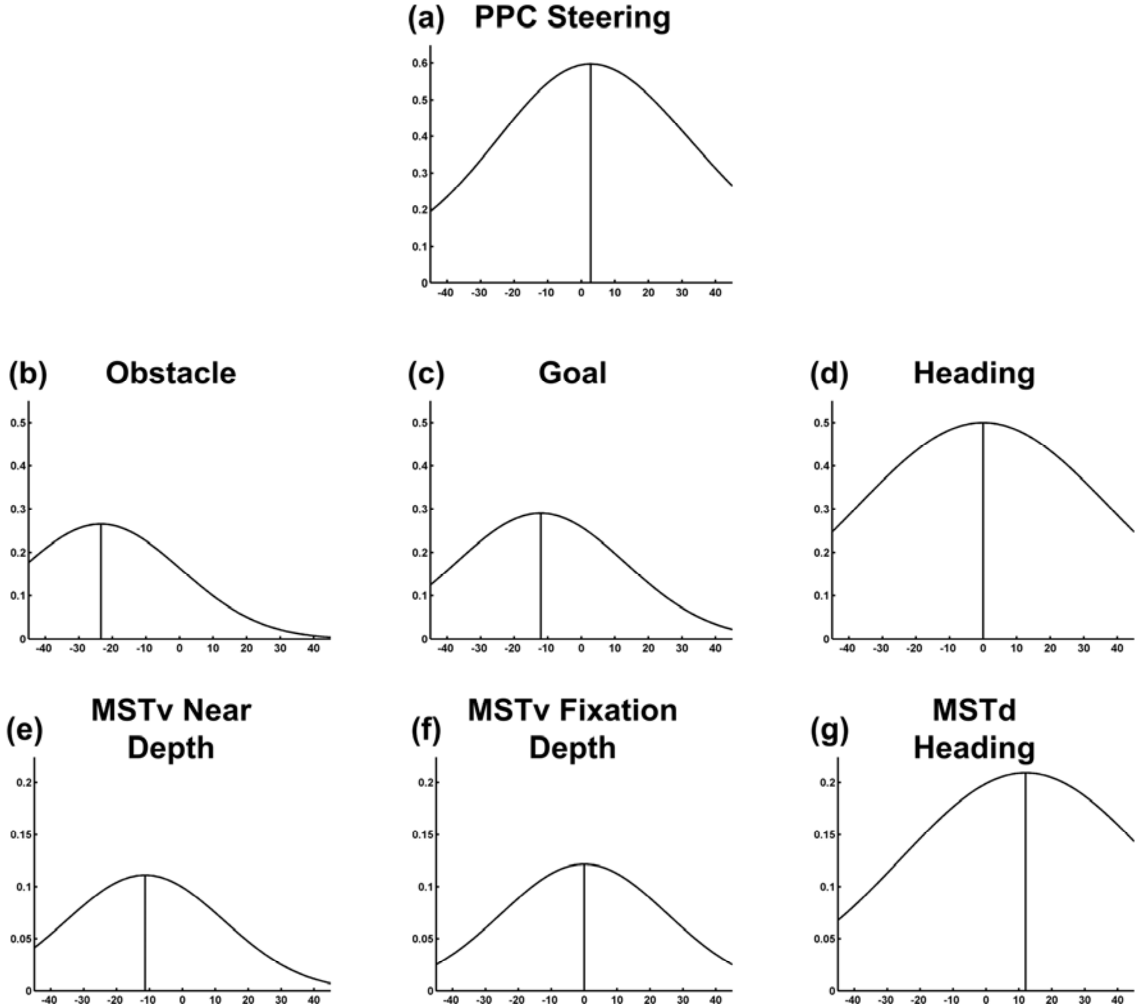


Figure 16. MSTv, MSTd, and PPC (Levels 6, 8b, & 10) simulation details. (a) shows profile of PPC steering peak at time step 20. This peak is the pointwise sum of the obstacle peak (b), goal peak (c), and heading peak (d). Note that the obstacle peak (b) is

subtracted from, not added to, the other peaks. The obstacle, goal, and heading representations in parts (b), (c), and (d) are in craniotopic coordinates, after applying gain fields. (e), (f), and (g) show the obstacle, goal, and heading direction in MSTv and MSTd, prior to their transformation into craniotopic coordinates via gain fields.

Goal, Obstacle, and Heading Processing. The MSTv near-depth and fixation-depth cell populations encode the location of the obstacle and goal in retinal coordinates, and the MSTd expansion flow cell population encodes the retinal direction of heading, after correcting for eye rotation. All three MST cell population responses are peaked activity distributions, with the location of the maximally active cell in the map encoding the position of the obstacle, goal, or heading direction. The population responses at time step 20 are shown in Figure 16. The outputs of the MST populations are gain-modulated and transformed to craniotopic coordinates. The results of gain field computation are shown in Figure 16. The craniotopic encodings of the goal, obstacle, and heading direction are then added point-wise to produce a steering peak in PPC. The location of the maximally active cell in this population encodes the direction of the steering movement. This is shown in Figure 16a. A shift in the PPC steering peak to the left causes the model to steer to the left, and a shift to the right causes the model to steer to the right. At time step 20, the model is steering to the right away from the obstacle, so the PPC steering peak in Figure 16 is shifted to the right.

4. Discussion

The STARS model provides an explanation for how motion processing mechanisms in cortical areas MT and MST can realize dynamic attractor-repeller control of steering. The local motion processing mechanisms in model area MT give rise to cortical representations of object positions in MSTv and heading in MSTd. The retinotopic position of objects and direction of heading are converted into craniotopic coordinates using eye position gain fields. A peak-shifting mechanism combines the craniotopic encodings of object positions and heading into a steering signal for the control of locomotion. The magnitude and direction of the peak shift signals the magnitude and direction that the model should steer. The model explains several challenging types of data regarding human steering performance in cluttered environments, including goal approach, obstacle avoidance, and the relative importance of heading and the perceived direction of the goal.

Model Design Considerations. The design and implementation of the STARS model were informed by a variety of factors, including psychophysical and neurophysiological data, as well as the need to maintain the focus of the model. We highlight some of the more important design considerations here. We have assumed that the head is immobile on the neck, and thus, that craniotopic coordinates are aligned with body-centered coordinates. Thus, under normal locomotor circumstances, the instantaneous representation of heading is centered. Since the heading peak is normally centered in craniotopic coordinates, one might wonder why the model expends considerable effort to extract heading from optic flow. Warren *et al.* (2001) demonstrated that optic flow is used to control steering in situations where the veridical direction of travel and the implied direction of travel from optic flow are misaligned. Although we have not simulated a motor control system, motor planning and control occurs in a body-centered coordinate frame, and steering computations are most natural in that frame. All the information needed to compute steering is available in retinal coordinates, so in principle a system could be designed to compare the positions of the goal and obstacle with the heading direction in retinal coordinates. A neural mechanism which could compare object position and heading direction at arbitrary locations in retinal space would be complex and is not known to exist physiologically. The steering system in the STARS model is simpler, relying only on the summation of peaked activity maps in craniotopic coordinates.

	Connection, cell type, or cell property in model	Functional interpretation	Selected references
1	Log-polar mapping in V1	Foveal expansion, peripheral compression	Daniel & Whitteridge (1961), Schwartz (1977, 1980)
2	Disparity-selective cells in V2	Partition scene into near, fixation, and far depth regions	Poggio & Fischer (1997), von der Heydt et al. (2002), Peterhans (1997)
3	V2 → MT pathway	Supply MT subtractive cells with disparity information	Maunsell & van Essen (1983a)
4	MT center-surround cells	Segment moving objects from background	Allman et al. (1985), Bradley & Andersen (1998), Born (2000), Born & Bradley (2005)
5	MT additive cells	Estimate local motion	Albright (1984), Allman et al. (1985), Born & Tootell (1992)
6	MT direction tuning	Cover range of motion directions	Albright (1984)
7	MT speed tuning	Logarithmic coverage of motion speeds	Nover et al. (2005)
8	MT center-surround cells → MSTv object motion cells	Supports object position and velocity encoding in MSTv	Berezovskii & Born (2000)
9	MT additive cells → MSTd optic flow cells	Supports local motion integration in MSTd	Berezovskii & Born (2000)
10	MSTv object motion cells	Encode direction and position of moving objects	Tanaka et al. (1993)
11	MSTv disparity selectivity	Encode objects (goal, obstacle) in different depth regions	Eifuku & Wurtz (1998)
12	MSTd radial flow cells	Heading from optic flow	Britten & van Wezel (1998, 2002), Duffy & Wurtz (1997)
13	MSTd planar flow cells	Background flow from eye rotation	Komatsu & Wurtz (1998), Duffy & Wurtz (1991)
14	Eye position gain fields	Integrate eye position signals for coordinate transformation	Andersen et al. (1985, 1987), Squatrito & Maioli (1997)
15	Eye velocity gain fields	Correct for eye rotation in MSTd heading response	Bradley et al. (1996), Shenoy et al. (2002)
16	VIP heading cells in craniotopic coordinates	Heading representation in craniotopic coordinates	Colby et al. (1993)
17	VIP object motion cells	Object motion in craniotopic coordinates	Zhang et al. (2004)
18	Smooth pursuit cells in FEF	Generate smooth pursuit eye movements	Gottlieb et al. (1993), Tian & Lynch (1996)

Table 3. Summary of anatomical and physiological evidence for the STARS model.

The STARS model is not primarily concerned with gaze behavior during navigation. We have made the assumption that the model fixates the goal throughout an approach trajectory. This assumption allows the model to use a simple smooth pursuit control system to track the target. Additionally, it provides a simple way to distinguish goal from obstacle by designating all objects in the near field closer than the fixation plane of the goal to be processed as obstacles. This avoids the complex issue of how the brain can maintain a stored representation of an intermittently fixated goal during gaze shifts to other objects, such as obstacles. It is also relevant to note that Wilkie & Wann (2002, 2003a, 2003b, 2005) have noted that fixating the goal, or a point on the desired future path, is a desirable control strategy that is employed in free gaze situations. Object recognition and the spatial allocation of attention in the presence of the full repertoire of eye movements are important and interesting topics in their own right, but are outside the scope of the present modeling work.

Neurological Bases of Steering and Obstacle Avoidance. To completely describe the cortical network for steering and obstacle avoidance, one would need to record from cells in the brain of a monkey that could freely navigate a controlled experimental environment. This is currently beyond the capabilities of experimental neuroscience, but several cortical areas have been identified which support computations useful for guiding locomotion. Table 3 summarizes anatomical and neurophysiological data that are consistent with STARS hypotheses. Area MSTd contains cells with large receptive fields, suitable for processing optic flow (Raiguel, Van Hulle, Xaio, Marcar, Lagae, & Orban, 1997). Tanaka and Saito (1989) identified cells in MSTd selective for optic flow expansion, contraction, rotation, and circular motion. Expansion and contraction, in particular, provide cues for determining the direction of forward or reverse motion. However, some studies have reported MSTd cells that are positionally invariant, meaning that they retain their selectivity for a particular optic flow stimulus regardless of stimulus location within the receptive field (Graziano, Andersen, & Snowden, 1994). If positional invariance held for all MSTd cells, they would not be ideal candidates for encoding the direction of heading because they cannot specify precisely where the focus of expansion is located. However, Duffy & Wurtz (1997) found that most MSTd cells responded to an optic flow pattern with a particular focus of expansion (Duffy & Wurtz, 1997). Stimuli used in this experiment were expansion flow fields with a planar flow field added point-wise. The addition of the planar flow field shifts the focus of expansion away from the direction of the planar flow. MSTd cells identified in this study have response properties consistent with signaling the direction of self-motion. Britten and van Wezel (1998, 2002) used microstimulation to determine whether the responses of MSTd cells actually influence monkeys' judgment of heading. Microstimulation biased heading judgments in the direction of the heading that best activated MSTd cells at the electrode site. The bulk of the research literature supports the notion that MSTd is specialized for processing optic flow and that it participates in extracting heading from flow. STARS mechanistically explicates how optic flow features, in particular heading, may be processed by cells in MSTd.

A second parietal area which may play a role in visual navigation is the area VIP. This area receives visual inputs from a number of areas, including V2, MT, and MSTd (Boussaoud, Ungerleider, & Desimone, 1990; Lewis & Van Essen, 2000), as well as polysensory inputs from other sources. The functional role of VIP in the cortical hierarchy has not been fully characterized, but the available evidence suggests that VIP may play a role in building a head-centered representation of space (Duhamel, Bremmer, BenHamed, & Graf, 1997), or help to coordinate hand-mouth movements (Colby, Duhamel, & Goldberg, 1993). The mechanism by which VIP builds a representation of head-centered space is not known, but could involve transformation of retinotopic signals via eye position gain fields, as described in this article. An intriguing recent study suggests that VIP encodes heading in head-centered coordinates (Zhang, Heuer, & Britten, 2004). This study found that VIP neurons have a stable, selective response for a particular expansion flow field regardless of the position of the eyes or the direction of eye movements. The study by Colby et al. (1993) identified cells in VIP that respond to the direction and speed of small moving stimuli, similar to cells in MT. It is not known whether these putative object motion cells also respond to optic flow, or whether the object motion and optic flow cells represent two distinct classes of cells in VIP. Related modeling work that builds on the current results clarifies how object motion cells and optic flow cells may interact to determine steering behavior in response to complex scenes (Browning, Grossberg, & Mingolla, 2008a, 2008b).

Smooth Pursuit Compensation Using Gain Fields: Use It, Then Lose It. Smooth pursuit eye movements cause distortions to the optic flow field which must be corrected for accurate heading perception (Warren & Hannon, 1990). An optic flow field caused by the combination of forward translation and an eye rotation has a shifted focus of expansion and is structurally similar to a flow field caused by translation along an oblique axis. Without information about ongoing eye movements, a computational system attempting to classify optic flow fields could incorrectly report a shifted heading direction when faced with an optic flow field containing an eye rotation. Considerable psychophysical evidence suggests that humans can accurately perceive heading direction during an active, but not simulated, eye rotation (Royden, Banks, & Crowell, 1992; van den Berg, 1993). Information about the depth structure of the scene may also be used (Crowell & Andersen, 2001).

The interactions between visual and oculomotor information are the subject of active research. Several studies suggest that compensation for eye rotation begins as early as area MSTd. Bradley and colleagues studied the role of MSTd in heading perception during eye rotation (Bradley et al., 1996). They found retinotopic and craniotopic heading responses in different MSTd cells, as well as intermediate types which have retinotopic heading responses, but which are gain-modulated during smooth pursuit. This gain-modulated cell type is hypothesized to represent an intermediate computational stage in which extraretinal information is nonlinearly mixed with visual information prior to pursuit compensation. A recent study by Shenoy, Crowell, and Andersen (2002) examined compensation across multiple pursuit speeds. They found that MSTd cells compensate for pursuit more at faster eye velocities, and seem to rely on extraretinal signals to facilitate compensation.

As in the data, STARS likewise compensates for smooth pursuit within MSTd. MSTd cells are driven by bottom-up visual inputs from MT^+. A model MT^+ cell cluster provides a distributed, map-like encoding of a motion vector. During smooth pursuit, the motion vector is shifted by the addition of rotational flow to the translational flow field generated by self-motion. This causes a shift in the activity pattern within a local MT cell cluster. This shift at a local level is removed by gain-modulating the MT cell activities during smooth pursuit. Introducing an eye-velocity-dependent gain signal adds an extra dimension to the signal arriving at MSTd heading cells, which allows for optically similar input flow fields to be encoded differently, depending on pursuit eye movements. The gain-modulated local motion signals arriving at MSTd heading cells are weighted, and the weights are self-organized as described in Appendix B.

This eye-rotation compensation mechanism is in the early layers of the MSTd radial flow cell population, rather than earlier in MT, because the MSTd planar flow cell population requires rotational flow to be encoded in the inputs from MT. The smooth pursuit mechanism relies on measurements of background flow from the MSTd planar flow cells, while the steering control mechanism requires a representation of heading that is not corrupted by rotational flow. The model thus needs to use rotational flow for eye movements, then lose it before computing heading, a principle we call *Use It Then Lose It.*

Gain Fields and Coordinate Transformation. Stably representing the external visual environment in the presence of gaze movements requires a coordinate transformation. Somehow the visual system incorporates information from the oculomotor system to construct a representation which is stable in the face of eye movements. The entire scenic representation need not update with every eye movement, but behaviorally relevant items, such as stationary saccade targets, should.

The exact mechanism by which the brain implements gain field coordinate transformation is still a topic of active research. Several studies have identified cells in the visual system whose visual receptive fields are fixed in retinotopic coordinates but are gain-modulated during eye movement. Such eye position gain fields have been described in area MSTd (Squatrito & Maioli, 1997) and area 7a (Andersen et al., 1985, 1987), and are believed to play an important role in coordinate transformation (Salinas & Thier, 2000).

Steering Control and Heading Perception. Fajen and Warren (2003) claim that steering through static environments can be explained by a completely reactive control model. The FW model steers through the environment using only on the instantaneous heading and goal and obstacle positions. STARS also steers without any active locomotion planning mechanism. Simply moving through the environment generates an optic flow field from which goal, obstacle, and heading information can be extracted.

However, there may be situations in which reactive steering control fails to account for human behavior. Fajen and Warren (2004) studied human steering dynamics while approaching a moving goal, and identified two possible control strategies. Using a *pursuit strategy*, a person would move toward the goal and attempt to reduce the angle between the goal and the current heading direction to zero. This strategy would yield a curved path which eventually arrives at the goal. On the other hand, an *interception strategy* would attempt to maintain the goal object at a fixed, constant angle relative to the heading angle. This strategy would yield a linear path intercepting the moving goal. The study found that humans tend to follow an interception strategy, except in the earliest part of the trajectory. Without modification, both the FW model and the STARS model would follow a pursuit strategy exclusively. Fajen and Warren have recently modified their model to intercept moving targets (Fajen & Warren, 2007). Switching to an interception strategy would require STARS to identify that the goal is moving independently in the environment. Besides the initial acquisition of the moving goal, both pursuit and interception are reactive control strategies and do not require explicit anticipation of the goal's future position.

Patla, Tomescu, and Ishac (2004) questioned the idea that steering control is purely reactive, even in a static environment. They had subjects walk through an arrangement of 12 randomly placed traffic cones toward a goal object. They compared the paths actually taken by subjects to the paths predicted by five different static path planning models and Fajen and Warren's dynamical model. The best-performing model plans paths statically by identifying and avoiding dense clusters of obstacles. Based on this model's performance, Patla and colleagues suggested that human locomotion uses static path planning. However, the researchers unnecessarily handicapped Fajen and Warren's model. The FW model can traverse environments containing multiple obstacles by summing the repulsive effects of individual obstacles, but Patla and colleagues only supplied the model with information about the obstacle closest to the path at any time. If the dynamical model had complete information about the entire obstacle array, it would have also avoided clusters of obstacles, as several nearby obstacles repel heading more than a single obstacle would.

Gaze and Steering. As mentioned above, the STARS model does not attempt to account for gaze behavior and the deployment of attention during steering. Control of gaze has important implications for steering because moving the eyes can significantly affect the quality of retinal flow information and because gaze angle provides a strong cue for the position of the fixated object relative to the direction of travel. A number of studies have shed light on aspects of gaze behavior during steering. Wilkie and Wann (2003a) studied the effects of fixed gaze versus free gaze on steering behavior. They found that tracking the center of a simulated roadway led to smaller steering errors than when gaze was artificially constrained to track a fixation cross. In the free gaze condition, subjects spent the majority of their time fixating the center of the roadway, allowing steering to be controlled by matching the curvature of the flow lines with the edges of the road. Wilkie and Wann (2002, 2003b, 2005) showed that steering control is governed by the interaction between retinal and extraretinal sources of information, specifically retinal flow, visual direction to the goal, and gaze direction. By perturbing each source of information and noting the degradation in steering performance, they conclude that steering relies on a weighted combination of the variables, and that the weighting can change depending on conditions such as lighting. Wilkie and Wann (2005) also found evidence that vestibular signals make a minimal contribution to steering control. The additive model proposed by Wilkie and Wann is consistent with some elements of the STARS model, although the STARS model does not differentiate between gaze direction and visual direction to the target.

Comparison with Other Models. STARS is the only neural network model capable of explaining the data of Fajen and Warren (2003) at the present time. However other models in the literature address heading perception from optic flow, coordinate transformation using gain fields, compensation for eye rotation, and steering control. Several of these alternative approaches are now compared STARS mechanisms.

Heeger and Jepson (1992) produced a mathematical analysis of the recovery of the components of optic flow, namely translation, rotation, and depth, from a set of discrete samples. Their subspace algorithm defines a least-squares optimization problem that solves for the components of flow field, without the use of explicit extraretinal information. In constrast, psychophysical studies have shown that extraretinal signals are needed for accurate heading perception during sufficiently fast eye rotation. The subspace algorithm solves for the global minimum in an error function. In contrast, biologically plausible synaptic learning laws only make use of local information about cell activities. The Lappe and Rauschecker (1993) model constructed cosine-tuned motion detectors in MT that produce a heading response in MSTd through precomputed weights. The weights were computed from Heeger and Jepson’s subspace solution. Although the Lappe and Rauschecker model can produce accurate heading responses in the absence of eye rotation, it requires hardwired weights and does not propose a method by which the weights could be trained. Lappe (1998) extended this model by incorporating extraretinal signals to aid in the removal of rotational flow. The Lappe (1998) model explains a wider range of psychophysical data, but still embeds the Heeger and Jepson algorithm in the weight matrix.

Perrone and Stone (1994) proposed a model of the MT-MSTd network in which weights from MT cells encode templates of optic flow fields corresponding to different translational heading directions. This model can account for heading perception during eye rotations, but Crowell (1997) identified situations under which it makes incorrect predictions. The Perrone and Stone model also requires a large number of templates to handle optic flow fields with different depth structures, and it does not explain how the templates could be learned. Zemel and Sejnowski (1998) showed how a model could be trained using gradient descent to mimic MST-like responses in a hidden layer and encode both object motion and heading. Gradient descent training methods for neural networks require non-local transport of learned weights, for which there is no known biological evidence. Royden (1997, 2002) developed a model using opponent motion cells in MT to build heading representations in MSTd. It can account for heading perception in the presence of independently moving objects. The opponent-motion cells in MT are similar to the MT center-surround cells in the STARS model, although STARS uses these cells for object segmentation.

Grossberg, Mingolla, and Pack (1999) modeled heading in the MT-MSTd network processing optic flow in log-polar coordinates. STARS extends this model with several new features. Notably, MSTd cells in STARS form a spatial map of heading direction, instead of by the directional tuning of the maximally activated MSTd cell. Wagner (2004) presented a computer vision analysis of optic flow in log-polar coordinates, which can extract translational heading direction from optic flow in the presence of moving objects, but not during an eye or camera rotation.

Several modeling studies have shown that eye position gain fields can transform from retinal to head-centered coordinates. Zipser and Andersen (1988) trained a backpropagation network to perform a coordinate transformation using retinal and extraretinal inputs. Hidden layer nodes exhibited eye position gain modulation. Salinas and Abbott (1995, 1996) proposed a model of gain fields in which gain-modulated retinal input cells project to motor output cells via trained weights. The weight training procedure is similar to that in the STARS model. The model observes a series of random motor commands and the corresponding retinal inputs and adapts the weights according to an associative learning law. Pouget and Sejnowski (1997) showed that sigmoidal gain fields can produce coordinate transforms via delta-rule learning. Their gain field network is similar to that in STARS, although the gain field functions and training rule differ. The Zipser and Andersen (1988) and Pouget and Sejnowski (1997) models use training methods that are not consistent with known neural constraints, although they demonstrate the utility of gain fields after training.

The Lappe (1998) model used extraretinal signals to correct for eye rotations in the heading responses of model MST cells. The explicit use of extraretinal signals differs from the Heeger and Jepson algorithm, but the weights from MT to MST and from the eye velocity cells to MST are computed using the subspace algorithm. Beintema and van den Berg (1998) extended the template model of Perrone and Stone (1994) to incorporate eye velocity gain fields to correct for eye rotations. While STARS also uses eye velocity gain fields, it does so at an earlier processing stage than this model. STARS applies eye velocity gain fields to the local motion inputs arriving at MST from MT. Beintema and van den Berg constructed a retinotopic representation of heading direction from motion templates and then uses eye velocity gain fields to shift the MST heading peak. This solution is viable because optic flow templates contain information about the depth structure of the scene generating the optic flow field. Applying gain fields to local motion inputs allows STARS to avoid the issue of compensating for scene depth.

Cameron, Grossberg, and Guenther (1998) presented a self-organizing neural model of optic flow-based navigation. The model learns an MSTd-like heading map by repeated exposure to optic flow fields consistent with different heading trajectories. It corrects for the effects of smooth eye movements by subtracting off the rotational component of flow using a Vector Associative Map. This was possible because the input to the heading map was the raw optic flow field, whereas STARS encodes optic flow using nonlinear equations for area MT. The navigational module in Cameron et al. (1998) detects obstacles using an estimate of time-to-contact derived from a depth map, and steers away when the distance to an obstacle falls below a threshold. STARS extracts the goal and obstacles from the optic flow field using motion and depth discontinuities, and does not require the estimation of time-to-contact. Certain components of the Cameron et al. model, especially the use of self-organizing maps for training, inspired the development of STARS.

All the models discussed above are capable of explaining psychophysical data or neurophysiological data on motion processing in areas like MT and MST. They make different design assumptions, and arrive at solutions with varying degrees of biological plausibility, particularly regarding the use of templates and weight matrices. These models are summarized in Table 4.

Paper	Type of model	Data simulated
Zipser & Andersen (1988)	Backpropagation	Eye position gain fields
Heeger & Jepson (1992)	Subspace algorithm	Optic flow decomposition
Lappe & Rauschecker (1993)	Neural net using subspace algorithm	MST & translational heading perception
Perrone & Stone (1994)	Template model	MST & heading perception during eye rotation
Salinas & Abbott (1995; 1996)	Hebbian learning	Linear eye position gain fields; coordinate transform
Crowell (1997)	Template model	Heading perception during eye rotation
Pouget & Sejnowski (1997)	Delta-rule learning	Sigmoidal eye position gain fields; coordinate transform
Royden (1997; 2002)	Opponent motion neural network	MST & heading perception in presence of moving objects
Beintema & van der Berg (1998)	Template model & extraretinal signal	Translational heading perception during eye rotation; eye velocity gain fields
Cameron, Grossberg, & Guenther (1998)	Self-organizing neural model using optic flow	Navigation using optic flow; correction for eye rotation using extraretinal information
Lappe (1998)	Neural net using subspace algorithm & extraretinal signal	MST & heading perception during eye rotation
Zemel & Sejnowski (1998)	Unsupervised learning	MST heading & object motion
Grossberg, Mingolla, & Pack (1999)	Neural net using log-polar optic flow	MST heading & optic flow classification; translational heading perception
Wagner (2004)	Computer vision, log-polar optic flow	Heading perception & spiral space flow classification
Elder, Grossberg, & Mingolla (2007)	Neural net using log-polar optic flow, extraretinal signals, & gain fields	Steering, obstacle avoidance, route selection; target tracking; MSTd heading, MSTv object motion

Table 4. Comparison with other models.

Future Model Extensions. The STARS model explains a range of data on steering, obstacle avoidance, and heading perception, but it is not a complete model of visually-based reactive navigation. Notably, it does not explain how the motion signals processed in MT are constructed from sequences of images on the retina. Our simulation system analytically computes a dense optic flow field from the scene geometry, and supplies this flow field to MT after transforming it into cortical coordinates. It does not explain the cortical mechanisms used to build the motion field. Chey, Grossberg, and Mingolla (1997), Grossberg, Mingolla, and Viswanathan (2001), and Berzhanskaya, Grossberg, and Mingolla (2007) have developed the 3D FORMOTION model of biological motion perception that could be adapted for this purpose. Related cortical motion extraction mechanisms were also proposed by Lidén and Pack (1999) and Bayerl and Neumann (2007).

Several changes would be needed to adapt 3D FORMOTION mechanisms to the present model. The 3D FORMOTION model operates in Cartesian spatial coordinates, which need to be transformed to log-polar cortical coordinates. The MT-MST pathway addressed in the 3D FORMOTION model is the MT^--MSTv pathway. These MT^- cells perform spatial competition and depth-selective formotion capture, mechanisms related to depth-selective motion opponency in STARS. The MSTv cells encode a single, uniform direction of object motion. However, since optic flow cells in MSTd respond to motion patterns that include many directions of motion, additional mechanisms are needed to explain the grouping of MT^+ cell activity into representations of full-field optic flow patterns. Browning, Grossberg, and Mingolla (2008) have reported such results.

STARS also does not attempt to explain how the depth signals from V2 are computed. Physiological data suggest that V2 and MT cells respond to relative disparity, not absolute depth (Thomas, Cumming, & Parker, 2002). STARS approximates V2 computations by segmenting the scene into coarse near-, fixation-, and far-depth regions that correspond to areas of crossed, zero, and uncrossed disparity. Modeling the binocular disparity mechanisms in the V1-V2 complex to produce signals needed by MT is a more challenging task. One of the most capable neural models of 3D vision and figure-ground separation is the 3D LAMINART model of Grossberg and colleagues (Fang and Grossberg, 2007; Grossberg & Howe, 2003; Cao & Grossberg, 2005; Grossberg & Swaminathan, 2004; Grossberg and Yazdanbakhsh, 2005). Among other features, 3D LAMINART can form disparity gradients along slanted surfaces using bipole grouping cells which link different disparities along continuous slanted edges. Their model could in principle be used in conjunction with the 3D FORMOTION model to extract depth and motion from a sequence of input images, and from there, feed signals to the steering mechanisms described in this paper.

The STARS model is one component in a larger navigational system. STARS can reactively control steering within a visible spatial environment, but cannot follow complex remembered routes because it does not have a planning system. The hippocampus and prefrontal cortex are known to play crucial roles in spatial and sequential memory and navigation (O'Keefe & Dostrovsky, 1971; O'Keefe & Nadel, 1978; Poucet et al., 2004; Hok et al., 2007). Neural modeling of the hippocampal place cell system, for instance, has shown how a spatial map can form during exploratory navigation and how animals can use that map to navigate to different remembered goal locations (e.g., Touretzky & Redish, 1996; Gerstner & Abbott, 1997; Fuhs, Redish, & Touretzky, 1998; Hasselmo, Hay, Ilyn, & Gorchetchnikov 2002; Stringer, Rolls, Trappenberg, & Araujo, 2002; Gorchetchnikov & Hasselmo, 2005). Recent modeling work has shown how grid cells in entorhinal cortex may give rise to place cells in hippocampus (Gorchetchnikov & Grossberg, 2007). These and other similar models typically focus on the encoding and retrieval of spatial information in the hippocampus, but do not address the nature of the sensory inputs to the hippocampus, the relationship of the hippocampus to the spatial vision systems in parietal cortex, or how navigational plans are learned and recalled. Other models have clarified how laminar cortical circuits in the prefrontal cortex can store sequential events in short-term working memory, encode them in learned sequential plans, and perform them at variable rates under volitional control (e.g., Grossberg and Pearson, 2008). The STARS model could supply visual information about the direction of heading and spatial layout of the scene to the hippocampus as it steers through environments. The hippocampal system could simultaneously construct a spatial map of the environment, including the location of goals, obstacles, and other salient features, while receiving planned read-out from prefrontal cortex.

Future modeling work will aim to unify these various types of complementary information into a more comprehensive neural theory of autonomous navigation.

Acknowledgments. D. M. E. was supported in part by the Air Force Office of Scientific Research (AFOSR F4960-01-1-0397), the National Geospatial-Intelligence Agency (NMA201-01-1-2016), the National Science Foundation (NSF SBE-0354378), and the Office of Naval Research (ONR N00014-01-1-0624). S. G. was supported in part by the National Science Foundation (NSF SBE-0354378) and the Office of Naval Research (ONR N00014-01-1-0624). E. M. was supported in part by the National Geospatial-Intelligence Agency (NMA201-01-1-2016), the National Science Foundation (NSF SBE-0354378), and the Office of Naval Research (ONR N00014-01-1-0624).

Appendix A: Model Equations

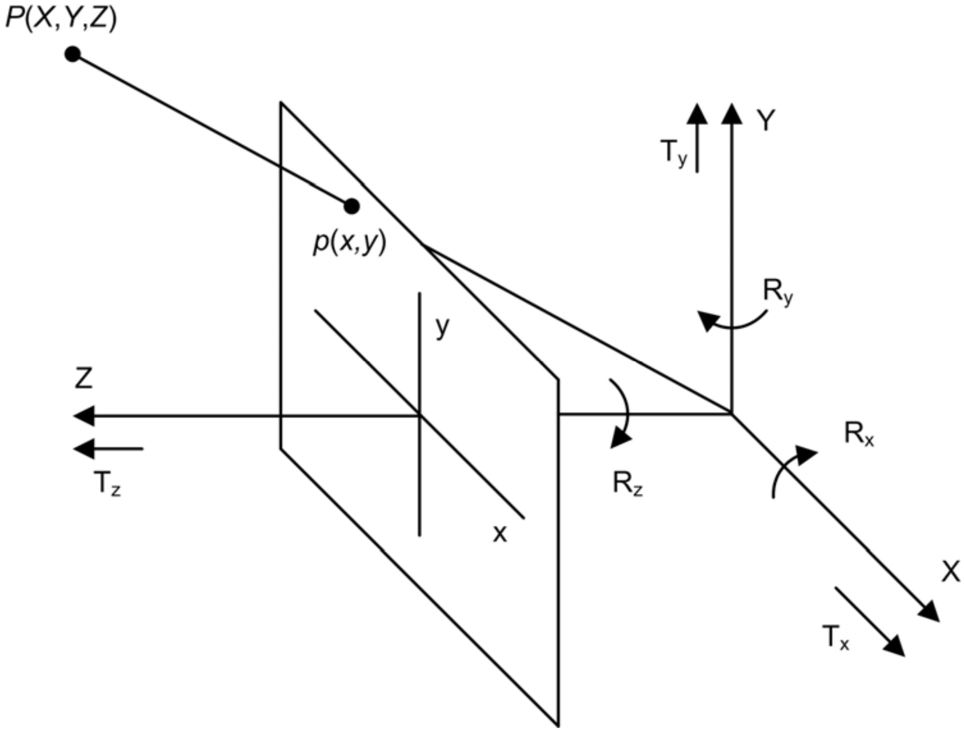


Figure 17. Geometry used to analytically compute optic flow. See text for details.

Level 1: Optic Flow Input. Analytic computation of the optic flow field relies on both the scene geometry and the translational and rotational motion of the viewpoint. A right-handed coordinate system is used with the viewpoint at the origin and the line-of-sight along the negative *z*-axis. The retina is approximated with a projection plane located at $z = -1$, parallel to the *xy*-plane, as illustrated in Figure 17. A 3D point $\mathbf{P}(X, Y, Z)$ in the environment projects to a point $\mathbf{p}(x, y) = (X/Z, Y/Z)$ in the retinal plane. Translation of the viewpoint is denoted $\mathbf{T} = (T_x, T_y, T_z)$, and rotation is denoted $\mathbf{R} = (R_x, R_y, R_z)$. Optic flow is computed using the equations

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

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

where Z is the depth of the point $\mathbf{P}$ (Longuet-Higgins & Prazdny, 1980).

Figure 16

Level 2: Log Polar Mapping of Optic Flow. A point **p**(x, y) on the retinal plane is equivalent to the complex number $z = x + iy$. The complex log-polar mapping transforms retinal points z into cortical points w. The mapping for the right and left hemifields is given by

$$w(z) = \begin{cases} \log(z+a) & \text{if } \operatorname{Re}(z) \geq 0 \\ 2\log(a) - \log(-z+a) & \text{if } \operatorname{Re}(z) < 0, \end{cases} \tag{3}$$

where the parameter a controls the size of the foveal region (Fischl, Cohen, & Schwartz, 1997). We set $a = 1$.

The log-polar equation defines a mapping from retinal coordinates **p**(x, y) to cortical coordinates **q**(u, v) = (Re(w), Im(w)). The optic flow vector field is mapped into log-polar coordinates using the Jacobian of the log-polar mapping:

$$J = \begin{pmatrix} \dfrac{\partial u}{\partial x} & \dfrac{\partial u}{\partial y} \\ \dfrac{\partial v}{\partial x} & \dfrac{\partial v}{\partial y} \end{pmatrix} = \begin{cases} \dfrac{1}{x^2 + 2ax + a^2 + y^2} \begin{pmatrix} x+a & y \\ -y & x+a \end{pmatrix} & \text{if } \operatorname{Re}(z) \geq 0 \\ \dfrac{1}{x^2 - 2ax + a^2 + y^2} \begin{pmatrix} -x+a & -y \\ y & -x+a \end{pmatrix} & \text{if } \operatorname{Re}(z) < 0, \end{cases} \tag{4}$$

for the right and left hemifields (Wagner, 2004). The optic flow vector (v_x, v_y) at the retinal location (x, y) is multiplied by the Jacobian J, evaluated at (x, y), and is denoted (v_u, v_v). The angular direction of the flow vector is $\hat{\theta}_{uv} = \operatorname{atan}\left(v_v / v_u \right)$ and the speed (magnitude) of the vector is $\hat{s}_{uv} = \sqrt{v_u^2 + v_v^2}$. The cortical coordinates (u, v) are continuous variables, but the simulations use a spatial discretization with indices (i, j). The log-polar optic flow variables are therefore written $\hat{\theta}_{ij}$ and $\hat{s}_{ij}$ in the following equations. The retinal domain has dimensions 512 × 512, and the log-polar domain has dimensions 512 × 194.

Level 3: Area V2 Depth Processing. The V2 depth signal is computed by partitioning the scene into near and fixation depth regions. Far depth is not used in the simulations. Let d_f denote the distance from the eye to the fixation point, and let z_{ij} denote the depth at point (i, j) in the log-polar depth map. The fixation-depth response function is a Gaussian centered on d_f and is given by:

$$D_{ij}^F = \exp\left(-\frac{\left(z_{\text{ij}} - d_f\right)^2}{2\sigma_{\text{Fix}}^2} \right), \tag{5}$$

where σ_{Fix} is the width of the Gaussian, set to 0.5.

The near-depth response function is a sigmoid that saturates for depths less than d_f, and is given by:

$$D_{ij}^N = \frac{1}{1 + \exp\left(z_{ij} - d_f + \sigma_{Near}\right)}, \tag{6}$$

where σ_{Near} is a bias term set to 0.5.

Level 4: Area MT^- subtractive Cells. Area MT^- subtractive cells have spatial receptive fields defined by a circular central region and an annular surround region. The model center-surround

MT cells are tuned for direction, speed, and depth. The receptive field centers are positioned in cortical log-polar coordinates, and the motion inputs to MT cells are transformed by the Jacobian of the log-polar mapping, as described above. The activity of a center-surround MT cell at location (*i, j*) and preferred direction *θ*, speed *s*, and depth *d* is thus denoted $MT_{ij}^{-}\left(\theta,s,d\right)$ and is computed by:

$$MT_{ij}^{-}\left(\theta,s,d\right)=\left[\alpha_{ij}^{c}\left(\theta,s,d\right)-\alpha_{ij}^{s}\left(\theta',s,d\right)\right]^{+}, \tag{7}$$

where $\alpha_{ij}^{c}\left(\theta,s,d\right)$ is the response function for the receptive field center; $\alpha_{ij}^{s}\left(\theta',s,d\right)$ is the response function for the receptive field surround; $\theta'=\left(\theta+180\right) \bmod 360$ is the anti-preferred direction that is 180° opposite direction *θ*, and $\left[x\right]^{+}=\max\left(x,0\right)$ denotes half-wave rectification. The depth variable *d* is either *F* for the fixation-depth cell population or *N* for the near-depth cell population. The receptive field center response function is given by

$$\alpha_{ij}^{c}\left(\theta,s,d\right)=\sum_{p,q}\Theta\left(\hat{\theta}_{pq},\theta\right)S\left(\hat{s}_{pq},s\right)D_{pq}^{d}P_{pqij}^{c}, \tag{8}$$

where *p*, *q* are dummy indices of summation, $\Theta\left(\hat{\theta}_{ij},\theta\right)$ is the Gaussian direction-tuning function, $S\left(\hat{s}_{ij},s\right)$ is the Gaussian speed-tuning function, D_{ij}^{d} is the depth tuning function, and P_{pqij}^{c} is the spatial kernel for the receptive field center. The direction-tuning function is defined by:

$$\Theta\left(\hat{\theta}_{ij},\theta\right)=\exp\left(-\frac{\left(\hat{\theta}_{ij}-\theta\right)^{2}}{2\sigma_{\theta}^{2}}\right), \tag{9}$$

where σ_{θ} is the width of the direction-tuning Gaussian. Angles $\hat{\theta}_{ij}$ wrap continuously such that -180° = +180°. The speed response function is defined by:

$$S\left(\hat{s}_{ij},s\right)=\exp\left(-\frac{\left(\hat{s}_{ij}-s\right)^{2}}{2\sigma_{s}^{2}}\right), \tag{10}$$

where σ_{s} is the width of the speed-tuning Gaussian. The depth response function is either D_{ij}^{F} or D_{ij}^{N} from equations (5) or (6), respectively. The spatial receptive field center is defined by:

$$P_{pqij}^{c}=\frac{1}{\sqrt{2\pi}\sigma_{c}}\exp\left(-\frac{\left(p-i\right)^{2}+\left(q-j\right)^{2}}{2\sigma_{c}^{2}}\right), \tag{11}$$

where σ_{c} is the width of the Gaussian kernel, set to 4.0 in the simulations.

The receptive field surround-response function is given by:

$$\alpha_{ij}^{s}\left(\theta',s,d\right)=\sum_{p,q}\Theta\left(\hat{\theta}_{pq},\theta'\right)S\left(\hat{s}_{pq},s\right)\bar{D}_{pq}^{d}P_{pqij}^{s}, \tag{12}$$

where the direction-tuning and speed-tuning functions are the same as above, and $\bar{D}_{ij}^{d}=1-D_{ij}^{d}$ is the depth response for the receptive field surround. This function is minimal for pixels (*i, j*) whose depth matches the preferred depth *d*. The spatial receptive field surround is defined by:

$$P_{pqij}^{s} = \frac{1}{\sqrt{2\pi}\sigma_u} \exp\left(-\frac{(p-i)^2 + (q-j)^2}{2\sigma_u^2} \right), \tag{13}$$

where σ_u is the width of the Gaussian kernel, set to 8.0 in the simulations.

In equations (11) and (13), the normalization term $1 / \sqrt{2\pi\sigma}$ is computed numerically. For many MT$^-$ cells, the 31 × 31 receptive field kernel falls partially outside the rectangular log-polar optic flow image or the valid range of the log-polar map function. We sum the kernel values evaluated at each pixel actually falling within the 31 × 31 receptive field kernel and divide by the total to compute the normalization constant. This ensures that cells having fewer pixels within their receptive fields generate signals similar in magnitude to those cells whose receptive fields are totally within the bounds of the log-polar range. A similar method is used to normalize the MT$^+$ additive cell receptive field kernels, defined below.

Level 5: Area MT$^+$ Additive Cells. Additive cells in area MT$^+$ have spatial receptive fields and Gaussian tunings for direction and speed of motion. The activity for an additive MT$^+$ cell at location (*i*, *j*) and preferred direction θ and speed s is denoted $MT_{ij}^{+}(\theta, s)$, and is computed by:

$$MT_{ij}^{+}(\theta, s) = \sum_{p,q} \Theta\left(\hat{\theta}_{pq}, \theta\right) S\left(\hat{s}_{pq}, s\right) P_{pqij}\,, \tag{14}$$

where $\Theta\left(\hat{\theta}_{ij}, \theta\right)$ is the Gaussian-tuned direction-tuning function, $S\left(\hat{s}_{ij}, s\right)$ is the speed-tuning function, and P_{pqij} is the 2D Gaussian kernel for the spatial receptive field. The direction-tuning and speed tuning functions are defined in equations (9) and (10), respectively. The spatial receptive field kernel is defined by:

$$P_{pqij} = \frac{1}{\sqrt{2\pi}\sigma_p} \exp\left(-\frac{(p-i)^2 + (q-j)^2}{2\sigma_p^2} \right), \tag{15}$$

where σ_p is the width of the Gaussian receptive field, set to 4.0 in the simulations.

Level 6: Area MSTv. Area MSTv cells sum inputs from MT$^-$ subtractive cells with like direction tunings. The contribution of an individual MT$^-$ cell to the activity of an MSTv cell is weighted by the preferred speed of the MT$^-$ cell, so that MSTv cells show an over-all increase in activity in response to faster stimuli; cf., Chey, Grossberg, & Mingolla (1998). Let $\bar{S}$ denote the set of preferred speeds in MT. Then the activity of an MSTv cell with preferred direction θ and depth *d* is given by:

$$MSTv_{ij}(\theta, d) = \sum_{p,q \in I} \sum_{s \in \bar{S}} s MT_{pq}^{-}(\theta, s, d), \tag{16}$$

where $I = \left\{ (p_1, q_1), \ldots, (p_i, q_i), \ldots, (p_{128,} q_{128}) \right\}$ is a set of indices into the MT$^-$ network chosen in the following way. Let (ϕ, ψ) denote the horizontal and vertical coordinates of the center of the MSTv cell's receptive field, in units of visual angle. We chose 128 additional coordinates (ϕ_i, ψ_i) from a 2D Gaussian distribution with mean $\boldsymbol{\mu}$ = (ϕ, ψ) and covariance matrix $\boldsymbol{\Sigma}$ $= \begin{bmatrix} \sigma_\phi & 0 \\ 0 & \sigma_\psi \end{bmatrix}$. Then we computed the log-polar transformation of the randomly chosen Cartesian coordinates, and found the indices $\left(p_i, q_i\right)$ of the nearest MT$^-$ cell cluster. These indices define the set of projections from MT- to MSTv. We set $\sigma_\phi = \sigma_\psi = 3$.

Level 7: Eye Position Gain Fields. Gain field modulation is modeled by a piecewise-linear function of eye position. Two such functions model the gain modulation when the eye is in the left or right half of the orbit. The left-saturating eye position gain fields have the form:

$$f_L\left(e; L_s, L_e\right) = \begin{cases} 1 & \text{if } e \le L_s \\ 1 - \dfrac{(e - L_s)}{(L_s - L_e)}, & \text{if } e > L_s \text{ and } e < L_e \\ 0 & \text{if } e \ge L_e \end{cases}, \tag{17}$$

where e is the eye position, and L_s and L_e are the endpoints of the linear region. Similarly, the right-saturating eye position gain fields are:

$$f_R\left(e; R_s, R_e\right) = \begin{cases} 0 & \text{if } e \le R_s \\ \dfrac{(e - R_s)}{(R_s - R_e)}, & \text{if } e > R_s \text{ and } e < R_e \\ 1 & \text{if } e \ge R_e \end{cases}, \tag{18}$$

where R_s and R_e are the endpoints of the linear region. The units of eye position e are degrees.

For the populations of near-depth MSTv, fixation-depth MSTv, and radial-flow MSTd cells, 6 populations of gain-modulated cells are simulated. Three populations have increasing gains when the eye is in the left half of the orbit, and three populations have increasing gains when the eye is in the right half of the orbit. The equations defining the three left eye position, gain-modulated near-depth MSTv populations are

$$MSTv_{ij}^{FL}\left(e, d_{\text{near}}\right) = \sum_{\theta \in \bar{\Theta}} f_L\left(e; -21, -9\right) MSTv_{ij}\left(\theta, d_{\text{near}}\right), \tag{19}$$

$$MSTv_{ij}^{ML}\left(e, d_{\text{near}}\right) = \sum_{\theta \in \bar{\Theta}} f_L\left(e; -12, 0\right) MSTv_{ij}\left(\theta, d_{\text{near}}\right), \tag{20}$$

and

$$MSTv_{ij}^{CL}\left(e, d_{\text{near}}\right) = \sum_{\theta \in \bar{\Theta}} f_L\left(e; -3, 9\right) MSTv_{ij}\left(\theta, d_{\text{near}}\right), \tag{21}$$

where e is the position of the eye in the orbit, FL means "far-left", ML means "middle-left", and CL means "central-left", which correspond to the gain fields (1), (2), and (3) in Figure **Error! Reference source not found.**. The gain-modulated functions sum over directions, so the gain-modulated MSTv cells encode the position of the object, not the direction of motion. The equations defining the three right eye position, gain-modulated near-depth MSTv cell populations are:

$$MSTv_{ij}^{CR}\left(e, d_{\text{near}}\right) = \sum_{\theta \notin \bar{\Theta}} f_R\left(e; -9, 3\right) MSTv_{ij}\left(\theta, d_{near}\right), \tag{22}$$

$$MSTv_{ij}^{MR}\left(e, d_{\text{near}}\right) = \sum_{\theta \notin \bar{\Theta}} f_R\left(e; 0, 12\right) MSTv_{ij}\left(\theta, d_{near}\right), \tag{23}$$

and

$$MSTv_{ij}^{FR}\left(e, d_{\text{near}}\right) = \sum_{\theta \notin \bar{\Theta}} f_R\left(e; 9, 21\right) MSTv_{ij}\left(\theta, d_{near}\right), \tag{24}$$

where CR means "central-right", MR means "middle-right", and FR means "far-right", which correspond to the gain fields (4), (5), and (6) in Figure **Error! Reference source not found.**.

The gain modulation equations for the fixation-depth MSTv and radial flow MSTd populations are similar, except that the MSTd equations do not include a summation over directions.

The retinotopic representations of goal, obstacle, and heading are transformed into craniotopic coordinates by computing a weighted sum of the 6 gain-modulated cell populations. The equation for the craniotopic representation of the obstacle, derived from the near-depth MSTv cell population, is:

$$\begin{aligned} O_{ij} &= \sum_{p,q}\left(w_{ijpq}^{FL} MSTv_{pq}^{FL} + w_{ijpq}^{ML} MSTv_{pq}^{ML} + w_{ijpq}^{CL} MSTv_{pq}^{CL} \right) \\ &\quad + \sum_{p,q}\left(w_{ijpq}^{CR} MSTv_{pq}^{CR} + w_{ijpq}^{MR} MSTv_{pq}^{MR} + w_{ijpq}^{FR} MSTv_{pq}^{FR} \right). \end{aligned} \tag{25}$$

The equations for the goal G_{ij} and heading H_{ij} are analogous. The weights projecting from the gain-modulated cell populations to the craniotopic cell populations are computed using a self-organizing map (SOM), described below.

Level 8a and 8b: MSTd Radial Flow Cells. MSTd radial flow cells encode the direction of heading from the distributed representation of the flow field in the additive MT^{+} cell population. The inputs to MSTd radial flow cells are outputs from MT^{+} cells which are gain-modulated by eye velocity. The eye velocity gain fields are piecewise-linear functions that saturate for leftward or rightward eye rotations. The left-saturating gain fields for leftward eye rotations are given by equation (17), and the right-saturating gain fields for rightward eye rotations are given by equation (18). The units of eye velocity v are deg/sec. We simulated 3 left-saturating and 3 right-saturating eye velocity gain fields. The 6 gain-modulated MT+ additive cell clusters at location (i, j), are specified by the following equations:

$$G_{ij}^{FL}\left(\theta,s\right) = f_L\left(v,-21,-9\right) MT_{ij}^{+}\left(\theta,s\right), \tag{26}$$

$$G_{ij}^{ML}\left(\theta,s\right) = f_L\left(v,-12,0\right) MT_{ij}^{+}\left(\theta,s\right), \tag{27}$$

$$G_{ij}^{SL}\left(\theta,s\right) = f_L\left(v,-3,9\right) MT_{ij}^{+}\left(\theta,s\right), \tag{28}$$

$$G_{ij}^{SR}\left(\theta,s\right) = f_R\left(v,-9,3\right) MT_{ij}^{+}\left(\theta,s\right), \tag{29}$$

$$G_{ij}^{MR}\left(\theta,s\right) = f_R\left(v,0,12\right) MT_{ij}^{+}\left(\theta,s\right), \tag{30}$$

and

$$G_{ij}^{FR}\left(\theta,s\right) = f_R\left(v,9,21\right) MT_{ij}^{+}\left(\theta,s\right), \tag{31}$$

where superscripts *FL*, *ML*, and *SL* denote fast-left, medium-left, and slow-left eye rotations, respectively, and superscripts *SR*, *MR*, and *FR* denote slow-right, medium-right, and fast-right eye rotations, respectively.

The connections from the gain-modulated MT^{+} additive cells to individual MSTd radial flow cells are weighted, and the activity of an MSTd cell is given by

$$MSTd_{ij}^{r} = \sum_{p,q\in I} \sum_{\theta\in\bar{\Theta}} \sum_{s\in\bar{S}} \sum_{\gamma\in\bar{\Gamma}} w_{ijpq\theta s\gamma} G_{pq}^{\gamma}\left(\theta,s\right), \tag{32}$$

where $w_{ijpq\theta s\gamma}$ is the weight from the gain-modulated MT^{+}additive cell $G_{pq}^{\gamma}\left(\theta,s\right)$ with preferred direction θ, preferred speed s, and gain field function γ. The summation indices (p, q) come from a randomly chosen subset of gain-modulated MT^{+} cells, θ ranges over the set of 8 MT preferred directions, s ranges over the set of 4 MT preferred speeds, and γ ranges over the set of 6 gain-field functions. The random subset of gain-modulated MT^{+} additive cells is chosen using the procedure described above for MSTv cells. Since MSTd cells have wider receptive fields, we set $\sigma_{\phi} = \sigma_{\psi} = 12$ in the 2D Gaussian distribution.

The weights are adapted using the self-organizing map learning law:

$$\frac{dw_{ijpq\theta s\gamma}}{dt} = \eta\left(w_{ijpq\theta s\gamma} - G^{\gamma}_{pq}\left(\theta, s\right)\right) N_{ij}$$

$$N_{ij} = \begin{cases} \dfrac{1}{N+1}, & \text{at } MSTd_{\max} \text{ and } N \text{ neighbors} \\ 0 & \text{otherwise,} \end{cases} \tag{33}$$

where η is a learning constant, $MSTd_{\max}$ is the cell with maximum activity in $MSTd$, and N_{ij} is a neighborhood function centered on $MSTd_{\max}$. The details of the training procedure are given in Appendix B.

Normally, several thousand training cycles are needed to converge the network weights to a steady state, which has a substantial run-time cost. However, the converged state of the weights can be computed directly using the above equations. As the training time t increases, N decreases to 0, at which point N_{ij} = 1 for the maximally activated cell and 0 otherwise (Grossberg, 1976; Kohonen, 1982). Solving the weight update equation at equilibrium yields:

$$\frac{dw_{ijpq\theta s\gamma}}{dt} = 0 \Rightarrow w_{ijpq\theta s\gamma} = G^{\gamma}_{pq}\left(\theta, s\right). \tag{34}$$

In words, the converged weights projecting to a given MSTd cell are equal to the cell's preferred input pattern. For a particular MSTd cell, we can assign it a 2D Cartesian heading direction, compute the corresponding optic flow field, and set its weights equal to the pattern of activity in the additive MT^+ cell population. This training shortcut is justified because the self-organizing feature map is a maximum inner-product classifier, and for normalized (x,y), the inner product $\langle x, y\rangle$ is maximized when $x = y$ (Grossberg, 1976; Kohonen, 1982).

Level 9: MSTd Planar Flow Cells. MSTd planar flow cells respond to the background flow field generated by smooth eye rotations. Two MSTd planar flow cells are simulated, $MSTd^L$ and $MSTd^R$, which respond to left and right background flow, respectively. These cells compute a sum of weighted inputs from the MT^+ additive cell population. Since only two cells are simulated, it is computationally tractable for each cell to receive inputs from the entire MT^+ cell population:

$$MSTd^R = \sum_{p,q} w^R_{ijpq\theta s} MT^+_{pq}\left(\theta, s\right) \tag{35}$$

and

$$MSTd^L = \sum_{p,q} w^L_{ijpq\theta s} MT^+_{pq}\left(\theta, s\right), \tag{36}$$

where p and q are summation indices ranging over MT^+, and the weights $w^R_{ijpq\theta s}$ and $w^L_{ijpq\theta s}$ are trained using a self-organizing map, as described above.

Level 10: PPC Steering. Cells in PPC receive input from the goal G_{ij}, obstacle O_{ij}, and heading H_{ij} cells. The equation defining the PPC steering field is:

$$S_{ij} = G_{ij} - O_{ij} + H_{ij}, \tag{37}$$

where the subtraction of the obstacle term O_{ij} causes the obstacle to repel heading. A winner-take-all network determines the maximum location of the steering. Let $\tilde{S}_{ij}$ denote the result of the winner-take-all procedure:

$$\tilde{S}_{ij} = \begin{cases} 1 & \text{if } ij = \operatorname{argmax}\left(S_{ij}\right) \\ 0 & \text{otherwise} \end{cases} . \tag{38}$$

The steering signal itself is generated from the activity of two steering cells encoding left turns (*L*) and right turns (R). The steering cells have activities defined by:

$$L = \sum_{i=1}^{N/2} \sum_{j} \left(w_L \left| i - \frac{N}{2} \right| \right) \tilde{S}_{ij} \tag{39}$$

and

$$R = \sum_{i=N/2+1}^{N} \sum_{j} \left(w_R \left| i - \frac{N}{2} \right| \right) \tilde{S}_{ij} , \tag{40}$$

where the outer summations in the L and R equations sum over the left and right halves of the winner-take-all field and $w_L = w_R = 4$. The absolute value term defines a V-shaped function with a trough at the center of the winner-take-all map. This ensures that larger peak shifts in the steering field S_{ij} away from the midpoint induce larger turning movements, and thereby indirectly controls angular acceleration. The steering signal is defined by:

$$\frac{d\phi}{dt} = \alpha\left(L - R\right), \tag{41}$$

where ϕ is the heading direction and $\alpha = 0.09375$ controls the steering rate.

Level 11: FEF Smooth Pursuit. The equations for the smooth pursuit system are adapted from Pack, Grossberg, and Mingolla (2001). These equations implement the network connectivity shown in Figure 5. The pursuit speed obeys the equation:

$$\frac{dp}{dt} = -p + \left(T_L - T_R\right), \tag{42}$$

where T_L and T_R are the left and right target tracking cell activities. The target tracking cells T_L and T_R are mutually inhibitory. They sum inputs from the fixation-depth MSTv cell populations encoding leftward and rightward motion of the goal (i.e., pursuit target). They also receive lateral excitatory projections from the background flow cells tuned for the opposite direction of motion. This ensures that leftward target motion is coupled with rightward background flow, and vice versa. The target tracking cell equations are given by:

$$\frac{dT_L}{dt} = -T_L + \left(1 - T_L\right)\left[B_R + M \sum_{ij} MSTv_{ij}\left(0°, Fix\right) - p \right] - JT_L T_R \tag{43}$$

and

$$\frac{dT_R}{dt} = -T_R + \left(1 - T_R\right)\left[B_L + M \sum_{ij} MSTv_{ij}\left(180°, Fix\right) + p \right] - JT_L T_R , \tag{44}$$

where $J = 20.0$ and $M = 0.5$ are parameters and B_L and B_R are activities of the two background flow cells. The background flow cells B_R and B_L are also mutually inhibitory. They receive bottom-up inputs from the MSTd planar flow cells encoding rightward and leftward background flow. They also receive lateral excitatory projections from the target tracking cells tuned for the opposite direction of motion. The background flow cell activities are defined by:

$$\frac{dB_R}{dt} = -B_R + \left(1 - B_R\right)\left[T_L + CMSTd^R - p \right] - FB_R B_L \tag{45}$$

and

$$\frac{dB_L}{dt} = -B_L + \left(1 - B_L\right)\left[T_R + CMSTd^L + p\right] - FB_R B_L, \tag{46}$$

where $C = 0.005$ and $F = 1.0$. As smooth pursuit begins, the pursuit system is driven primarily by retinal slip of the target. Once the pursuit speed matches the relative motion of the target, there is no retinal slip and the system is driven primarily by the background flow. The target tracking and background flow cells also receive extraretinal feedback from the pursuit cell, which helps to mainten predictive of smooth pursuit.

Route Selection Parameters. In the simulation of the route selection data from Fajen and Warren (2003), we multiply the term O_{ij} in Equation (37) by 0.5 to reduce the repulsion of the obstacle. This is similar to the parameter change in Fajen and Warren's simulation, and could be interpreted as a reallocation of attention to favor the goal over the obstacle in these situations.

Appendix B: Simulation techniques

The STARS model was implemented in a simulation environment written in C++ and OpenGL on an NVIDIA 6800 Ultra graphics processing unit (GPU), using general-purpose GPU techniques. All model equations are implemented as OpenGL Shading Language (GLSL) fragment shaders, and cell activities are stored in texture maps. This simulation method exploits the parallelism of the GPU, and executes the model efficiently. The field of view spanned 90º horizontally and vertically. The scene screenshots and Cartesian optic flow field were computed at a 512 × 512 resolution. The log-polar optic flow field and the MT additive and subtractive populations were computed at a resolution of 512 × 194 pixels. The MSTv, MSTd, and PPC cell populations were computed at a resolution of 256 × 8. The vertical dimension was reduced to improve computational efficiency in the simulations. The model only makes horizontal steering movements and eye movements, so information in the vertical dimension is not needed.

STARS employs self-organizing maps in two places, first in training the weights projecting from the eye-velocity gain-modulated MT^+ additive cells to the MSTd expansion cells, and second in training the eye-position gain-modulated outputs of the MSTv near-depth and fixation-depth cells and the MSTd expansion cells to the craniotopic representations of goal, obstacle, and heading in PPC. The full-scale model simulations used a short-cut to precompute the converged state of the weights directly. Here we describe a training procedure that could be used to adapt the weights according to the self-organizing map learning rule. The MT^+-MSTd weights are initially randomized with a uniform (0,1) distribution. First a horizontal heading trajectory is chosen randomly and the optic flow field and MT^+ additive cell activity is computed. Initially the eye velocity is zero, and we determine the eye-velocity gain-modulation (for velocity zero) and compute the gain-modulated MT^+ additive cell responses, followed by the MSTd expansion cell responses, using the randomized weights. We then maintain the activity of the maximally active cell in the MSTd expansion cell map and its neighbors while the model makes a series of smooth pursuit eye movements while moving along the chosen trajectory. The MT^+ additive cell activities are recomputed and new gain modulation is applied for each smooth pursuit eye movement. The weights projecting to the winning MSTd cell and its neighbors are adapted during the eye movements. After a number of eye movements are made, a new heading trajectory is chosen and the procedure begins again. This method allows MSTd cells to associate optic flow fields (encoded in MT^+) containing varying amounts of rotational flow with a particular retinotopic heading direction. Maintaining the activity of a group of MSTd cells during repeated eye movements is a form of *self-supervised learning*. Learning is reset when a new heading trajectory is chosen. A similar training procedure can be applied to the weights projecting from the eye-position gain-modulated cell populations in MSTv and MSTd.

References

Albright, T. D. (1984). Direction and orientation selectivity of neurons in visual area MT of the macaque. *Journal of Neurophysiology*, *52*, 1106-1130.

Allman, J. M., & Kaas, J. H. (1975). The dorsomedial cortical visual area: a third tier area in the occipital pole of the owl monkeys (*Aotus trivirgatus*), *Brain Research, 100*, 473-487.

Allman, J., Miezin, F., & McGuinness, E. (1985). Direction- and velocity-specific responses from beyond the classical receptive field in the middle temporal visual area (MT). *Perception*, *14*, 105-126.

Andersen, R. A., Essick, G. K., & Siegel, R. M. (1985). Encoding of spatial location by posterior parietal neurons. *Science*, *230*, 456-458.

Andersen, R. A., Essick, G. K., & Siegel, R. M. (1987). Neurons of area 7 activated by both visual stimuli and oculomotor behavior. *Experimental Brain Research*, *67*, 316-322.

Atchley, P., & Andersen, G. J. (1999). The discrimination of heading from optic flow is not retinally invariant. *Perception & Psychophysics*, *61*, 387-396.

Bayerl, P., & Neumann, H. (2007). A fast biologically inspired algorithm for recurrent motion estimation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *29*, 246-260.

Beall, A. C. & Loomis, J. M. (1997). Optic flow and visual analysis of base-to-final turn. *Perception*, *25*, 481-494.

Beintema, J. A., & van den Berg, A. V. (1998). Heading detection using motion templates and eye velocity gain fields. *Vision Research*, *38*, 2155-2179.

Berezovskii, V. K., & Born, R. T. (2000). Specificity of projections from wide-field and local motion-processing regions within the middle temporal visual area of the owl monkey. *The Journal of Neuroscience*, *20*, 1157-1169.

Berzhanskaya, J., Grossberg , S. and Mingolla , E. (2007). Laminar cortical dynamics of visual form and motion interactions during coherent object motion perception. *Spatial Vision, 20*,337-395.

Born, R. T., & Tootell, R. B. H. (1992). Segregation of global and local motion processing in primate middle temporal visual area. *Nature*, *357*, 497-499.

Born, R. T. (2000). Center-surround interactions in the middle temporal visual area of the owl monkey. *Journal of Neurophysiology*, *84*, 2658-2669.

Born, R. T., & Bradley, D. C. (2005). Structure and function of visual area MT. *Annual Review of Neuroscience*, *28*, 157-189.

Boussaoud, D., Ungerleider, L. G., & Desimone, R. (1990). Pathways for motion analysis: cortical connections of the medial superior temporal and fundus of the medial superior temporal visual areas in the macaque. *Journal of Comparative Neurology*, 296, 462-495.

Bradley, D. C., & Andersen, R. A. (1998). Center-surround antagonism based on disparity in primate area MT. *Journal of Neuroscience*, *18*, 7552-7565.

Bradley, D. C., Maxwell, M., Andersen, R. A, Banks, M. S., & Shenoy, K. V. (1996). Mechanisms of heading perception in primate visual cortex. *Science*, *273*, 1544-1547.

Britten, K. H. (1998). Clustering of response selectivity in the medial superior temporal area of extrastriate cortex in the macaque monkey. *Visual Neuroscience*, *15*, 553-558.

Britten, K. H., & van Wezel, R. J. A. (1998). Electrical microstimulation of cortical area MST biases heading perception in monkeys. *Nature Neuroscience*, *1*, 59-63.

Britten, K. H., & van Wezel, R. J. A. (2002). Area MST and heading perception in macaque monkeys. *Cerebral Cortex*, *12*, 692-701.

Browning, A., Grossberg, S., & Mingolla, E. (2008a). A neural model of how the brain computes heading from optic flow in realistic scenes. Submitted for publication.

Browning, A., Grossberg, S., & Mingolla, E. (2008b). Cortical dynamics of navigation and steering in natural scenes: Motion-based object segmentation, heading, and obstacle avoidance. Submitted for publication.

Bullock, D., & Grossberg, S. (1991). Adaptive neural networks for control of movement trajectories invariant under speed and force rescaling. *Human Movement Science*, ***10***, 3-53.

Cameron, S., Grossberg, S., & Guenther, F. H. (1998). A self-organizing neural network architecture for navigation using optic flow. *Neural Computation*, *10*, 313-352.

Cao, Y., & Grossberg, S. (2005). A laminar cortical model of stereopsis and 3D surface perception: Closure and da Vinci stereopsis. *Spatial Vision*, *18*, 515-578.

Chey, J., Grossberg, S., & Mingolla, M. (1997). Neural dynamics of motion grouping: From aperture ambiguity to object speed and direction. *Journal of the Optical Society of America*, 14, 2570-2594.

Chey, J., Grossberg, S., & Mingolla, E. (1998). Neural dynamics of motion processing and speed discrimination. *Vision Research*, *38*, 2769-2786.

Colby, C. L., Duhamel, J.-R., & Goldberg, M. E. (1993). Ventral intraparietal area of the macaque: anatomic location and visual response properties. *Journal of Neurophysiology*, *69*, 902-914.

Contreras-Vidal, J.L., Grossberg, S., and Bullock, D. (1997). A neural model of cerebellar learning for arm movement control: Cortico-spino-cerebellar dynamics. *Learning and Memory, **3***, 475-502.

Crowell, J. A. (1997). Testing the Perrone and Stone (1994) model of heading estimation. *Vision Research, 37,* 1653-1671

Crowell, J. A. & Andersen, R. A. (2001). Pursuit compensation during self-motion. *Perception, 30,* 1465-1488.

Crowell, J. A., & Banks, M. S. (1993). Perceiving heading with different retinal regions and types of optic flow. *Perception & Psychophysics*, *53,* 325-337.

Cutting, J. E., Vishton, P. M., & Braren, P. A. (1995). How we avoid collisions with stationary and moving obstacles. *Psychological Review, 102*, 627-651.

Daniel, P. M., & Whitteridge, D. (1961). The representation of the visual field on the cerebral cortex in monkeys. *Journal of Physiology*, *159*, 203-221.

Desimone, R. & Ungerleider, L. (1986). Multiple visual areas in the caudal superior temporal sulcus of the macaque. *Journal of Comparative Neurology*, *248*, 164-189.

Duffy, C. J., & Wurtz, R. H. (1991). Sensitivity of MST neurons to optic flow stimuli. I. A continuum of response selectivity to large-field stimuli. *Journal of Neurophysiology*, *65*, 1329-1345.

Duffy, C. J., & Wurtz, R. H. (1997). Planar directional contributions to optic flow responses in MST neurons. *Journal of Neurophysiology*, *77*, 782-796.

Duhamel, J.-R., Bremmer, F., Ben Hamed, S., & Graf, W. (1997). Spatial invariance of visual receptive fields in parietal cortex neurons. *Nature*, *389*, 845-848.

Eifuku, S., & Wurtz, R. H. (1998). Response to motion in extrastriate area MSTl: center-surround interactions. *Journal of Neurophysiology, 80*, 282-296.

Fajen, B. R., & Warren, W. H. (2003). Behavioral dynamics of steering, obstacle avoidance, and route selection. *Journal of Experimental Psychology: Human Perception and Performance*, *29*, 343-362.

Fajen, B. R., & Warren, W. H. (2004). Visual guidance of intercepting a moving target on foot. *Perception*, *33*, 689-715.

Fajen, B. R., & Warren, W. H. (2007). Behavioral dynamics of intercepting a moving target. *Experimental Brain Research* [Electronic publication ahead of print]. Available online February 2, 2007, at http://www.springerlink.com/content/w0hku376748k1450/.

Fang, L. & Grossberg, S. (2007). From stereogram to surface: How the brain sees the world in depth. Submitted for publication.

Fischl, B., Cohen, M. A., & Schwartz, E. L. (1997). The local structure of space-variant images. *Neural Networks*, *10*, 815-831.

Fuhs, M. C., Redish, A. D., & Touretzky, D. S. (1998). A visually driven hippocampal place cell model. In J. Bower (Ed.), *Computational Neuroscience: Trends in Research, 1998*, pp. 379-384. New York: Plenum Publishing.

Gerstner, W., & Abbott, L. F. (1997). Learning navigational maps through potentiation and modulation of hippocampal place cells. *Journal of Computational Neuroscience*, *4*, 79-94.

Gibson, James J. (1950). Perception of the visual world. Boston, MA: Houghton-Mifflin.

Gibson, J. J., Olum, P., & Rosenblatt, F. (1955). Parallax and perspective during aircraft landings. *American Journal of Psychology*, *51,* 453-471.

Gibson, James J. (1979). The ecological approach to visual perception. Boston, MA: Houghton-Mifflin.

Gnadt, W., & Grossberg, S. (2008). SOVEREIGN: An autonomous neural system for incrementally learning planned action sequences to navigate towards a rewarded goal. *Neural Networks, 21*, 699-758 .

Gorchetchnikov, A., & Grossberg, S. (2007). Space, time, and learning in the hippocampus: How fine spatial and temporal scales are expanded into population codes for behavioral control. *Neural Networks*, *20*, 182-193.

Gorchetchnikov, A., & Hasselmo, M. E. (2005). A biophysical implementation of a bidirectional graph search algorithm to solve multiple goal navigation tasks. *Connection Science*, *17*, 145-164.

Gottlieb, J. P., Bruce, C. J., & MacAvoy, M. G. (1993). Smooth eye movements elicited by microstimulation in the primate frontal eye field. *Journal of Neurophysiology*, *69*, 786-799.

Grossberg, S. (1976). Adaptive pattern classification and universal recoding. I: Parallel development and coding of neural feature detectors. *Biological Cybernetics*, *23,* 121-134.

Grossberg, S. (2000). The complementary brain: Unifying brain dynamics and modularity. *Trends in Cognitive Sciences, 4*, 233-246.

Grossberg, S., & Howe, P. D. L. (2003). A laminar cortical model of stereopsis and three-dimensional surface perception. *Vision Research*, *43*, 801-829.

Grossberg, S., Kuhlmann, L., and Mingolla, E. (2007). A neural model of 3D shape-from-texture: Multiple-scale filtering, boundary grouping, and surface filling-in. Vision Research, 47(5):634-672.

Grossberg, S., & Kuperstein, M. (1986). Neural dynamics of adaptive sensory-motor control: Ballistic eye movements. Amsterdam: North-Holland.

Grossberg, S., Mingolla, E., & Pack, C. (1999). A neural model of motion processing and visual navigation by cortical area MST. *Cerebral Cortex*, *9*, 878-895.

Grossberg, S., Mingolla, E., & Viswanathan, L. (2001). Neural dynamics of motion integration and segmentation within and across apertures. *Vision Research*, *41*, 2521-2553.

Grossberg, S., & Pearson, L. (2008). Laminar cortical dynamics of cognitive and motor working memory, sequence learning and performance: Toward a unified theory of how the cerebral cortex works. *Psychological Review, 115,* 677-732.

Grossberg, S., & Swaminathan, G. (2004). A laminar cortical model for 3D perception of slanted and curved surfaces and of 2D images: development, attention, and bistability. *Vision Research*, *44*, 1147-1187.

Grossberg, S., & Yazdanbakhsh, A. (2005). Laminar cortical dynamics of 3D surface perception: Stratification, transparency, and neon color spreading. *Vision Research, 45*, 1725-1743.

Hasselmo, M. E., Hay, J., Ilyn, M., & Gorchetchnikov, A. (2002). Neuromodulation, theta rhythm, and rat spatial navigation. *Neural Networks*, *15*, 689-707.

Heeger, D. J., & Jepson, A. D. (1992). Subspace methods for recovering rigid motion. I. Algorithm and implementation. *International Journal of Computer Vision*, *7,* 95-117.

Hildreth, H. C., Beusmans, J. M. H., Boer, E. R., & Royden, C. S. (2000). From vision to action: experiments and models of steering control during driving. *Journal of Experimental Psychology: Human Perception and Performance*, *26*, 1106-1132.

Howard, I. P. and Rogers, B. J. (1995). *Binocular Vision and Stereopsis.* New York: Oxford University Press.

Hubel, D. H., & Wiesel, T. N. (1968). Receptive fields and functional architecture of monkey striate cortex. *Journal of Physiology*, *194*, 215-243.

Kohonen, T. (1982). Self-organized formation of topologically correct feature maps. *Biological Cybernetics*, *43*, 59-69.

Komatsu, H., & Wurtz, R. H. (1988). Relation of cortical areas MT and MST to pursuit eye movements. III. Interaction with full-field visual stimulation. *Journal of Neurophysiology*, *60*, 621-644.

Land, M. & Horwood, J. (1995). Which parts of the road guide steering? *Nature*, *377*, 339-340.

Lappe, M. (1998). A model of the combination of optic flow and extraretinal eye movement signals in primate extrastriate visual cortex. Neural model of self-motion from optic flow and extraretinal cues. *Neural Networks*, *11*, 397-414.

Lappe, M., & Rauschecker, J. P. (1993). A neural network for the processing of optic flow and ego-motion in man and higher mammals. *Neural Computation*, *5*, 374-391.

Lee, D. N. (1976). A theory of visual control of braking based on information about time-to-collision. *Perception*, *5*, 437-459.

Lewis, J. W., & Van Essen, D. C. (2000). Corticocortical connections of visual, sensorimotor, and multimodal processing areas in the parietal lobe of the macaque monkey. *Journal of Comparative Neurobiology*, *428*, 112-137.

Lidén, L. & Pack, C. (1999). The role of terminators and occlusion cues in motion integration and segmentation: a neural network model. *Vision Research*, *39*, 3301-3320.

Longuet-Higgins, H. C. & Prazdny, K. (1980). The interpretation of a moving retinal image. *Proceedings of the Royal Society of London B*, *208*, 385-397.

Maunsell, J. H. R., & Van Essen, D. C. (1983a). The connections of the middle temporal visual area (MT) and their relationship to a cortical hierarchy in the macaque monkey. *Journal of Neuroscience*, *3*, 2563-2586.

Maunsell, J. H. R., & Van Essen, D. C. (1983b). Functional properties of neurons in middle temporal visual area of the macaque monkey. I. Selectivity for stimulus direction, speed, and orientation. *Journal of Neurophysiology*, *49*, 1127-1147.

Nover, H., Anderson, C. H., & DeAngelis, G. C. (2005). A logarithmic, scale-invariant representation of speed in macaque middle temporal area accounts for speed discrimination performance. *Journal of Neuroscience, 25*, 10049-10060.

O'Keefe, J. M., & Dostrovsky, J. (1971). The hippocampus as a spatial map. Preliminary evidence from unit activity in the freely moving rat. *Brain Research, 34*, 171-175.

O'Keefe, J., & Nadel, L. (1978). *The Hippocampus as a Cognitive Map*. Oxford, UK: Oxford University Press.

Ohzawa, I., DeAngelis, G. C. & Freeman, R. D. (1990). Stereoscopic depth discrimination by the visual cortex: neurons ideally suited as disparity detectors. *Science, 249*, 1037-1041.

Pack, C., Grossberg, S., & Mingolla, E. (2001). A neural model of smooth pursuit control and motion perception by cortical area MST. *Journal of Cognitive Neuroscience*, *13*, 102-120.

Pack, C. C., Livingstone, M. S., Duffy, K. R., & Born, R. T. (2003). End-stopping and the aperture problem: two-dimensional motion signals in macaque V1. *Neuron*, *39*, 671-680.

Patla, A. E., Tomescu, S. S., & Ishac, M. G. A. (2004). What visual information is used for navigation around obstacles in a cluttered environment? *Canadian Journal of Physiological Pharmacology*, *82*, 682-692.

Perrone, J. A., & Stone, L. S. (1994). A model of self-motion estimation within primate extrastriate visual cortex. *Vision Research, 34*, 2917-2938.

Peterhans, E. (1997). Functional organization of area V2 in the awake monkey. *Cerebral Cortex*, *12,* 335-357.

Poggio, G. F., & Fischer, B. (1977). Binocular interaction and depth selectivity in striate and prestriate cortex of behaving rhesus monkey. *Journal of Neurophysiology, 40,* 1392-1405.

Ponce, C.F., Lomber, S.G., & Born, R. T. (2008). Integrating motion and depth via parallel pathways. *Nature Neuroscience, 11*, 216-223.

Poucet, B., Lenck-Santini, P. P., Hok, V., Save, E., Banquet, J. P., Gaussier, P., & Muller, R. U. (2004). Spatial navigation and hippocampal place cell firing: The problem of goal encoding. *Reviews in the Neurosciences*, *15,* 89-107.

Pouget, A., & Sejnowski, T. J. (1997). Spatial transformations in the parietal cortex using basis functions. *Journal of Cognitive Neuroscience*, *9,* 222-237.

Raiguel, S., Van Hulle, M. M., Xaio, D.-K., Marcar, V. L., Lagae, L., & Orban, G. A. (1997). Size and shape of receptive fields in the medial superior temporal area (MST) of the macaque. *Neuroreport, 8*, 2803-2808.

Royden, C. S. (1997). Mathematical analysis of motion-opponent mechanisms used in the determination of heading and depth. *Journal of the Optical Society of America A*, *14,* 2128-2143.

Royden, C. S. (2002). Computing heading in the presence of moving objects: a model that uses motion-opponent operators. *Vision Research, 42,* 3043-3058.

Royden, C. S., Banks, M. S., & Crowell, J. A. (1992). The perception of heading during eye movements. *Nature, 360*, 583-585.

Royden, C. S., Crowell, J. A., & Banks, M. S. (1994). Estimating heading during eye movements. *Vision Research, 34,* 3197-3214.

Royden, C. S. & Hildreth, E. C. (1996). Human heading judgments in the presence of moving objects. *Perception & Psychophysics, 58,* 836-856.

Salinas, E., & Abbott, L. F. (1995). Transfer of coded information from sensory to motor networks. *Journal of Neuroscience, 15,* 6461-6474.

Salinas, E., & Abbott, L. F. (1996). A model of multiplicative neural responses in parietal cortex. *Proceedings of the National Academy of Sciences, 93,* 11956-11961.

Salinas, E. & Thier, P. (2000). Gain modulation: A major computational principle of the central nervous system. *Neuron, 27,* 15-21.

Schwartz, E. L. (1977). Spatial mapping in the primary sensory projection: analytic structure and relevance to perception. *Biological Cybernetics, 25,* 181-194.

Schwartz, E. L. (1980). Computational anatomy and functional architecture of striate cortex: a spatial mapping approach to perceptual coding. *Vision Research, 10,* 645-669.

Shenoy, K. V., Crowell, J. A., & Andersen, R. A. (2002). Pursuit speed compensation in cortical area MSTd. *Journal of Neurophysiology, 88,* 2630-2647.

Sincich, L. C., & Horton, J. C. (2003). Independent projection streams from macaque striate cortex to the second visual area and middle temporal area. *Journal of Neuroscience, 23,* 5684-5692.

Squatrito, S., & Maioli, M. G. (1997). Encoding of smooth pursuit direction and eye position by neurons of area MSTd of macaque monkey. *Journal of Neuroscience, 17,* 3847-3860.

Stringer, S. M., Rolls, E. T., Trappenberg, T. P., & Araujo, I. E. T. (2002). Self-organizing continuous attractor networks and path integration: Two-dimensional models of place cells. *Network, 13,* 429-446.

Tanaka, K., & Saito, H. (1989). Analysis of motion 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 brain. *Journal of Neurophysiology, 62,* 626-641.

Tanaka, K., Sugita, Y., Moriya, M., & Saito, H.-A. (1993). Analysis of object motion in the ventral part of the medial superior temporal area of the macaque visual cortex. *Journal of Neurophysiology, 69,* 128-412.

Tian, J. R., & Lynch, J. C. (1996). Functionally defined smooth and saccadic eye movement subregions in the frontal eye field of Cebus monkey. *Journal of Neurophysiology, 76,* 2740-2753.

Thomas, O. M., Cumming, B. G., & Parker, A. J. (2002). A specialization for relative disparity in V2. *Nature Neuroscience, 5,* 472-478.

Touretzky, D. S., & Redish, A. D. (1996). Theory of rodent navigation based on interacting representations of space. *Hippocampus*, *6*, 247-270.

van den Berg, A. V. (1992). Robustness of perception of heading from optic flow. *Vision Research*, *32*, 1285-1296.

van den Berg, A. V. (1993). Perception of heading. *Nature*, *365*, 497-498.

von der Heydt, R., Zhou, H., & Friedman, H. S. (2002). Representation of stereoscopic edges in monkey visual cortex. *Vision Research*, *40*, 1955-1967.

Wagner, R. (2004). Stereo active vision and peripheral optical flow: computer vision applications of the wide-field human visual representation. PhD thesis, Boston University, Boston, MA,

Wallis, G. Chatziastros, A. & Bülthoff, H. (2002). An unexpected role for visual feedback in vehicle steering control. *Current Biology*, *12*, 295-299.

Warren, W. H., Blackwell, A. W., Kurtz, K. J., Hatsopoulos, N. G., & Kalish, M. L. (1991). On the sufficiency of the velocity field for perception of heading. *Biological Cybernetics*, *65*, 311-320.

Warren, W. H. & Hannon, D. J. (1988). Direction of self-motion is perceived from optical flow. *Nature*, *336*, 162-163.

Warren, W. H. & Hannon, D. J. (1990). Eye movements and optical flow. *Journal of the Optical Society of America A*, *7*, 160-169.

Warren, W. H., Kay, B. A., Zosh, W. D., Duchon, A. P., & Sahuc, S. (2001). Optic flow is used to control human walking. *Nature Neuroscience*, *4*, 213-216.

Warren, W. H., & Kurtz, K. J. (1992). The role of central and peripheral vision in perceiving the direction of self-motion. *Perception & Psychophysics*, *51*, 443-454.

Warren, W. H., Morris, M. W., & Kalish, M. (1988). Perception of translational heading from optic flow. *Journal of Experimental Psychology: Human Perception and Performance*, *14*, 646-660.

Wilkie, R. M., & Wann, J. P. (2002). Driving as night falls: the contribution of retinal flow and visual direction to the control of steering. *Current Biology*, *12,* 2014-2017.

Wilkie, R. M., & Wann, J. P. (2003a). Eye-movements aid the control of locomotion. *Journal of Vision*, *3*, 677-684.

Wilkie, R. M., & Wann, J. P. (2003b). Controlling steering and judging heading: Retinal flow, visual direction, and extraretinal information. *Journal of Experimental Psychology: Human Perception and Performance*, *29*, 363-378.

Wilkie, R. M., & Wann, J. P. (2005). The role of visual and nonvisual information in the control of steering. *Journal of Experimental Psychology: Human Perception and Performance*, *31*, 901-911.

Warren, W. H., & Saunders, J. A. (1995). Perceiving heading in the presence of moving objects. *Perception*, *24*, 315-331.

Zeki, S. M. (1971). Convergent input from the striate cortex (area 17) to the cortex of the superior temporal sulcus in the rhesus monkey. *Brain Research*, *28*, 338-340.

Zeki, S. M. (1974). Functional organization of a visual area in the posterior bank of the superior temporal sulcus of the rhesus monkey. *Journal of Physiology*, *236*, 549-573.

Zemel, R. S., & Sejnowski, T. J. (1998). A model for encoding multiple object motions and self-motion in area MST of primate visual cortex. *Journal of Neuroscience*, *18*, 531-547.

Zhang, T., Heuer, H. W., & Britten, K. H. (2004). Parietal area VIP neuronal responses to heading stimuli are encoded in head-centered coordinates. *Neuron*, *24*, 993-1001.

Zipser, D., & Andersen, R. A. (1988). A back-propagation programmed neural network that simulates response properties of a subset of posterior parietal neurons. *Nature*, *331*, 679-684.