

COMPUTATIONAL MECHANISMS FOR OPTIC FLOW ANALYSIS IN PRIMATE CORTEX

Markus Lappe

Department of Zoology and Neurobiology, Ruhr University Bochum, Bochum, Germany

- I. Introduction
- II. Foundations and Goals of Modeling
- III. Models of Optic Flow Processing in Primates
 - A. Models Based on Learning Rules
 - B. Template Matching Models
 - C. Differential Motion Parallax
 - D. Optimal Approximation: The Population Heading Map Model
- IV. Comparisons with Physiology: Optic Flow Representation in Area MT
- V. Comparisons with Physiology: Optic Flow Selectivity in Area MST
 - A. Selectivity for Multiple Optic Flow Patterns
 - B. Selectivity for the Location of the Focus of Expansion
 - C. Optic Flow Selectivity during Eye Movements: Integration of Visual and Extraretinal Signals
- VI. Receptive Fields of Optic Flow Processing Neurons
- VII. The Population Heading Map
 - A. Properties of the Population Heading Map
 - B. Analysis of Population Data from Area MST
- VIII. Conclusion
- References

I. Introduction

In the visual cortex of the primate, the information-processing steps necessary to analyze optic flow occur in a hierarchical system of specialized motion-sensitive areas. Computational models of optic flow processing that employ neural network techniques are useful to interpret the neuronal data obtained from these areas. This is important because behaviorally relevant parameters are not encoded in the single neuron activity but rather distributed across a neuronal population. The goal of such models is twofold. On the one hand, they have to reproduce physiological data from single unit studies, and ultimately strive to explain the mechanisms underlying the neuronal properties. On the other hand, they have to concern themselves with the generation of behavior and show how the properties of individual neurons in turn relate to psychophysical measurements at the system level.

During the last ten years, psychophysical studies have described how humans perceive their direction of heading from the patterns of optic flow and how the various visual and nonvisual cues are functionally combined to solve this task (see e.g. the chapter by van den Berg, this volume). At the same time, experimental physiological studies have provided an account of optic flow processing in the primate visual cortex (see chapters by Bremmer *et al.*, Andersen *et al.*, and Duffy, this volume). It is well documented that visual motion information proceeds from the primary visual cortex (V1) to the middle temporal area (MT) and then to the medial superior temporal (MST) area and other areas in the parietal cortex. Area MT contains a preprocessed representation of the optic flow field that is well suited to serve as a basis for flow-field analysis. Area MST subsequently analyzes the flow field to estimate self-motion. In the course of this transformation from local image motion to global self-motion, additional signals that support self-motion estimation are combined with the optic flow. These are oculomotor signals, retinal disparity, and vestibular signals. To understand the complex information processing that occurs along the pathway of motion analysis in the primate cortex, it is useful to complement single-unit neurophysiology and behavioral/psychophysical observations with theoretical and computational considerations. This requires the formulation and evaluation of biologically plausible models.

II. Foundations and Goals of Modeling

The primary goal of the modeling approaches described in this chapter is to understand at a computational level how neurons and neuronal populations analyze optic flow and contribute to the control and perception of self-motion. By mathematically devising biologically plausible models, one can summarize and formalize experimental findings, provide ways to test hypotheses about the function of neurons and cortical areas quantitatively, and develop unifying concepts of how the brain solves complex computational problems.

Computational models are important for interpreting neuronal data and formulating testable predictions. Hence they must directly interact with physiological experiments. Empirical findings are the basis on which models are built and constrain the elements that can be used in the construction of a model. A valid neurobiological model must capture as much of the physiological and anatomical properties of the structure it wants to model. To demonstrate that this is the case, it is necessary to compare the model to experimental findings. The comparison of model behavior with physiological data will therefore be a central part of this chapter.

However, a useful model must not only reproduce and predict neuronal properties but also show how these properties contribute to a behavioral function or task. The models that will be discussed here address the task of heading estimation from retinal flow. Heading estimation is an important part of successful goal-directed movement and is involved in many daily activities. Visual computation of heading from retinal flow is a difficult task. During straightforward self-movement, heading is indicated by the focus of expansion. In general, however, heading detection is more complicated than a simple search for the focus of expansion. This is because we perform smooth pursuit eye movements to track a visual target during self-motion. Such tracking eye movements induce retinal image slip. When combined with the optic flow during self-translation, image slip induced by eye rotation transforms the structure of the retinal flow field and obscures the focus of expansion (see Lappe and Hoffmann, this volume, for more information). The primate visual system is thus often confronted with a very complex flow pattern. It needs mechanisms to analyze efficiently retinal flow fields that are perturbed by eye movements.

Mathematically, the task of heading estimation can be formulated as follows. The visual motion seen by the eye of a moving observer can—like any rigid body motion—be described by its translational and rotational components. Each of these has in principle three degrees of freedom. Heading detection requires the determination of the direction of translation. This is a problem with many unknown parameters. These are the six degrees of freedom in the self-motion plus the distances of all elements of the flow field from the eye. The latter are involved because the visual motion of an element of the flow field depends on the parameters of the self-motion and on the distance of the element from the eye. For translational self-movements, the visual speed of each element scales with inverse distance. This is known as motion parallax. Motion parallax is an important cue to segregate translational from rotational motion because rotational motion induces equal angular speed in all image points, independent of distance. Accurate measurement of the retinal flow provides information to solve the heading task, namely the direction and speed of every moving point. This allows us to decompose the flow mathematically into translational and rotational components and to determine the direction of heading if more than six moving points are registered (Longuet-Higgins and Prazdny, 1980). Many computational algorithms have been developed for the computational of self-motion from optic flow (see Heeger and Jepson, 1992b, for an overview; see Sinclair *et al.*, 1994; Fermüller and Aloimonos, 1995, for more recent work). Several models that compute heading from optic flow using neuronal elements have been proposed. The next section provides an overview of different classes of models. The subsequent sections will compare these models to experimental findings from monkey neurophysiology.

III. Models of Optic Flow Processing in Primates

The models that are discussed in this paper consist of neuronlike elements that respond to optic flow and that can be compared to optic flow processing neurons in primate visual motion pathway. The typical model layout consists of two layers of neurons, representing areas MT and MST, respectively. The properties of the neurons in the second (MST) layer mainly depend on their synaptic connections with the first (MT) layer neurons. Different models can be distinguished by the way in which these neuronal elements and their connections are constructed.

A. MODELS BASED ON LEARNING RULES

One class of models uses learning rules that originate from artificial neural networks theory to specify synaptic connections. These are back-propagation networks (Hatsopoulos and Warren, 1991; Beardsley and Vaina, 1998) or unsupervised learning mechanisms (Zhang *et al.*, 1993; Wang, 1995; Zemel and Sejnowski, 1998). The synaptic connections between the neuronal elements are generated by repetitively presenting a learning set of optic flow stimuli as input to the model and each time adjusting the synaptic connections according to the learning rule. The properties of the second-layer neurons then depend on the choice of input flow fields and learning rule. For instance, the basic response properties of MST neurons to expansion, rotation, and translation of large-field random-dot patterns can directly be generated by presenting various combinations of such flow patterns using unsupervised learning techniques (Zhang *et al.*, 1993; Wang, 1995; Zemel and Sejnowski, 1998). However, the neurons in this case only learn to form associations between input patterns but not between the input and a function or behavior, as would be required for the determination of self-motion. Therefore, such models do not directly address the issue of how optic flow is analyzed nor in which way the neurons contribute to the processing of optic flow. Zemel and Sejnowski (1998) used a learning procedure to generate a sparse encoding of typical flow fields obtained from moving scenes. They then demonstrated that heading can be estimated from this encoding. This required a further computational step, however, which was not part of the original learning procedure. Hatsopoulos and Warren (1991) trained a back-propagation network to determine the location of the focus of expansion. This covers the case of heading direction during simple translation. However, this approach did not generalize to the case of combined observer transla-

tion and eye rotation. The models that are presented next address the task of heading estimation from retinal flow in the general case.

B. TEMPLATE MATCHING MODELS

Template matching models attempt to solve the task of heading estimation by constructing a priori (i.e., without learning) neurons that are tuned to individual optic flow patterns (Perrone, 1992; Perrone and Stone, 1994; Warren and Saunders, 1995). In these models, each neuron forms a template for a specific flow pattern; hence, these models are called template models. The response of an individual neuron in a template model depends on the match between the input flow field and the template of that neuron. Sensitivity for the direction of heading is obtained by building templates for all flow fields that could possibly occur for any given direction of heading. This immediately results in a major problem because infinitely many flow patterns could arise from a single direction of heading. This is because eye rotations and the structure of the visual environment modify the pattern of flow on the retina but leave the direction of heading unchanged. The original template model of Perrone (1992) hence suffered from the fact that an unrealistically large number of templates would be required. Later work has attempted to cut down the number of templates. Perrone and Stone (1994) opted to do this by constraining the eye movements. Their model considered only those eye movements which stabilize gaze on an environmental target. This is the most prominent natural oculomotor behavior. However, complete reliance on this constraint is not consistent with human psychophysical data (Crowell, 1997).

A different approach toward fewer templates was taken by van den Berg and Beintema (1997, 1998). Instead of constructing individual templates for any combination of observer translation and eye rotation, they proposed to approximate such templates by the combination of only two first-order templates. The first template would be tuned to pure observer translation. The second template represents the derivative of the first template with respect to the amount of rotation in the flow. Formally, this is equivalent to approximating a mathematical function by the first two terms of its Taylor series. The activity of the derivative template is used to compensate changes in the activity of the pure translation template when the eye rotates. Such a combination of templates is tuned to heading because it always prefers the same observer translation irrespective of eye rotation. The benefit of this approach clearly is that fewer templates are needed. However, the approximation induces systematic errors for high

rotation rates, because the compensation is done to the first order only. Yet, a dependence of the error on rotation rate is also often seen in human psychophysical data (van den Berg, this volume). The error can be overcome by the inclusion of an extraretinal eye movement signal. This signal modulates the activity of the derivative template and extends the effective range of rotations for which compensation is successful.

C. DIFFERENTIAL MOTION PARALLAX

Other models draw on research in computer vision algorithms for the recovery of camera motion. These models implement computer vision algorithms with physiologically plausible neural processing elements. Two computational procedures have received particular attention. They are presented in this and the next section. The first algorithm originated from work of Rieger and Lawton (1985) who determined heading from differential motion parallax. The differential-motion-parallax algorithm uses not the individual motion vectors in the optic flow but rather the difference vectors between adjacent flow vectors. During a combination of translation and rotation, these difference vectors always point toward the direction of heading, much like the optic flow vectors during pure translation point toward the focus of expansion. This procedure has been used to model human psychophysical data (Hildreth, 1992a, b). Recently it has also been put into a neurobiological framework (Royden, 1997). This framework proposes that MT neurons compute the motion parallax field by center-surround mechanisms. This appears possible because the center and the surround of the receptive field of MT neurons indeed show opposite motion sensitivities (Allman *et al.*, 1985; Raiguel *et al.*, 1995).

D. OPTIMAL APPROXIMATION: THE POPULATION HEADING MAP MODEL

Another technique often used in computer vision procedures employs optimization methods. In the case of self-motion, this means finding a set of motion parameters (translation and rotation) that optimally predict a measured flow field. Mathematically, this is achieved by minimizing the mean squared difference between the measured flow field and all flow fields constructed from any possible combination of observer translation and rotation (Bruss and Horn, 1983; Koenderink and van Doorn, 1987). Finding the self-motion parameters that minimize this difference is equivalent to finding the actual self-motion. However, because of motion parallax, any candidate flow field also depends on the 3-D

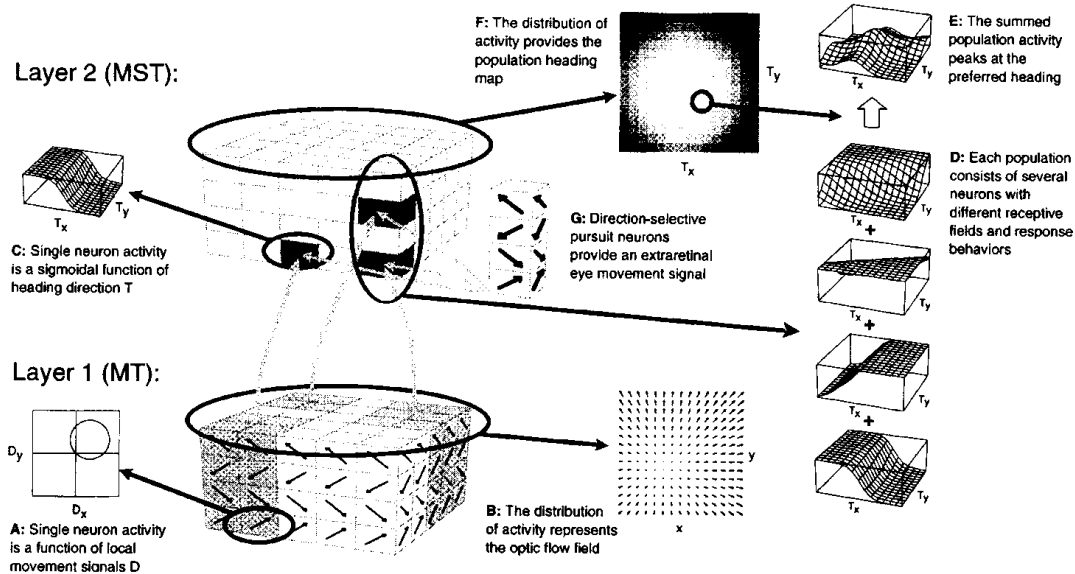


FIG. 1. The population heading map model (Lappe and Rauschecker, 1993a,b; Lappe *et al.*, 1996) consists of two layers of neurons. They correspond to areas MT and MST in monkey cortex. The first (MT) layer contains neurons that are selective to local speed and direction of motion (A). Their receptive fields are arranged in a retinotopic map. Each map position consists of a hypercolumn containing neurons with many different selectivities to local visual motion. Together, the neurons of one hypercolumn encode the visual motion at the corresponding part of the visual field. The distribution of the activities of all hypercolumns in the first layer represents the optic flow field (B).

The second (MST) layer contains neurons that analyze the optic flow and determine self-motion. This layer also contains a topographic map. But unlike the first layer, this is a map of heading directions. The map is constructed in two steps. First, the response of individual neurons to optic flow fields depends on the direction of heading. The dependency has a sigmoidal shape (C). Second, groups of neurons are collected into columns that each represent a specific heading. Each column of neurons receives input from different parts of the visual field (shaded areas in the first layer) and contains cells with different optic flow response properties (D). The activities of all neurons within one such column are summed into a population activity (E). The population activity is maximal when the preferred heading of the population and the true heading of the optic flow are the same. The distribution of the activities of all populations in the second layer provides a computational map of heading (F). The activity peak in this map signals the true heading of the observer.

Recently, this model has been modified to include extraretinal eye movement signals (Lappe, 1998). Pursuit neurons, which form a separate population in the MST layer (G), are active during pursuit eye movements and are selective for the direction of the eye movement. Their activity reflects an extraretinal (i.e., nonvisual) input such as an efference copy. This signal is fed into the optic flow processing neurons in the MST layer. The optic flow processing neurons hence can use both visual and extraretinal signals to compensate for eye movements in heading estimation.

structure of the visual scene. Therefore, the minimization must include not only all possible self-motion parameters but also all possible 3-D scene layouts. As with the template-matching method discussed earlier, this amounts to a very large number of possibilities—six degrees of freedom of the observer's motion plus one for each visible point in the scene. Fortunately, if one is only interested in determining the translational heading then the number of parameters can be dramatically reduced. Heeger and Jepson (1992b) have presented a modified version of the least-squares optimization method that is much more economical and can be easily implemented in parallel processing elements.

Based on this method, Lappe and Rauschecker (1993a, b) have developed a population-based neural network model of optic flow processing in areas MT and MST (Fig. 1). In this model, populations of optic flow processing neurons compute the mean-squared differences for many individualized headings in parallel. Each heading is represented by a small population of neurons. The activity of this population defines the momentary likelihood of this specific heading. This results in a heading map in which each position in the map contains a set of neurons that act together as a population, representing a predefined heading direction. No single neuron of a population can determine heading alone. Only the combination of individual neuron responses into the population gives the appropriate signal. Any individual neuron in a single population contributes only to part of the calculation. Therefore, the properties of a neuronal population in the map and of its constituent neurons are not directly equal. This is a distinctive feature of this model which will become important in Section VII.

IV. Comparisons with Physiology: Optic Flow Representation in Area MT

Area MT is the first area in the primate visual pathway that is dedicated specifically to the processing of motion. Most models assume that area MT forms the cortical representation of the optic flow field. Area MT contains a retinotopic map of the contralateral visual field (Albright and Desimone, 1987). Most neurons in area MT are direction- and speed-selective (Maunsell and Van Essen, 1983). Cells in area MT are organized in direction columns similar to the orientation columns in V1 (Albright *et al.*, 1984). The MT analogon of a V1-hypercolumn could provide all essential information for encoding local motion at a single visual field location. The combination of many such hypercolumns would encode the full optic flow field. Typically, models employ such

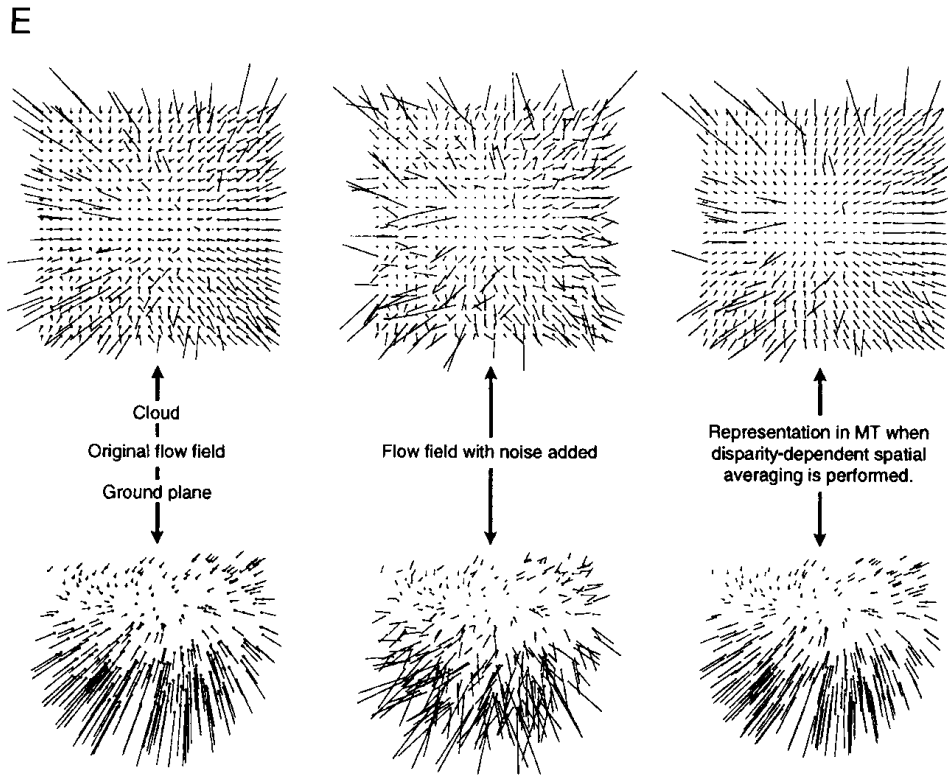
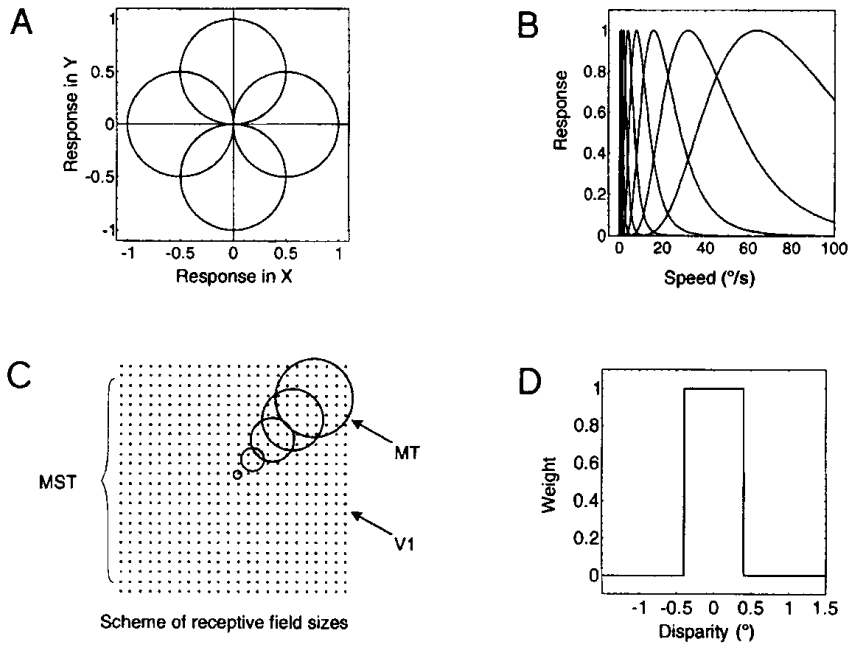
MT hypercolumns as the starting point of the representation of the flow field.

Modeling studies have shown how some aspects of the global organization of MT can be used for optic flow processing. These include properties of the retinotopic mapping in MT, the antagonistic center/surround organization of the receptive fields, and the disparity sensitivity of MT neurons.

A relation to optic flow is already apparent in the retinotopic mapping in area MT. Preferred speeds of MT neurons increase with the eccentricity of their receptive field (Maunsell and Van Essen, 1983) similarly to the way optic flow speeds naturally do. The number of direction-sensitive neurons preferring motion away from the fovea is significantly higher than the number of neurons preferring motion toward the fovea (Albright, 1989). This property is well adapted to the centrifugal structure of the flow field under natural self-motion conditions (Lappe and Rauschecker, 1994, 1995b).

The receptive fields of many MT neurons consist of a direction-selective central region complemented by an antagonistic surround (Allman *et al.*, 1985; Raiguel *et al.*, 1995). Surround motion in the same direction as in the center of the receptive field inhibits the neuron. Such an arrangement of selectivities could yield detectors that compute the local motion parallax field. From such detectors, heading could be estimated using the differential-motion-parallax algorithm (Royden, 1997) or the optimization algorithm of Heeger and Jepson (1992a), which forms the computational basis of the population heading map model.

A detailed model of the representation of visual motion in area MT showed that two further receptive field properties of MT neurons benefit the representation of optic flow (Lappe, 1996). First, receptive field sizes in area MT grow with eccentricity of the receptive field center (Albright and Desimone, 1987). Second, MT neurons respond to motion signals from within their receptive field in a disparity-selective way (Bradley *et al.*, 1995). The combination of these two factors provides an effective way to enhance the representation of the flow field in the presence of noise (Fig. 2). The extended receptive fields of MT neurons provide a spatial smoothing of the flow field which reduces motion noise. Because the structure of the flow field is very fine in the center but much coarser and more uniform in the periphery, it makes sense to vary the scale of the smoothing (the sizes of the receptive fields) with eccentricity. However, extensive smoothing of the flow field might remove signals that are necessary for heading direction. Especially important in this regard is motion parallax, i.e., the difference in speed of objects at different depths. Motion parallax carries important information to separate



translational and rotational flow components. Such a separation is required for heading detection during eye movements. A loss of motion parallax information because of extensive smoothing is therefore not desirable. Figure 2 shows how this unwanted effect is overcome when the spatial smoothing is made disparity-dependent, as it is in MT neurons. Spatial smoothing is performed only within depth planes. In this way, noise is reduced, while motion parallax information is retained. This model of disparity-dependent spatial averaging of the flow field can explain the enhanced robustness of heading perception from stereoscopic flow stimuli observed in humans (Lappe, 1996; van den Berg and Brenner, 1994).

V. Comparisons with Physiology: Optic Flow Selectivity in Area MST

Area MST receives major input from area MT and is thought to analyze optic flow and determine self-motion (Bremmer *et al.*, this volume; Anderson *et al.*, this volume; Duffy, this volume). Most effort in modeling optic flow processing has focussed on area MST and on comparing the properties of model neurons with those of neurons recorded from this area. In comparing the behavior of model neurons with their MST counterparts, several key properties of MST neurons must be consid-

FIG. 2. Model of the representation of the optic flow field in MT (Lappe, 1996). For each position in a retinotopic map, the model assumes a set of direction-selective neurons with different tuning properties. These contain four cosine direction tunings along the cardinal axes (A) and eight Gaussian speed tunings on a logarithmic scale between 0.5° and 64° (B). Direction and speed of local motion are determined from the activities of these neurons by a population code.

The size of receptive fields in area MT varies with eccentricity (Albright and Desimone, 1987). The model assumes that the response of a neuron is determined by the spatial average of the visual motion inside its receptive field (C). Consistent with electrophysiological findings (Bradley *et al.*, 1995), the response also depends on the disparity of motion signals. The contribution of each motion signal to the spatial summation in the receptive field is weighted according to its disparity (D). Only motion signals with a disparity close to the preferred disparity of a neuron contribute.

Such a representation is very robust against noise in the flow field (E). On the left, two flow fields which are experienced during a combination of observer translation and eye rotation are shown. The upper flow field depicts motion in a random 3-D environment (a cloud of random dots). The lower flow field depicts motion across a horizontal ground plane. The central column presents noisy versions of the same flow fields. When these noisy flow fields are processed by the MT model (right column), much of the noise is removed. The representation in MT is very close to the original flow patterns.

ered: the selectivity for several types of flow patterns such as expansion, contraction, rotation, and unidirectional motion (Duffy and Wurtz, 1991a) or spirals (Graziano *et al.*, 1994), the dependence of the response on the location of the focus of expansion (Duffy and Wurtz, 1995; Lappe *et al.*, 1996), the position invariance of the selectivity (Graziano *et al.*, 1994; Lagae *et al.*, 1994), and the combination with extraretinal eye movement signals (Erickson and Thier, 1991; Bradley *et al.*, 1996; Page and Duffy, 1999).

Several of these characteristics are also found in area VIP (Schaafsma and Duysens, 1996; Bremmer *et al.*, 1997, this volume). At present, it is not clear how these two areas differ with respect to optic flow processing. Here we will focus on comparisons between model and MST neurons. This is done with the understanding that similar comparisons can be made to VIP neurons. As more experimental findings will become available, differences between the two areas might become apparent.

A. SELECTIVITY FOR MULTIPLE OPTIC FLOW PATTERNS

When Tanaka and Saito (1989a, b) first systematically investigated the optic flow response properties of MST neurons, they used a set of “elementary” optic flow stimuli. This stimulus set consisted of unidirectional motion, rotation, expansion, and contraction. For each of these elementary flow fields, they found neurons that responded selectively to only a single one. Tanaka and Saito required a genuine optic-flow neuron to respond to expansion, contraction, or rotation but not to unidirectional motion. Later studies found that the majority of MST neurons responded to several different flow stimuli (Duffy and Wurtz, 1991a, b). Duffy and Wurtz proposed a classification into triple, double, and single component neurons, with the explicit understanding that this reflects a continuum of selectivities in MST. Triple-component neurons respond selectively to one direction of unidirectional motion, one sense of rotation, and either expansion or contraction. Double-component neurons respond to unidirectional motion and to either rotation or expansion/contraction. Very few neurons in Duffy and Wurtz’s study responded to rotation and expansion/contraction but lacked direction selectivity. The most selective, but least populous group consists of single-component neurons that responded only to one of the stimuli: 9% of all neurons respond to expansion or contraction; 4% respond to rotation. The predominance of triple- and double-component neurons and the relative scarcity of single-component neurons has since been confirmed in two subsequent studies (Duffy and Wurtz, 1995; Lappe *et al.*, 1996).

In models, such a variety of selectivities can arise in two conceptually different ways. Either the model begins with highly selective neurons that for instance respond only to pure expansions, and then adds mechanisms that induce sensitivity also for other flow patterns. This is the approach proposed by the template models (Perrone, 1992; Perrone and Stone, 1994). Or alternatively, the model starts with broadly selective neurons such as the triple-component neurons in MST and refines its selectivity successively until more selective response properties are reached. This is the approach of the optimization and learning models (Lappe and Rauschecker, 1993a, b; Lappe *et al.*, 1996; Zemel and Sejnowski, 1998). Probably in favor of the latter is the observation that many more MST neurons respond to multiple flow patterns than to only a single one. This would seem to suggest that these single-component neurons represent a higher degree of abstraction, which might be achieved by internal convergent connections in MST. Lappe *et al.* (1996) have suggested such a two-step convergence. A large number of broadly selective neurons are combined in populations in the heading map. Higher, more selective neurons could read-out the population heading map. This arrangement would have some analogy to the way complex cells in V1 are constructed from converging inputs of simple cells. It would result in a substantially smaller number of highly selective (single-component) neurons than broadly selective (triple- and double-component) neurons.

In the heading map model, different degrees of component selectivity can be obtained by assuming different selectivities for combinations of self-motion and eye movements (Lappe and Rauschecker, 1993a; Lappe *et al.*, 1996). This applies to triple- and double-component neurons. Single-component radial (expansion/contraction) neurons can be generated at a higher level by convergence. The template model of Perrone and Stone contains single-component radial and unidirectional neurons as well as one type of double-component selectivity (radial + unidirectional). But it fails to predict the other types of selectivity found in MST (Perrone and Stone, 1998), which is actually the majority of neurons. Both models, however, do not generate true single-component rotation selectivity.

Selective responses to many different flow patterns can be observed in MST. Yet, some flow patterns clearly appear more related to self-motion than others. It is easy to see why expansion selectivity could be required for heading detection. It corresponds to forward movement. This is much less clear for contraction selectivity (backward movement) or even rotation selectivity. Full field rotation would require to rotate the head around the axis of gaze which is unlikely to occur frequently in the

natural behavior of ground living animals. These neurons might be involved in tasks other than self-motion estimation (e.g., the perception of the motion of objects) (Graziano *et al.*, 1994; Zemel and Sejnowski, 1998). On the other hand, very many MST neurons combine responses to expansion and rotation (and to unidirectional motion), suggesting that such a combination might be useful also for self-motion processing. Some MST cells even prefer vectorial combinations of rotation, and expansion/contraction over the two separate patterns, and display a selectivity for spiral motion (Graziano *et al.*, 1994). The question arises as to why such combinations might be useful for optic flow processing.

Theoretical considerations and models have provided a possible answer to this question. It is based on the structure of naturally occurring flow patterns and the benefits of broad selectivity in a population code. Spiraling retinal motion patterns are quite common when the motion of the observer consists of combined self-movement and eye movement (Lappe *et al.*, 1998; Warren and Hannon, 1990; Lappe and Rauschecker, 1994). Thus, responses to spiraling flow patterns might be expected from neurons that process self-motion in the presence of eye movements. Therefore, most models of heading detection contain neurons that respond to spiraling flow patterns. When the selectivities of such neurons are broad, they also include selective responses for pure rotations as do the triple-component neurons in MST. However, these model neurons are not designed to respond selectively to rotational flow patterns. Rather the rotational or spiral selectivity is a consequence of their selectivity for heading in the presence of eye movements.

B. SELECTIVITY FOR THE LOCATION OF THE FOCUS OF EXPANSION

Another important question is the dependence of optic flow responses on spatial parameters of the stimulus. This concerns especially the position invariance of the selectivity to optic flow components and the sensitivity for the location of the focus of expansion. Position invariance describes the observation that many MST neurons preserve their selectivity for a small optic flow pattern even when this pattern is moved to another location within the receptive field of the neuron (Duffy and Wurtz, 1991a; Graziano *et al.*, 1994; Lagae *et al.*, 1994). This is a very illustrative feature of the complexity of the response behavior of MST neurons. It gives a clear demonstration that these neurons truly respond to entire optic flow patterns and not just to part of the local motions within the patterns. However, it has been troublesome for models of heading detection from optic flow. Shifting a radial motion pattern to a different location also

shifts the focus of expansion along with it. Yet, the response of any heading detection neuron should vary with the location of the focus of expansion. Models deal with this problem in two ways. The first is to observe that retaining a preference for expansion versus contraction at different positions in the receptive field does not imply that the response strength to expansion has to be equal at these positions. Hence a variation of response strength with the location of the focus of expansion can occur despite a conserved preference of expansion over contraction. Secondly, position invariance occurs more frequently for restricted variations of spatial location of small stimuli (Lagae *et al.*, 1994; Graziano *et al.*, 1994) and less often for larger stimuli and displacements (Duffy and Wurtz, 1991b; Lappe *et al.*, 1996). Positional invariance within restricted regions of the visual field is often found in models (Lappe *et al.*, 1996; Perrone and Stone, 1998; Zemel and Sejnowski, 1998).

The neural specificity for the location of the focus on expansion in full-field optic flow stimuli is more directly related to self-motion processing. Lappe *et al.* (1996) tested predictions of the population heading map model. They investigated neuronal sensitivity for the focus of expansion in area MST and compared the results to computer simulations of model neurons. In these experiments, large field (90° by 90°) computer-generated optic flow stimuli were presented. They simulated forward (expansion) and backward (contraction) self-motion in different directions with respect to a random cloud of dots in three-dimensional space. The dependence of neuronal responses on the focus of expansion were determined for up to 17 different locations. The response profile of an example MST neuron is shown in Fig. 3A. The neuronal responses to expansion and contraction vary smoothly with the position of the singular point, saturating as the focus of expansion is moved into the visual periphery. Responses to both expansion and contraction can be elicited by proper placement of the singular point. These experimental findings are predicted by computer simulations of a model neuron (Fig. 3B).

Different models predict different shapes of the response profile for variations of the focus of expansion. In the population heading map, sigmoidal response profiles like the ones depicted in Fig. 3 are used for the constituent neurons of the heading populations (Fig. 1D). They form the majority of cells in the model. Neurons that read out the population activities—necessarily a minority of model neurons—retain a different selectivity profile and show a peaked tuning curve for the focus of expansion (Fig. 1E). Indeed, some MST neurons have a peak-shaped response profile rather than a sigmoidal one (Duffy and Wurtz, 1995; Lappe *et al.*, 1996). The proportion of these neurons varies somewhat between studies. Lappe *et al.* (1996) reported very few peak-shaped responses,

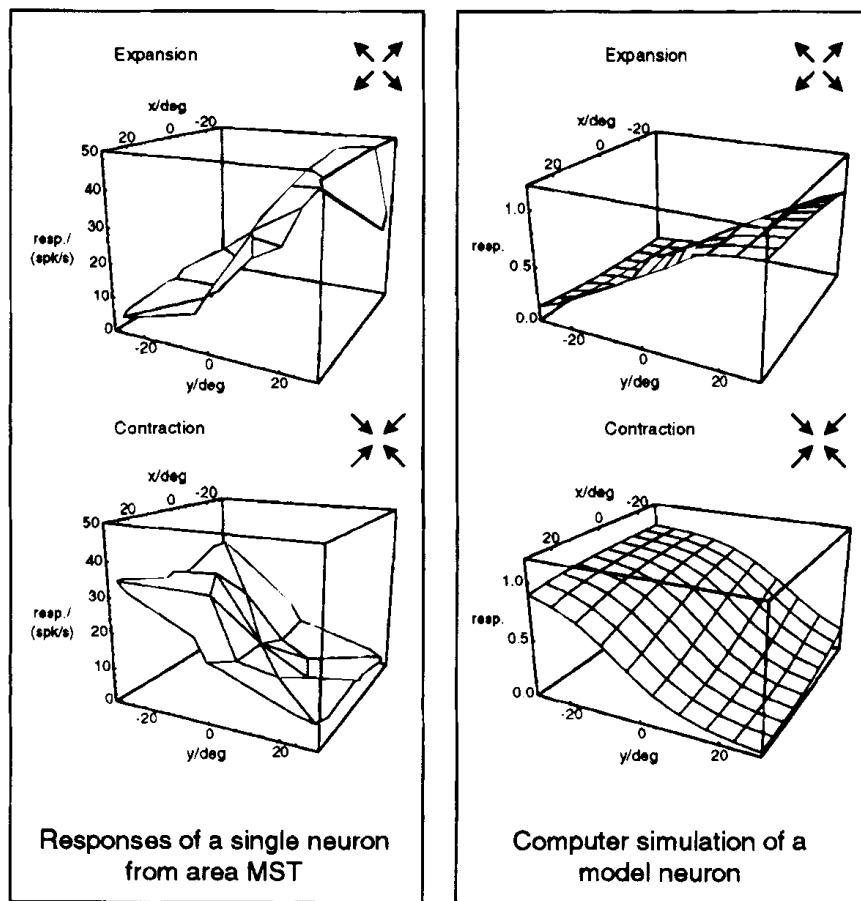


FIG. 3. Electrophysiological experiments in area MST (left) and computer simulations of a single neuron from the population heading map model (right) show very similar results (Lappe *et al.*, 1996). In both cases, the stimuli consisted of full field expanding or contracting optic flow patterns in which the retinal position of the focus was systematically varied. The modulations of response strength with the retinal position of the focus are displayed by 3-D surface plots. The (x, y) -plane represents the positions of the focus, the z -axis neuronal activity.

whereas the data of Duffy and Wurtz (1995) indicates a higher proportion. But in both studies such neurons are a minority. This is consistent with the predictions of the population model. The template model of Perrone and Stone (Perrone, 1992; Perrone and Stone, 1994) exclusively uses Gaussian (peak)-shaped tuning curves. This results in difficulties to account for the sigmoid tuning properties often observed in MST. Gaussian-tuned model neurons with centers very far in the periphery might resemble the sigmoidal tuning curves in MST (Perrone and Stone, 1998). However, in the model they would not be expected to make up

a substantial proportion of cells, as in fact they do in MST. The template model of van den Berg and Beintema (1997, 1998) arrives at one class of neurons with a tuning curve very similar to the sigmoidal tuned neurons described by Lappe *et al.* (1996). In their model, these neurons are required for the estimation of eye rotation and the compensation of rotational motion for heading detection.

C. OPTIC FLOW SELECTIVITY DURING EYE MOVEMENTS: INTEGRATION OF VISUAL AND EXTRARETINAL SIGNALS

For a complete and unambiguous representation of self-motion, optic flow needs to be combined with other sensory signals (van den Berg, this volume). The combination with stereoscopic disparity (Lappe, 1996) has already been mentioned. Even more important is the combination of optic flow with signals related to eye movements. Eye movements add visual motion to the optic flow field on the retina and severely complicate its analysis. Extraretinal eye movement signals can be used to overcome this problem. Interestingly, areas MT and MST have been shown to play a major role not only in optic flow analysis but also in the control of eye movements. Areas MT and MST are involved in the generation of smooth pursuit movements (Dürsteler and Wurtz, 1988; Komatsu and Wurtz, 1989), and in optokinetic (Hoffmann *et al.*, 1992; Ilg, 1997) and ocular following (Kawano *et al.*, 1994; Inoue *et al.*, 1998) reflexes. Eye-movement-related signals in area MST contain extraretinal components. Visual tracking neurons continue to fire during ongoing pursuit when the tracked target is briefly extinguished (Sakata *et al.*, 1983; Newsome *et al.*, 1988). They even fire during pursuit of a figure that never directly stimulates their visual receptive field (Ilg and Thier, 1997). Other MST neurons show a contrary behavior: they respond to motion of a visual stimulus during fixation but do not respond when the same retinal image motion is induced by an eye movement (Erickson and Thier, 1991). In these neurons, extraretinal eye movement information acts to differentiate true motion in the world from self-induced motion.

Could these extraretinal signals be used by optic flow processing neurons? Lappe *et al.* (1994, 1998) have argued that pursuit neurons in MST could provide a signal of current eye velocity that might allow optic flow processing neurons to compensate to some degree for the visual disturbance by the eye movement. In this model, the pursuit signal is used to subtract the eye-movement-induced visual motion component from the total retinal flow. However, this process is likely to be incomplete. For instance, the speed of eye movement seems to be less well rep-

resented by the pursuit neurons than its direction. Precise heading detection would still have to rely on the visual signal but would be supplemented by extraretinal compensation. Such a hybrid model can account for the most prevalent conditions in which human heading detection has been shown to rely on extraretinal input (Lappe *et al.*, 1994). Moreover it directly leads to neurons that implement the extraretinal influences observed in MST (Erickson and Thier, 1991) with unidirectional motion patterns (Lappe, 1998).

The use of pursuit signals can also be seen in actual recordings of MST neuronal responses to optic flow stimuli during ongoing smooth pursuit. Bradley *et al.* (1996) recorded selectivity of single MST cells to the focus of expansion while the monkey performed a smooth pursuit eye movement. These experiments compared the neuronal responses in two situations. In one condition, an expanding optic flow pattern was presented while the monkey tracked with his eyes a small dot that moved in the preferred or antipreferred pursuit direction of the neuron. In the second condition, the monkey was required to keep its eyes stationary and the optic flow pattern now included not only the expansion but also an overall translation. This combination of expansion and translation visually simulated the effects of the previously performed eye movement. Therefore, the two conditions presented identical visual but different extraretinal input. Many optic flow selective MST neurons responded differently in the two conditions. In the condition that included the active eye movement, the neurons responded to the focus of expansion of the stationary optic flow pattern, even though the eyes moved. Thus, an extraretinal signal in MST can compensate for the disturbances of the optic flow field due to eye movements. However, this compensation is far from complete for most neurons. On average, neurons compensated for only about 50% of the eye movement. Individual neurons strongly undercompensated or even in some cases overcompensated the eye-movement induced retinal motion.

The results of Bradley *et al.* have been reproduced in two models that are concerned with the combination of visual and extraretinal signals in MST (Lappe, 1998; Beintema and van den Berg, 1998). Figure 4 shows results from the model of Lappe (1998). The eye-movement-induced retinal image motion is compensated by extraretinal input. Then heading is determined from the compensated flow field. Different levels of compensation in individual neurons are modeled by varying the strength of the extraretinal signal. The model produces accurate heading judgments even with the incomplete compensation observed in MST. This is possible because the model uses a visual backup mechanism parallel to the extraretinal compensation. The model of van den Berg and Beintema

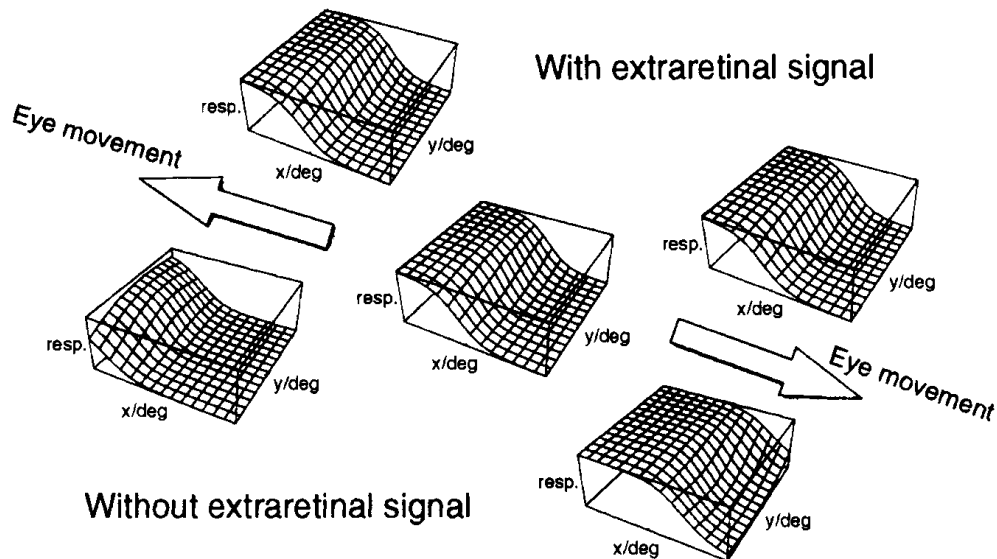


FIG. 4. A simulation of the experiment of Bradley *et al.* (1996). The plots show responses of an optic flow selective neuron to combined optic flow and simulated eye movement stimuli (after Lappe, 1998). The central plot shows the responses to pure expansional flow stimuli. It is dependent on the location (x, y) of the focus of expansion. The outer plots show the responses during eye movements. In the absence of an extraretinal signal, the response curve shifts along the direction of the eye movement. In contrast, when extraretinal input is available, the response curve stays fixed, independent of the eye movement. The response is the same as to an undisturbed pure expansion.

(1997, 1998) uses two types of templates. Retinal motion templates respond to the retinal position of the focus of expansion and only partially compensate for eye movements using the properties of the flow field. Heading templates receive an additional extraretinal eye movement signal and respond to the direction of heading thus compensating for eye movements. This model can also account for the broad range of extraretinal compensation strength in area MST.

The experiments of Bradley *et al.* (1996) tested the selectivity of MST neurons to heading along one axis of pursuit. Complete mappings of two-dimensional heading selectivity during pursuit in different directions were obtained by Page and Duffy (1999). Most MST neurons in their experiments showed significant pursuit influences on the tuning for heading. Many neurons responded to different headings during pursuit rather than during fixation. Page and Duffy concluded that individual MST neurons do not accurately compensate for the influences of pursuit and cannot directly signal heading. But correct heading estimates could be obtained from a population analysis. This is consistent with models that propose a population code for heading in area MST (Lappe and

Rauschecker, 1993b; Lappe *et al.*, 1996; Zemel and Sejnowski, 1998) but not with models that use single neurons as heading detectors (Perrone, 1992; Perrone and Stone, 1994; Beintema and van den Berg, 1998).

VI. Receptive Fields of Optic Flow Processing Neurons

Optic flow response properties of MST neurons must be generated at least in part from the combination of selectivities for local 2-D motion signals from area MT. The question is how the local 2-D motion sensitivities are organized within the receptive field of optic flow processing neurons. This is an important issue that has so far escaped a satisfactory answer. An analysis of the receptive field structure of neurons in models of optic flow processing might provide some help in understanding this relationship and guiding further experiments. It turns out that several models independently arrive at similar predictions for the structure of optic flow processing receptive fields.

Let us first consider the simple idea that the arrangement of local motion selectivities is directly equal to the preferred optic flow pattern. This is the basic assumption of early template models (Perrone, 1992; Perrone and Stone, 1994, 1998). They construct, for instance, expansion-selective cells by arranging 2-D motion sensitivities in a radial expanding pattern inside the receptive field. Experimental studies in MST have clearly disproved such a simple arrangement. Duffy and Wurtz termed it the "direction mosaic" hypothesis. They tested it directly by comparing optic flow selectivity to the selectivity to small 2-D motion stimuli in different parts of the receptive field (Duffy and Wurtz, 1991b). Lagae *et al.* (1994) performed a similar test. Clearly, for true expansion-selective cells, the 2-D motion selectivities in subparts of the receptive field did not match the hypothesis of a radial arrangement as proposed by the template models. According to the template model, single-component expansion cells possess the most basic and most simple tuning. In contrast, in the experimental data, the tuning of single-component neurons to optic flow is most difficult to explain directly from 2-D motion inputs in the subfields (Lagae *et al.*, 1994; Duffy and Wurtz, 1991b). It is somewhat easier for the triple-component neurons. These findings are again consistent with the view that single-component neurons form the highest level of optic flow analysis in MST, as proposed by the population model (Lappe and Rauschecker, 1993b; Lappe *et al.*, 1996).

Several other models make different predictions for the structure of the receptive fields of optic flow processing neurons. These models em-

ploy differences between flow vectors as the main computational element. Differences between flow vectors are useful because of the properties of motion parallax. For translational self-movement, points at different distances move at different visual speeds. For rotational (eye) movements, all image points move at the same angular speed, independent of their distance from the observer. Motion parallax is therefore a major cue to separate rotational and translational components in the flow field. Because the rotational component of all flow vectors is identical, the difference between any two flow vectors depends only on the translational component. For instance, differences between neighboring flow vectors can be used to compute a rotation-independent local-motion-parallax field which reconstructs the focus of expansion (Rieger and Lawton, 1985). This procedure is used in the models of Hildreth (1992a,b) and Royden (1997). It could be implemented by motion-selective neurons with center-surround opponent-motion selectivity which are found in area MT. Sensitivity for local opponent motions is also an important mechanism in the model of Beintema and van den Berg. The neurons that represent the derivative templates (i.e., the neurons that subserve the decomposition of translation and rotation) actually compute local motion parallax (Beintema and van den Berg, 1998).

Finally, the population heading map model, too, uses differences between flow vectors as its basic computations. The construction of the synaptic connections consists of comparisons between small groups of two to five flow vectors (Lappe and Rauschecker, 1995b). If these flow vectors are near each other, then the comparison is equivalent to an opponent motion detector (Heeger and Jepson, 1992a).

However, the usefulness of differences between flow vectors is not restricted to local motion parallax. One might argue that parallax information from widely separated regions of the visual field could often even be more useful because most local image regions contain only small depth variations (i.e., limited motion parallax). In the population heading map model, the comparisons need not be local. Comparisons might be performed between any two motion vectors. If these are far apart in the visual field, the computation becomes more complex, however, requiring the comparison of both the speed and direction of motion. Local opponent motion can therefore be regarded as a special case of a more global flow analysis in this model.

In summary, several current models contain similar operators despite the fact that they start out from very different computational approaches. This suggests that these operators reflect a common principle of optic flow processing. As of yet, the structure of the receptive field of flow field processing neurons in MST is unknown. It will be interesting

to see whether the prediction derived from these models can be found in the properties of real optic flow processing neurons.

VII. The Population Heading Map

In many areas of the brain, behaviorally relevant parameters are represented in the form of a topographic map. By analogy, such a map has been hypothesized also for the representation of heading (Lappe and Rauschecker, 1993b; Perrone and Stone, 1994; Duffy and Wurtz, 1995). Questions then arise about the structure of this map and the properties of its constituents. Area MST has only a very crude receptive field retinotopy (Desimone and Ungerleider, 1986). Instead, there is some indication that neurons with similar optic flow sensitivities are clustered together (Duffy and Wurtz, 1991a; Lagae *et al.*, 1994; Geesaman *et al.*, 1997; Britten, 1998). The model of Lappe and Rauschecker (1993b) was the first to propose a computational map of heading direction as the functional organization of flow-field-sensitive neurons in area MST. This map is based on a distributed representation of heading by populations of neurons. The population heading map has a number of features that will be discussed in this section.

A. PROPERTIES OF THE POPULATION HEADING MAP

The population heading map proposes a retinotopic encoding (or map) of heading. This is different from the typical retinotopy of the receptive fields in other visual areas. This retinotopic encoding associates each heading direction in visual field coordinates with a corresponding map location. Heading leftward of the direction of gaze is represented on the left and rightward to the direction of gaze, on the right. The preferred heading of an individual neuron (i.e., its place in the heading map) does not necessarily coincide with the spatial location of its receptive field. The receptive field consists of the manifold of all inputs to the neuron. The preferred heading, on the other hand, is the neuron's contribution to the functional organization of the map. The two must not be confused. A neuron-encoding leftward heading can have a receptive field in the right field of view. Yet, its place in the heading map is in the left hemifield. This dissociation is very important from the computational perspective because information about any given heading can be gathered from any part of the visual field. This is necessary to ensure that all visual input contributes to the determination of heading.

However, the retinotopic representation of a higher motion parameter such as heading precludes the existence of a retinotopic visual field representation in the same neural structure. This is consistent with the poor retinotopy of the receptive fields in MST (Desimone and Ungerleider, 1986) and the clustering of optic flow selectivities instead (Duffy and Wurtz, 1991a; Lagae *et al.*, 1994; Geesaman *et al.*, 1997).

The population heading map is composed not of individual neurons but of populations of neurons that jointly represent a specific heading. Thus, there is an additional computational step between the individual neuron's sensitivity to optic flow and the representation of the direction of heading. In this intermediate step, the responses of several neurons are combined into a population response. These neurons share the preference for the same direction of heading and occupy the same location in the population heading map. The sum of their respective responses, i.e., the population response, is proportional to the current likelihood of the preferred heading of this population. However, the optic flow sensitivities of the constituent neurons need not be all alike. For instance, different neurons receive input from different parts of the visual field (i.e., possess different receptive fields). More importantly, they also possess different response selectivities to optic-flow stimuli. This is evident from Fig. 1D. The response curves for heading are quite different for the four neurons. Yet, they share one heading to which all neurons of the population respond. This is the heading associated with the population and with the map location. The aggregation of the four activities into the population response yields a peak-tuning for the preferred heading.

There are two important consequences of the two-step process by which the map is constructed. First, it is the flanks of the response profiles of the individual neurons and not their respective best response that contribute to the heading map. Figure 1E shows that peak activity of the population coincides with the intersection of the response flanks of the individual neurons in Fig. 1D. Second, the behavior of individual neurons can be quite different from the behavior of the population. Both properties are important for relating experimental findings to model predictions and for an interpretation of the contribution of MST neuronal activity to the task of heading estimation.

B. ANALYSIS OF POPULATION DATA FROM AREA MST

We want to compare the features of the population heading map with the properties of optic flow processing in area MST. The first question is whether the MST population activity can estimate the direction

of heading. To demonstrate this, Lappe *et al.* (1996) computed maps of the location of the focus of expansion from the activity of a set of recorded neuronal responses. They used a least-mean-square minimization scheme to derive the position of the focus of expansion from the actual recording data. The procedure is similar in spirit to the population heading map. The grayscale maps of Fig. 5 present the mean-squared heading error for nine focus positions within the central 15° of the visual field. From it, the location of the focus can be retrieved with an average precision of 4.3° of visual angle. Clearly, the MST population provides enough information to locate the focus of expansion.

Recent results by Page and Duffy (1999) show that the same conclusion holds for heading estimation during eye movements. The responses of single neurons are not invariant against eye movements. But the population provides sufficient information to determine heading.

Further experimental support for a population heading map in area MST is provided by the analysis of a perceptual illusion in which the focus of expansion appears to shift in the field of view. This illusion occurs when a radially expanding random-dot pattern is transparently overlapped by unidirectional motion of a second pattern of random dots (Duffy and Wurtz, 1993). The focus of expansion is perceived not at its original location but displaced in the direction of the overlapping motion. This illusion is reproduced in computer simulations of the model (Lappe and Rauschecker, 1995a). Recently, Charles Duffy and I have compared the responses of individual MST neurons and of individual model neurons to the illusory stimulus (Lappe and Duffy, 1999). We used a paradigm in which two sets of optic flow stimuli were compared (Fig. 6A). One set contained the transparent overlap stimuli which give rise to the illusory shift of the focus of expansion. The other set contained stimuli in which the two motion patterns were vector summed. Vector summation yields a pure expansion pattern with a true eccentric focus. Hence in both cases the focus positions (true or illusory) are eccentric. However, these eccentric positions are 180° opposite for each pair from the two stimulus sets!

Different models predict different response behavior of individual neurons in this case. Template matching predicts that transparently overlapping unidirectional motion should cause neurons encoding the true center of expansion to stop firing, whereas neurons encoding the illusory center of expansion should start firing. The population model predicts that transparently overlapping unidirectional motion should cause graded changes in the responses of all neurons to alter the aggregate response in a manner that shifts the net center of motion. In the

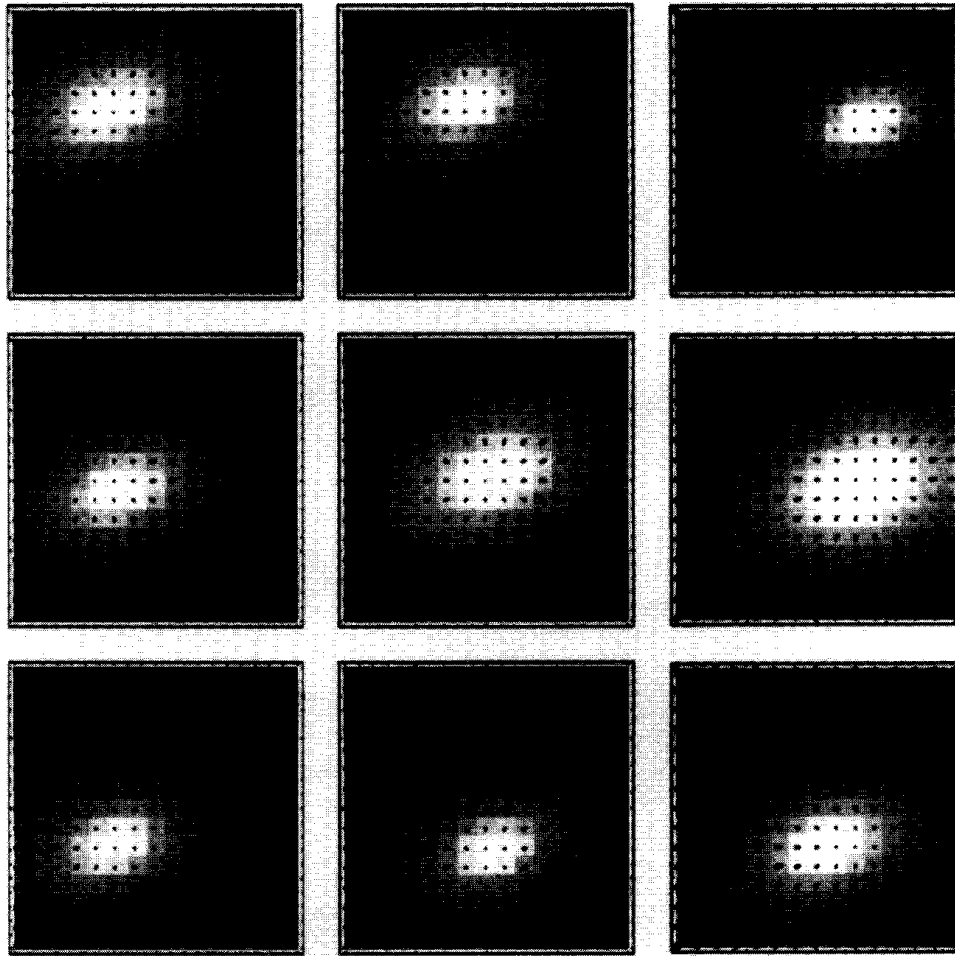
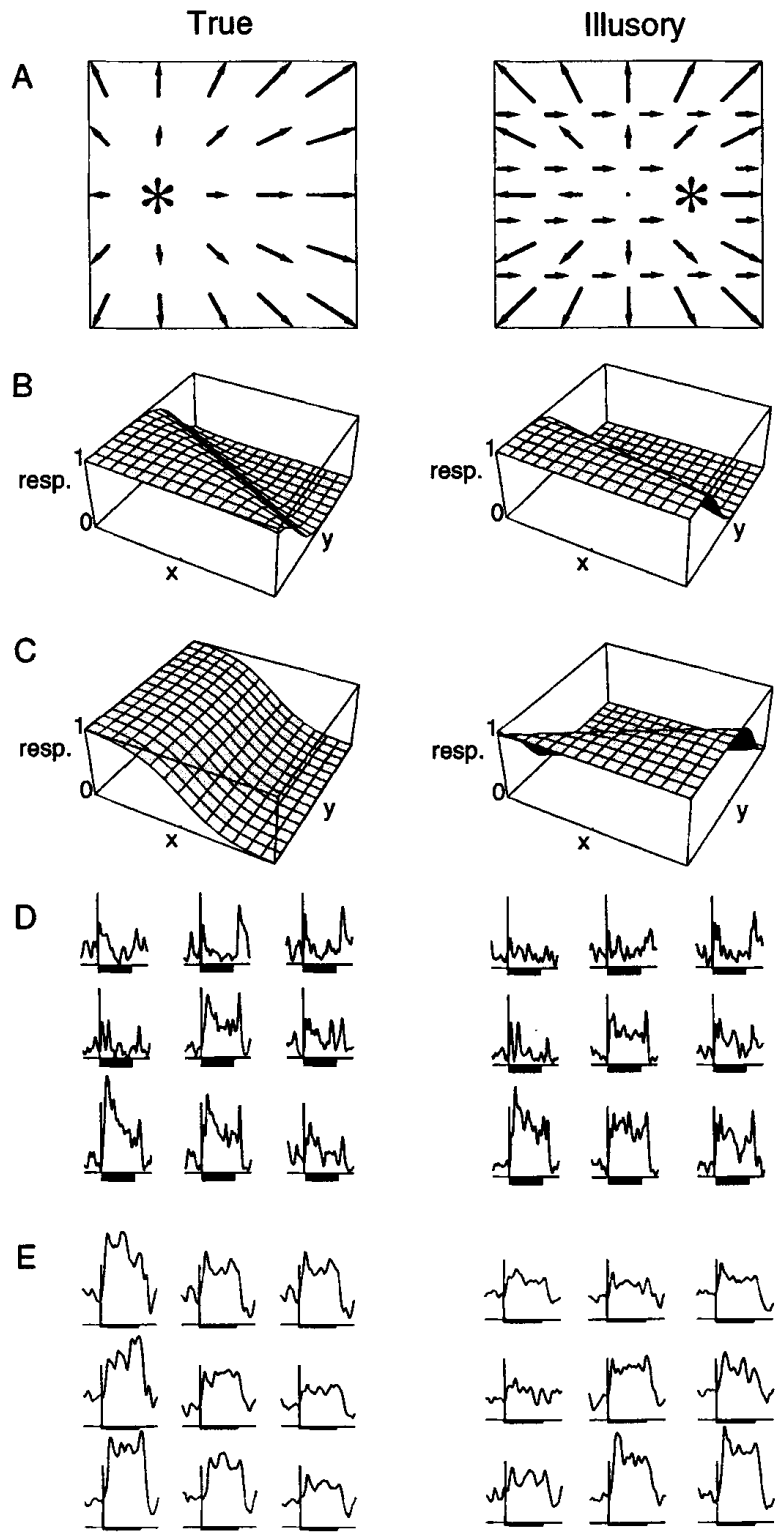


FIG. 5. Computational heading maps obtained from the recorded neuronal activities of 31 MST neurons (Lappe *et al.*, 1996). Each panel shows a retinotopic heading map. The brightness gives the likelihoods of each specific heading. For comparison with the true location of the focus of expansion, the optic flow stimuli are plotted on top of the grayscale maps.

population heading map, the behavior of single-model neurons is actually very different from the behavior seen in the human perception and at the population level in the model. Rather than shift or reverse their response behavior between the two types of stimuli, individual neurons rotate their response gradients with variable amounts of rotation (Figs. 6B and 6C). In the model, such a rotation of the response profiles of single neurons is sufficient to result in the observed shift of the focus of expansion at the population level. The population activity is derived from



the overlap of the response profiles of individual neurons. The shift of the population response is therefore the result of the combination of many individual response profiles (Fig. 7). Model neurons are grouped in populations that encode different directions of heading, i.e., different locations of the focus of expansion. For the vector-summed stimuli, which contain a true eccentric focus of expansion, the response profiles of all neurons in one such population are arranged such that the neuronal responses cohere maximally when the focus of expansion is at that position. In the transparent overlap condition, the individual response profiles are rotated. Maximum coherence now occurs for the opposite location in the visual field (i.e., at the illusory focus position).

We represented the vector sum and transparent overlap stimuli to 228 MST neurons and compared the results to the analysis of 300 model neurons using comparable stimuli (Lappe and Duffy, 1999). The model not only predicted the illusory shift of the center of expansion but also predicted the behavior of individual MST neurons (Figs. 6B–6E). The findings are compatible with a population code for heading but not with the prediction from template matching. For the illusory stimuli, the behavior of single MST neurons is qualitatively different from the perceptual findings. The perceived center of motion in the vector sum and the transparent overlap conditions shift, whereas the response profiles of

FIG. 6. Neuronal responses to true and illusory displacements of the focus of expansion (Lappe and Duffy, 1999). (A) The illusion occurs when a radial and a unidirectional motion pattern are presented simultaneously as a transparent overlapping motion (right panel) (Duffy and Wurtz, 1993). The focus of the expansion is perceived to be displaced in the direction of the overlapping unidirectional motion, which is to the right in this example. This asterisk indicates the perceived location of the focus of expansion. The illusory focus position is opposite the focus position obtained when the two motion patterns are simply vector summed (left). In this case, a pure expansion stimulus is generated with a true focus on the left. (B, C) Single-model neurons respond differently in the two conditions but their responses do not exhibit a similar shift. Neuron B responds maximally when the true center of expansion is located in the lower-left hemifield. In the illusory condition, the neuron responds strongest when upward motion is presented transparently overlapping (right). Perceptually, and also at the population level in the model, this motion pattern results in an upward displacement of the center of expansion, opposite from the location of the focus in the pure expansion stimuli. In contrast, the response profile of the neuron is merely slightly rotated between the two conditions. C shows an example of a larger rotation of the response profile but still not a complete reversal. (D, E) Spike density histograms show mean responses over six presentations of each stimulus for two different MST neurons. Neuron D responded similarly to the two stimulus sets, much like the model neuron in B. Neuron E shows different behavior. It responded best to left centers of motion in the pure expansion stimuli and for up- and leftward motion in the transparent overlap condition. The response behaviors of model and MST neurons are very similar.

single neurons in the two conditions are rotated against each other. The population heading map reconciles this apparent mismatch. It reproduces both the single neuron behavior and the perceptual shift. This shows that the graded response profile rotations observed in MST can provide enough modulation to the distribution of neural activity to induce the illusory shift. The population heading map suggests that the distribution of rotation angles, not the individual rotation of individual neurons, subserves the perceptual effect. We therefore compared the distribution of rotations of the response profiles of single MST neurons to that in the model. The distribution of rotation angles in MST closely matched the model prediction (Lappe and Duffy, 1999).

The illusory stimuli contain large differences in direction and speed of individual dot motions. This is similar to strong motion parallax. The rotation of the response profiles of single model neurons is related to their analysis of motion parallax. A similar response profile rotation in model and MST neurons is also observed with more realistic motion parallax stimuli which simulate self-motion in an depth-rich environment containing (Lappe *et al.*, 1996; Pekel *et al.*, 1996). Motion parallax is an important cue for compensating eye-movement-induced perturbations of the retinal flow. Such perturbations displace the retinal projection of the center of expansion (see van den Berg, this volume, or Lappe and Hoffmann, this volume). Model neurons use motion parallax to shift the perceived center of expansion (the population response) back to the correct position. The magnitude of the illusory shift in humans is consistent with such a visual compensation mechanism (Lappe and Rauschecker, 1995a). The similarity between the behavior of single MST and model neurons could suggest that MST can use motion parallax cues to compensate for eye movements, but this compensation occurs at the population level.

FIG. 7. Schematic illustration of the population encoding of the perceived shift of the center of expansion (Lappe and Duffy, 1999). We consider four individual neurons that form a single population tuned to leftward heading (see also Figs. 1D and 1E). (A) Individual response profiles for true displacements of the focus of expansion. Grayscale maps correspond to viewing 3-D surface plots as in Fig. 6 from above. Brightness represents response activity. (B) Population response profile obtained from summing the responses of the four neurons. The individual response profiles in A are differently oriented such that they maximally overlap at a position left of the center. The population reaches peak activity when the center of motion is at that point (i.e., heading is to the left). (C) Response profiles of the same four neurons for a presentation of the illusory stimuli. Individual response gradients are rotated by different amounts. (D) The summation results in a different population response profile. Maximum activity (i.e., optimum overlap of the rotated profiles) now occurs on the right.

The population heading map might also explain puzzling results of microstimulation in MST (Britten and van Wezel, 1998). A monkey had to discriminate between leftward and rightward heading in optic-flow stimuli. Electrical stimulation was applied in order to influence heading judgments and thereby demonstrate an involvement of area MST in this task. Stimulation at sites within area MST that contained many neurons which responded best to the same heading (e.g., leftward) biased the behavioral choices of the monkey toward the predicted (leftward) heading in only 67% of the cases. In the remaining 33%, there was instead a bias in the opposite direction (Britten and van Wezel, 1998). How can the population heading map explain such a result? The prediction of the effect of microstimulation in the experiment was estimated from the preferences of the stimulated neurons. When several neurons within a recording penetration respond strongest to leftward heading, it was assumed that this site only contributed to the percept of leftward heading. However, neuronal populations in the heading map are not formed by neurons with the same best response (the neurons in Fig. 1D clearly have different best responses). Instead they are formed by neurons with optimum overlap of their response profiles. In the population heading map, the effects of microstimulation cannot be predicted from observations of the tuning of individual neurons but only from the populations. Microstimulation in the population heading map could shift the population response in a direction different from the best response of individual neurons. When only two directional choices are available (left/right), a considerable amount of responses in the apparently opposite direction would be expected. Further experiments will establish more precisely this global organization of optic flow processing in population of neurons.

VIII. Conclusion

The neuronal analysis of optic flow is a task which can only be understood by a multidisciplinary approach combining theoretical and experimental work. A number of models of optic flow processing in primate cortex that can replicate the basic response properties of neurons in areas MT and MST have been proposed. Several of these models appear to use similar basic mechanisms for the analysis of motion parallax. They differ, however, in the way by which these basic measurements are combined. Experimental findings together with modeling considerations suggest that heading in area MST is represented by a population map.

Acknowledgments

I am grateful to Jaap Beintema for comments on the manuscript and to the Human Frontier Science Program and the Deutsche Forschungsgemeinschaft for financial support.

References

- Albright, T. D. (1989). Centrifugal directionality bias in the middle temporal visual area (MT) of the macaque. *Vis. Neurosci.* **2**, 177–188.
- Albright, T. D., and Desimone, R. (1987). Local precision of visuotopic organization in the middle temporal area (MT) of the macaque. *Exp. Brain Res.* **65**, 582–592.
- Albright, T. D., Desimone, R., and Gross, C. G. (1984). Columnar organization of directionally selective cells in visual area MT of the macaque. *J. Neurophysiol.* **51**, 16–31.
- Allman, J. M., Miezin, F., and McGuinness, E. (1985). Stimulus specific responses from beyond the classical receptive field: Neurophysiological mechanisms for local-global comparisons in visual neurons. *Ann. Rev. Neurosci.* **8**, 407–430.
- Beardsley, S. A., and Vaina, L. M. (1998). Computational modelling of optic flow selectivity in MSTd neurons. *Network* **9**, 467–493.
- Beintema, J., and van den Berg, A. V. (1998). Heading detection using motion templates and eye velocity gain fields. *Vision Res.* **38**, 2155–2179.
- Bradley, D., Maxwell, M., Andersen, R., Banks, M. S., and Shenoy, K. V. (1996). Mechanisms of heading perception in primate visual cortex. *Science* **273**, 1544–1547.
- Bradley, D., Qian, N., and Andersen, R. (1995). Integration of motion and stereopsis in middle temporal cortical area of macaques. *Nature* **373**, 609–611.
- Bremmer, F., Duhamel, J.-R., Ben Hamed, S., and Graf, W. (1997). The representation of movement in near extrapersonal space in the macaque ventral intraparietal area (VIP). In: "Parietal Lobe Contributions to Orientation in 3D-Space," (P. Thier and H.-O. Karnath, Eds.), vol. 25 of Exp. Brain Res. Ser., pp. 619–630. Springer, Heidelberg.
- Britten, K. H. (1998). Clustering of response selectivity in the medial superior temporal area of extrastriate cortex in the macaque monkey. *Vis. Neurosci.* **15**, 553–558.
- Britten, K. H., and van Wezel, R. J. A. (1998). Electrical microstimulation of cortical area MST biases heading perception in monkeys. *Nat. Neurosci.* **1**, 59–63.
- Bruss, A. R., and Horn, B. K. P. (1983). Passive navigation. *Comp. Vis. Graph Image Proc.* **21**, 3–20.
- Crowell, J. (1997). Testing the Perrone and Stone (1994) model of heading estimation. *Vision Res.* **37**, 1653–1671.
- Desimone, R., and Ungerleider, L. G. (1986). Multiple visual areas in the caudal superior temporal sulcus of the macaque. *J. Comp. Neurol.* **248**, 164–189.
- Duffy, C. J., and Wurtz, R. H. (1991a). Sensitivity of MST neurons to optic flow stimuli. I. A continuum of response selectivity to large-field stimuli. *J. Neurophysiol.* **65**, 1329–1345.
- Duffy, C. J., and Wurtz, R. H. (1991b). Sensitivity of MST neurons to optic flow stimuli. II. Mechanisms of response selectivity revealed by small-field stimuli. *J. Neurophysiol.* **65**, 1346–1359.
- Duffy, C. J., and Wurtz, R. H. (1993). An illusory transformations of optic flow fields. *Vision Res.* **33**, 1481–1490.

- Duffy, C. J., and Wurtz, R. H. (1995). Response of monkey MST neurons to optic flow stimuli with shifted centers of motion. *J. Neurosci.* **15**, 5192–5208.
- Dürsteler, M. R., and Wurtz, R. H. (1988). Pursuit and optokinetic deficits following chemical lesions of cortical areas MT and MST. *J. Neurophysiol.* **60**, 940–965.
- Erickson, R. G., and Thier, P. (1991). A neuronal correlate of spatial stability during periods of self-induced visual motion. *Exp. Brain Res.* **86**, 608–616.
- Fermüller, C., and Aloimonos, Y. (1995). Direct perception of three-dimensional motion from patterns of visual motion. *Science* **270**, 1973–1976.
- Geesaman, B. J., Born, R. T., Andersen, R. A., and Tootell, R. B. H. (1997). Maps of complex motion selectivity in the superior temporal cortex of the alert macaque monkey: A double-label 2-deoxyglucose study. *Cereb. Cortex* **7**, 749–757.
- Graziano, M. S. A., Andersen, R. A., and Snowden, R. (1994). Tuning of MST neurons to spiral motions. *J. Neurosci.* **14**(1), 54–67.
- Hatsopoulos, N. G., and Warren, Jr., W. H. (1991). Visual navigation with a neural network. *Neural Networks* **4**(3), 303–318.
- Heeger, D. J., and Jepson, A. (1992a). Recovering observer translation with center-surround motion-opponent mechanisms. *Invest. Ophthalmol. Vis. Sci. Suppl.* **32**, 823.
- Heeger, D. J., and Jepson, A. (1992b). Subspace methods for recovering rigid motion I: Algorithm and implementation. *Int. J. Comput. Vision* **7**, 95–117.
- Hildreth, E. C. (1992a). Recovering heading for visually-guided navigation. *Vision Res.* **32**, 1177–1192.
- Hildreth, E. C. (1992b). Recovering heading for visually-guided navigation in the presence of self-moving objects. *Philos. Trans. Roy. Soc. Lond. B* **337**, 305–313.
- Hoffmann, K.-P., Distler, C., and Ilg, U. (1992). Callosal and superior temporal sulcus contributions to receptive field properties in the macaque monkey's nucleus of the optic tract and dorsal terminal nucleus of the accessory optic tract. *J. Comp. Neurol.* **321**, 150–162.
- Ilg, U. J., and Thier, P. (1997). MST neurons are activated by pursuit of imaginary targets. In: "Parietal Lobe Contributions to Orientation in 3D-Space," P. Thier, and H.-O. Karnath, Eds.), pp. 173–184. Springer, Berlin, Heidelberg, New York.
- Ilg, U. (1997). Responses of primate area MT during the execution of optokinetic nystagmus and afternystagmus. *Exp. Brain Res.* **113**, 361–364.
- Inoue, Y., Takemura, A., Kawano, K., Kitama, T., and Miles, F. A. (1998). Dependence of short-latency ocular following and associated activity in the medial superior temporal area (MST) on ocular vergence. *Exp. Brain Res.* **121**, 135–144.
- Kawano, K., Shidara, M., Watanabe, Y., and Yamane, S. (1994). Neural activity in cortical area MST of alert monkey during ocular following responses. *J. Neurophysiol.* **71**, 2305–2324.
- Koenderink, J. J., and van Doorn, A. J. (1987). Facts on optic flow. *Biol. Cybern.* **56**, 247–254.
- Komatsu, H., and Wurtz, R. H. (1989). Modulation of pursuit eye movements by stimulation of cortical areas MT and MST. *J. Neurophysiol.* **62**, 31–47.
- Lagae, L., Maes, H., Raiguel, S., Xiao, D.-K., and Orban, G. A. (1994). Responses of macaque STS neurons to optic flow components: A comparison of areas MT and MST. *J. Neurophysiol.* **71**, 1597–1626.
- Lappe, M. (1996). Functional consequences of an integration of motion and stereopsis in area MT of monkey extrastriate visual cortex. *Neural Comp.* **8**, 1449–1461.
- Lappe, M. (1998). A model of the combination of optic flow and extraretinal eye movement signals in primate extrastriate visual cortex. *Neural Networks* **11**, 397–414.
- Lappe, M., Bremmer, F., and Hoffmann, K.-P. (1994). How to use non-visual information

- for optic flow processing in monkey visual cortical area MSTd. *In: ICANN 94—Proceedings of the International Conference on Artificial Neural Networks*, 26–29 May 1994, Sorrento, (M. Marinaro and P. G. Morasso, Eds.), pp. 46–49. Springer, Berlin, Heidelberg, New York.
- Lappe, M., Bremmer, F., Pekel, M., Thiele, A., and Hoffmann, K.-P. (1996). Optic flow processing in monkey STS: A theoretical and experimental approach. *J. Neurosci.* **16**, 6265–6285.
- Lappe, M., & Duffy, C. (1999). Optic flow illusion and single neuron behavior reconciled by a population model. *Eur. J. Neurosci.* **11**, 2323–2331.
- Lappe, M., Pekel, M., and Hoffmann, K.-P. (1988). Optokinetic eye movements elicited by radial optic flow in the macaque monkey. *J. Neurophysiol.* **79**, 1461–1480.
- Lappe, M., and Rauschecker, J. P. (1993a). Computation of heading direction from optic flow in visual cortex. *In: “Advances in Neural Information Processing Systems,”* vol. 5 (C. L. Giles, S. J. Hanson, and J. D. Cowan, Eds.), pp. 433–440. Morgan Kaufmann, San Mateo, CA.
- Lappe, M., and Rauschecker, J. P. (1993b). A neural network for the processing of optic flow from ego—motion in man and higher mammals. *Neural Comp.* **5**, 374–391.
- Lappe, M., and Rauschecker, J. P. (1994). Heading detection from optic flow. *Nature* **369**, 712–713.
- Lappe, M., and Rauschecker, J. P. (1995a). An illusory transformation in a model of optic flow processing. *Vision Res.* **35**, 1619–1631.
- Lappe, M., and Rauschecker, J. P. (1995b). Motion anisotropies and heading detection. *Biol. Cybern.* **72**, 261–277.
- Longuet-Higgins, H. C., and Prazdny, K. (1980). The interpretation of a moving retinal image. *Proc. Roy. Soc. Lond. B* **208**, 385–397.
- Maunsell, J. H. R., and Van Essen, D. C. (1983). Functional properties of neurons in middle temporal visual area of the macaque monkey. I. Selectivity for stimulus direction, speed, and orientation. *J. Neurophysiol.* **49**(5), 1127–1147.
- Newsome, W. T., Wurtz, R. H., and Komatsu, H. (1988). Relation of cortical areas MT and MST to pursuit eye movements. II. Differentiation of retinal from extraretinal inputs. *J. Neurophysiol.* **60**(2), 604–620.
- Page, W. K., and Duffy, C. J. (1999). MST neuronal responses to heading direction during pursuit eye movements. *J. Neurophysiol.* **81**, 596–610.
- Pekel, M., Lappe, M., Bremmer, F., Thiele, A., and Hoffmann, K.-P. (1996). Neuronal responses in the motion pathway of the macaque to natural optic flow stimuli. *NeuroReport* **7**, 884–888.
- Perrone, J. A. (1992). Model for the computation of self-motion in biological systems. *J. Opt. Soc. Am* **9**, 177–194.
- Perrone, J. A., and Stone, L. S. (1994). A model of self-motion estimation within primate extrastriate visual cortex. *Vision Res.* **34**, 2917–2938.
- Perrone, J. A., and Stone, L. S. (1998). Emulating the visual receptive field properties of MST neurons with a template model of heading estimation. *J. Neurosci.* **18**, 5958–5975.
- Raiguel, S., Van Hulle, M., Xiao, D., Marcar, V., and Orban, G. (1995). Shape and spatial distribution of receptive fields and antagonistic motion surrounds in the middle temporal area (V5) of the macaque. *Eur. J. Neurosci.* **7**, 2064–2082.
- Rieger, J. H., and Lawton, D. T. (1985). Processing differential image motion. *J. Opt. Soc. Am. A* **2**, 354–360.
- Royden, S. C. (1997). Mathematical analysis of motion-opponent mechanisms used in the determination of heading and depth. *J. Opt. Soc. Am. A* **14**, 2128–2143.

- Sakata, H., Shibutani, H., and Kawano, K. (1983). Functional properties of visual tracking neurons in posterior parietal association cortex of the monkey. *J. Neurophysiol.* **49**, 1364–1380.
- Schaafsma, S., and Duysens, J. (1996). Neurons in the ventral intraparietal area of awake macaque monkey closely resemble neurons in the dorsal part of the medial superior temporal area in their responses to optic flow patterns. *J. Neurophysiol.* **76**, 4056–4068.
- Sinclair, D., Blake, A., and Murray, D. (1994). Robust estimating of egomotion from normal flow. *Int. J. Computer Vision* **13**, 57–69.
- Tanaka, K., and Saito, H.-A. (1989a). 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 monkey. *J. Neurophysiol.* **62**, 626–641.
- Tanaka, K., and Saito, H.-A. (1989b). Underlying mechanisms of the response specificity of expansion/contraction and rotation cells in the dorsal part of the medial superior temporal area of the macaque monkey. *J. Neurophysiol.* **62**, 642–656.
- van den Berg, A. V., and Beintema, J. (1997). Motion templates with eye velocity gain fields for transformation of retinal to head centric flow. *NeuroReport* **8**, 835–840.
- van den Berg, A. V., and Brenner, E. (1994). Why two eyes are better than one for judgements of heading. *Nature* **371**, 700–702.
- Wang, R. (1995). A simple competitive account of some response properties of visual neurons in area MSTd. *Neural Comp.* **7**, 290–306.
- Warren, W. H., and Saunders, J. A. (1995). Perceiving heading in the presence of moving objects. *Perception* **24**, 315–331.
- Warren, W. H., and Hannon, D. J. (1990). Eye movements and optic flow. *J. Opt. Soc. Am. A* **7**, 160–169.
- Zemel, R. S., and Sejnowski, T. J. (1998). A model for encoding multiple object motions and self-motion in area MST of primate visual cortex. *J. Neurosci.* **18**, 531–547.
- Zhang, K., Sereno, M. I., and Sereno, M. E. (1993). Emergence of position-independent detectors of sense of rotation and dilation with hebbian learning: An analysis. *Neural Comp.* **5**, 597–612.