PERGAMON

Neural Networks 11 (1998) 397–414

Neural Networks

Contributed article

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

Markus Lappe*

Department of Zoology and Neurobiology, Ruhr University, Bochum 44780, Germany

Received 27 December 1996; accepted 7 January 1998

Abstract

The determination of the direction of heading from optic flow is a complicated task. To solve it the visual system complements the optic flow by non-visual information about the occurrence of eye movements. Psychophysical studies have shown that the need for this combination depends on the structure the visual scene. In a depth-rich visual environment motion parallax can be exploited to differentiate self-translation from eye rotation. In the absence of motion parallax, i.e. in the case of movement towards a frontoparallel plane, extraretinal signals are necessary for correct heading perception (Warren and Hannon, 1990). Lappe and Rauschecker (1993b) have proposed a model of visual heading detection that reproduces many of the psychophysical findings in the absence of extraretinal input and links them to properties of single neurons in the primate visual cortex. The present work proposes a neural network model that integrates extraretinal signals into this network. The model is compared with psychophysical and neurophysiological data from experiments in human and non-human primates. The combined visual/extraretinal model reproduces human behavior in the case of movement towards a frontoparallel plane. Single model neurons exhibit several similarities to neurons from the medial superior temporal (MST) area of the macaque monkey. Similar to MST cells (Erickson and Thier, 1991) they differentiate between self-induced visual motion that results from eye movements in a stationary environment, and real motion in the environment. The model predicts that this differentiation can also be achieved visually, i.e. without extraretinal input. Other simulations followed experiments by Bradley et al. (1996), in which flow fields were presented that simulated observer translation towards a frontoparallel plane plus an eye rotation. Similar to MST cells, model neurons shift their preference for the focus of expansion along the direction of the eye movement when extraretinal input is not available. They respond to the retinal location of the focus of expansion which is shifted by the eye movement. In the presence of extraretinal input the preference for the focus of expansion is largely invariant to eye movements and tied to the location of the focus of expansion with regard to the visual scene. The model proposes that extraretinal compensation for eye movements need not be perfect in single neurons to achieve accurate heading detection. It thereby shows that the incomplete compensation found in most MST neurons is sufficient to explain the psychophysical data.

Keywords: Modelling; Visual motion; Self-motion; Heading; Optic flow; Cortex; Monkey; Human

1. Introduction

In higher mammals and humans, the perception of space and of motion through space is largely based on vision. Movements of the eyes in the head massively change the visual input though we continue to perceive a stable world around us. Questions about the use of extraretinal signals for the perception of such spatial stability during eye movements have long been raised and much discussed. Two classical theories present the two opposing approaches to this issue. The 'inferential' theory proposes that, in order to attain stability of the visual world during eye movements, an extraretinal signal, e.g. an efference copy of the motor command driving the eye muscles, is used to compensate for the visual effects of the eye movement (von Helmholtz, 1896; Mittelstaedt, 1990; Sperry, 1950: Von Holst and Mittelstaedt, 1950). An opposite point of view is taken by the theory of 'direct perception' (Gibson, 1950). In this view, no extraretinal signal is required, and the perception of the movement of objects in the world or of the observer himself is solely mediated from the visual information alone. In recent years, however, it has become evident that both theoretical views were for one reason or another

* Requests for reprints should be sent to M. Lappe. Fax: 0049 234 709 4278; e-mail: lappe@neurobiologie.ruhr-uni-bochum.de

0893–6080/98/$19.00
PII: S0893-6080(98)00013-6

inappropriate, and that aspects from both sides have to be considered (see Wertheim (1994) for a review). One important point in this regard is that the extraretinal signal typically underestimates the speed of the eye movement (Wertheim, 1987; Honda, 1990) and operates with a variable gain (Haarmeier and Thier, 1996).

A special case of this discussion concerns the detection of the direction of heading in the presence of eye movements. Eye movements here refer to slow eye movements such as smooth pursuit or the slow phases of the optokinetic and vestibulo-ocular reflexes. Since the original suggestion by Gibson (1950) that the optic flow field, i.e. the visual motion induced by self-motion in a stable world, provides all necessary information for navigation in an unknown environment, several computational (Bruss and Horn, 1983; Heeger and Jepson, 1992; Hildreth, 1992; Koenderink and van Doorn, 1987; Longuet-Higgins and Prazdny, 1980; Rieger and Lawton, 1985) and psychophysical (Rieger and Toet, 1985; van den Berg, 1992; Warren et al., 1988; Warren and Hannon, 1990) studies have supported this view. In the psychophysical experiments, the performance of human subjects in two conditions were compared. In the first condition, an optic flow stimulus simulating observer translation was presented while the subjects tracked a small pursuit target with his eyes. In the second condition, the subject kept his eyes stationary, while the stimulus simulated the visual effect of the eye movement together with the observer translation. Often there were few differences between the two conditions. There are, however, experimental situations which do require extraretinal signals. Most notably, a movement towards a frontoparallel plane (Regan and Beverly, 1982; Warren and Hannon, 1990) gives rise to a characteristic misjudgment of the direction of heading when only optic flow is available as a cue. However, also for movements on top of a ground plane, extraretinal signals are required, depending on the type of the eye movement. When eye movements simulate the pursuit of a horizontally moving target, humans cannot correctly judge the direction of heading (Royden et al., 1994). When the same eye movements are actually performed, i.e. when extraretinal information is available, the judgments are correct. In contrast, when eye movements simulate the stabilization of gaze on an environmental point, no extraretinal signal is required for a correct perception (van den Berg, 1993). The different simulated eye movements in these two situations give rise to fundamentally different flow patterns on the retina (Lappe and Rauschecker, 1995b). The different results can be linked to this difference in structure of the optic flow and to the similarity of the stimulus to the optic flow experienced during movement on a curved path (Lappe and Rauschecker, 1994; Royden, 1994).

Taken together, the present results from psychophysical studies suggest that the human visual system can make use of the information available in the optic flow field to determine the direction of heading. However, extraretinal input can improve the performance in many cases and is especially required to disambiguate the visual information in certain problematic situations.

The medial superior temporal area, or area MST, is regarded as a cortical site where the determination of egomotion parameters might take place (Duffy and Wurtz, 1991a, 1995; Lagae et al., 1994; Pekel et al., 1996). Lappe and Rauschecker proposed a neural network model of heading detection from optic flow that achieves the visual response properties of MST neurons from the output of MT-like neurons (Lappe and Rauschecker, 1993b) and links these properties to the psychophysics of human heading detection from optic flow (Lappe and Rauschecker, 1995a, Lappe and Rauschecker, 1995b). Recent experimental studies (Duffy and Wurtz, 1995; Pekel et al., 1996; Lappe et al., 1996) have confirmed the predictions from this model.

However, besides its specificity for optic flow analysis, area MST is also linked to eye movements. ‘Pursuit’ or ‘visual tracking’ neurons in MST, discharge during ongoing smooth pursuit (Erickson and Dow, 1989; Komatsu and Wurtz, 1988) and receive extraretinal signals about the occurrence of a pursuit eye movement (Newsome et al., 1988; Ilg and Thier, 1996b). While area MST is involved in the generation of smooth pursuit (Dürsteler and Wurtz, 1988; Komatsu and Wurtz, 1988), these neurons might also provide area MST with a signal about the presence of an eye movement (Erickson and Thier, 1991). As with the extraretinal input to the pursuit neurons (Newsome et al., 1988) this distinction is not found earlier in the visual motion pathway, neither in area MT (Erickson and Thier, 1991), nor in area V1 (Ilg and Thier, 1996a).

This article presents a model of heading detection that includes not only the processing of visual, but also of extraretinal signals. In order to account for the psychophysically observed capabilities of the human visual system in the task of heading detection, neither of the two classical theories outlined above, the ‘inferential’ or the ‘direct’ theory, is sufficient. Instead, the psychophysical results suggest a situation-dependent combination of visual and extraretinal signals. Royden (1994) has proposed that the visual system uses an extraretinal signal to remove the eye-movement-induced motion from the retinal flow field and subsequently estimate translation and rotation on a curved path. The model presented here follows a similar approach, since it also first removes the eye-movement induced visual motion and then analyses the resulting ‘purified’ flow field. However, different from Royden (1994) the subsequent analysis of the purified flow field still allows for residual influences of eye movements that might result from imperfect compensation. Moreover, the model provides a neuronal basis for this proposed combination of visual and extraretinal signals. Also, the paper will not be concerned with motion on a curved path, but rather with self-translations along a straight line.

The proposed model assumes that an extraretinal signal is first used to estimate the eye movement induced retinal

image motion and then compensate for it. Consistent with experimental evidence (Haarmeier and Thier, 1996; Honda, 1990; Wertheim, 1987) this compensation mechanism must be assumed to be incomplete since it accounts only for a certain proportion of eye speed. Therefore the resulting purified flow field is nevertheless processed by a heading detection mechanism that allows for the presence of eye rotation and takes visual evidence of eye movements into account. The visual heading detection is based on an earlier developed model (Lappe and Rauschecker, 1993b). This earlier model does not include any extraretinal input but retrieves heading directions visually from retinal flow even in the presence of eye movements. It is successful in many situations and consistent with human psychophysical data in a way that the earlier model fails in the case of approaching a frontoparallel plane, in which extraretinal input is required. This paper describes how extraretinal input can be combined with this model such as to also retrieve heading in this more complicated situation. The interaction of visual and extraretinal signals in the model is based on known properties of MST neurons. The combination of visual and extraretinal signals is performed by optic flow processing MST-like model neurons that receive input from direction-selective MT-like model neurons and from MST-like model pursuit neurons. The basic mechanisms of this interaction, i.e. the requirements for the appropriate synaptic connections, have been briefly presented earlier (Lappe et al., 1994). This article will, after presenting the combined model in Section 2, compare it to psychophysical and neurophysiological data (Sections 3 and 4), and discuss the results and implications (Section 5). The comparison to psychophysical findings will focus on the case of an observer approaching a frontoparallel plane. While this is a highly specific situation in a natural environment, it is also a situation that is especially difficult for human subjects and clearly benefits from extraretinal input as indicated in the psychophysical literature. Section 3 will present simulations and analytical results of the model behavior in this situation. The comparison with neurophysiological findings will concern three aspects: the ability of neurons to differentiate real motion in the world from visual motion induced by eye movements (Section 4.1), the visual responses of optic flow neurons when eye movements are superimposed (Section 4.2), and the influence of extraretinal input on optic flow selectivity (Section 4.3). The model yields testable predictions for neurophysiological experiments.

2. The model

The goal of this work is to provide a neural network model of a specific functionality of the primate visual system. For the scope of this paper, we strive for biological plausibility which comprises three requirements. First (I), all computations have to be performed by neuronal elements that have only a limited processing power, like e.g. simple input–output neurons. Second (II), the model needs to conform with known anatomical and physiological properties of those parts of the visual system that are primarily involved in the task at hand. This requirement has two components. On the one hand (IIa) the structure of the model, i.e. the arrangement of neurons and connections, has to be consistent with the relevant structures of the visual system. On the other hand (IIb), the properties of the individual model neurons need to be similar to the properties of real neurons in the respective visual areas. Note that while the goal here is to explain the mechanisms of the human visual system, most modelling constraints in IIa, b have to come in fact from animal studies. Third (III), the behavior of the model as a whole needs to be consistent with human behavior as tested in psychophysical experiments. Requirements I and IIa influence the design of the model from the outset; requirements IIb and III have to be tested later in simulations.

The basic layout of the model is schematically drawn in Fig. 1. A representation of the optic flow field by direction selective, MT-like neurons serves as the input. In a second stage, the direction of heading is computed by populations of optic flow sensitive neurons, presumably located in area MST. The extraretinal input is provided by a separate population of pursuit neurons. It is combined with the optic flow information from the MT layer at the level of MST. Thus the extraretinal compensation for the eye movement is done simultaneously with the optic flow computations. An intermediate stage explicitly containing a purified optic flow representation is not necessary.

In the following the actual computations carried out by the various cell populations are described. Section 2.1, Section 2.2 and Section 2.3 review the model of Lappe and Rauschecker (1993b) which provides the visual component of the combined visual/extraretinal model. Section 2.1 describes the computations underlying the visual heading detection. Section 2.2 and Section 2.3 describe the neural implementation of these computations in the MT and MST layers of the model. Section 2.4 finally introduces extraretinal input and presents the new combined model.

2.1. Visual computation of egomotion

This section describes a computational method for estimating the direction of heading from the optic flow which is a purely mathematical formalism that leads to an optimization process. As such, it does not specify how the optimization shall be performed. The next section will then show how these computations can be implemented by neuronal processes.

Mathematically, the optic flow field is obtained by perspectively projecting the motion of points (X, Y, Z) in the three-dimensional world onto a two-dimensional image plane (x, y) (Prazdny, 1980; Heeger and Jepson, 1992). This image plane is located at distance f from the nodal

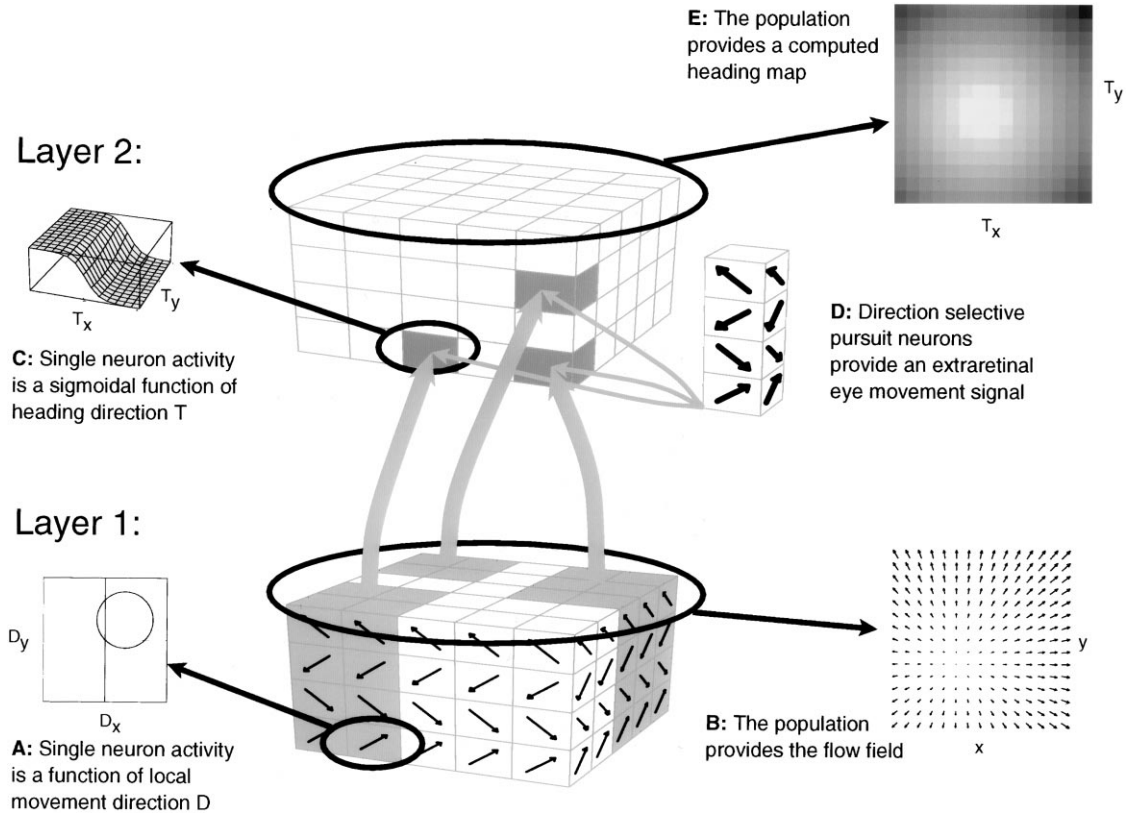

Fig. 1. Schematic representation of the model. The model consists of two layers of neurons. They correspond to areas MT and MST in monkey-cortex. The first layer, the model equivalent to area MT, contains neurons that are selective to local speed and direction of motion (A). The combined responses of all of these neurons give a population encoding of the retinal flow field (B). The second layer contains two types of neurons. Pursuit neurons are active during pursuit eye movements and are selective for the speed and direction of eye movement (C). Their eye movement related activity reflects an extraretinal, i.e. non-visual signal. Optic flow processing neurons (D) in the second layer receive input from the first layer as well as from the pursuit neurons. They use these inputs to compute the direction of heading. Individual neurons vary their activity as a function of the position of the focus of expansion and the extraretinal eye movement signal. The combined activity of all neurons results in a population code that signals the direction of heading (E).

point of the observer's eye, which serves as the center of the coordinate system. The optic flow equation is best expressed as the sum of two terms. The first term depends on the observer's three-dimensional translation $\mathbf{T} = (T_X, T_Y, T_Z)^t$ and on the depth structure $Z(x, y)$ of the visible scene. The second term depends on the eye-rotation $\mathbf{\Omega} = (\Omega_x, \Omega_y, \Omega_z)^t$. The optic flow at a given point (x, y) on the image plane is

$$\boldsymbol{\theta}(x,y) = \frac{1}{Z(x,y)}\mathbf{A}(x,y)\mathbf{T} + \mathbf{B}(x,y)\mathbf{\Omega} \tag{1}$$

with the matrices

$$\mathbf{A}(x,y) = \begin{pmatrix} -f & 0 & x \\ 0 & -f & y \end{pmatrix} \text{ and}$$

$$\mathbf{B}(x,y) = \begin{pmatrix} xy/f & -(f+x^2/f) & y \\ f+y^2/f & -xy/f & -x \end{pmatrix}$$

The task of heading detection is to estimate $\mathbf{T}$ given only the optic flow $\boldsymbol{\theta}(x, y)$, i.e. when $Z(x, y)$ and $\mathbf{\Omega}$ are not known. Heeger and Jepson (1992) have introduced a least-squared-error algorithm that computes the most likely direction $\mathbf{T}$ by minimizing the residual function:

$$R(\mathbf{T}) = \sum_{i=1}^{m} \left| \boldsymbol{\theta}(x_i,y_i) - \frac{1}{Z(x_i,y_i)}\mathbf{A}(x_i,y_i)\mathbf{T} + \mathbf{B}(x_i,y_i)\mathbf{\Omega} \right|^2 \tag{2}$$

This can be written in matrix form as

$$R(\mathbf{T}) = ||\mathbf{\Theta} - \mathbf{C}(\mathbf{T})\mathbf{q}||^2 \tag{3}$$

with the following definitions: $\mathbf{\Theta} = (\boldsymbol{\theta}_1, \cdots, \boldsymbol{\theta}_m)^t$ is a $2m$ vector consisting of the components of m optic flow vectors $\boldsymbol{\theta}_i$ from different visual field locations; $\mathbf{q} = (1/Z(x_1,y_1), \cdots, 1/Z(x_m, y_m), \Omega_X, \Omega_Y, \Omega_Z)^t$ is an $(m+3)$ vector containing the unknown depth structure and rotation; $\mathbf{C}(\mathbf{T})$ is a $2m \times (m+3)$ matrix,

$$\mathbf{C}(\mathbf{T}) = \begin{pmatrix} \mathbf{A}(x_1,y_1)\mathbf{T} & \cdots & 0 & \mathbf{B}(x_1,y_1) \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & \mathbf{A}(x_m,y_m)\mathbf{T} & \mathbf{B}(x_m,y_m) \end{pmatrix} \tag{4}$$

which depends only on the visual field locations (x_i, y_i) and on the translation $\mathbf{T}$. Heeger and Jepson (1992) then show

that minimizing Eq. (3) is equivalent to minimizing

$$R(\mathbf{T}) = ||\mathbf{\Theta}^t \mathbf{C}^{\perp}(\mathbf{T})||^2 \quad (5)$$

where the unknown parameters $Z(x, y)$ and $\mathbf{\Omega}$ have been eliminated. In this equation the matrix $\mathbf{C}^{\perp}(\mathbf{T})$ is the orthogonal complement of the matrix $\mathbf{C}(\mathbf{T})$. It is defined in the following way: provided that $\mathbf{C}(\mathbf{T})$ is full rank, the columns of $\mathbf{C}(\mathbf{T})$ span an $(m + 3)$-dimensional subspace of the $\mathcal{R}^{2m}$. The matrix $\mathbf{C}^{\perp}(\mathbf{T})$ is the matrix that spans the remaining orthogonal $(2m - (m + 3))$-dimensional subspace of the $\mathcal{R}^{2m}$. $\mathbf{C}^{\perp}(\mathbf{T})$ can be directly calculated from $\mathbf{C}(\mathbf{T})$ by orthogonalization (Heeger and Jepson, 1992).

The minimization of R($\mathbf{T}$) provides a way to determine the direction of heading solely from visual information (the optic flow field). It does so even in the presence of arbitrary rotations of the eye. Lappe and Rauschecker have modified R($\mathbf{T}$) in order to impose constraints on the rotation (Lappe and Rauschecker, 1993a, b). They have argued that such constraints make the model more consistent with biological data for several reasons. The first is that human eye movements are indeed constrained to move in a way that minimizes eye torsion, i.e. rotation around the Z-axis (Listing's law, see Carpenter, 1988). This effectively reduces the degrees of freedom from three to two. Moreover, during normal self-movement vestibulo-ocular and optokinetic reflexes induce stereotyped eye movements that stabilize gaze on a foveated environmental target (Baloh et al., 1988; Paige and Tomko, 1991; Solomon and Cohen, 1992; Lappe, et al., 1998). These reflexes thereby couple the eye movement to the self-motion and the environment, eventually reducing the degrees of freedom even further. Furthermore, eye movement constraints make the model more consistent with biological data because such constraints have implications for the model neurons that allow them to explain a wider range of MST response properties (Lappe and Rauschecker, 1993a).

Both constraints listed above can be formalized by modifying Eq. (1). For the first constraint (no torsion) Ω_Z is assumed to be zero and the optic flow becomes

$$\hat{\boldsymbol{\theta}}(x,y) = \frac{1}{Z(x,y)}\mathbf{A}(x,y)\mathbf{T} + \hat{\mathbf{B}}(x,y)\begin{pmatrix}\Omega_X\\ \Omega_Y\end{pmatrix}$$

where

$$\hat{\mathbf{B}}(x,y) = \begin{pmatrix} xy/f & -(f + x^2/f)\\ f + y^2/f & -xy/f\end{pmatrix}$$

now is a 2×2 matrix. As a result

$$\hat{\mathbf{C}}(\mathbf{T}) = \begin{pmatrix} \mathbf{A}(x_1,y_1)\mathbf{T} & \cdots & 0 & \hat{\mathbf{B}}(x_1,y_1)\\ \vdots & \ddots & \vdots & \vdots\\ 0 & \cdots & \mathbf{A}(x_m,y_m)\mathbf{T} & \hat{\mathbf{B}}(x_m,y_m)\end{pmatrix}$$

now is a $2m \times (m + 2)$ matrix. This in turn leads to a modification of the residual function (Lappe and Rauschecker, 1993a):

$$\hat{\mathrm{R}}(\mathbf{T}) = \|\mathbf{\Theta}^t \hat{\mathbf{C}}^{\perp}(\mathbf{T})\|^2 \quad (6)$$

The second constraint assumes that gaze is stabilized with regard to a foveated environmental target which is projected onto the center (0, 0) of the visual field. The distance of the target from the eye, however, is not known. Therefore we assume it to be some unknown distance $Z(0, 0)$ away from the eye. The eye-rotation necessary to stabilize the position of the target in the visual field center can then be derived from the condition that the retinal movement of the target has to vanish:

$$\boldsymbol{\theta}(0,0) = \frac{1}{Z(0,0)}\begin{pmatrix}-fT_X\\ -fT_Y\end{pmatrix} + \begin{pmatrix}-f\Omega_Y\\ +f\Omega_X\end{pmatrix} = \begin{pmatrix}0\\ 0\end{pmatrix} \quad (7)$$

with $\Omega_z = 0$ (no torsion), we find

$$\mathbf{\Omega} = \frac{1}{Z(0,0)}(T_Y, -T_X, 0)^t \quad (8)$$

The optic flow Eq. (1) in this case becomes

$$\tilde{\theta}(x,y) = \frac{1}{Z(x,y)}\mathbf{A}(x,y)\mathbf{T} + \frac{1}{Z(0,0)}\tilde{\mathbf{B}}(x,y)\mathbf{T}$$

with

$$\tilde{\mathbf{B}}(x,y) = \begin{pmatrix} f + x^2/f & xy/f & 0\\ xy/f & f + y^2/f & 0\end{pmatrix}$$

Therefore

$$\tilde{\mathbf{C}}(\mathbf{T}) = \begin{pmatrix} \mathbf{A}(x_1,y_1)\mathbf{T} & \cdots & 0 & \tilde{\mathbf{B}}(x_1,y_1)\mathbf{T}\\ \vdots & \ddots & \vdots & \vdots\\ 0 & \cdots & \mathbf{A}(x_m,y_m)\mathbf{T} & \tilde{\mathbf{B}}(x_m,y_m)\mathbf{T}\end{pmatrix}$$

and the residual function, Eq. (5), becomes:

$$\tilde{\mathrm{R}}(\mathbf{T}) = \|\mathbf{\Theta}^t \tilde{\mathbf{C}}^{\perp}(\mathbf{T})\|^2 \quad (9)$$

Note, however, that since $Z(0, 0)$ is not known, the gaze stabilization constraint only requires that

$$\mathbf{\Omega} \propto (T_Y, -T_X, 0)^t \quad (10)$$

Hence this proportionality also includes the case of no eye rotation ($\mathbf{\Omega} = 0$) in the limit $Z(0, 0) \rightarrow \infty$.

The three variants of the residual function estimate the direction of heading with different constraints on the eye movements. However, most eye movements that occur during normal locomotion satisfy all of the above constraints. Therefore, it is possible to use any of the three residual functions in the model. The goal in all three cases is to find that instance of $\mathbf{T}$ that minimizes R($\mathbf{T}$). Such a minimization can be performed in various ways, for instance by a gradient descent iteration, but such an iteration is slow and not necessarily biologically plausible. The model instead uses a different approach. It assumes a two-dimensional map of possible (candidate) heading directions. For each of these heading directions, a function inversely proportional to R($\mathbf{T}$) is evaluated. Then the optimum $\mathbf{T}$ is found

within this heading map by locating the map position in which R(**T**) is minimal. This implementation of the above algorithm is parallel, fast, and, as will be shown below, can be realized with simple neuronal elements that resemble cells from areas MT and MST.

2.2. Optic flow encoding by direction selective neurons in area MT

In the visual cortex of macaque monkeys, a successive analysis of visual motion is performed by a series of interconnected visual areas. Already area V1 contains cells sensitive for the direction of movement of bars through their receptive field, but the first area which is explicitly specialized for visual motion is the middle temporal area, or area MT. It contains a high percentage of direction selective cells, i.e. cells that respond to local unidirectional motion at a certain position in the visual field (Maunsell and Van Essen, 1983). Experimental (Movshon et al., 1985; Newsome et al., 1990; Rodman and Albright, 1989) as well as theoretical (Bülthoff et al., 1989; Nowlan and Sejnowski, 1995: Wang et al., 1989; Wilson and Kim, 1994) studies have presented evidence that area MT contains a distributed encoding of the optic flow field.

The first layer of the network represents the optic flow input by direction selective neurons similar to cells from area MT. Several neurons with different preferred velocities $\mathbf{e}_{i\alpha}$, and velocity tuning functions $s_{i\alpha}$ are assumed to form a population encoding of an optic flow vector at image location i

$$\boldsymbol{\theta}_i = \sum_{\alpha=1}^{n} s_{i\alpha}\mathbf{e}_{i\alpha} \tag{11}$$

Since speed and direction tuning of single MT cells are independent from one another (Rodman and Albright, 1987), velocity tuning functions can be modelled as the product of speed (θ) tuning and direction (ϕ) tuning:

$$s_{i\alpha}(\theta,\phi) = s_{i\alpha}^{\text{speed}}(\theta)s_{i\alpha}^{\text{dir}}(\phi) \tag{12}$$

The same holds for the preferred velocities: the velocity preference of an individual neuron is

$$\mathbf{e}_{i\alpha} = e_{i\alpha}^{\text{speed}}\mathbf{e}_{i\alpha}^{\text{dir}} \tag{13}$$

For the encoding of speed, eight speed preferences are used at each visual field location:

$$e_{i\alpha}^{\text{speed}} = \theta_{\text{pref}}(i,\alpha) \text{ with } \theta_{\text{pref}}(i,\alpha) \in \{0.5°/s, 1°/s, 2°/s, 4°/s, 8°/s, 16°/s, 32°/s, 64°/s\}$$

The speed tuning (Fig. 2B) is modelled as a Gaussian of the logarithm of the ratio between actual speed θ and preferred θ_{pref}

$$s_{i\alpha}^{\text{speed}}(\theta) = \exp(-(\log_2(\theta/\theta_{\text{pref}}(i,\alpha)))^2) \tag{14}$$

For the encoding of direction, four direction preferences ϕ_{pref} are used at each visual field location.

$$\mathbf{e}_{i\alpha}^{\text{dir}} = \begin{pmatrix}\cos(\phi_{\text{pref}}(i,\alpha))\\ \sin(\phi_{\text{pref}}(i,\alpha))\end{pmatrix}$$

with $\phi_{\text{pref}}(i,\alpha)\epsilon\{\phi_0, \phi_0+\pi/2, \phi_0+\pi, \phi_0+3\pi/2\}$.

Here, the individual preferred directions θ_{pref} are equally spaced around the direction ϕ_0 of a line running from the center of the receptive field of the neuron to the fovea. The direction tuning is modelled as a rectified cosine function (Fig. 2A). A neuron's direction specific response to a movement into an arbitrary direction ϕ is

$$s_{i\alpha}^{\text{dir}}(\phi) = \begin{cases}\cos(\phi-\phi_{\text{pref}}(i,\alpha)) & \text{if } \cos(\phi-\phi_{\text{pref}}(i,\alpha)) > 0\\ 0 & \text{otherwise}\end{cases} \tag{15}$$

Therefore each visual field location i contains a total of $4 \times 8 = 32$ neurons, each with a different combination of speed and direction selectivity.

In area MT the distribution of preferred directions is not isotropic. Foveofugal motion, i.e. motion away from the fovea, is more often preferred than foveopetal motion (Albright, 1989). This anisotropy can be included in the model by removing neurons with preferred directions ϕ_0, i.e. foveopetal motion, from the input layer, leaving only three preferred directions at each location (Lappe and

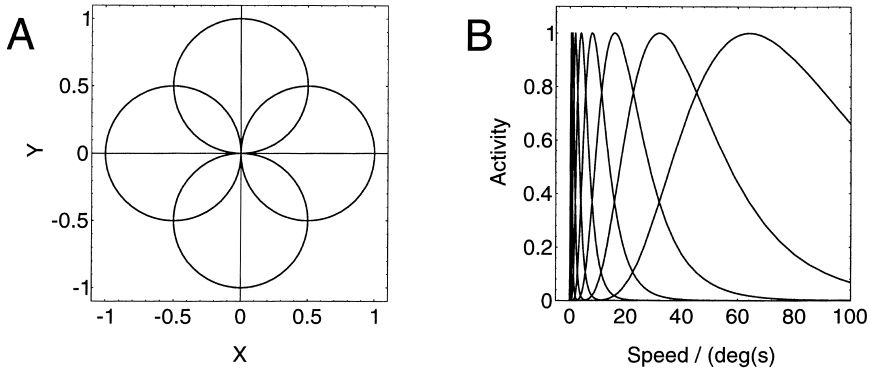

Fig. 2. Tuning curves of first layer model neurons for direction and speed of visual motion. (A) Polar plot of the direction tuning. The direction tuning curve of an individual neuron follows a rectified cosine function. To obtain a population encoding of movement direction, four neurons with different preferred directions are used. (B) The speed tuning is modelled as a Gaussian on a logarithmic speed scale. Eight speed preferences are used.

Rauschecker, 1995b). In this case, motion directions with a foveopetal component will be misrepresented by the first layer. However, in most simulations performed in this paper no foveopetal motion components occur and both input distributions lead to the same result. An exception will be given in Section 4.1 where the responses of individual neurons to general motion stimuli are described. In this case, examples with both variants of the input neuron distribution are presented.

2.3. Optic flow processing neurons in area MST

Once the optic flow field is determined, the next step for the visual system would be to estimate egomotion parameters and parameters of the structure of the visual scene from the optic flow. Electrophysiological studies suggest that this is achieved in area MST which directly follows area MT in the visual motion pathway (Boussaoud et al., 1990). It contains cells that respond selectively to large-field random dot optic flow patterns such as expansions, contractions, or rotations (Saito et al., 1986;Tanaka et al., 1986; Tanaka and Saito, 1989; Duffy and Wurtz, 1991a, b). However, studies testing whether a mathematical decomposition of the optic flow into the elementary flow components divergence, deformation, and curl (Koenderink and van Doorn, 1975) is performed in area MST did not find any evidence for this (Graziano et al., 1994; Lagae et al., 1994). Instead, a continuum of response selectivities was found (Duffy and Wurtz, 1991a; Graziano et al., 1994) suggesting a distributed encoding of egomotion parameters by a neural network of broadly selective individual elements (Duffy and Wurtz, 1995; Lappe et al., 1996).

In the model, the determination of the direction of heading is performed by the synaptic projection to a second layer. This layer represents a two-dimensional map of possible heading directions. Each map position contains a cell population that represents one heading direction $\mathbf{T}_j$. Such a population shall become maximally excited when $\mathrm{R}(\mathbf{T}_j) = 0$. If this is the case, the map computes in parallel the residual function for a large number of possible heading directions. Thus, the peak of activity in the map signals the direction of heading that optimally fits with the input flow field. The activity peak is achieved in two computational steps. First, the single neuron β within population j evaluates part of the argument of $\mathrm{R}(\mathbf{T}_j)$. It is assumed to receive input from a random subset of first layer cell populations. By definition these form the receptive field of the second layer neuron. The synaptic weights with neurons $s_{i\alpha}$ from the first layer (Eq. (11)) are chosen such that the input to the neuron equals $\mathbf{\Theta}^t \mathbf{C}_\beta^\perp(\mathbf{T}_j)$. The input/output relationship of a single second layer neuron follows the classical sigmoidal form:

$$u_{j\beta} = \mathrm{g}\left(\sum_{i=1}^{m} \sum_{\alpha=1}^{n} J_{j\beta,i\alpha} s_{i\alpha} - \mu \right) \tag{16}$$

where g() is a sigmoid function, μ a threshold, and $J_{j\beta,i\alpha}$ the synaptic connection between second layer neuron $j\beta$ and first layer neuron $i\alpha$. Thus, to ensure that the input to the neuron equals $\mathbf{\Theta}^t \mathbf{C}_\beta^\perp(\mathbf{T}_j)$ the synaptic strengths need to satisfy

$$\sum_{i=1}^{m} \sum_{\alpha=1}^{n} J_{j\beta,i\alpha} s_{i\alpha} = \mathbf{\Theta}^t \mathbf{C}_\beta^\perp(\mathbf{T}_j) \tag{17}$$

or

$$\sum_{i=1}^{m} \sum_{\alpha=1}^{n} J_{j\beta,i\alpha} s_{i\alpha} = \sum_{i=1}^{m} \boldsymbol{\theta}_i^t \begin{pmatrix} C_{\beta,2i-1}^\perp(\mathbf{T}_j) \\ C_{\beta,2i}^\perp(\mathbf{T}_j) \end{pmatrix} \tag{18}$$

The individual synaptic values can be calculated by inserting Eq. (11):

$$\sum_{i=1}^{m} \sum_{\alpha=1}^{n} J_{j\beta,i\alpha} s_{i\alpha} = \sum_{i=1}^{m} \sum_{\alpha=1}^{n} s_{i\alpha} \mathbf{e}_{i\alpha}^t \begin{pmatrix} C_{\beta,2i-1}^\perp(\mathbf{T}_j) \\ C_{\beta,2i}^\perp(\mathbf{T}_j) \end{pmatrix} \tag{19}$$

Pairwise comparison yields that the synaptic strengths for a single second layer neuron have to be:

$$J_{j\beta,i\alpha} = \mathbf{e}_{i\alpha}^t \begin{pmatrix} C_{\beta,2i-1}^\perp(\mathbf{T}_j) \\ C_{\beta,2i}^\perp(\mathbf{T}_j) \end{pmatrix} \tag{20}$$

The second step involves generating population activities that peak when $\mathrm{R}(\mathbf{T}_j) = 0$, i.e. when $\mathbf{\Theta}^t \mathbf{C}_\beta^\perp(\mathbf{T}_j) = 0$. To this end, the activities from all neurons within population j are simply summed. The threshold μ is chosen slightly negative so that the population activity

$$U_j = \sum_\beta u_{j\beta} \tag{21}$$

peaks when $\mathrm{R}(\mathbf{T}_j) = 0$ (Lappe and Rauschecker, 1993b).

In Section 2.1 three different variants of the residual function have been presented and it has been argued that all three would be consistent with normal oculomotor behavior. Previous work (Lappe and Rauschecker, 1993a) has shown that these three variants lead to different selectivities of the second layer neurons with respect to rotational flow stimuli and that all three are necessary to cover the spectrum of MST response properties. Therefore, all three variants of the residual function will be used concurrently here. For each second layer neuron, one residual function is picked at random for the generation of all its synaptic connections. Each residual function will thus effectively be implemented by a third of the neurons.

2.4. Combining visual and extraretinal input: pursuit and optic flow neurons in area MST

Area MST receives a variety of extraretinal signals in addition to the visual input provided mainly by area MT. These signals are related to eye position (Bremmer et al., 1997), smooth pursuit eye movement (Newsome et al., 1988), and to vestibulo-ocular (Thier and Erickson, 1992) and optokinetic (Kawano and Sasaki, 1984) eye movements. For the purpose of the model, only the signals related to eye

movements are important. A subpopulation of MST neurons, the 'pursuit' or 'visual tracking' neurons, discharge during ongoing smooth pursuit (Erickson and Dow, 1989; Komatsu and Wurtz, 1988) in a direction selective manner. Their responses result in part from the visual flow signal induced by the eye movement, but some of these neurons also retain their firing activity even in the absence of visual stimulation (Newsome et al., 1988) or during pursuit of an imaginary target (Ilg and Thier, 1996b), thereby clearly exhibiting evidence of extraretinal input.

In the model, the visual-only heading detection system outlined above is therefore complemented by a separate population of pursuit neurons that independently signal the direction and speed of an ongoing eye movement. Their activity is assumed to result from an extraretinal signal and is thus independent from the flow field. Similar to the visual neurons in the first layer of the network, the pursuit neurons each possess different preferred velocities $\mathbf{e}_\lambda$ and velocity tuning functions p_λ for eye movements. They provide the second layer cells with a population encoding of the eye-rotation:

$$\mathbf{\Omega} = \gamma \sum_{\lambda=1}^{k} p_\lambda \mathbf{e}_\lambda \tag{22}$$

However, in this equation it is assumed that the representation of the eye rotation is imperfect, i.e. that the speed of the eye movement is only signaled up to a scalar gain factor γ. This assumption is in line with many experimental observations indicating that the gain of the extraretinal reference signal is smaller than 1.0 (Haarmeier and Thier, 1996; Honda, 1990; Wertheim, 1987).

The second layer neurons use the signal from the pursuit neurons to subtract the eye movement related visual motion from the input flow field. The component of the retinal flow field that results from the pursuit eve movement is obtained by inserting Eq. (22) in Eq. (1) with $\mathbf{T} = 0$. At a single image location i, the eye movement induced retinal motion is

$$\boldsymbol{\theta}_i^p = \gamma \sum_{\lambda=1}^{k} p_\lambda \mathbf{B}_i \mathbf{e}_\lambda \tag{23}$$

with $\mathbf{B}_i = \mathbf{B}(x_i, y_i)$. With the input from the pursuit neurons, the activity of a single layer neuron becomes

$$u_{j\beta} = \mathrm{g}\left(\sum_{i=1}^{m} \sum_{\alpha=1}^{n} J_{j\beta,i\alpha} s_{i\alpha} + \sum_{\lambda=1}^{k} J^p_{j\beta,\lambda} p_\lambda - \mu \right) \tag{24}$$

instead of Eq. (16). Here $J^p_{j\beta,\lambda}$ denotes the synaptic connections between the second layer neuron $j\beta$ and the pursuit neuron λ. Then, the calculation of the synaptic connections starts from

$$\sum_{i=1}^{m} \sum_{\alpha=1}^{n} J_{j\beta,i\alpha} s_{i\alpha} + \sum_{\lambda=1}^{k} J^p_{j\beta,\lambda} p_\lambda = (\mathbf{\Theta} - \mathbf{\Theta}^p)^t \mathbf{C}^\perp_\beta(\mathbf{T}_j) \tag{25}$$

$$= \sum_{i=1}^{m} \left[(\boldsymbol{\theta}_i - \boldsymbol{\theta}_i^p)^t \begin{pmatrix} C^\perp_{\beta,2i-1}(\mathbf{T}_j) \\ C^\perp_{\beta,2i}(\mathbf{T}_j) \end{pmatrix} \right] \tag{26}$$

instead of Eq. (17). Substituting Eq. (11) and Eq. (23) yields

$$\sum_{i=1}^{m} \sum_{\alpha=1}^{n} J_{j\beta,i\alpha} s_{i\alpha} + \sum_{\lambda=1}^{k} J^p_{j\beta,\lambda} p_\lambda$$

$$= \sum_{i=1}^{m} \left[\left(\sum_{\alpha=1}^{n} s_{i\alpha} \mathbf{e}^t_{i\alpha} - \gamma \sum_{\lambda=1}^{k} p_\lambda (\mathbf{B}_i \mathbf{e}_\lambda)^t \right) \begin{pmatrix} C^\perp_{\beta,2i-1}(\mathbf{T}_j) \\ C^\perp_{\beta,2i}(\mathbf{T}_j) \end{pmatrix} \right] \tag{27}$$

Therefore the synaptic weights between the first layer neurons and the second layer neurons remain unchanged, still given by Eq. (20). The synaptic weights between the pursuit neurons and the second layer neurons have to be

$$J^p_{j\beta,\lambda} = -\gamma \sum_{i=1}^{m} (\mathbf{B}_i \mathbf{e}_\lambda)^t \begin{pmatrix} C^\perp_{\beta,2i-1}(\mathbf{T}_j) \\ C^\perp_{\beta,2i}(\mathbf{T}_j) \end{pmatrix} \tag{28}$$

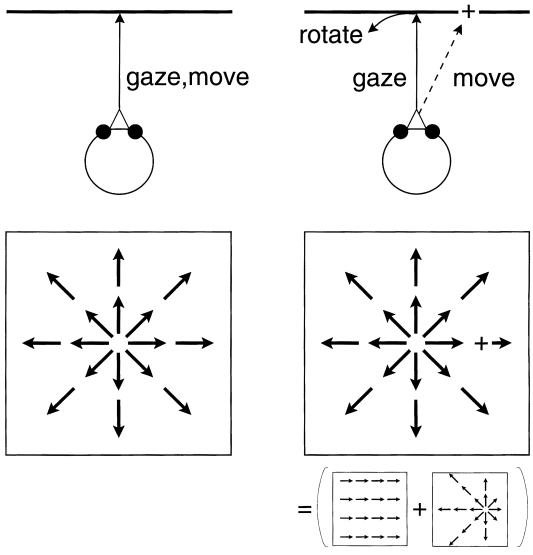

Fig. 3. Two different self-motions yield the same retinal flow field. On the left side, an observer approaches a frontoparallel plane. The resulting flow field is that of a uniform expansion, with the center of expansion located in the direction of heading. The right side shows the approach of the same plane with a direction of self-motion, tilted towards the right. His point of impact on the plane is marked by a cross. In addition, the observer performs an eye rotation towards the left in order to keep his gaze fixed on the straight-ahead element of the plane. The resulting retinal flow field is the same as the one on the left panel. It is the vectorial superposition of two flow field components, shown in parentheses. The first component is the eye rotation induced retinal flow which approximates a unidirectional frontoparallel motion towards the right. The second component is a uniform expansion with the center of expansion located on the right, matching the direction of heading. This component corresponds to the optic flow in a world-centered coordinate system. The vectorial superposition of the two components, i.e. the *retinal* flow field, is a uniform expansion with a pseudo focus of expansion located straight ahead, at the direction of gaze instead of the direction of heading. Thus, the direction of heading, i.e. the point of impact on the plane (+), is not identified by a focus of expansion in the retinal flow field in this situation.

3. Extraretinal input disambiguates flow field information

During movement with respect to a plane, the instantaneous optic flow field is inherently ambiguous (Longuet-Higgins, 1984). A particular situation occurs when an observer approaches a frontoparallel plane while performing an eye movement that is intended to stabilize gaze on a single point on the frontoparallel plane (Fig. 3). When such an optic flow field is simulated, human subjects cannot accurately determine their direction of heading (Regan and Beverly, 1982). Instead, they report the direction of gaze, i.e. the visual field center, as the direction of heading. However, when real eye movements are performed, i.e. when extra retinal signals are available, the task is solved easily (Warren and Hannon, 1988, 1990; Royden, 1994). In this situation, motion parallax, which is an important cue for the determination of egomotion parameters, is missing. The absence of motion parallax also constitutes a very difficult situation for the previous model (Lappe and Rauschecker, 1993b). This section will concentrate on the ambiguity arising when an observer approaches a frontoparallel plane and show how it is removed by extraretinal input.

An example simulation is shown in Fig. 4. Fig. 4A shows the flow field input to the network. The flow field was constructed by assuming a certain visual environment and observer movement. Then the motion vectors that occurred at the inputs of the network were calculated from the assumed imaging geometry following Eq. (1). The visual environment consisted of a plane parallel to the image plane. The simulated movement was a translation towards the plus (+) and an eye rotation appropriate to keep the direction of gaze on a single point on the plane. Hence this point is projected onto the center of the visual field ($\times$). As a result, the center of the visual field contains a singular point (Lappe and Rauschecker, 1995b) that resembles a focus of expansion. Without extraretinal input ($\gamma = 0$), the output of the network (Fig. 4B) shows two activity peaks. One is close to the correct direction of heading (+). The second one is in the visual field center ($\times$), where the singular point of the flow field is located. This second peak is consistent with the incorrect responses of human subjects in the absence of extraretinal information. Fig. 4C displays the

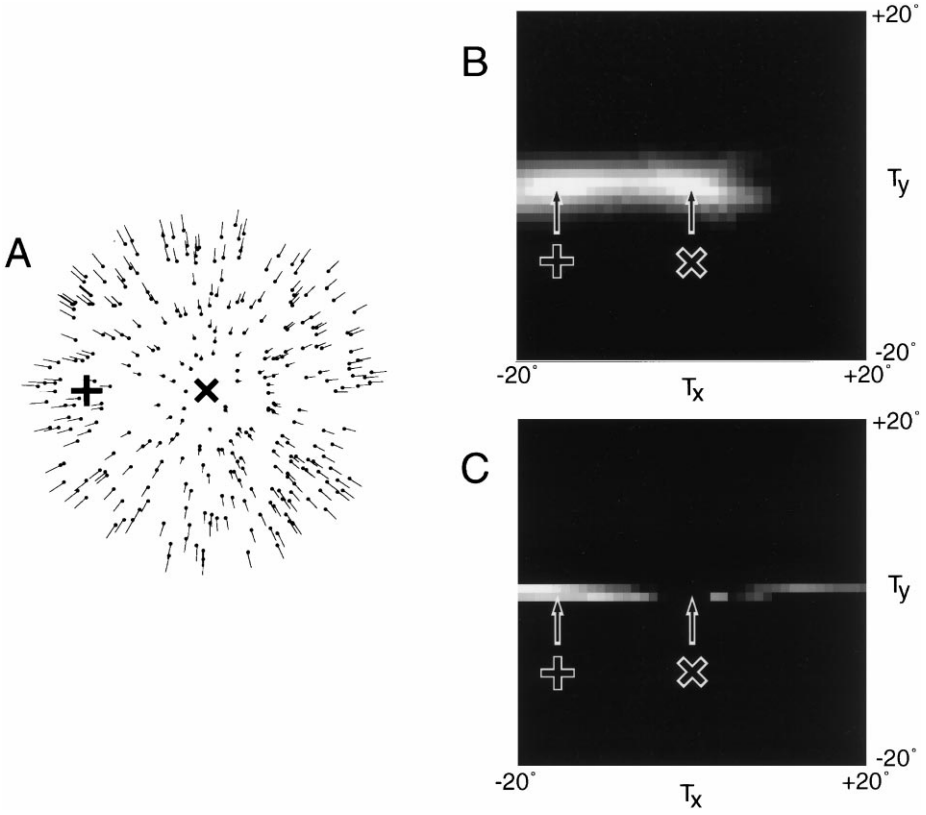

Fig. 4. Influence of the extraretinal signal on heading detection. The figures show the influence of extraretinal input in a situation where an observer moves towards a wall and keeps his gaze directed towards a fixed point on the wall (see Fig. 3). The left side shows the optic flow field. The right side shows the output of the network as a logarithmic grayscale map. It represents a 40 $\times$ 40 grid of heading directions. A brightness peak in the map indicates that this direction of heading has been found by the network. The true direction of heading, i.e. the point of impact on the plane, is indicated by a plus sign. The direction of gaze is indicated by an $\times$. Keeping gaze on target requires to perform an eye movement during forward translation. Because of this eye movement and the planar structure of the wall, the direction of gaze contains a pseudo focus of expansion (A). In this case the flow field is ambiguous. Both the direction of gaze and the direction of heading could correspond to possible self-motions. Without extraretinal input (B) the model output consequently exhibits two peaks of activity. When extraretinal input is available (C), the spurious second peak at the direction of gaze is largely suppressed.

case where extraretinal input is available ($\gamma = 1$). In this situation, the activity peak in the visual field center is strongly suppressed and much smaller than in Fig. 4B.

This suggests that the model behavior might be consistent with that of human subjects in these two conditions. To quantify the performance of the model in the presence and absence of extraretinal input a series of simulations was carried out. In these simulations, the direction of heading was determined from the network-output by locating the neuronal population that was maximally active. This corresponds to the single brightest pixel in Fig. 4B, C. Simulations either included an extraretinal signal or did not. However, the assumption of a perfect unity gain of the extraretinal signal is certainly not biologically plausible. Several psychophysical observations show that the gain is usually different from 1.0 (Wertheim, 1987; Honda, 1990; Wertheim, 1994; Haarmeier and Thier, 1996). In the model, a perfect extraretinal signal is not required. To show this, a number of simulations were performed in which the gain of the extraretinal signal was systematically varied. For each gain value, 100 simulation trials were run with different directions of heading. Then the average error over these trials was computed. Without the extraretinal input the network exhibits a large error. Inspection of individual trials indicated that most often the model determined the central peak as the direction of heading. This is consistent with the behavior of human subjects in the absence of extraretinal input (Warren and Hannon, 1990). When extraretinal input was available the average error was much smaller. The model often determined the correct direction of heading, also consistent with human behavior. Fig. 5 also demonstrates that a gain less than one, which is a physiologically sensible assumption, is actually slightly preferable for the model. A sharp drop of the error occurs as soon as some extraretinal input is available. In comparison, a variation of the gain of the extraretinal input leads to only modest variations of the average error. This observation suggests a qualitative difference between the presence and absence of extraretinal input. The sharp drop that is caused as soon as even weak extraretinal input is added can be understood with the following considerations: as mentioned above, the flow field in Fig. 4A is inherently ambiguous with regard to the direction of heading. Mathematically, exchanging the movement direction **T** and the plane normal **E** could yield an identical flow field, provided that the eye rotation is appropriately modified (Longuet-Higgins, 1984). This explains the two peaks in the network output. However, the same applies also to the flow field after the rotation has been partially removed by an extraretinal signal. In this case, however, the network output only has a single peak (see Fig. 4C). The fact that the model behaves differently in the two situations is related not only to the existence of an extraretinal signal, but also to the previously introduced assumptions about the types of eye movement that may occur. A third of the model neurons in the second layer of the network implement in their synaptic connections a constraint on the eye movement, which assumes that the eye movement is due to the stabilization of gaze onto an environmental object (Eq. (10)). Therefore, for any valid estimation of the egomotion parameters by the network, the rotation inherent in the flow field has to satisfy Eq. (10). This requirement is fulfilled by both solutions in the case of Fig. 4B but only in one solution in the case of Fig. 4C. This can be proven analytically, by explicitly considering the two possible solutions in turn. For this, it it necessary to consider sets of self-motion parameters that can create such a flow field. The flow field in Fig. 4A was constructed by assuming a certain visual environment and observer motion. The following set of parameters was used:

$$\mathbf{E} = (0,0,1)^t,\ Z = \mu,\ \mathbf{T} = (T_X, T_Y, T_Z)^t,\ \mathbf{\Omega} = \frac{1}{\mu}(T_Y, -T_X, 0)^t \tag{29}$$

However, in the presence of an extraretinal signal with gain $\gamma > 0$ part of the rotation is actually removed from the flow field before it reaches the heading detection stage. In this case, the effective flow field parameters are

$$\mathbf{E} = (0,0,1)^t,\ Z = \mu,\ \mathbf{T} = (T_X, T_Y, T_Z)^t,$$

$$\mathbf{\Omega} = (1-\gamma)\frac{1}{\mu}(T_Y, -T_X, 0)^t \tag{30}$$

Obviously, this solution fulfills Eq. (10). But a different set of parameters,

$$\hat{\mathbf{E}} = \frac{1}{T}(T_X, T_Y, T_Z)^t,\ Z(x,y) = \frac{f\mu T}{xT_X + yT_Y + fT_Z},$$

$$\hat{\mathbf{T}} = (0,0,T)^t,\ \hat{\Omega} = -\gamma\frac{1}{\mu}(T_Y, -T_X, 0)^t \tag{31}$$

results in an identical flow field. This can easily be seen by inserting both sets of parameters in Eq. (1). However, in the second case, if $\gamma \neq 0$, the rotation violates Eq. (10):

$$\hat{\Omega} \neq \frac{T_Z}{\mu}(\hat{T}_Y, -\hat{T}_X, 0)^t = 0 \tag{32}$$

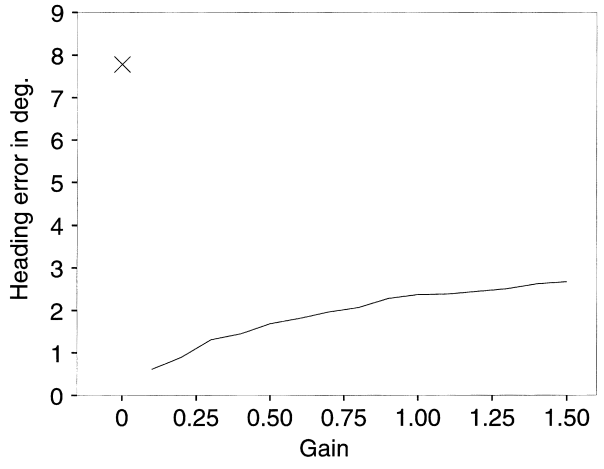

Fig. 5. Mean error in computing the direction of heading in the situation shown in Fig. 4 when the gain of the extraretinal signal is varied.

Thus, only the first set of parameters comprises a valid solution, but if $\gamma = 0$, i.e. no extraretinal signal is present, Eq. (10) is again satisfied since

$$\hat{\Omega} = 0 = \frac{T_Z}{\mu}(\hat{T}_Y, -\hat{T}_X, 0)^t \quad (33)$$

Thus, the network is subject to the ambiguity only when no extraretinal input is available. This explains the sharp drop of the average error between the conditions of no extraretinal input and of weak extraretinal input in Fig. 5. Weak extraretinal input already suffices to resolve the ambiguity.

4. Consequences of extraretinal input for single neurons in the model of area MST

Single neurons in the second layer of the model receive extraretinal input from the pursuit neurons in addition to the visual input they receive from the first layer neurons. The consequences of the interaction between the two signals for the response properties of individual second layer neurons were tested in simulations of single neurons. When only the visual input is considered, neurons in the second layer of the model respond selectively to various optic flow patterns (Lappe and Rauschecker, 1993b). These response properties are similar to the properties of cells in area MST (Duffy and Wurtz, 1991a, 1995; Lagae et al., 1994; Lappe et al., 1996). In the model, different degrees of selectivity which have been described in area MST (Duffy and Wurtz, 1991a) are achieved by the different constraints on the eye movements as outlined above (Lappe and Rauschecker, 1993a), but an important point for the model neurons, as well as for most MST cells (Duffy and Wurtz, 1991a), is that a single cell might not only respond to a single preferred flow stimulus such as for instance an expansional flow field. The same cell might also respond in a direction selective manner to other motion stimuli such as for instance global frontoparallel unidirectional motion. In fact, these neurons fulfill all criteria for standard direction selectivity to frontoparallel motion. Extraretinal eye movement information enables the second layer neurons to distinguish such visual motion from the visual motion that results from slow eye movements over a lit background.

4.1. Distinction between self-induced and externally induced visual motion

Many neurons in area MST carry the ability to distinguish between active, self-induced stimulation, resulting from smooth pursuit over a lit background, and passive, externally induced stimulation when a visual pattern is moved across the receptive field (Erickson and Thier, 1991). This property most likely reflects the use of an extraretinal input. It has also been hypothesized that the distinction could also be performed from visual signals alone (Erickson and Thier, 1992), although the mechanisms are not well understood.

The same behaviour can be observed in the model neurons. Fig. 6 shows simulation results of nine individual neurons. Each neuron was tested in three paradigms. Full-field unidirectional frontoparallel motion (Fig. 6A) was used to determine direction selectivity. To simulate the visual effects of a pursuit eye movement, full-field rotational motion (Fig. 6B) was generated by assuming a rotation around a frontoparallel axis. This input was used in two conditions, $\gamma = 0$, simulating the response without extraretinal input, and $\gamma = 0.9$, simulating the response with high-gain extraretinal input. These simulations are similar to three types of visual motion: object motion, simulated self-induced motion without extraretinal input, and real self-induced motion with extraretinal input. To illustrate the neuronal response properties in these three conditions, polar plots of the responses as a function of movement directions were generated. To compare unidirectional translation with simulated eye rotation, both polar plots were arranged such that the average direction of motion in the stimuli were identical. Examples are shown in Fig. 6A, B. Fig. 6A is a translation towards the left. Fig. 6B is a rotation around a vertical axis. In both figures the overall motion is towards the left and both would be mapped to leftward motion in the polar plots below.

How do the nine example neurons differ from each other? First of all, they possess different receptive fields because they receive input from different random subsets of first layer neuronal populations. Second, as described in Section 2.1, different eye movement constraints are imposed on different neurons which lead to different synaptic connections. Furthermore, as described in Section 2.2, an anisotropic distribution of direction preferences can be used for the first layer neurons, following evidence for a similar anisotropy in area MT. This distribution also influences the properties of the second layer neurons. To visualize the spectrum of response properties in the second layer of the model the nine neurons in Fig. 6C–K exemplify different combinations of constraints and input distributions. For neurons C and D, an isotropic distribution of preferred directions in layer one has been used. Neuron C implements the gaze-stabilization constraint. Neuron D does not constrain the eye movements at all. Neurons E–K have an anisotropic distribution of direction preferences in the MT layer. Neurons G, H, and J do not impose any constraints on the eye movements. Neurons E and F use the gaze-stabilization constraint. Neurons I and K implement the no-torsion constraint,

A number of observations can be made from Fig. 6. First, most neurons respond direction selective to frontoparallel translational motion. An exemption is neuron J, which does not respond to translational motion. Also, directional tuning curves are fairly broad, although some neurons (E, F, G) have somewhat narrower tuning curves. These observations are consistent with the known properties of MST neurons.

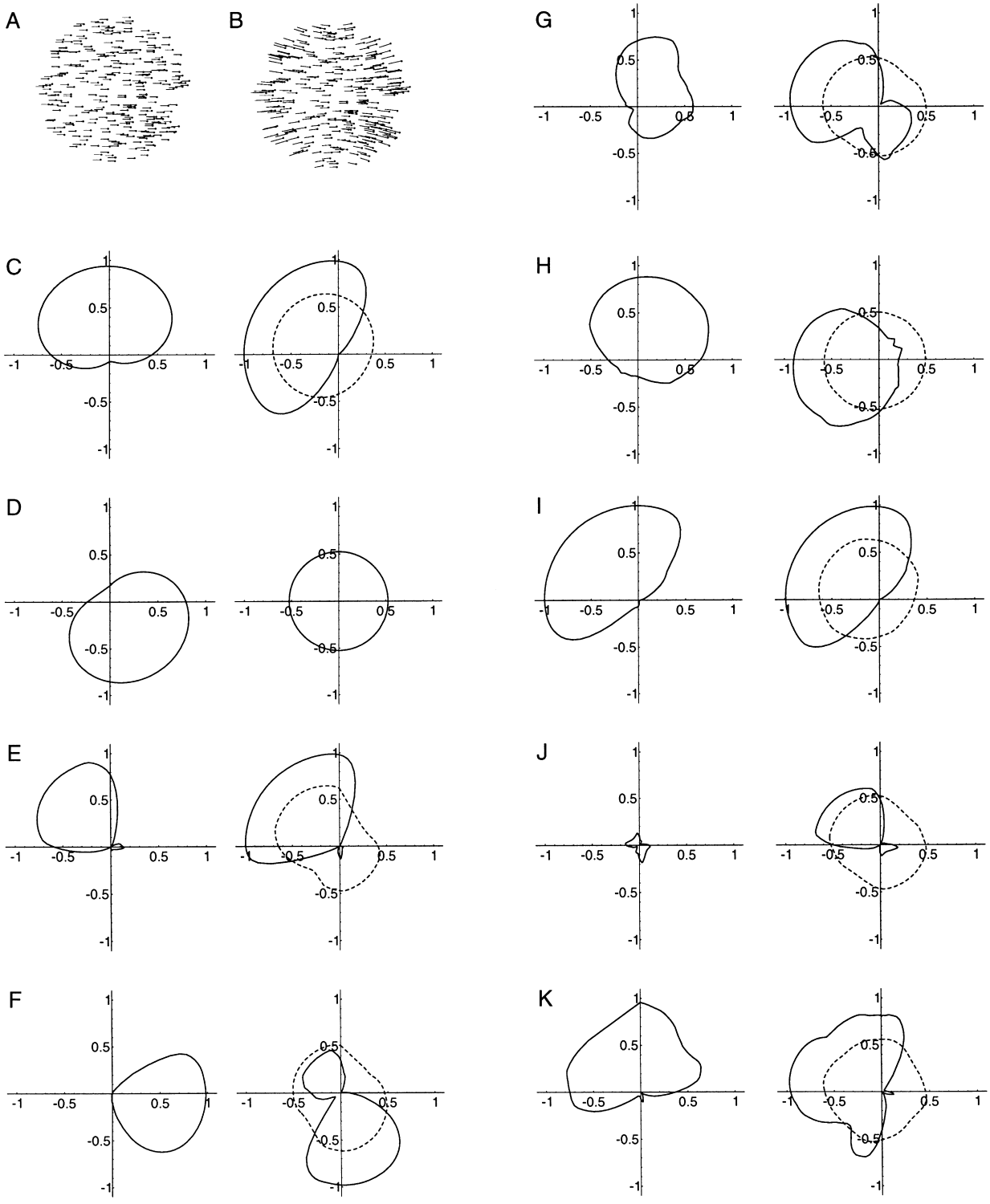

Fig. 6. Examples of the effect of extraretinal input on the direction selectivity of individual second layer neurons. For each neuron C-J two polar plots are shown. The left ones give the passive direction selectivity of the neuron when stimulated with full-field frontoparallel translation. The corresponding visual stimulus is shown in (A). The right panels present responses to an optic flow pattern simulating the visual effects of a smooth pursuit eye movement over a structured background (this stimulus is shown in B). The plots show simulations without (solid lines) as well as with (broken lines) an extraretinal signal of gain 0.9.

Second, when the responses to simulated eye rotations without extraretinal input ($\gamma = 0$) are considered, most neurons again display a directional modulation, but this directional modulation is not necessarily congruent with the direction selectivity observed with translational motion. However the differences in preferred direction are small for most neurons, seldom exceeding 45°. Neuron H is the sole exception. Nevertheless, this shows that the model neurons take the full pattern of retinal motion into account and also react to the comparatively small differences between the two stimuli in Fig. 6A, B. A systematical experimental comparison between these two types of motion in MST has not yet been performed. Duffy and Wurtz (Duffy and Wurtz, 1991a) briefly noted that a small group of neurons they tested with four directions each, showed no strong differences. However, to experimentally test whether neurons could visually differentiate rotation from translation requires the use of very large stimuli. The differences between the two motion patterns are more prominent in the periphery of the visual field. Thus, large stimuli are necessary to allow a distinction. In the simulations, stimulus size was 90 by 90° of visual angle. Most experimental studies were performed with smaller stimuli, usually matched to the receptive field sizes of individual neurons which average up to 10 by 10° in area MT and between 10 by 10 and 60 by 60° in area MST. In these conditions, observations of differences between translation and rotation would necessarily be difficult because the differences in the stimuli are only marginal.

As a third point, when extraretinal signals are included in the model ($\gamma = 0.9$) the directional modulation is strongly diminished. All of the examples in Fig. 6 would be considered to respond not direction selective anymore. This behavior is well in line with the results of Erickson and Thier (1991). By using extraretinal input, the model cells are able to distinguish self induced from external motion in the same way as neurons in MST. However, extraretinal signals might not always be necessary. The model neuron shown in D distinguishes between translational and rotational motion even in the absence of extraretinal input ($\gamma = 0$). It does so by using only the visual differences between the two motion types. Such a possibility has also been hypothesized for MST neurons on the basis of experimental results (Erickson and Thier, 1992). The complementary behavior is displayed by neuron J. In the absence of extraretinal input, this neuron responds more strongly and more selectively to rotational than to translational motion. This cell signals global eye rotation on the basis of purely visual input and distinguishes this from global translations. It can be used as a visual backup for the pursuit neurons. In conclusion, the model can generate cells with the ability to discriminate self-induced from externally induced visual motion and predicts

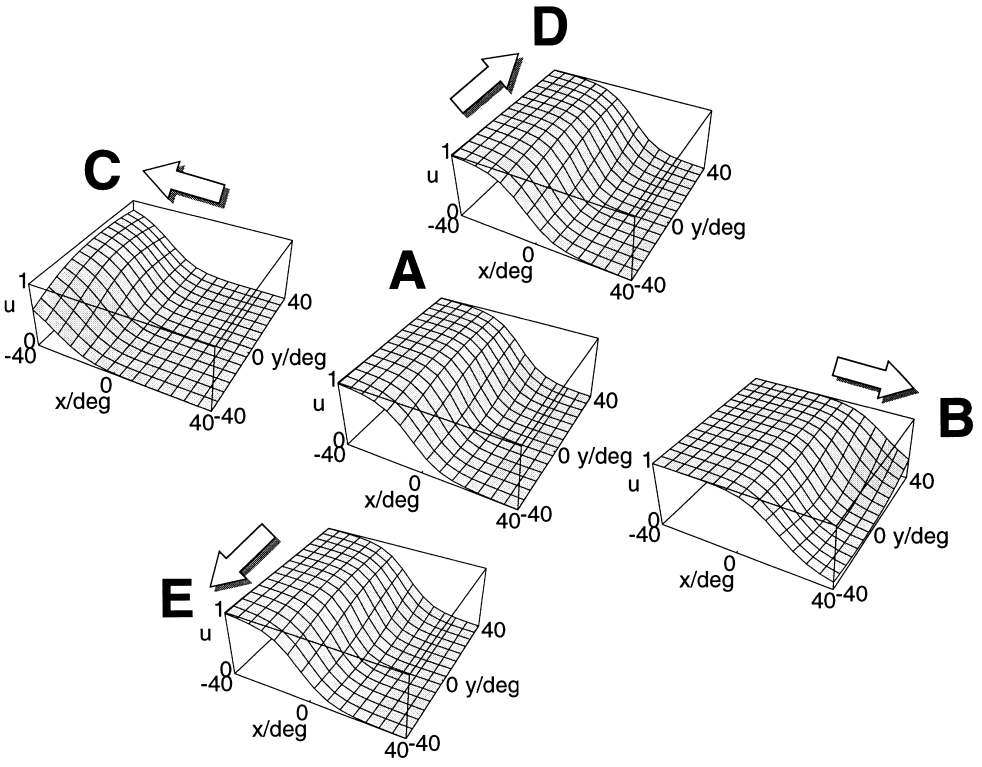

Fig. 7. Responses of an optic flow selective neuron to combined optic flow and simulated eye movement stimuli. (A) The response to a pure expansional optic flow pattern depends in a sigmoidal manner on the position of the focus of expansion. The three-dimensional surface plot shows the response u as a function of the location (x, y) of the focus of expansion in the visual field. (B) Neuronal response as a function of the direction of heading when an eye movement in the preferred direction is added. No extraretinal signal is available. (C) Neuronal response as a function of the direction of heading when an eye movement in the anti-preferred direction is added. (D, E) Responses when eye movements in orthogonal directions are added.

that this ability must not necessarily depend on extraretinal input.

4.2. Interaction between optic flow and eye movements

An important question in MST neurophysiological experiments is the influence of eye movements on optic flow responses. Model simulations might be useful to develop hypothesis and predictions that can be tested experimentally. Fig. 7 shows how the optic flow responses of a model neuron are altered by visually superimposing an eye rotation. The central panel (Fig. 7A) depicts the pure optic flow responses to expansional motion stimuli. The responses are displayed as three-dimensional surface plots. Along the x and y axes, the visual field position of the singular point (i.e. the focus of expansion) is varied. The neuronal activity depends in a sigmoidal manner on the position of the singular point. The neuron responds strongly when the focus of expansion is located in the left visual hemifield. It does not respond when the focus of expansion is located in the right hemifield. The inflection point of the sigmoidal curve roughly lies along the vertical meridian. From the functional standpoint of the model this is the location where the greatest sensitivity for the location of the focus of expansion occurs. Such a sigmoidal dependence of neuronal responses on the position of the singular point has recently also been reported for part of the neurons in area MST (Duffy and Wurtz, 1995; Lappe et al., 1996). The other four panels in Fig. 7 display the response profiles of the same cell when simulated eye movements are added to the flow stimuli. Eye movements in the preferred direction (Fig. 7B), in the anti-preferred direction (Fig. 7C), and in two orthogonal directions (Fig. 7D, E) were simulated. In all cases, the sigmoidal profile remains, but its position is horizontally shifted when eye movements in the preferred or anti-preferred direction occur. This shift is in the direction of the eye movement. In contrast, when orthogonal eye movements are simulated, no shift occurs.

The shift of the response curve can be more easily interpreted when the structure of the combined expansion and eye movement stimuli are considered. In these simulations, the expansion stimuli consisted of uniform expansion that simulated the approach towards a frontoparallel plane. Adding simulated eye rotations to such a flow stimulus is equivalent to shifting the center of expansion in the direction of the eye movement. This is so, because to a first approximation simulated eye rotation adds a unidirectional flow component to all flow vectors (see Fig. 3). The shift of the response curve of the neuron in fact reflects the shift of the center of expansion in the stimulus. Therefore, in all cases the inflection point of the sigmoidal response curve is in the vicinity of the center of expansion. When eye movements orthogonal to the preferred direction are simulated, the response curve does not shift, simply because the center of expansion now is displaced alongside the inflection zone, i.e. upwards in Fig. 7D and downwards in Fig. 7E.

The shift of the response profile is not limited to expansion responses. It also occurs for responses to contraction or rotation. In fact, like many MST cells, many of the model neurons respond to rotational flow stimuli in addition to responses to expansion or contraction. In these cases the response curve also shifts with the position of the singular point. For contraction, it shifts in the direction opposite to the eye movement. For rotation it shifts orthogonal to the direction of the eye movement.

4.3. Influence of extraretinal input on the interaction between optic flow and eye movements

So far, in these simulations no extraretinal signal was included. The responses in Fig. 7 thus correspond to an experimental situation of simulated eye movement, rather than real, active eye movement. If an extraretinal signal would have been included, the transformations of the response curves would have been weaker or would not occur at all. This is illustrated in Fig. 8. In Fig. 8 cross sections through the response curve of a single neuron are shown when extraretinal input is available (Fig. 8B) and when it is not (Fig. 8A). Clearly, the shift of the response curve is smaller, when extraretinal input is available. In this case, the inflection point of the response curve cannot be aligned with the position of the focus of expansion in the retinal flow field, since this position shifts when the eye is moved. Instead, it is aligned with the direction of heading, which remains at the same simulated position in all cases. This position corresponds to the focus of expansion when no eye movements are simulated. In this case, the retinal and the external position of the focus of expansion is the same. When eye movements are simulated the direction of heading corresponds to the location of the focus of expansion in the outside world, i.e. in a world-centric reference frame. Only the retinotopic position of the focus of expansion is altered by the eye movement. Thus, with extraretinal input (Fig. 8B) the neuronal response depends on the position of the focus of expansion in the outside world.

However, the same is true for model neurons that are immune to eye movements per se, such as the neuron in Fig. 6D. This neuron by itself would respond to the world-centered position of the focus of expansion, whether eye movements are simulated or not. In summary, the prediction from these simulations is that many MST neurons, most likely those classified as triple component selective by Duffy and Wurtz (Duffy and Wurtz, 1991a, 1995), should shift their areas of best response to expansion and contraction in the presence of simulated pursuit eye movements, but not during real, actively performed smooth pursuit.

Bradley et al. (1996) have recorded single neurons in area MST in an equivalent paradigm. They found indeed that many MST cells did show a significant shift of their response curves when simulated eye movements were

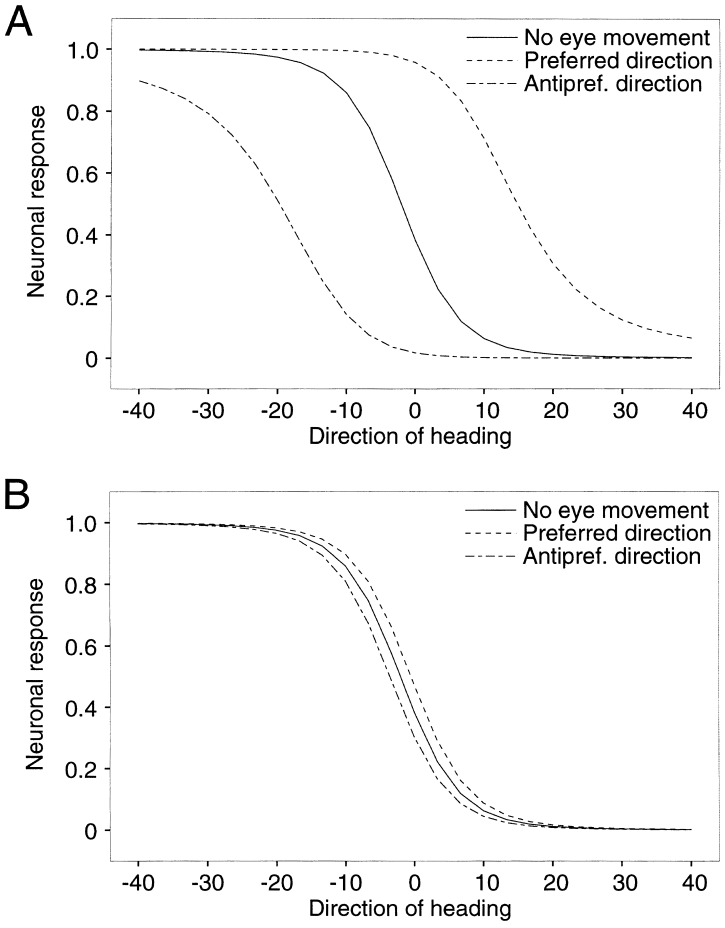

Fig. 8. Influence of extraretinal input on the shifts of the optic flow response curves. (A) Cross sections through the response curves shown in Fig. 7A-C. The solid curve shows the response to a pure expansion, when the direction of heading, i.e. the world-centered location of the focus of expansion is varied (Fig. 7A). The dotted lines show the responses to combined optic flow and simulated eye movements in the preferred and anti-preferred directions (Fig. 7B and C). In these cases no extraretinal input is available. The response curves shift laterally in conjunction with the shift of the retinal position of the focus of expansion. (B) Same situation when extraretinal input is available (gain 0.9). In this case, the response curves stay close to the pure optic flow curve (solid line). Thus, their position is aligned with the direction of heading, i.e. the focus of expansion in the external world, despite the fact that the retinal position of the focus of expansion has shifted.

compared to real eye movements. In the case of simulated eye movements, the response curves were shifted and aligned with the retinal focus of expansion. In contrast, in the case of real eye movements the response curves were better aligned with the focus of expansion in world-centric coordinates, i.e. the shift was much smaller. These findings fit very well with the model predictions.

Another puzzling outcome of their study is also consistant with the model. They found that the degree to which an individual cell's response curve stays aligned with the focus of expansion in world-centric coordinates varies for different neurons. In fact, they found a continuum with regard to the strength of extraretinal compensation. In the model, the strength of the extraretinal signal is governed by its gain γ. As has been demonstrated in Fig. 5, the influence of the actual value of γ on the model performance is rather small. A variation of γ in individual neurons on the other hand leads to intermediate or exaggerated shifts of the response curve similar to the experimental findings (Fig. 9). Furthermore, in the experiments the median shift of the response curve in the full neuronal population amounted to about half the required perfect compensation. This corresponds to a gain of $\gamma = 0.5$ in the model. As Fig. 5 shows, a value of about $\gamma = 0.5$ is actually preferable to a unity gain, because errors in model simulations are lower. Therefore, the rather low average extraretinal compensation found in area MST could be optimal for the system.

5. Discussion

Research in human optic flow processing has shown that the visual system benefits from extraretinal eye movement signals. These signals aid in the process of the detection of the direction of heading during combined self-motion and eye movements. In this paper a neural network model of heading detection was developed that uses both optic flow and extraretinal eye movement signals. This model takes as

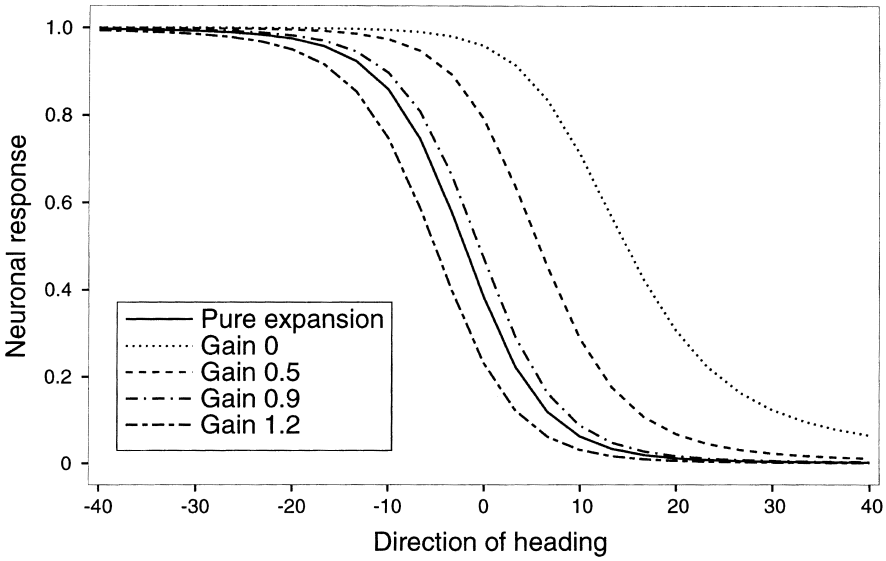

Fig. 9. Influence of extraretinal gain on the shifts of the optic flow response curves. Cross sections through the response curves are shown as in Fig. 8. The solid curve shows the response to a pure expansion, when the direction of heading, i.e. the world-centered location of the focus of expansion is varied. The dotted and broken lines show the responses to combined optic flow and simulated eye movements in the preferred direction for various values of the extraretinal gain γ. The resulting shifts of the response curves could result in no compensation ($\gamma = 0$), partial ($\gamma = 0.5$) to nearly complete ($\gamma = 0.9$) compensation, or even over-compensation ($\gamma = 1.2$). The latter leads to a shift of the response curve in the opposite direction. All of these shifts have been observed in area MST. The model can tolerate these variations because it also uses visual information for heading detection in addition to the extraretinal signal (see Fig. 5).

input a retinal flow field that results from combined self-motion and eye movements. It uses an extraretinal signal to form an estimate of that part of the retinal motion that is induced by the eye movement. Subtracting this estimate from the retinal flow extracts the self-motion induced retinal flow. The direction of heading is subsequently determined from this self-motion induced retinal flow. Consistent with psychophysical (Wertheim, 1987; Honda, 1990; Haarmeier and Thier, 1996) and electrophysiological (Bradley et al., 1996) data the model assumes that the extraretinal signal underestimates the speed of the eye rotation. Because of this, the estimate of the self-motion induced retinal flow still contains residual eye movement induced retinal motion. Therefore, the subsequent determination of the direction of heading also has to take eye rotations into account, i.e. it requires a visual backup for deficiencies of the extraxetinal signal. This visual backup is based on an earlier model by Lappe and Rauschecker (1993b) which determined heading direction visually without using an extraretinal signal.

The model implements the computational steps involved in the determination of the direction of heading from visual and extraretinal cues with neuronal processing elements in a biologically plausible way. It proposes a specific neuronal mechanism of the interaction of visual and extraretinal signals in primate visual cortical area MST. However, in this specific implementation the computational steps outlined above are not performed in the serial manner suggested by the above description. Rather, all of them are performed concurrently in one step of parallel synaptic transmissions. These transmissions originate from motion processing cells in area MT and from extraretinal pursuit neurons within area MST and terminate on optic flow processing cells in area MST. An intermediate representation of the flow field after the subtraction of the eye movement induced motion is not necessary.

The origin of the extraretinal input that is received by the pursuit neurons in the model has not been specified. The reason is that the origin of the extraretinal input is not important for the model to function. Either an efference copy or eye muscle proprioception could serve as effective signals. Another possibility has been suggested by van den Berg (1996). In psychophysical experiments he found that the instruction given by the experimenter can also modify the degree to which human subjects attribute visual rotation inherent in the flow field stimulus to eye movements. This suggests that the focussing of attention or the expectation of the subject can also contribute to flow field analysis. This is especially interesting since cells in area MST already show effects of attentional modulation (Newsome et al., 1988; Andersen et al., 1990; Treue and Maunsell, 1996) of their responses.

The performance of the model has been tested so far only in simulations of an observer approaching a frontoparallel plane. This special condition has been chosen because it has in the past received much attention from psychophysicists and most clearly demonstrates the need for extraretinal input in human observers (Warren and Hannon, 1990; Royden et al., 1994). The model reproduces the characteristic misjudgments of human subjects in the absence of extraretinal input. Similar to humans it is able to determine the direction of heading correctly when extraretinal input is

available. Moreover, the simulations show that the model is robust against variations of the gain of the extraretinal signal. This is important because the gain of extraretinal signal is variable both for human subjects (Wertheim, 1987; Honda, 1990; Haarmeier and Thier, 1996) and for single MST cells (Bradley et al., 1996).

Simulations show that several properties of the model neurons are consistent with experimental observation in area MST. Two types of extraretinal modulation of neuronal motion responses are present in the model as well as in area MST. First, pursuit neurons are active during a smooth pursuit and thus can signal the presence of a pursuit eye movement. These cells carry an excitatory extraretinal signal that enhances their response during a pursuit eye movement. Second, cells that take direct part in the processing of visual information use this pursuit signal to differentiate true visual motion from motion induced on the retina by an eye movement. These neurons receive an inhibitory extraretinal signal in the sense that their response to visual motion is reduced when the visual motion results from an eye movement. Both types of extraretinal modulation have been observed in area MST (Newsome et al., 1988; Erickson and Thier, 1991).

Model simulations also show interactions between optic flow responses and eye movement induced visual motion at the single neuron level. Single model neurons respond to the location of the singular point in a flow field, such as the focus of expansion in a radial flow pattern or the center of rotation in a rotational flow pattern. Their response function is similar to that of MST neurons (Duffy and Wurtz, 1995; Lappe et al., 1996). When eye movements are superimposed on such a flow stimulus the singular point is shifted in the direction of the eye movement. The behavior of the model neuron in this case depends on the extraretinal input and on the direction of the eye movement. When extraretinal input is absent the direction of the eye movement is in the preferred direction of the neuron, the response of the neuron shifts with the shift of the singular point. A similar shift is observed in MST neurons (Bradley et al., 1996). When extraretinal input is available, this shift is diminished. The magnitude of the shift depends on the gain of the extraretinal signal. For gain $\gamma = 1$, the neuronal response is invariant to the presence of eye movements. For other gain values, the shift either under- or over-compensates for the eye movement. Both behaviors have been described in area MST (Bradley et al., 1996). The model proposes that such imperfect compensation is already sufficient for accurate heading detection when visual and extraretinal signals are used in combination.

Acknowledgements

This research was funded by the Deutsche Forschungsgemeinschaft

References

Albright T.D. (1989). Centrifugal directionality bias in the middle temporal visual area (MT) of the macaque. *Vis. Neurosci., 2*, 177–188.

Andersen R., Graziano M., & Snowden R. (1990). Translational invariance and attentional modulation of MST cells. *Soc. Neurosci. Abstr., 16*, 7.

Baloh R.W., Beykirch K., Honrubia V., & Yee R.D. (1988). Eye movements induced by linear acceleration on a parallel swing. *J. Neurophysiol., 60*, 2000–2013.

Boussaoud D., Ungerleider L.G., & Desimone R. (1990). Pathways for motion analysis: cortical connections of the medial superior temporal visual areas in the macaque. *J. Comp. Neurol., 296*, 462–495.

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

Bremmer F., Ilg U.J., Thiele A., Distler C., & Hoffmann K.-P. (1997). Eye position effects in monkey cortex. I: visual and pursuit related activity in extrastriate areas MT and MST. *J. Neurophysiol., 77*, 944–961.

Bruss A.R., & Horn B.K.P. (1983). Passive navigation.. *Comput. Vis. Graph. Image Proc., 21*, 3–20.

Bülthoff H., Little J., & Poggio T. (1989). A parallel algorithm for real-time computation of optical flow. *Nature, 337*, 549–553.

Carpenter, R.H.S., 1988. Movement of The Eyes, 2nd ed. Pion , London.

Duffy C.J., & 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 (6)*, 1329–1345.

Duffy C.J., & 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 (6)*, 1346–1359.

Duffy C.J., & 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., & 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., & Dow B.M. (1989). Foveal tracking cells in the superior temporal sulcus of the macaque monkey. *Exp. Brain Res., 78*, 113–131.

Erickson R.G., & Thier P. (1991). A neuronal correlate of spatial stability during periods of self-induced visual motion. *Exp. Brain Res., 86*, 608–616.

Erickson R.G.L., & Thier P. (1992). Responses of direction-selective neurons in monkey cortex to self-induced visual motion. *Ann. N.Y. Acad. Sci., 656*, 766–774.

Gibson, J. J., (1950). The Perception of the Visual World. Houghton Mifflin, Boston.

Graziano M.S.A., Andersen R.A., & Snowden R. (1994). Tuning of MST neurons to spiral motions. *J. Neurosci., 14 (1)*, 54–67.

Haarmeier T., & Thier P. (1996). Modification of the Filehne illusion by conditioning visual stimuli. *Vision Res., 36*, 741–750.

Heeger D.J., & Jepson A. (1992). Subspace methods for recovering rigid motion I: algorithm and implementation. *Int. J. Computer Vision, 7 (2)*, 95–117.

Hildreth E.C. (1992). Recovering heading for visually-guided navigation. *Vision Res., 32*, 1177–1192.

Honda H. (1990). The extraretinal signal from the pursuit eye movement system: Its role in the perceptual and the egocentric localization systems. *Percep. Psychophys., 48*, 509–515.

Ilg U.J., & Thier P. (1996a). Inability of rhesus monkey area V1 to discriminate between self-induced and externally induced retinal image slip. *Eur. J. Neurosci., 8*, 1156–1166.

Ilg, U. J. & Thier, P., (1996b). MST neurons are activated by pursuit of imaginary targets. In: Thier, P., Karnath, H.-O. (Eds.), Parietal Lobe Contributions to Orientation in 3D-Space. Experimental Brain Research Supplement, Springer, Berlin.

Kawano K., & Sasaki M. (1984). Response properties of neurons in

posterior parietal cortex of monkey during visual-vestibular stimulation II. Optokinetic neurons. *J. Neurophysiol., 51*, 352–360.

Koenderink J.J., & van Doorn A.J. (1975). Invariant properties of the motion parallax field due to the movement of rigid bodies relative to an observer. *Opt. Acta, 22 (9)*, 773–791.

Koenderink J.J., & van Doorn A.J. (1987). Facts on optic flow. *Biol. Cybernet., 56*, 247–254.

Komatsu H., & Wurtz R.H. (1988). Relation of cortical areas MT and MST to pursuit eye movements. I. Localization and visual properties of neurons. *J. Neurophysiol., 60 (2)*, 580–603.

Lagae L., Maes H., Raiguel S., Xiao D.-K., & 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., & Rauschecker, J. P., (1993a). Computation of heading direction from optic flow in visual cortex. In: Giles, C. L., Hanson, S. J., Cowan, J. D. (Eds.), Advances in Neural Information Processing Systems, Vol. 5. Morgan Kaufmann, San Mateo, CA, pp. 433–440.

Lappe M., & Rauschecker J.P. (1993b). A neural network for the processing of optic flow from egomotion in man and higher mammals. *Neural Comput., 5*, 374–391.

Lappe M., & Rauschecker J.P. (1994). Heading detection from optic flow. *Nature, 369*, 712–713.

Lappe M., & Rauschecker J.P. (1995a). An illusory transformation in a model of optic flow processing. *Vision Res., 35*, 1619–1631.

Lappe M., & Rauschecker J.P. (1995b). Motion anisotropies and heading detection. *Biol. Cybernet., 72*, 261–277.

Lappe, M., Bremmer, F., Hoffmann & K.-P., (1994). How to use non-visual information for optic flow processing in monkey visual cortical area MSTD. In: Marinaro, M., Morasso, P. G. (Eds.), ICANN 94—Proceedings of the International Conference on Artificial Neural Networks, Springer, Berlin, pp. 46–49.

Lappe M., Bremmer F., Pekel M., Thiele A., & Hoffmann K.-P. (1996). Optic flow processing in monkey STS: a theoretical and experimental approach. *J. Neurosci., 16*, 6265–6285.

Lappe, M., Pekel, M., & Hoffmann, K.-P. (1998). Optokinetic eye movements elicited by radial optic flow in the macaque monkey. *J. Neurophysiol. 79*, 1461–1480.

Longuet-Higgins H.C. (1984). The visual ambiguity of a moving plane. *Proc. R. Soc. Lond. B, 223*, 165–175.

Longuet-Higgins H.C., & Prazdny K. (1980). The interpretation of a moving retinal image. *Proc. R. Soc. Lond. B, 208*, 385–397.

Maunsell J.H.R., & 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.

Mittelstaedt, H., (1990). Basic solutions to the problem of head-centric visual localization. In: Warren, R., Wertheim, A.H. (Eds.), The Perception and Control of Self-Motion. Erlbaum, Hillsdale, NJ.

Movshon, J. A., Adelson, E. H., Gizzi, M. S. & Newsome, W. T., (1985). The analysis of moving visual patterns. In: Chagas, C., Gattass, R., Gross, C. (Eds.), Pattern Recognition Mechanisms. Springer, Berlin.

Newsome W.T., Wurtz R.H., & 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.

Newsome, W. T., Britten, K. H., Salzman, C. D., Movshon, J. A., (1990). Neuronal mechanisms of motion perception. Cold Spring Harbor Symposia on Quantitative Biology.

Nowlan S., & Sejnowski T. (1995). A selection model for motion processing in area MT of primates. *J. Neurosci., 15*, 1195–1214.

Paige G.D., & Tomko D.L. (1991). Eye movement responses to linear head motion in the squirrel monkey I. Basic characteristics. *J. Neurophysiol., 65*, 1170–1183.

Pekel M., Lappe M., Bremmer F., Thiele A., & Hoffmann K.-P. (1996). Neuronal responses in the motion pathway of the macaque monkey to natural optic flow stimuli. *Neuroreport, 7 (4)*, 884–888.

Prazdny K. (1980). Egomotion and relative depth map from optical flow. *Biol. Cybernet., 36*, 87–102.

Regan D., & Beverly K.I. (1982). How do we avoid confounding the direction we are looking and the direction we are moving?. *Science, 215*, 194–196.

Rieger J.H., & Lawton D.T. (1985). Processing differential image motion. *J. Opt. Soc. Am. A, 2*, 354–360.

Rieger J.H., & Toet L. (1985). Human visual navigation in the presence of 3-D rotations. *Biol. Cybernet., 52*, 377–381.

Rodman H.R., & Albright T.D. (1987). Coding of visual stimulus velocity in area MT of the macaque. *Vision Res., 27 (12)*, 2035–2048.

Rodman H.R., & Albright T.D. (1989). Single-unit analysis of pattern-motion selective properties in the middle temporal visual area MT. *Exp. Brain Res., 75*, 53–64.

Royden C.S. (1994). Analysis of misperceived observer motion during simulated eye rotations. *Vision Res., 34*, 3215–3222.

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

Saito H.-A., Yukie M., Tanaka K., Hikosaka K., Fukada Y., & Iwai E. (1986). Integration of direction signals of image motion in the superior temporal sulcus of the macaque monkey. *J. Neurosci., 6 (1)*, 145–157.

Solomon D., & Cohen B. (1992). Stabilization of gaze during circular locomotion in light: I. compensatory head and eye nystagmus in the running monkey. *J. Neurophysiol., 67 (5)*, 1146–1157.

Sperry R.W. (1950). Neural basis of the spontaneous optokinetic response produced by visual inversion. *J. Comput. Physiol. Psychol., 43*, 482–489.

Tanaka K., & Saito H.-A. (1989). Underlying mechanisms of the response specificity of expansion/contraction and rotation cells in the dorsal part of the medial superior temporal area of the macaque monkey. *J. Neurophysiol., 62 (3)*, 642–656.

Tanaka K., Hikosaka K., Saito H.-A., Yukie M., Fukada Y., & Iwai E. (1986). Analysis of local and wide-field movements in the superior temporal visual areas of the macaque monkey. *J. Neurosci., 6 (1)*, 134–144.

Thier P., & Erickson R.G. (1992). Resonses of visual-tracking neurons from cortical area MST-1 to visual, eye and head motion. *Eur. J. Neurosci., 4*, 539–553.

Treue S., & Maunsell J. (1996). Attentional modulation of visual motion processing in cortical areas MT and MST. *Nature, 382*, 539–541.

van den Berg A.V. (1992). Robustness of perception of heading from optic flow. *Vision Res., 32 (7)*, 1285–1296.

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

van den Berg A.V. (1996). Judgements of heading. *Vision Res., 36*, 2337–2350.

von Helmholtz, H. (1896). Handbuch der Physiologischen Optik. Leopold Voss, Hamburg.

Von Holst E., & Mittelstaedt H. (1950). Das Reafferenzprinzip (Wechselwirkung zwischen Zentralnervensystem und Peripherie). *Naturwissenschaften, 37*, 464–476.

Wang H.T., Mathur B.P., & Koch C. (1989). Computing optical flow in the primate visual system. *Neural Comput., 1*, 92–103.

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

Warren W.H. Jr., & Hannon D.J. (1990). Eye movements and optical flow. *J. Opt. Soc. Am. A, 7 (1)*, 160–169.

Warren W.H., Morris M.W., & Kalish M. (1988). Perception of translational heading from optical flow. *J. Exp. Psychol. Hum. Percept. Perform., 14 (4)*, 646–660.

Wertheim A.H. (1987). Retinal and extraretinal information in movement perception: how to invert the Filehne illusion. *Perception, 16*, 299–308.

Wertheim A.H. (1994). Motion perception during self-motion. *Behav. Brain Sci., 17*, 293–355.

Wilson H., & Kim J. (1994). A model for motion coherence and transparency. *Vis. Neurosci., 11*, 1205–1220.