Centric-minded templates for self-motion perception

Paolo Cavalleri, Silvio P. Sabatini *, Fabio Solari, Giacomo M. Bisio

Department of Biophysical and Electronic Engineering, University of Genoa, Via all'Opera Pia 11/A, 16145 Genova, Italy

Received 15 May 2002; received in revised form 21 January 2003

Abstract

We propose a two-layer neuromorphic architecture by which motion field pattern, generated during locomotion, are processed by template detectors specialized for gaze-directed self-motion (expansion and rotation). The templates provide a gaze-centered computation for analyzing motion field in terms of how it is related to the fixation point (i.e., the fovea). The analysis is performed by relating the vectorial components of the act of motion to variations (i.e., asymmetries) of the local structure of the motion field. Notwithstanding their limited extension in space, such centric-minded templates extract, as a whole, global information from the input flow field, being sensitive to different local instances of the same global property of the vector field with respect to the fixation point; a quantitative analysis, in terms of vectorial operators, evidences this property as tuning curves for heading direction. Model performances, evaluated in several situations characterized by conditions of absence and presence of pursuit eye movements, validate the approach. We observe that the gaze-centered model provides an explicit testable hypothesis that can guide further explorations of visual motion processing in extrastriate cortical areas.
© 2003 Elsevier Science Ltd. All rights reserved.

Keywords: Visual motion; Optic flow; Motion field; Heading; Computational model; MST

1. Introduction

Motion flow field, which refers to the changes in the optic array induced by relative motion between the observer and the environment, has long been considered a useful representation for visual motion information (Gibson, 1950). These changes are a rich source of information for guiding navigation and in particular for indicating the direction in which the observer is moving, i.e., his heading. The problem of recovering navigational information from motion field can be tackled mathematically in two different ways: by evaluating the differential properties of the motion field, or by solving the equations that describe the relationship between the observer's motion and the motion field. It is possible to evaluate the first-order differential properties of the motion field (Koenderink, 1986; Verri, Straforini, & Torre, 1992) and hence to reveal the presence of expansion, rotation and translation patterns which correspond to particular situations of ego-motion. Alternatively, from the measure of velocity in a bunch of points on the image plane it is possible to algebraically solve the set of equations describing motion field and to determine observer's motion parameters (Heeger & Jepson, 1990; Longuet-Higgins & Prazdny, 1980).

A different *structural* approach consists in the detection of specific motion field configurations by means of template matching techniques (Dahmen, Franz, & Krapp, 2001; Lappe, Bremmer, Pekel, Thiele, & Hoffmann, 1996; Perrone & Stone, 1994). It is possible to find a relationship between this approach and the others, in the sense that template matching can be seen as a way to approximate vectorial differentials by means of integral operations (cf. Poggio, Verri, & Torre, 1991). The template output combines the responses of sets of subunits tuned to specific values of visual motion parameters (velocity and directions), in such a way to respond maximally to the motion field corresponding to particular observer's motion parameters. Choosing in a proper way the configuration of the templates and the subunits' tuning functions, the space invariance properties of the differential analysis are reproduced and therefore templates allow one to approximate differential operators like divergence and rotor and to obtain their properties of space invariance (Verri et al., 1992; Zhang, Sereno, & Sereno, 1993).

* Corresponding author. Tel.: +39-10-3532092; fax: +39-10-3532289.

E-mail address: silvio@dibe.unige.it (S.P. Sabatini).

Alternatively, templates can be defined to be selective to particular configurations, which are associated to specific acts of motion. In both situations it is possible to establish a relationship between the template output and the 3-D motion parameters of the observer.

The template approach agrees with the experimental observations made on the properties of the neurons in a region of the medial superior temporal (MST) area of the parietal cortex of the macaque, evidencing specific selectivity for large field motion stimuli (Duffy & Wurtz, 1991; Graziano, Andersen, & Snowden, 1994; Orban et al., 1992; Raiguel et al., 1997; Tanaka & Saito, 1989) (cf. also Morrone et al., 2000). Specifically, their sensitivities to shifts in the focus of expansion of radial motion (Duffy & Wurtz, 1995) and to the differences between the movement speed in the center and the periphery (Duffy & Wurtz, 1997b; Orban, Lagae, Raiguel, Xiao, & Maes, 1995) make these neurons ideal candidates for involvement in the analysis of motion field and, in particular in heading determination. MST cells collect inputs from the middle temporal (MT) area that, in its turn, gathers inputs from the primary visual cortex (V1) and all together constitute the cortical motion pathway (Orban, 1992). Several feed-forward models of visual processing from the MT area to MST (Lappe & Rauschecker, 1993; Perrone, 1992; Wang, 1996) have examined the emergence of flow pattern tuning based on the computational pooling of single direction selective units spatially arranged according to the desired templates.

In this paper, to define the template properties of our MST-like units, we introduce a gaze-centered computational hypothesis by which the motion field is analyzed with respect to the fovea. The adoption of such a strategy implies that the model yields information about the deviations between the angle of heading and the angle of gaze, rather than a direct heading estimation. Accordingly, the templates are portions of the same motion field pattern observed when the center of motion is located in the fovea. Since heading directions are not coded in the output of single MST-like neurons, such as in classical template models, but in the whole population, each template does not imply a specific scheme of interconnections from MT-like units for each heading direction. A log-polar transformation is adopted for the topology of the gaze-centered templates to account for the retino-cortical mapping of the motion field.

The properties of the gaze-centered approach are studied from an analytical viewpoint, evidencing how motion parameters are represented in the response distribution of neuron units. First, templates are described by means of vectorial operators (Section 3), and then they are approximated by selectivity functions that are characterized by a greater biological plausibility (Section 4). The performances of the model are evaluated in Section 5. Heading estimation is tested in several situations characterized by the conditions of passive and active navigation. In the case of passive navigation (absence of eye movements), since motion field is the consequence of translational motion only, the model guarantees a very accurate heading estimation. Instead, if gaze is fixed on a static point in the scene (active navigation), eye or head movements introduce ambiguities in the motion field analysis, since effects of eye rotations can be confused with translations along frontoparallel directions. Yet, by exploiting the motion parallax principle (Longuet-Higgins & Prazdny, 1980; Rieger & Lawton, 1985) and by building the templates through MT units sensitive to local velocity differences, the model can cope with these ambiguities and its performances are comparable to those obtained by other models of human heading computation based on visual information, only.

2. Self-motion, motion field and space-variant sensing

The motion of an observer in a static environment can be described at each instant t as a rigid-body motion, by means of two vectors (i.e., kinetic characteristics): the translational velocity $\mathbf{T} = (T_X, T_Y, T_Z)^t$, and the angular velocity $\mathbf{\Omega} = (\Omega_X, \Omega_Y, \Omega_Z)^t$.

If one considers a viewer-centered coordinate frame (see Fig. 1) (i.e., fixed to the observer), each point $\mathbf{P} = (X, Y, Z)^t$ in the 3-D space has an apparent motion given by

$$\dot{\mathbf{P}} = -(\mathbf{T} + \mathbf{\Omega} \wedge \mathbf{P}). \tag{1}$$

According to the inverse perspective mapping, this point projects on the image plane as

$$x = f\frac{X}{Z},$$

Fig. 1. Viewer-centered coordinate frame.

$$y = f\frac{Y}{Z},$$

where f is the focal length of the optical system. Thus, the motion field can be related to the motion vectors of the observer as follows:

$$(\dot{x}, \dot{y})^t = \frac{1}{Z}\mathbf{A}(x,y)\mathbf{T} + \mathbf{B}(x,y)\mathbf{\Omega} \overset{\text{def}}{=} \mathbf{v}_T(x,y) + \mathbf{v}_\Omega(x,y), \quad (2)$$

where $\mathbf{v}_T$ and $\mathbf{v}_\Omega$ are the translational and the rotational components of the velocity field, respectively,

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

$$\mathbf{B}(x,y) = \begin{bmatrix} \frac{xy}{f} & -\left(f+\frac{x^2}{f}\right) & y \\ f+\frac{y^2}{f} & -\frac{xy}{f} & -x \end{bmatrix}$$

and $Z = Z(x,y)$ is the depth of the object projecting in (x,y) at current time.

Given the presence of a wide visual field and sufficient depth changes in the environment, it has been theoretically and experimentally demonstrated (Koenderink & van Doorn, 1987; Warren & Hannon, 1990) that it is possible to recover information about the direction of heading, and thus determine the motion vector $\mathbf{T}$, through a direct analysis of the motion field. In particular, if no eye/head rotations occur and the observer moves along a straight line in a static scene (i.e., $\mathbf{\Omega} = 0$), the motion field shows a radial pattern flowing out from a singularity commonly named *focus of expansion* (FOE) (see Fig. 2(a) and (b)).

Under these conditions, this singularity corresponds to the projection of the heading direction on the retinal plane, defined by the spatial location $(f(T_X/T_Z),$

$f(T_Y/T_Z))$ where $\mathbf{v}_T = 0$. Therefore, the problem of heading estimation, can be solved by analyzing the motion field and locating the position of that singularity. More generally, rotations caused by eye and head movements or self-motion along a curved path may significantly alter the retinal motion field, which results as the superposition of a centrifugal (i.e., radial) component, generated by the forward translation, with a perturbation component due to the head and eye rotations. In such a situation, the FOE shifts or ceases to exist, and, in general, no longer coincides with heading direction (see Fig. 2(c)). Though, if one makes use of information about motion parallax, the confounding rotation flow components cancel out and only information about heading and scene depth survives. The resulting differential motion field still presents a radial structure with difference vectors that approximately point directly toward or away from the image point that corresponds to the observer's heading direction. Differential motion approaches, grounded on the mathematical analysis by Longuet-Higgins and Prazdny (Hanada & Ejima, 2000; Hildreth, 1992; Longuet-Higgins & Prazdny, 1980; Rieger & Lawton, 1985), have been recently proved to be easily implemented by biological vision systems by exploiting the motion-opponent properties found in many cells in area MT, ideally suited to compute differential motion (Born, 2000; Marcar, Xiao, Raiguel, Maes, & Orban, 1995; Raiguel, Hulle, Xiao, Marcar, & Orban, 1995; Xiao, Raiguel, Marcar, Koenderink, & Orban, 1995; Xiao, Raiguel, Marcar, & Orban, 1997).

On the basis of these considerations, the translational velocity $\mathbf{T}$ can be related to a residual velocity field $\mathbf{v}(x,y)$ that represents the "recovered" $\mathbf{v}_T(x,y)$ (if $\mathbf{\Omega}$ is

(a)   (b)   (c)
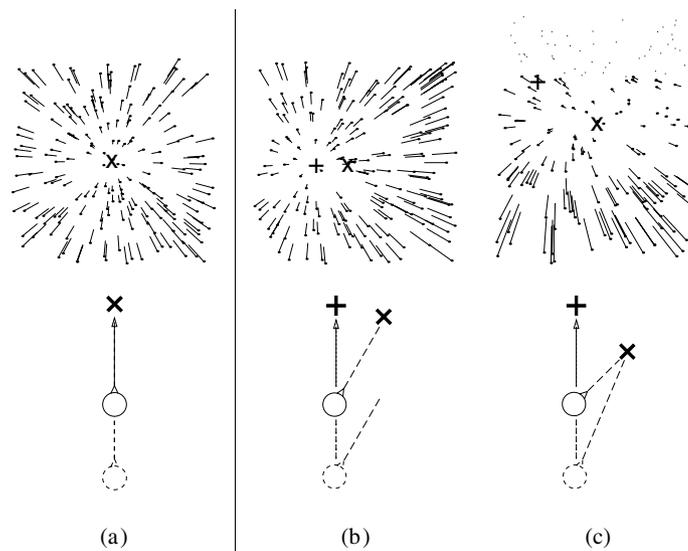
Fig. 2. The flow field induced on the retina has a qualitatively centrifugal structure not only when heading and gaze directions coincide (a) but also during passive navigation with distinct heading and gaze directions (b) or when eye/head movements take place to stabilize the gaze on ground targets (c). The + sign indicates the heading direction and the × sign indicates the gaze direction.

cancelled) or the "actual" $\mathbf{v}_T(x, y)$ (if the effect of $\mathbf{\Omega}$ is measured or null). In the following, we will also refer to $\mathbf{v}(x, y)$ simply as the motion field. Accordingly, Eq. 2 can be rewritten as

$$\mathbf{v}(x, y) = \frac{1}{\tilde{Z}} \mathbf{A}(x, y) \mathbf{T}, \tag{3}$$

where $\tilde{Z} = \tilde{Z}(x, y)$ is a function of the depths of points in the scene that project to a local neighborhood in the image plane.

The intrinsic radial character of the residual velocity field $\mathbf{v}(x, y)$ suggests one to consider radial symmetries centered around the fovea as reference situations with respect to which to analyze the global properties of the motion field (see Section 3). This assumption can be related to a perceptual bias for centrifugal motion which has been reported in psychophysical studies on humans (Ball & Sekuler, 1980; Raymond, 1994; Perrone, 1986), as well as to a centrifugal anisotropy displayed by the distribution of direction selectivities in the motion-processing areas of the visual cortex of cats and monkeys (Albright, 1989; Maunsell & Van Essen, 1983; Rauschecker, Grünau, & Poulin, 1987) suggesting a special preference for movements away from the fovea.

From an architectural point of view, the space-variant characteristics of retino-cortical mapping well adapt to these radial expansion patterns of the motion field. In this way, indeed, fovea-centered expansive and rotatory flow fields are mapped into space-invariant translational velocity patterns according to a criterion of uniform allocation of resources on the cortical surface (Johnston & Wright, 1983; Johnston & Wright, 1986). Functional implications of this geometric transformation reside in the symmetry of the visual field and of the visual system which is radial and not cartesian, as recognized earlier by Gibson (Gibson, 1950, 1966) (see also Zetzsche, Krieger, & Wegmann, 1999). It is therefore natural to adopt a gaze-centered reference frame to represent "ecological" retinal flows, by choosing the position of the fixation point as the center for the representation.

By introducing the retino-cortical transformation we can express the motion field with respect to cortical coordinates, modeling the topological transformation from the retinal plane (x, y) to the cortical plane (ξ, θ), by a log-polar mapping (cf. (Schwartz, 1977) (Tistarelli & Sandini, 1993))

$$\begin{cases} \xi = \log(r + a) = \log(\sqrt{x^2 + y^2} + a), \\ \theta = \theta = \arctan(y/x), \end{cases}$$

where a is a small real constant to make the mapping regular at the origin. By explicating the transformation of velocity from the cartesian to the log-polar plane

$$\begin{bmatrix} \dot{\xi} \\ \dot{\theta} \end{bmatrix} = \begin{bmatrix} e^{-\xi} \cos \theta & e^{-\xi} \sin \theta \\ -\dfrac{\sin \theta}{e^{\xi} - a} & \dfrac{\cos \theta}{e^{\xi} - a} \end{bmatrix} \begin{bmatrix} \dot{x} \\ \dot{y} \end{bmatrix}, \tag{4}$$

it is possible to rewrite the motion field relation (3) as

$$\upsilon(\xi, \theta) = \frac{1}{\tilde{Z}} \mathbf{\Phi}(\xi, \theta) \mathbf{T}, \tag{5}$$

where $\upsilon(\xi, \theta)$ is the residual velocity field in the log-polar coordinates and

$$\mathbf{\Phi}(\xi, \theta) = \begin{bmatrix} -fP(\xi) \cos \theta & -fP(\xi) \sin \theta & P(\xi)/Q(\xi) \\ fQ(\xi) \sin \theta & -fQ(\xi) \cos \theta & 0 \end{bmatrix} \tag{6}$$

with $P(\xi) = e^{-\xi}$ and $Q(\xi) = (e^{\xi} - a)^{-1}$. Such polar mapping separates locally the angular and radial components of image velocities and allows, through its log-scaling property, an almost invariant representation of the global flow patterns.

3. Gaze-centered operators for motion field analysis

In a retinotopic frame of reference, the detection of the direction of heading means specifically the determination of the direction of motion of the observer with respect to the direction of gaze. From a computational point of view, direction of heading can be recovered by the degree of deviation of the motion field from a radial flow centered on the fovea. Such deviations result in radial asymmetries of the velocity vector field in both direction and magnitude. In fact, if the direction of gaze and the direction of movement are coincident, resulting in a null gaze-movement angle, then the velocity flow field is symmetric with respect to both the directions and the magnitudes [1] of image element trajectories. As gaze-movement angle increases, the asymmetry in the distribution of image velocities also increases and such asymmetries can be used to determine heading. In particular, several authors (Dyre & Andersen, 1997; Richards, 1975; Telford & Howard, 1996) suggested that the observer could use the asymmetries in the retinal flow pattern to guide a sequence of fixations until the distribution of image velocity is symmetric, resulting in the coincidence of the directions of movement and gaze. This approach has received relatively little attention in the literature about heading, though it may have a number of interesting functional implications for purposive active vision in relation to the problem of understanding how and by which cortical mechanisms the visual system contributes to proprioception.

In this perspective, we adopt a gaze-centered approach, in which the motion field is analyzed in terms of how it is related to the fixation point. Accordingly, asymmetries in the field distribution are analyzed in the neighborhood of each location $\mathbf{r} = (r, \theta)$ of the retinal

---

[1] It should be observed that scene with non symmetric depth changes would affect the symmetry of velocity magnitudes.

plane through two independent sets of operators sensitive to radial and rotational patterns around the fovea, respectively,

$$e_{\mathrm{EXP}}(\mathbf{r}) = \int_{S(\mathbf{r})} \mathbf{h}_{\mathrm{EXP}}(\mathbf{r}') \cdot \mathbf{v}(\mathbf{r}') \, \mathrm{d}\mathbf{r}',$$

$$e_{\mathrm{ROT}}(\mathbf{r}) = \int_{S(\mathbf{r})} \mathbf{h}_{\mathrm{ROT}}(\mathbf{r}') \cdot \mathbf{v}(\mathbf{r}') \, \mathrm{d}\mathbf{r}', \qquad (7)$$

These two operators represent abstract cortical cells that integrate, over a limited spatial domain $S(\mathbf{r})$ (i.e. their receptive fields), the dot product between the motion field and a set of weight vector fields of unitary expansion and rotation rate, defined as

$$\mathbf{h}_{\mathrm{EXP}}(\mathbf{r}) = \frac{1}{2}\mathbf{r},$$

$$\mathbf{h}_{\mathrm{ROT}}(\mathbf{r}) = \frac{1}{2}\mathbf{n} \wedge \mathbf{r}, \qquad (8)$$

where $\mathbf{n}$ is the unit vector perpendicular to the image plane. These operators act as matched filters detecting the presence of a particular directional pattern within the perceived motion field and producing their maximal output when the local flow vectors are parallel to the weight vector fields (see Fig. 3).

They contribute with a local measure of how much the motion field is radial and tangential, respectively.

3.1. Direction tuning in the retinal plane

Notwithstanding their limited extension in space, such centric-minded operators, extract, as a whole, global information from the input flow field, being sensitive to different local instances of the same global property of the vector field with respect to the fixation point. To clear up

Fig. 3. During passive navigation motion field expands out from a point O', which represents the projection of heading direction on the retinal plane. Expansion and rotation operators are described as portion of radial and circular flow fields centered on the fovea (point O). Each operator integrates on a circular sector, $S(\mathbf{r})$, to obtain the degree of matching between the input velocity vector and those that define the filter operator.

this point, let us consider the case of passive navigation towards a planar surface perpendicular to the direction of gaze. When an observer looks at a direction different from the one towards which he or she is moving, the motion field expands out from a point O' not coinciding with the fovea (see Fig. 3). Formally, denoting with γ the expansion rate and with $\mathbf{p} = (p, \theta_{\mathrm{p}})$ the position vector of the focus of expansion, and substituting the velocity field $\mathbf{v}(\mathbf{r}') = \gamma(\mathbf{r}' - \mathbf{p})$, in Eqs. (7) and (8) we have

$$e_{\mathrm{EXP}}(\mathbf{r}; \mathbf{p}) = \frac{\gamma}{2}\left(\int_{S(\mathbf{r})} \mathbf{r}'^2 \mathrm{d}\mathbf{r}' - \mathbf{p} \cdot \int_{S(\mathbf{r})} \mathbf{r}' \, \mathrm{d}\mathbf{r}' \right),$$

$$e_{\mathrm{ROT}}(\mathbf{r}; \mathbf{p}) = -\frac{\gamma}{2}(\mathbf{p} \wedge \mathbf{n}) \cdot \int_{S(\mathbf{r})} \mathbf{r}' \, \mathrm{d}\mathbf{r}'.$$

Assuming that the centroid of the plane region $S(\mathbf{r})$ that defines the receptive field coincides with $\mathbf{r}$: [2]

$$e_{\mathrm{EXP}}(r, \theta; p, \theta_{\mathrm{p}}) = \frac{\gamma}{2}(\mu - A\mathbf{p} \cdot \mathbf{r})$$

$$= \frac{\gamma}{2}[\mu - Apr\cos(\theta_{\mathrm{p}} - \theta)] \qquad (9)$$

$$e_{\mathrm{ROT}}(r, \theta; p, \theta_{\mathrm{p}}) = -\frac{\gamma}{2}A(\mathbf{p} \wedge \mathbf{n}) \cdot \mathbf{r}$$

$$= -\frac{\gamma}{2}Apr\cos\left(\theta_{\mathrm{p}} - \theta - \frac{\pi}{2}\right), \qquad (10)$$

where A is the area of the receptive field and μ is its moment of inertia about an axis through the origin perpendicular to the retinal plane. Expressions (9) and (10) can be interpreted as sinusoidal direction tuning curve, whose preferred direction depends on the angular position θ of the specific integration domain considered (i.e., on the angular position of the cell's receptive field). Though, by considering the relative angle $\theta_{\mathrm{p}} - \theta$ as the independent variable, the tuning curves of the cells result invariant to their position and present a maximum at $\theta_{\mathrm{p}} - \theta = \pi$ for an expansion selective cell and at $\theta_{\mathrm{p}} - \theta = \pi/2$ for a rotation selective cell. This means that this "relative" tuning is a property that is shared by *all* the cells of the population, independently of their absolute position.

As an example, let us consider $S(\mathbf{r})$ as an almost "square" circular sector

$$S(\mathbf{r}) = \left\{ r - \frac{\Delta r}{2} < r' < r + \frac{\Delta r}{2} \right\} \bigcap \left\{ \theta - \frac{\Delta\theta}{2} < \theta' \right.$$

$$\left. < \theta + \frac{\Delta\theta}{2} \right\} \quad \text{with} \quad \Delta\theta \sim \frac{\Delta r}{r},$$

the excitation for an expansion selective cell located in $\mathbf{r}$ is

$$e_{\mathrm{EXP}}(r, \theta; p, \theta_{\mathrm{p}}) = \gamma r^3\left[\lambda\frac{p}{r}\cos(\theta_{\mathrm{p}} - \theta - \pi) + \kappa \right] \qquad (11)$$

---

[2] This hypothesis is not too restrictive since we can reasonably choose an integration domain S(**r**) symmetric with respect to its center.

and, similarly, for a rotation selective cell

$$e_{ROT}(r, \theta; p, \theta_p) = \gamma r^3 \left[\lambda \frac{p}{r} \cos \left(\theta_p - \theta + \frac{\pi}{2} \right) \right] \quad (12)$$

with

$$\kappa = \left(1 + \frac{1}{12} \frac{\Delta r^2}{r^2} \right) \frac{\Delta r}{r} \frac{\Delta \theta}{2},$$

$$\lambda = \frac{\Delta r}{r} \sin \left(\frac{\Delta \theta}{2} \right),$$

Fig. 4. Direction tuning curves of the expansion and rotation operators for a fixed relative size of the receptive field ($\Delta r/r = \pi/3$) and for different values of the relative position of the center of motion with respect to the receptive field center ($p/r = 0.5, 1, 2, 4$). Expansion operators have their maximal response when $\theta_p - \theta = \pi$, i.e. when the focus of expansion is located in the region of the visual field opposed to the integration domain of the operator. Rotation operators have their maximal response when $\theta_p - \theta = \pi/2$. As the center of motion moves toward the fovea the direction tuning curves broaden, evidencing a loss of sensitivity when the gaze is almost aligned with the heading direction.

where $\Delta r/r$ indicates the relative size of the receptive field respect to its radial position and p/r is the relative position of the center of motion with respect to the receptive field center.

The resulting direction tuning curves are shown in Fig. 4a–b, for a fixed ratio of $\Delta r/r = \pi/3$ and for increasing values of p/r.

Such curves evidence that cells respond in a consistent way, to a global property of the motion field, irrespective to relative position of their receptive field centers, i.e., independently of the fact that on a local basis, such a property shows itself in different ways in their receptive fields. In the next Section we will show that the activity distribution of a population of such cells can be used to gain information about global motion parameters.

3.2. Heading representation in the cortical plane

In general, we can express the cells' responses in terms of the translational velocity components. In these cases, it is convenient to take advantage of the log-polar mapping and to define the gaze-centered weight vector fields (Eqs. (8)) in the cortical plane

$$\mathbf{h}_{EXP}(\xi, \theta) = \left[\frac{1}{2}(1 - ae^{-\xi}), 0 \right] \simeq \left(\frac{1}{2}, 0 \right)$$

$$\mathbf{h}_{ROT}(\xi, \theta) = \left(0, \frac{1}{2} \right),$$

evidencing that their specific tuning for radial and angular (i.e., tangential) components of the motion field result in the cortical plane in uniform fields directed along the ξ- and θ-axes, respectively. [3] In the cortical plane Eq. 7 become

$$e_{EXP}(\xi, \theta) = \int_C \mathbf{h}_{EXP}(\xi', \theta') \cdot \boldsymbol{v}(\xi', \theta') \, d\xi' \, d\theta',$$

$$e_{ROT}(\xi, \theta) = \int_C \mathbf{h}_{ROT}(\xi', \theta') \cdot \boldsymbol{v}(\xi', \theta') \, d\xi' \, d\theta', \quad (13)$$

where $\boldsymbol{v}(\xi, \theta)$ is the residual velocity field in cortical coordinates (see Eq. 5).

In the case of passive navigation, one has

$$e_{EXP} \left(\xi, \theta; \frac{T_X}{T_Z}, \frac{T_Y}{T_Z} \right)$$
$$= \frac{T_Z}{\tilde{Z}} \left[Q_1(\xi) \left(-\frac{T_X}{T_Z} \cos \theta - \frac{T_Y}{T_Z} \sin \theta \right) + Q_2(\xi) \right], \quad (14)$$

$$e_{ROT} \left(\xi, \theta; \frac{T_X}{T_Z}, \frac{T_Y}{T_Z} \right) = \frac{T_Z}{\tilde{Z}} \left[P_1(\xi) \left(\frac{T_X}{T_Z} \sin \theta - \frac{T_Y}{T_Z} \cos \theta \right) \right] \quad (15)$$

---

[3] Since a is a small constant the approximation of the radial component of $\mathbf{h}_{EXP}$ by a constant is acceptable.

with $P_1(\xi)$, $Q_1(\xi)$ and $Q_2(\xi)$ quantities depending on the position and spatial extension of the receptive fields, their detailed expressions are reported in the Appendix. $\tilde{Z}$ is an average function of the depth of objects in the domain specified by C.

From Eqs. (14) and (15) one can directly analyze the tuning properties of expansion and rotation selective cells to heading direction $(T_X/T_Z, T_Y/T_Z)$. Each cell response, as a function of $(T_X/T_Z, T_Y/T_Z)$, is described by a linear expression parametrized by coefficients that depend on the cell's cortical location (ξ, θ), where the dependence on θ is the most relevant one. Accordingly, for each position on the cortical plane, cells are selective to a specific combination of T_X/T_Z and T_Y/T_Z. Fig. 5 shows the responses of a set of expansion cells, whose receptive field centers vary along a spiral on the retinal plane. Cells' responses increase steadily with the intensity of the velocity components; keeping fixed $(T_X^2 + T_Y^2)$ and T_Z, each cell has a maximal response for a preferred heading direction, as evidenced in the related coordinate frames shown in the figure.

By construction, the preferred heading direction of each cell is always along the conjunction between its receptive field center and the gaze. By analyzing their
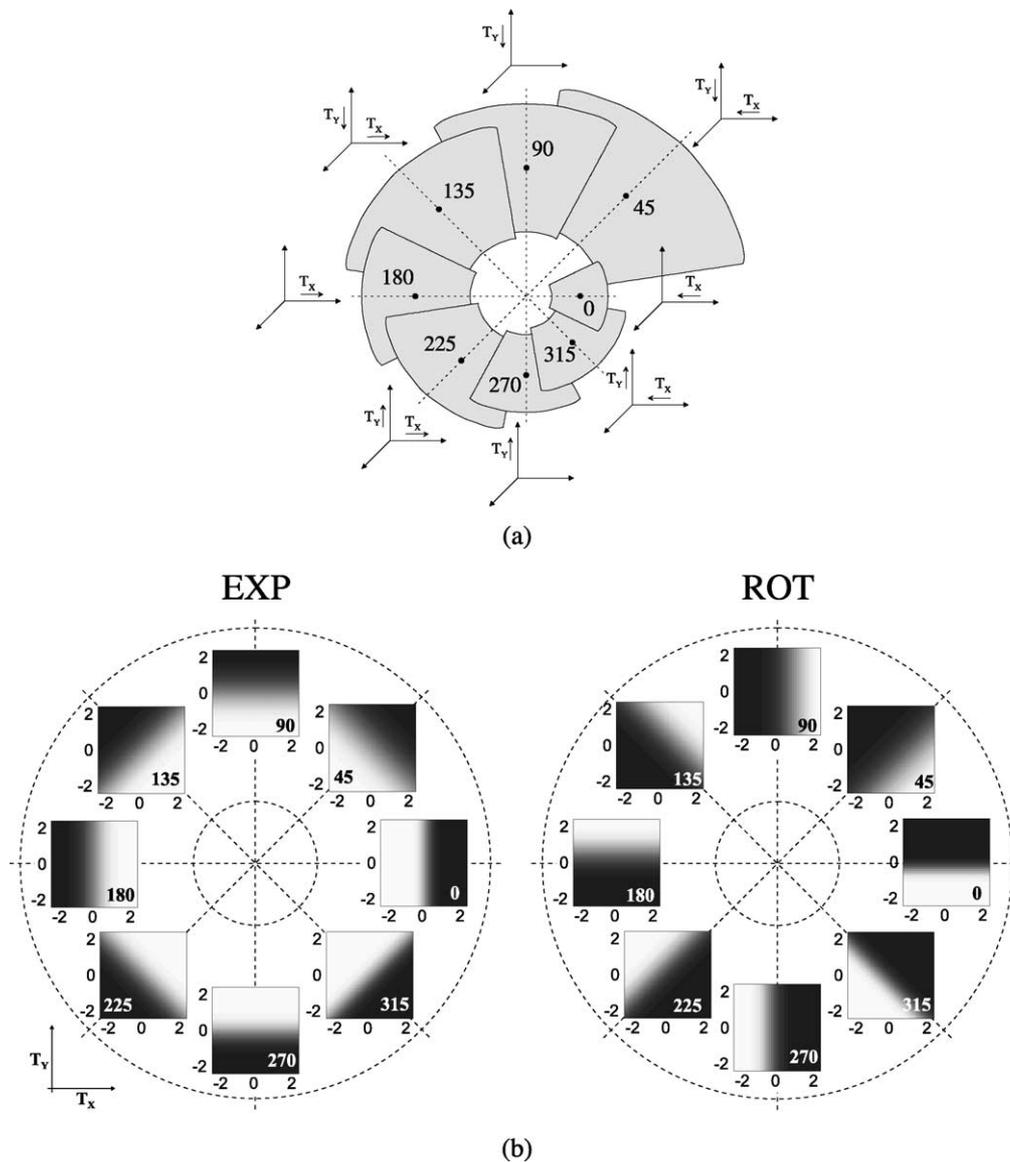
Fig. 5. (a) Schematic representation of how expansion cells are selective to different heading directions, according to the position of their receptive fields. For the sake of the representation we displayed the receptive fields as grayed sectors along a spiral on the retinal plane. The preferred heading direction (T_X, T_Y) of a cell is evidenced in the corresponding 3-D coordinate frame; (b) responses for the expansion and rotation selective cells whose receptive field centers are located as in (a), when the scene is constituted by a frontoparallel plane. Each response is shown as a grayscale map: the brightest region indicates, through the pairs $(\frac{T_X}{T_Z}, \frac{T_Y}{T_Z})$, the position of the center of motion which elicits the highest cell response. Expansion cells maximally respond when the center of motion is located in the opposed quadrant. The responses of rotation selective cells are rotated by $\pi/2$ with respect to those of expansion selective cells.

responses as a function of T_X/T_Z and T_Y/T_Z we can evidence how single cell responses in the model contribute to the population encoding of the perceived heading. The responses for expansion and rotation selective cells, whose receptive field centers are localized as in Fig. 5(a), are shown as grayscale maps in Fig. 5(b). Brightest regions indicate, through the pairs $(T_X/T_Z, T_Y/T_Z)$ the positions of the centers of motion which elicit the highest cells' responses. According to their receptive field locations, cells are selective to centers of motion that are shifted in different positions in the visual field. By example an expansion selective cell with its receptive field at $\theta = 45°$ maximally responds when the center of motion is located in the lower left hemifield. For rotation selective cells the response maps appear rotated by $\pi/2$ with respect to those of expansion selective cells.

Eqs. (14) and (15) evidence a sort of *aperture problem* for heading estimation, similar to that one has usually to face when attempting to estimate the velocity components of a moving object observed through a limited aperture (Marr & Ullman, 1981). The locus of headings compatible with a given response level of a cell is indeed a straight line in the $(T_X/T_Z, T_Y/T_Z)$ plane; this line represents a sort of heading constraint. Therefore, this implies that the response of a single model cell is not sufficient to fully estimate heading direction. In other words, given the response of a cell located in (ξ, θ) there is a line of possible solutions for $(T_X/T_Z, T_Y/T_Z)$ and for each different position there will generally be a different constraint line. The combination of the responses of different cells gives the intersection of these constraints thus disambiguating the true direction of heading. It is also worth noting that if one considers, for a fixed cortical location, a pair of quadrature cells (i.e. one expansion selective and the other rotation selective) their responses define a pair of perpendicular axes in the plane of heading that represent a complete orthogonal basis for representing heading.

4. Centric-minded templates

The vectorial formulation presented in Section 3 provides a formal way to describe how the motion field can be analyzed with respect to the gaze direction. Here, we want to define a biologically-plausible template model that approximates such functionalities on the basis of the computational resources available in the cortical areas of the primate visual system. The model, formalized as a two-layer network, that represent the MT and the MST areas, respectively, still preserves the qualitative conclusions drawn in the vectorial analysis. The cells of the first layer, characterized by a specific selectivity to velocity vectors, approximate the dot product in the weight vector field operators, whereas, the cells of the second layer provide the global motion

field analysis. To this purpose, we adopted grounded models for MT velocity tuning (Lagae, Raiguel, & Orban, 1993; Maunsell & Van Essen, 1983; Rodman & Albright, 1987). In particular, for the extraction of the differential motion information we followed the model proposed by Royden (1997) based on MT-like units characterized by motion-opponent receptive fields, similar to the ones experimentally observed in primates (Born, 2000; Marcar et al., 1995; Raiguel et al., 1995; Xiao et al., 1995; Xiao et al., 1997). On this basis, we built our own centric-minded MST template model through the collection of MT cell outputs, according to appropriate spatial arrangements of the afferent contributions.

More specifically, for MT cells, the direction and speed selectivity to local image motion in each position of the receptive field is defined as a product of two separable Gaussian tuning functions that approximate the dot product between the actual velocity vector and a neuron's preferred direction of motion

$$S_{\text{speed}}(v) = \frac{1}{\sqrt{2\pi}\sigma_v} \exp\left[-\frac{1}{2}\left(\frac{v - v_0}{\sigma_v}\right)^2 \right], \tag{16}$$

$$S_{\text{dir}}(d) = \frac{1}{\sqrt{2\pi}\sigma_d} \exp\left[-\frac{1}{2}\left(\frac{d - d_0}{\sigma_d}\right)^2 \right], \tag{17}$$

where $v \stackrel{\text{def}}{=} \sqrt{\dot{x}^2 + \dot{y}^2}$ and $d \stackrel{\text{def}}{=} \arctan \dot{y}/\dot{x}$ are the speed and the direction of the local velocity in the retinal plane; v_0 and d_0 are the cell's preferred speed and direction. The overall model MT neuron's activation is computed by a weighted sum of spatially pooled contributions in the cell's receptive field:

$$
\begin{aligned}
e_{\text{MT}}(\xi, \theta) = \int_C & W(\xi - \xi', \theta - \theta') \\
& \times S_{\text{dir}}[d(x(\xi', \theta'), y(\xi', \theta')) - d_0(\xi, \theta)] \cdot \\
& S_{\text{speed}}[v(x(\xi', \theta'), y(\xi', \theta')) - v_0(\xi, \theta)] \\
& \times \mathrm{d}\xi' \mathrm{d}\theta'.
\end{aligned}
\tag{18}
$$

Assuming that a mapped representation of the retinal velocity vector field is available as cortical input on the cortical plane, the weighting function $W(\xi, \theta)$ is modeled, in cortical coordinates, as a radially symmetric Gaussian, for non-opponent (i.e., synergistic) operators, and as a directional derivative of a Gaussian for opponent (i.e., antagonist) operators

$$W = W_{\text{no-opp}}(\xi, \theta) = \frac{1}{2\pi\sigma^2} \exp\left[-\frac{1}{2\sigma^2}\left(\xi^2 + \theta^2\right) \right]$$

for non-opponent cells, or

$$W = W_{\text{opp}}(\xi, \theta) = \frac{\partial}{\partial \xi} W_{\text{no-opp}}(\xi, \theta)$$

for opponent cells (see Fig. 6).
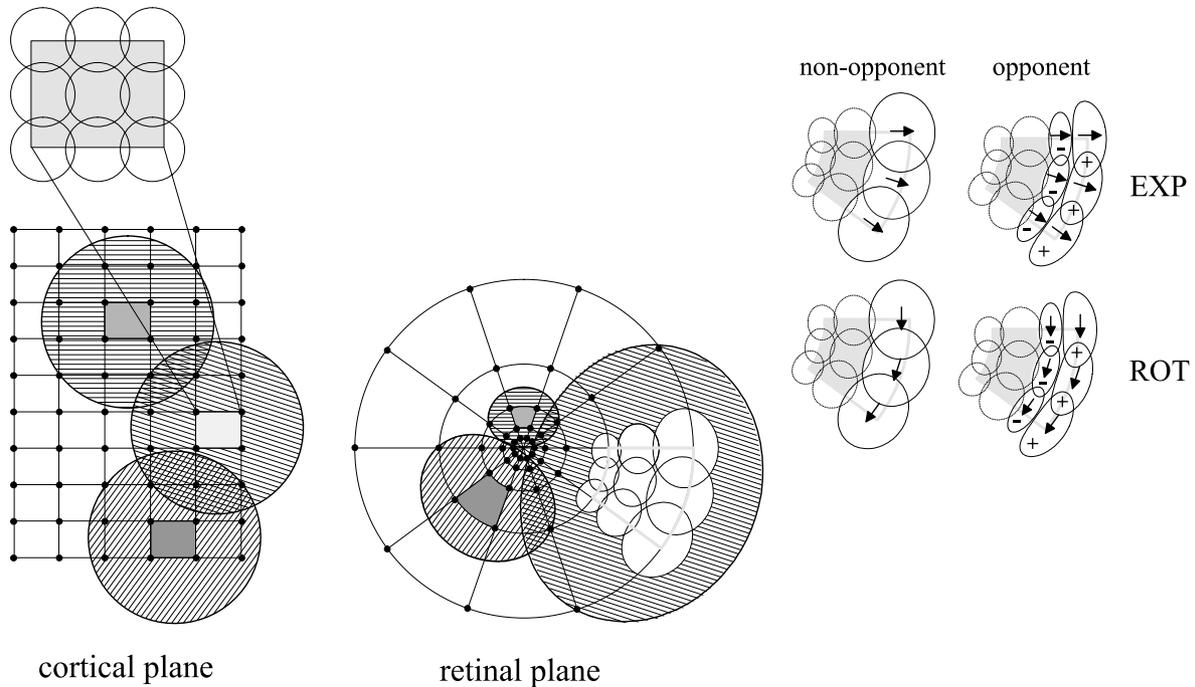
Fig. 6. Schematic representation of the two-layer network architecture: (a) cortical plane sampling in cartesian coordinates. For the sake of representation, the shown nodes represent only a subset of the model cortical cells. The large hatched circles indicate MST receptive fields, whereas the small circles in the zoom are MT receptive fields; (b) mapping of MT and MST receptive fields on the retinal plane. It is worthy to note the space variance in the receptive field size and (c) arrangement, in the retinal plane, of opponent and non-opponent MT cell receptive fields to build expansion and clockwise rotation MST templates.

To keep the complexity of the model to a minimal, we have considered opponent operators in which the angle of the axis between the excitatory and inhibitory regions is always orthogonal to the radius.

To perform as centric-minded motion field templates, the MST-like cells of the second layer analyze a portion of the visual field and collect the outputs of a set of cells from the first layer tuned to proper direction and speed values in relation to the position of their receptive field centers with respect to the fovea

$$e_{\mathrm{MST}}(\xi, \theta) = \int_C G(\xi - \xi', \theta - \theta') e_{\mathrm{MT}}(\xi', \theta') \, \mathrm{d}\xi' \, \mathrm{d}\theta' \qquad (19)$$

where G localizes the cell's receptive field in the cortical plane C. The resulting selectivity of a MST-like cell to expansion and rotation is essentially a result of the arrangement of the MT-like units within its receptive field. By example, an expansion cell in the second layer collects contributions from cells, in the first layer, selective to velocity vectors radially directed out from the fovea, and to speed values increasing with the eccentricity of their receptive fields: the functions $d_0(\xi, \theta)$ and $v_0(\xi, \theta)$ specify just these dependences of the direction and speed tuning parameters of MT-like cells in relation to their receptive field positions in the retinal plane. The spatial dependences of d_0 and v_0, have to be very specific to implement the functionality of radial and rotational templates (see Fig. 6). For the radial template $d_0(\xi, \theta)$ is

the direction of the radius going from the coordinate origin to the point $[x(\xi, \theta), y(\xi, \theta)]$, while for the rotational template $d_0(\xi, \theta)$ is the direction perpendicular to such radius. For what concerns the velocity modulus, in the case of both the radial and the rotational template, $v_0(\xi, \theta)$ increases linearly with the distance of point $[x(\xi, \theta), y(\xi, \theta)]$ from the coordinate origin depending on a specific rate. Furthermore, we note that σ_v will increase linearly with the radius to keep constant the relative bandwidth of the tuning curve S_{speed}.

Granting that the spatial dependences of $d_0(\xi, \theta)$ and $v_0(\xi, \theta)$ have to be very specific to implement the radial and rotational centric-minded templates, it is worthy to note that the sampling strategy of the MT afferents is considerably simplified in the log-polar cortical coordinates, changing from space-variant radial geometries, in the retinal plane, to space-invariant cartesian geometries in the cortical plane (see Fig. 6).

Since in the model both non-opponent and opponent MT receptive fields are available, the resulting MST template is able to handle motion field characterized by non-null rotational components $(\boldsymbol{\Omega} \neq 0)$. Indeed, by adopting MST cells that collect afferent contributions from MT motion-opponent units, we can approximate the matching operation between the residual differential flow and the centric-minded template, that has been analytically investigated for ideal vector templates. More precisely, motion-opponent operators provide

information about the orientation line of the difference velocity vector, only. To disambiguate its direction additional information on the 3-D structure of the scene should be used. The behavior of these cells suggests that in MST area are present two kinds of cells: one dedicated to the analysis of the motion of a translating passive observer and one for the analysis of self-motion with tracking eye movement (Roy, Komatsu, & Wurtz, 1992; Roy & Wurtz, 1990).

In our simulations, the retinal plane is sampled by a 60×40 grid in cortical coordinates. For each location we consider one MT cell and a triplet of MST cells (one selective to expansion, one to clockwise rotation and one to counterclockwise rotation). The receptive field diameter of MT cells is fixed to five cortical units, whereas it is fixed to 13 cortical units for MST cells. This corresponds in the retinal plane to a variable receptive field size increasing with eccentricity (average square root of receptive field area of 17° and 45°, for MT and MST cells, respectively). Each receptive field is spaced by one cortical unit, thus resulting in a significative overlap between adjacent operators. The motion field is sampled on a log-polar grid and weighted by a Gaussian receptive field. In the simulations 13×13 samples of velocity are analyzed within each template for each retinal eccentricity.

5. Results

With these specifications on the computational properties of centric-minded templates we intend to demonstrate how they can provide heading estimation. The vectorial analysis (see Section 3) evidenced that the use of local gaze-centered operators allows one to recover information about the global properties of the motion field, and thence on heading direction. All the local responses can indeed be directly compared among each other, being all derived by a "common node analysis". The spatial distribution of the operators' outputs is indeed characterized by a marked anisotropy when the focus of expansion is peripheral, and the location of the peak of activity can be related to the location of the focus of expansion, and thereof to the heading. The distribution of activity turns to be isotropic (i.e., independent of the angular coordinate of the receptive field) as the focus moves towards the center of the visual field. More specifically, the layer of the expansion-selective cells will be isotropically activated, whereas the layer of the rotation-selective cells will be isotropically silent. For vector operators, such behaviour allowed us to derive real sinusoidal direction tuning that can be used to guide a sequence of fixations to align the direction of heading and the direction of gaze. It can be verified that, in the case of Gaussian templates, similar specificities occur.

Let us at first consider the case of motion toward a frontoparallel plane. In this case of pure translation towards a frontoparallel plane, the motion field has a radial structure centered on the heading direction. Since the depth is constant for all the points in the scene, the speed increases linearly from the focus of expansion to the periphery (see Eq. (2)). As shown in Fig. 7, the activity distribution of expansion cells reaches a peak in the region where the motion field is maximally radial in the foveal frame, i.e., in the region opposed to the position of the focus of expansion respect to the fovea. For the cells selective to rotational flow the distribution of activity peaks in the region where the motion field is maximally tangential. We can observe that, in the case of expansion and rotation vector operators, the peaks would occur along orthogonal radial directions (cf. Eqs. 9 and 10), but do not occur necessarily so for templates. Gaussian tuning differs from dot product, since it does not accomplish a vector decomposition of velocities, but, for a given modulus, searches for the best direction matching. Consequently, the two peaks are no more constrained in their relative positions to be located along orthogonal radial directions. In general, more complex is the scene (objects at different depth that fall in different portions of the visual field), more independent is the information supplied by expansion cells respect to rotation cells. The two classes of cells are, indeed, designed to be sensitive to different instances of the motion field originated by the same cause (e.g., ego-motion), resulting from the apparent velocities of objects in different positions in the visual field.

On the basis of these considerations, it is possible to derive a simple and direct strategy for the control of the heading, based on the complementary use of the responses of expansion and rotation cells. This kind of strategy can be very likely implemented at cortical level, relying upon a mere localization of the peak in the distribution of activity of MST-like cells, rather than resorting to algebraic resolution of linear equations or to the detection of flow singularities (this latter approach would be very sensitive to noise). The implementation of such a control strategy in a closed perception-action loop, would not require an explicit estimation of the heading direction, rather information to guide a correction saccade to align the direction of heading with the direction of gaze. In these conditions, the system "knows" which was the direction of heading only when it has already reached the alignment of the gaze in that direction. Though, it is possible to define a procedure to analyze the "open-loop" model's performance by direct observation of the spatial distribution of activity of the MST-like cells. In particular, the activity distributions for clockwise (CW) and counterclockwise (CCW) rotation cells provide sufficient information to localize heading direction through a triangulation procedure (see Fig. 7). In such a way, we can directly evaluate model's

FRONTOPARALLEL PLANE
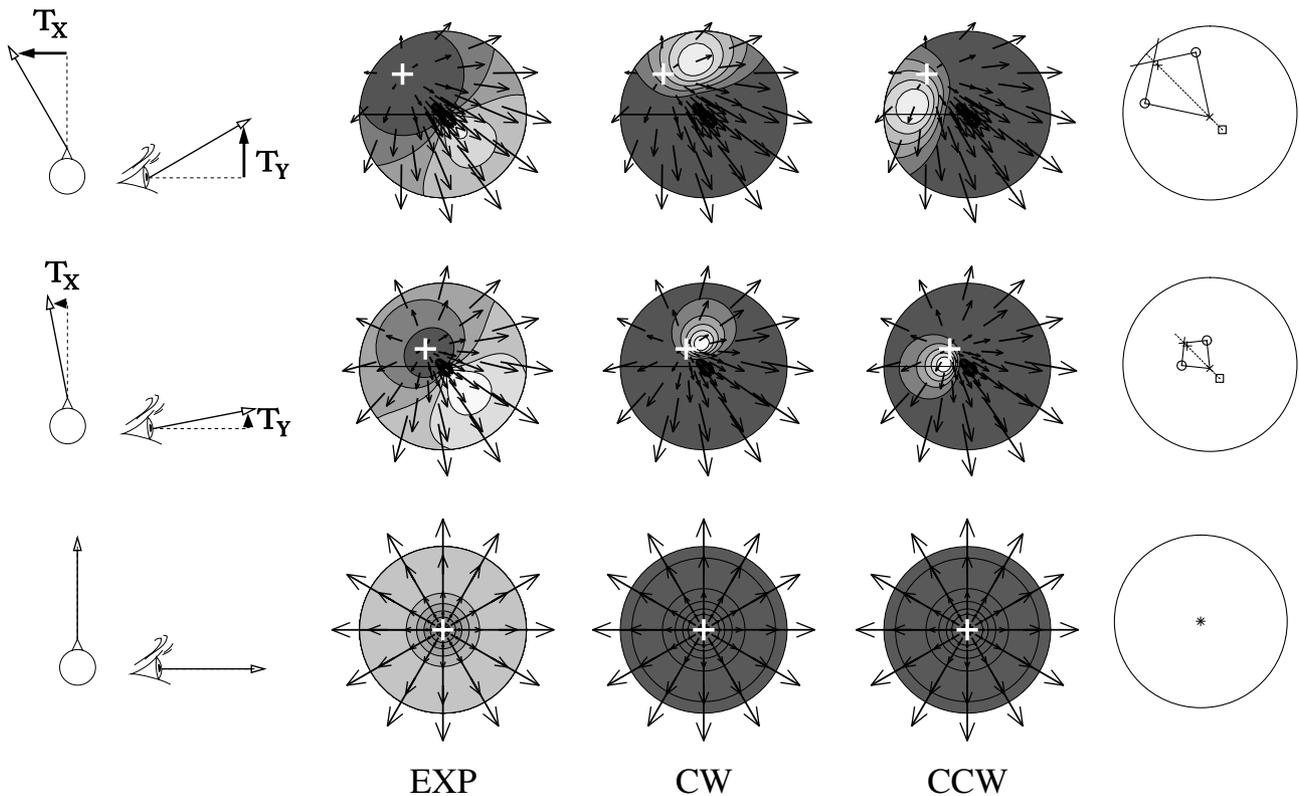
Fig. 7. Distributions of activity of expansion (EXP) and rotation selective cells (CW and CCW) for three situations of passive navigation towards a frontoparallel plane, characterized by different values of the heading-gaze angle. In each retinal location the gray level indicates the activity level of the corresponding cortical unit. The white cross evidences the position of heading. When heading is distant from the fovea the cells' responses are peaked, while, when heading and gaze coincide, the responses are maximally uniform. In the rightmost column the triangulation procedure used to localize heading direction is schematized: (×) is the fovea location; (+) is the heading direction, the small square locates the maximum of the responses for the expansion cells, the two small circles locate the maxima for the CW and CCW rotation cells.

performances, and defer to a future work the implementation of the control algorithm. In quantitative terms the resulting estimation of heading is shown in Fig. 8. A quite accurate estimation is obtained for the frontoparallel case, that represents the most simple test. In the following, we report on the performance of the model in more complex situations (straight paths towards a cloud of dots and over a ground floor), for both passive and active navigation.

5.1. Passive navigation

As already recalled in Section 2, in absence of rotation (i.e., under passive navigation), the motion field radiates outward from a single location (FOE) determined by the direction of heading, independently of the structure of the scene. Only for the frontoparallel case the distribution of image velocity magnitudes is radially symmetric and linearly increasing from the FOE. Here we want to investigate how the centric-minded model copes with the field distributions originated by motion through more complex scenes.

5.1.1. Cloud of dots

Let us assume that the depth of the objects present in the scene is not constant. To simulate the motion field experienced in such a situation, we considered an observer moving towards a cloud of dots. The cloud of dots has been obtained considering the depth of objects varying randomly in a range [0.5–11 m], the speed of the observer is 1 m/s. Fig. 9(a) shows one instance of the activity distributions of the cortical units. The location of the maxima are still well defined, the irregularities and the slight asymmetry of the motion field give rise to minor differences in the activation patterns of CW and CCW rotation cells, but do not impair the estimation procedure (see Fig. 10).

Even if the motion field vectors do not match anymore the linear speed selectivity of the model cells, errors remain small in the wide range of heading direction considered.

5.1.2. Ground floor

Another natural situation of ego-motion is represented by an observer translating over a ground floor
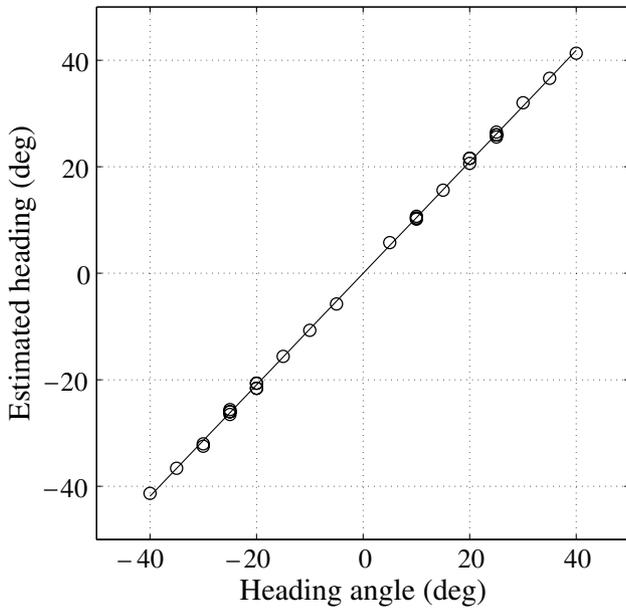
Fig. 8. Estimated heading as a function of the heading angle. Both quantities are with reference to the fixation direction. The simulations refer to the motion fields experienced by an observer moving without rotation towards a frontoparallel plane on which there is a random distribution of points. For each heading angle five different random distributions of points were considered.
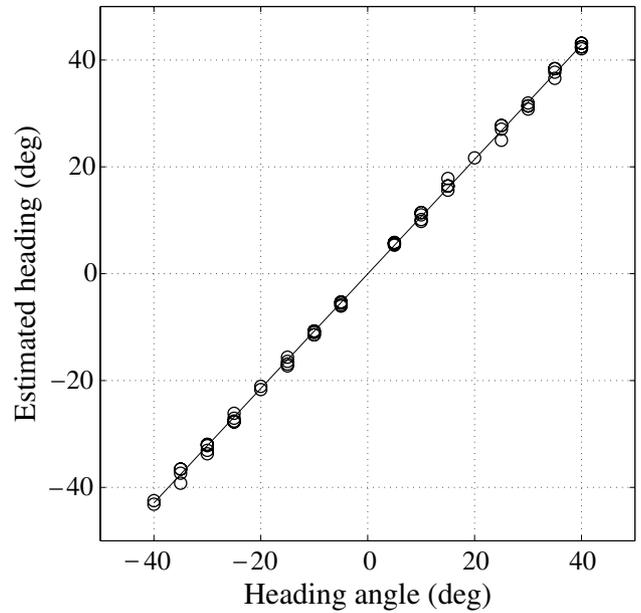
Fig. 10. Estimated heading as a function of the heading angle. The simulations refer to the motion fields experienced by an observer moving without rotation towards a cloud of dots. For each heading angle five different random distributions of points were considered.

clear of obstacles. In this condition, the only region of the visual field where the stimulus is present is the low hemifield, that is limited by the horizon at the top. Since the stimulus does not cover all the visual field, the peaks of the cell activity distributions cannot occur in the upper hemifield, and this will influence the precise de-

tection of heading. Though, the perceptual task is simplified by the fact that the observer can move only on the plane and thus the prediction should be about left- or right-ward motion only. The observer's eye is located at 1 m over the ground and the down-ward vertical angle of the Z-axis is equal to arctan(0.1), and the Z component of the speed of the observer is 1 m/s.
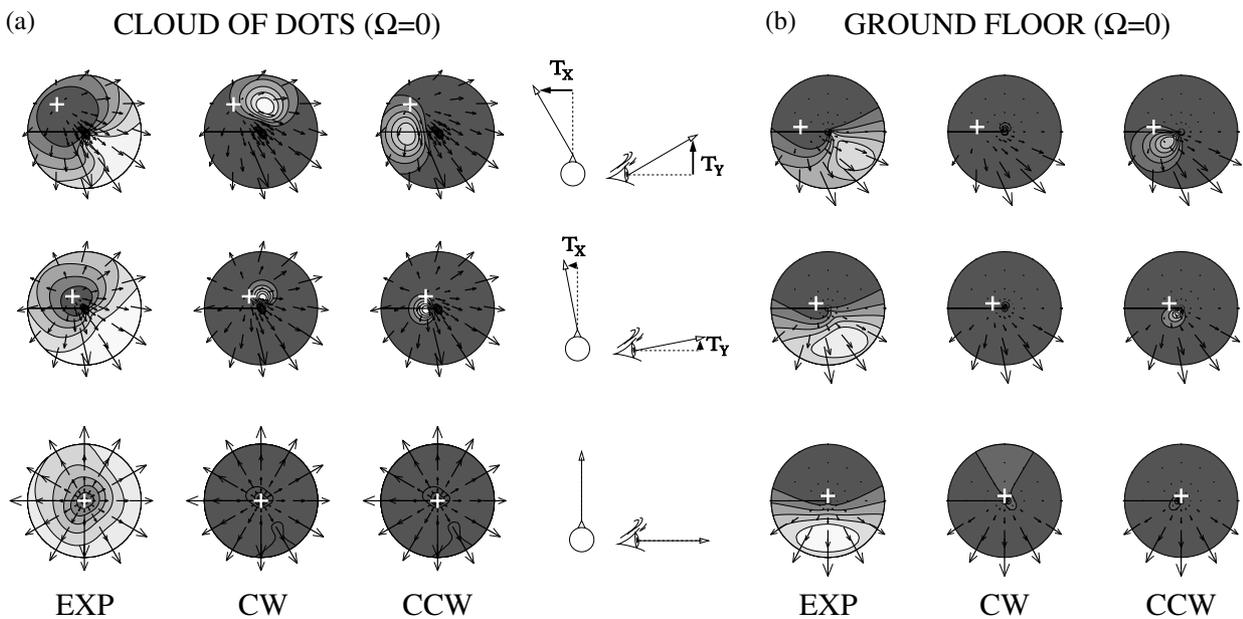
Fig. 9. Distributions of activity of expansion (EXP) and rotation selective cells (CW and CCW) in the case of passive navigation towards a cloud of dots (a) and on a ground floor (b), for three different values of the heading-gaze angle.

If one considers the activity patterns of expansion and rotation cells (see Fig. 9(b)), one observes that the maximum of the activity of expansion cells maintains information about the direction along which to look for heading. One of the rotation selective cell population (the clockwise in the case shown) is always silent, whereas the other one responds properly. On the whole, the three populations still provide sufficient information to locate heading by triangulation with the horizontal line (see Fig. 11).

5.2. Active navigation

Here we want to analyze how the model treats motion field in presence of rotation components ($\Omega \neq 0$) due to pursuit eye movements during locomotion, when the observer still moves along a straight path but fixates a point in the scene. By using the functionalities of templates, based on motion-opponent MT units (see Section 4), the model provides information to locate heading direction independently of the value of rotation, without any explicit decomposition of the motion field in its translation and rotation components.

5.2.1. Cloud of dots

The structure of the cloud of dots and the speed of the observer are the same of the one described for the passive navigation situation. The observer moves toward the cloud fixating the point where the Z-axis intersects the frontoparallel plane located at an intermediate distance within the cloud. Fig. 12(a) shows one instance of the activity distributions of the cortical units. The activity patterns could be qualitatively compared
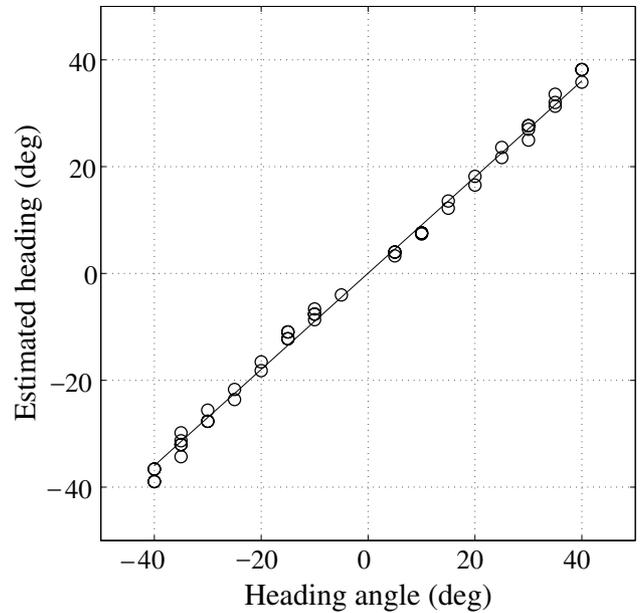
Fig. 11. Estimated heading as a function of the heading angle. The simulations refer to the motion fields experienced by an observer moving without rotation on a ground floor. For each heading angle five different random distributions of points were considered.

with those obtained in absence of rotation (cf. Fig. 9(a)), and the estimated heading is still strongly correlated to the heading angle (see Fig. 13), even though with larger errors.

5.2.2. Ground plane

The observer moves on a ground floor fixating the point where the Z axis intersects the ground plane. The
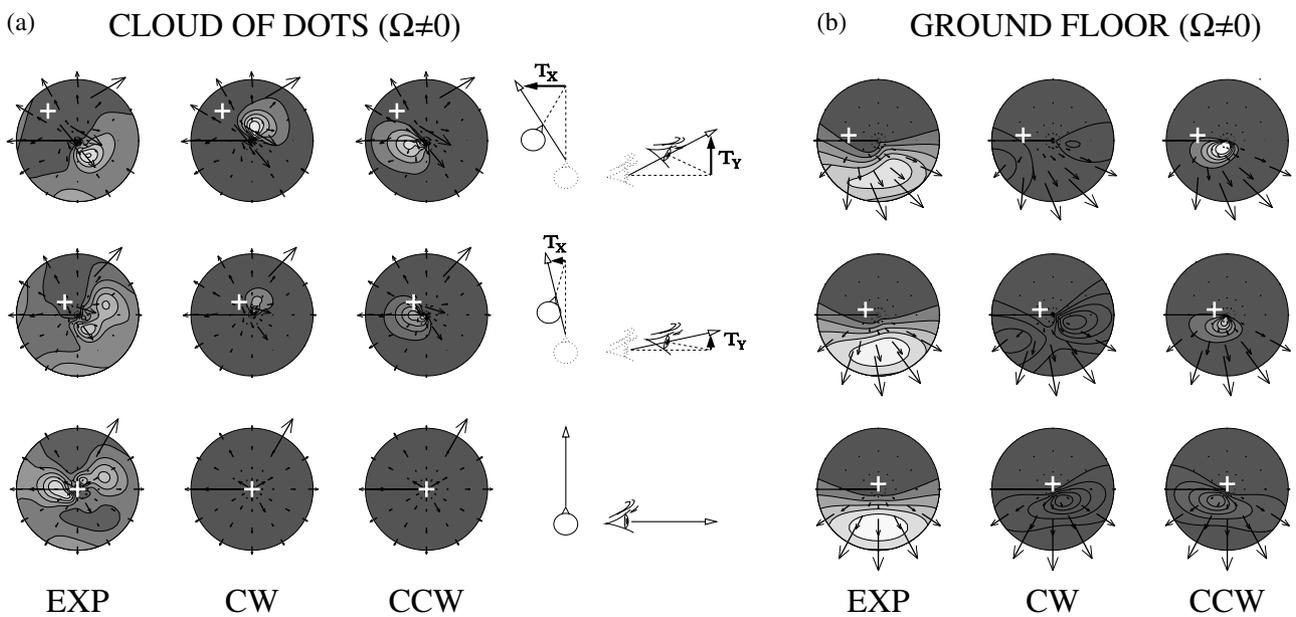
Fig. 12. Distributions of activity of expansion (EXP) and rotation selective cells (CW and CCW) in the case of active navigation towards a cloud of dots (a) and on a ground floor (b), for three different values of the heading-gaze angle.
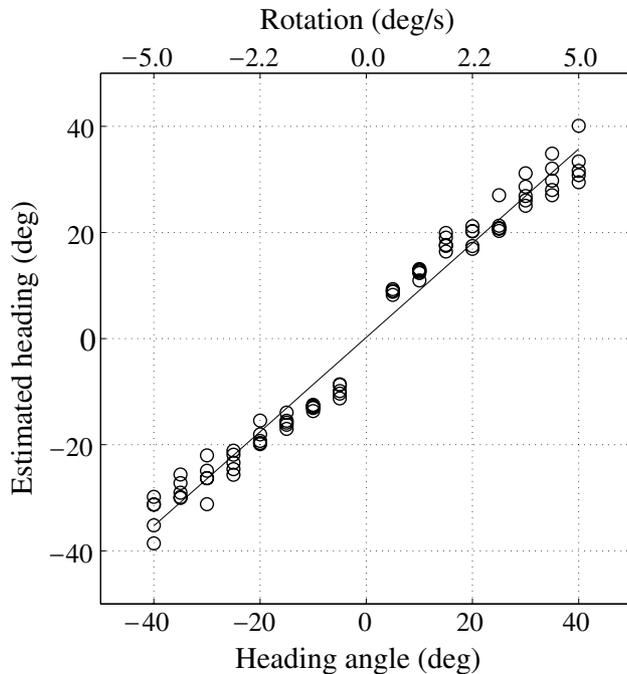
Fig. 13. Estimated heading as a function of the heading angle. The simulations refer to the motion fields experienced by an observer moving towards a cloud of dots while fixating a point in the middle of the cloud. For each heading angle five different random distributions of points were considered.
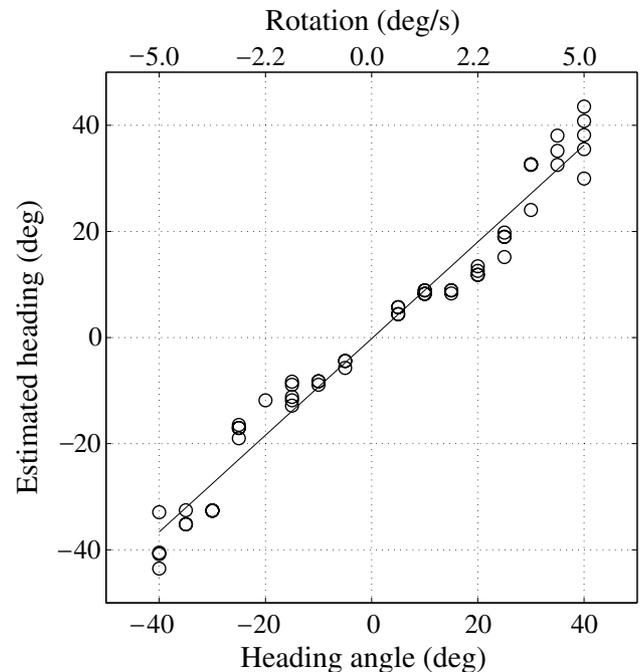
Fig. 14. Estimated heading as a function of the heading angle. The simulations refer to the motion fields experienced by an observer moving on a ground floor while fixating a point on the floor. For each heading angle five different random distributions of points were considered.

starting position and gaze of the observer is the same of the one described for the passive navigation situation. In the present case (see Fig. 12(b)) the activity patterns are very similar to those obtained in absence of rotation (cf. Fig. 9(b)), even though a significant increase in error estimation occurs (see Fig. 14).

5.3. Influences of depth asymmetries

To some extent, we have already shown, with the 3-D cloud, that non-symmetric depth changes do not adversely affect the accuracy of the model. Here, we specifically address this issue with some further tests in conditions of marked depth asymmetries. Since our model is based on centrally symmetric templates—respect to the fovea—these conditions represent a severe benchmark. In particular, we considered step changes in depth as well asymmetric ranges of depth. Results can be summarized as follows.

- The activity patterns observed for expansion and rotation cells (see Figs. 15 and 17) closely resemble those obtained for motion towards the frontoparallel plane and the cloud of dots (see Figs. 7 and 9(a), respectively), still allowing one to properly detect the direction of the correction saccade to align heading and gaze directions.

- At a finer analysis, the asymmetric structure of the two scenes along the horizontal coordinate introduces asymmetries in the magnitudes of image velocities, superimposed to those due to a non-null heading-gaze angle, which lead to slight errors in heading estimation (see Figs. 16 and 18). This situation is also evident when gaze and heading direction coincide, in which case there is a residual minor anisotropy in activity distribution patterns. [4]
- The precision of heading estimation presents an asymmetry related to the stimulus asymmetry, as evidenced by the different slopes of the regression lines for positive and negative heading angles.

5.4. Noisy flows

As previously shown, the scene structure influences the motion field in the speed layout, but not in the di-

---

[4] We can also observe that, if one considers expansion and rotation cells with the same speed tuning, when gaze and heading directions coincide the distribution of activity for expansion and rotation cells are anisotropic in the same way. Such anisotropy depends, indeed, on the magnitude asymmetries in the velocity pattern (depending on the depth of the scene only) to which both expansion and rotation populations are identically sensitive. These considerations could be used to discriminate anisotropic activity patterns due to depth asymmetries from those due to non-null heading-gaze angle.

STEP DISCONTINUITY IN DEPTH
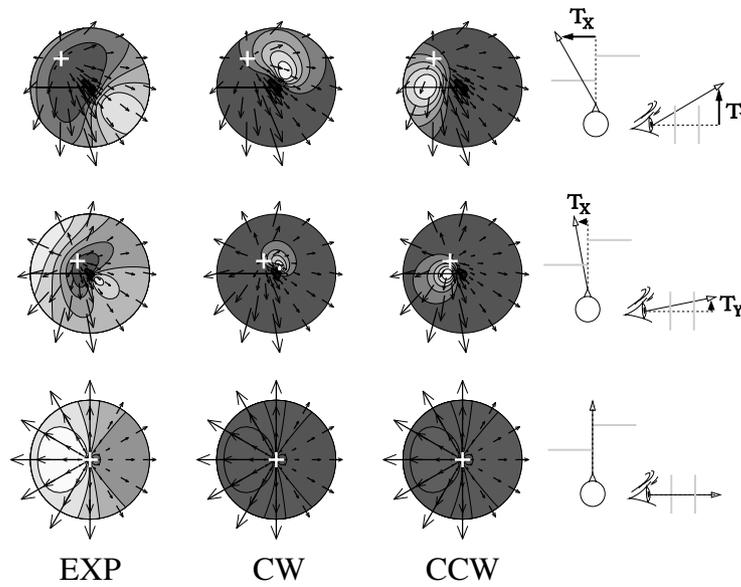
Fig. 15. Influence of scene asymmetries: step discontinuity in depth. Distributions of activity of expansion (EXP) and rotation selective cells (CW and CCW) in the case of passive navigation towards two frontoparallel hemiplanes located a different depths (1:3 ratio). The depth edge is vertical and occurs in correspondence of the gaze direction. Three different values of the heading-gaze angle are considered.
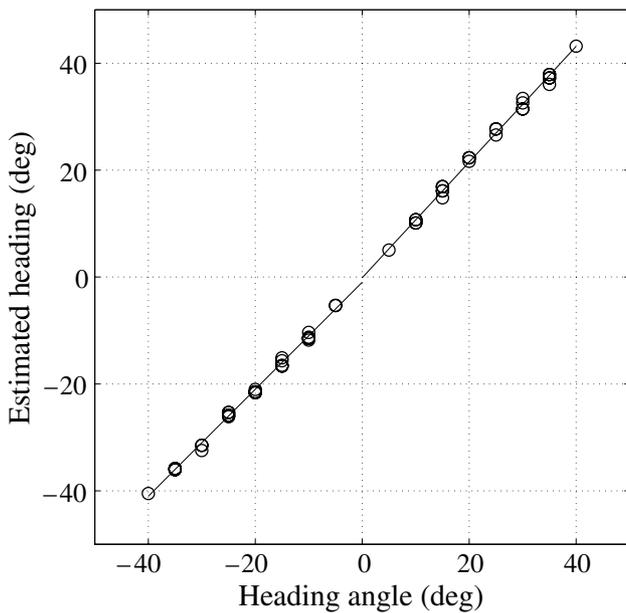
Fig. 16. Estimated heading as a function of the heading angle. The simulations refer to the motion fields experienced by an observer moving towards two frontoparallel hemiplanes located a different depths. For each heading angle five different random distributions of points were considered. Due to the asymmetries of the scene, the correlation between the estimated heading and heading angle is determined separately for positive and negative angles. Only a minor difference can be detected between the slopes of the two regression lines.

rection layout, that substantially maintains its radial morphology. This suggests that to determine his or her heading an observer relies on the radial structure of the flow pattern, rather than on the increasing speed pattern. To test whether in our model the radial pattern of vector directions provides the effective information to estimate heading, we considered two artificial cases of motion field with noisy directions and speeds, respectively. If the model relies on the radial structure of motion vectors, it makes accurate heading judgment for the flow fields with correct directions and altered speeds but not for those with the correct speeds and altered directions.

The two noisy flows has been obtained by adding noise to the speed or to the direction of the motion field experienced in the case of motion towards a frontoparallel plane. The flows obtained adding increasing amount of noise to the vector speeds are very similar to those generated with the cloud of dots, thus producing similar results: the heading estimation is quite accurate even for high values of noise. On the contrary, in the case of noisy directions the heading judgment is no more accurate: even for small percentage of noise the error committed by the model becomes significant. Hence, it is possible to conclude that the main source of information for the model to make prediction about heading direction resides in the directions of the flow vectors (Warren, Blackwell, Kurtz, Hatsopoulos, & Kalish, 1991).

6. Discussion and conclusions

The main property that characterizes our model is the gaze-centered hypothesis for motion field analysis: all

ASYMMETRIC RANGES OF DEPTH

EXP   CW   CCW
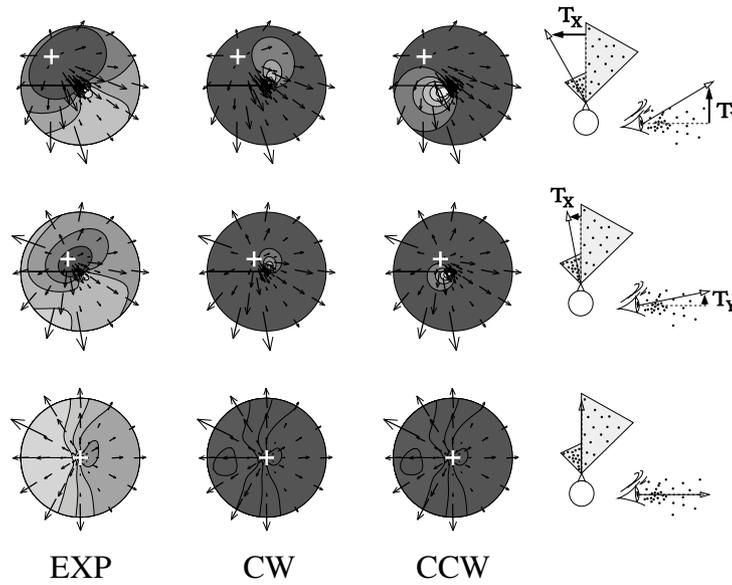
Fig. 17. Influence of scene asymmetries: asymmetric ranges of depth in a cloud of dots. Distributions of activity of expansion (EXP) and rotation selective cells (CW and CCW) in the case of passive navigation towards a cloud of dots for which the range of depth across the visual field varies discontinuously across the vertical plane of gaze (1:3 ratio). Three different values of the heading-gaze angle are considered.
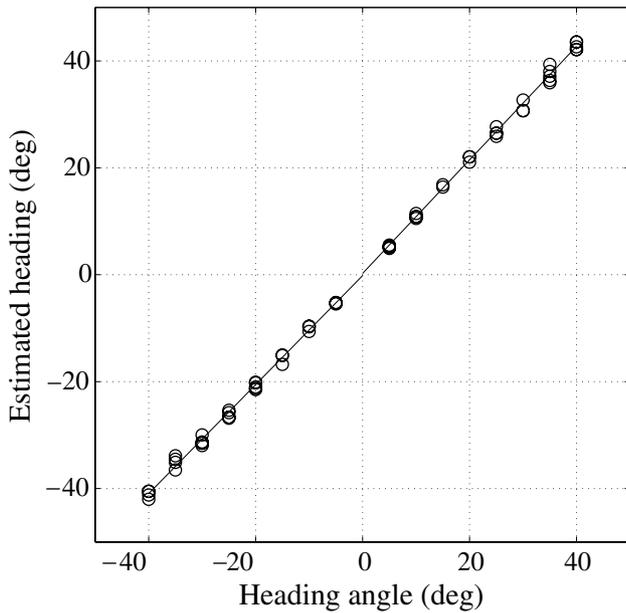
Fig. 18. Estimated heading as a function of the heading angle. The simulations refer to the motion fields experienced by an observer moving towards a cloud of dots for which the range of depth across the visual field varies discontinuously across the vertical plane of gaze. For each heading angle five different random distributions of points were considered. Similar consideration to those presented in Fig. 16 apply for the regression lines.

cells perform on the input motion field a local measure of the presence of radial and rotational patterns symmetric with respect to the current fixation point, and the fovea can be therefore considered a common reference point respect to which the properties of each template pattern have been defined. This approach implies that the result of the analysis performed by each cell has to be interpreted in relation to the position of its receptive field: taking advantage of their topological organization, it is possible to gain global information on the motion field and therefore on heading direction by means of local operators. Although our templates have been designed with reference to the circumstance of motion towards a frontoparallel plane with $\mathbf{\Omega} = 0$, we showed how the model is able to handle real-world situations with $\mathbf{\Omega} \neq 0$, by supposing to make use, as input, of a differential motion information, as those attainable by front-end MT motion-opponent operators to cancel out rotational components. We verified that, provided that the scene contains enough depth variations, the residual velocity field includes adequate information for the determination of the the observer's heading direction.

The choice of the gaze as a reference point assumes a relevant role during self-motion, in fact when an observer is moving in a structured environment a typical behaving task is to direct the attention (i.e. the gaze) towards the translation direction to perceive the presence of obstacles. Observing the mechanisms that guide attention of an observer moving in a structured scene it is possible to evidence two mechanisms, in fact on one side the observer, periodically, fixates details of the scene, like road signs, and on the other side his attention is mostly attracted by the future trajectory to have full control of heading and to perceive possible obstacles in front of him. Estimating heading with respect to gaze

direction helps to solve this task, in fact it provides a direct guide to move the gaze on the motion direction. The importance of defining a strategy to guide the attention towards the heading direction has been also evidenced by Cutting, Springer, Braren, and Johnson (1992). This strategy privileges the indication on which side of the gaze the heading direction lay, rather than its precise localization. Similarly to Cutting et al., our approach suggests a closed-loop control strategy to align heading and gaze direction based on an active scanning pattern and a set of rules by which the observer can infer heading. Cutting's rules for guiding fixation towards heading have a nature similar to that we use for determining where to look for heading on the basis of where expansion cells give their maximum response respect to the fovea. Unlike Cutting's, our model exploits a distributed neural representation of the motion field, rather than directly object's retinal images, and relies on information over the whole field and not limited to the points around the fixation.

6.1. Advantages of polar and log-polar mapping for visual navigation

The log-polar image geometry was first introduced (Braccini, Gambardella, Sandini, & Tagliasco, 1982; Chaikin & Weiman, 1980; Fischer, 1973; Schwartz, 1980) to model the space-variant topology of the human retina receptors—with increasing density towards the fovea and decreasing density from the fovea towards the periphery—in relation to the data compression it achieves.

When compared to cartesian representations, the log-polar ones allow higher sampling rates on artificial vision systems without reducing the size of the field of view and the resolution in the foveal region. In the last years, however, it has been noticed that the log-polar geometry also provides important algorithmic benefits in recognition tasks (rotation and scaling invariance) (Wechsler, 1992; Yeshurun & Schwartz, 1989) as well as navigational tasks (ego-motion and time to collision estimation, motion detection) (Jain, Bartlett, & O'Brien, 1987; Sandini & Tistarelli, 1992).

In our model, the adoption of a log-polar mapping of the motion field is a natural consequence of the gaze-centered hypothesis, and leads to a substantial simplification of the templates' structure which can be described through cardinal directions in the polar cortical plane (cf. Burr, Badcock, & Ross, 2001; Grossberg, Mingolla, & Pack, 1999). By example, fovea-centered expansions are mapped on the cortical plane as vector fields with null angular components, therefore all the templates of the expansion selective cells can be described by the same relation, independently of the position of their receptive field centers. Furthermore, the space variant representation, given by the logarithmic scale, supports

an efficient control of the gaze towards heading direction. In fact, although the range of correctly perceived headings is not directly influenced by the width of the visual field, the space variant sampling allows, for a constant number of samples, a representation of a wide visual field and thence a correct estimation of heading with a high degree of eccentricity. The accuracy in heading estimation is connected to the resolution of the retinal plane, and therefore the judgment is coarser when heading direction falls in the periphery of the visual field. Interpreting the model output as a guide in the attentive process to align the gaze with the heading direction, the space variant resolution implies that head and eye movements are coarser when the angle between heading and gaze is larger and finer when the two directions are nearly coincident. In this way, the attention is brought on the future path as quickly and accurately as possible.

6.2. Influence of eye and head rotations

The pattern of motion field can be highly complex when to the translational component it is added a rotational component due to eye or head movement of the observer. The presence of rotation discriminates between a passive observer and an active fixating observer. To estimate heading accurately, observers must be able to disambiguate the motion field, extracting the translational component. The ability to perform this task during tracking eye movements has been recently investigated (Royden, Banks, & Crowell, 1992; Warren & Hannon, 1990) and it has been discussed if extra-retinal signals like vestibular signals and proprioceptive feedback from the extra-ocular muscles, are used. In this research conflicting results have been obtained, in fact while there is general agreement that the visual system is not able to solve the problem in ambiguous conditions, like observer motion towards a frontoparallel plane, there is no evidence that observers are able to estimate heading relying only on visual information. Comparing the result of tests with real and simulated eye movements Warren and Hannon concluded that humans can make an accurate estimation only with visual information (Warren & Hannon, 1990), while Royden and colleagues concluded that it is possible only in presence of slow eye rotations (Royden et al., 1992).

Numerous models of heading estimation have been proposed and they can be divided in two categories (Banks, Ehrlich, Backus, & Crowell, 1996): retinal image models (Heeger & Jepson, 1990; Longuet-Higgins & Prazdny, 1980; Perrone & Stone, 1994) and extra-retinal models (Beintema & van den Berg, 1998; Royden, Crowell, & Banks, 1994). Retinal models rely on the properties of motion parallax, which evidences how flows due to translation depend on the three dimensional structure

of the scene, while flows due to rotation are insensitive to the depth of the scene objects.

The model we proposed has been basically designed to deal with situations of pure ego-motion without taking care of possible eye movements, that add a rotational component to the motion field. Yet, the availability of motion-opponent MT cells allows us to exploit information of motion parallax and thus to handle situations with $\Omega \neq 0$. In Section 5.2 we reported the results we obtained for combination of translations and rotations experienced by an observer that moves fixating a static point in the scene, though, the results could be generalized for situations with arbitrary rotations.

6.3. Comparison with other models

Following the classification proposed in Lappe, Bremmer, and van den Berg (1999) computational models of heading perception can be subdivided in two main categories: *population heading map models* and *templates for specific flow patterns*. The first group comprises those models that are composed by two layers (one for MT-like units and the other one for MST-like units) which have the retinal flow as input and the estimated heading as output; the computation of heading depends on the interactions between the first and the second layer. By example, in the model proposed by Lappe et al. (1996) the required connection strength are precalculated from a mathematical formalization of the underlying heading detection algorithm (Lappe & Rauschecker, 1993); the network finds among the possible fields, the one that minimizes the mean squared difference with the measured flow field. The possible flow fields are chosen in a subspace of all possible combinations of translation and rotation. Each population of neurons computes the current likelihood of a specific heading giving raise to a heading map, whose peak represents the most likely heading direction.

In the second group of models, the *template models*, the response of an individual neuron depends on the match between the input flow field and the template associated to that neuron. Theoretically, one needs a high-level processing neuron for each possible heading and therefore, to achieve a sufficiently accurate sampling of infinite combinations of observer translation and rotation, a very large number of templates is required. The main problem associated to template models is therefore how to reduce the number of templates. Perrone and Stone (1994) proposed to consider only eye rotations that occur in fixation tasks (the "gaze stabilization eye movement" hypothesis). Under this hypothesis the number of parameters necessary to describe motion is reduced. The rotation is constrained to the translation, and thus possible headings are sampled in a 3-D space (T_X, T_Y, T_Z) rather than in a 5-D space $(T_X, T_Y, T_Z, \Omega_X, \Omega_Y)$. Another solution adopted to mini-

mize the number of templates is to combine visual information with extra-retinal signals to subtract rotational components to the motion field. In the model proposed by Beintema and van den Berg (1998) two sets of templates are used, one selective to pure translation and the other one selective to combinations of translation and head/eye rotation. An oculomotor signal is used to combine the two classes of templates to remove the term of rotation. In any case, as in the Perrone and Stone's model (Perrone & Stone, 1994), a large set of templates is required to sample all possible translation direction.

Grossberg et al. (1999) proposed a model which relies on the transformation of the retinal flow on the cortical plane described in log-polar coordinates. Heading neurons are selective to the elementary components, in which the flow can be decomposed with respect to the fovea: expansion and rotation. By measuring and comparing the response of the templates selective to the two components an estimate of the observer motion direction can be made. This model deals only with the situation in which the observer is translating in absence of eye rotation. Hence, to face the problem of corruption of the motion field by eye movements it is required to consider the presence of extra-retinal signals. The reduction of the number of templates to only two templates (radial and circular templates) is made possible by the simplifying hypothesis of observer motion over a ground plane which generates particular flows structures in the lower visual hemifield.

The model presented in this paper can be considered belonging to the group of template models, where the templates are represented by radial and rotational flow patterns centered on the fovea. There is not a specific template for each possible heading direction, but all the neuron units, belonging to the same population, are characterized by portions of the same template. There is no need to sample the possible heading direction space, and the space resolution of the estimate is given by the coarse coding representation by MST cells. Since each unit is not selective to a particular heading, the output distribution do not generate a heading map, but they provide a measure of how much the motion field is radially or rotationally symmetric. Thus to compute heading, it is necessary to look at the response of the entire population of model neurons to the given stimulus, by combining the gaze centered template approach with population coding it is possible to drastically reduce the number of the necessary templates.

6.4. Comparison with human sensitivity to the structure of the motion field

Considering the *structural* character of our model it could be interesting to compare its performance with the data available in the literature on the human sensitivity

to the structure of the motion field in relation with the task of heading estimation. Human performance in perceiving the direction of motion have been analyzed considering the locus of information in the visual field and the type of motion field observed (radial vs lamellar or, in other terms, heading eccentricities respect to the center of the stimulus). On the basis of their data, Warren and Kurtz (1992) postulated a "functional sensitivity hypothesis", in which self-motion is perceived on the basis of optical information rather than retinal locus of stimulation. Specifically, this hypothesis states that central vision is more sensitive to radial flow than the periphery which is more sensitive to lamellar flow. Arguing that Warren and Kurtz confounded the retinal eccentricity of stimulation with the visual angle between the visible flow and the heading, Crowell and Banks (1993) explored this issue using flow fields containing radial flow and lamellar flow, both presented to a wide variety of retina positions. On the basis of their data, they proposed a "retinal invariance hypothesis", where the perception of heading is largely independent of retinal position (but see Atchley & Andersen (1999)). Due to their centric-minded character, the model proposed in this paper presents, by construction, a "retinally invariant functional sensitivity", in the sense that the criterion adopted to design the templates is retinally invariant, being templates selective to central (fovea centered) stimulation independently of their position in the visual field. By the complementary use of expansion- and rotation-selective units the system shows, as a whole, various behaviours in relation to how the optical stimulus is parcelled. Indeed, we observed an eccentricity dependence of the heading error in accordance with Warren and Kurtz (1992) for all the tested situations (see Section 5). Moreover, by masking central vision when it contains the center of the radial flow, the model's performance would worsen (cf. Warren & Kurtz (1992)) since the maximum distribution of activity used for heading estimation is mainly confined within the foveal and parafoveal regions.

To test our model in the experimental conditions of Crowell and Banks (1993), we measured the differential sensitivity in the responses of our MST units to small field (10°) patterns of moving clouds of dots displayed at different retinal eccentricities. The patterns present different heading eccentricities respect to the center of the stimulus. For each condition, we measured the minimum angular shift (threshold), respect to a given heading eccentricity, necessary to obtain a noticeable difference in the cells' response with respect to a reference situation (both for expansion- and rotation-selective populations). Specifically, we evaluated the average of the absolute value of the relative differences. Simulations are performed for retinal eccentricities of 0°, 5°, 10°, and 20° and for increasing heading eccentricities ranging from 0° to ±40°. The "V" shape dependence of the threshold angle on heading eccentricities (cf. Fig. 4 in Crowell & Banks, 1993) is reproduced in our simulations, with small variations with retinal eccentricities. Minor differences in the sensitivity are observed for expansion- and rotation-selective cells. The threshold values, determined by our simulations, are slightly higher than those of Crowell and Banks, but this is not a meaningful difference since it depends on the relative difference criterion adopted. In conclusion, taking into account that we are testing the model in a very peculiar condition (the stimulus is present on a very limited portion of the visual field), we can observe that no contradictory results show up respect to Crowell and Banks' conclusions: the retinal locus of stimulation is less important than the structure of the motion field, being the sensitivity higher for radial flows both for central and peripheral vision.

Another issue concerns the influence of image velocity magnitudes on heading perception (Dyre & Andersen, 1997). Specific simulations with optical stimuli presenting velocity asymmetries due to not symmetric depth changes in the scene, showed slight errors in heading estimation (see Section 5.3). In general, such errors are comparable with those reported by Dyre and Andersen (1997) on human performance.

6.5. Implications to MST cell modeling

Though it is not the aim of the present paper to propose a detailed model of MST cells, we have many times remarked how the functionalities of our template operators can be related to the response properties of those cells. On this basis, preliminary results obtained by our model do not contradict the experimental data currently available in the literature, regarding: (1) the inability of performing a mathematical decomposition of the motion field in its elementary components (Orban et al., 1992), (2) the center-of-motion tuning (Duffy & Wurtz, 1995) (Duffy & Wurtz, 1997a), and (3) the possibility of obtaining a continuum of selectivity from expansion to rotation through all spiral flows (Geesaman & Andersen, 1996; Graziano et al., 1994). Yet, it is very difficult to operate a systematic comparison considering the heterogeneity of experimental data and the fact that in many experiments it has not been paid attention to find a relationship between the position of the MST receptive field and the position of the stimulus with respect to the fovea. However, these considerations, and the fact that our model reproduces some experimental evidences, suggest that the numerous and sometimes contradictory physiological cell properties relating MST cell response to the placement of motion field stimuli in the visual field, including position invariance, center-of-motion tuning, direction reversals, could be re-examined from this new gaze-centered perspective.

Appendix A

Here, we derive the detailed coefficient expressions in the output of the radial and rotational operators reported in Eqs. (14) and (15) for passive navigation

$$Q_1(\xi) = f \sin\frac{\Delta\theta}{2} \int_{\xi-\frac{\Delta\xi}{2}}^{\xi+(\Delta\xi/2)} e^{-\xi'} d\xi'$$

$$= f \sin\frac{\Delta\theta}{2} \left(-e^{-\xi-(\Delta\xi/2)} + e^{-\xi+(\Delta\xi/2)} \right)$$

$$Q_2(\xi) = \frac{\Delta\theta}{2} \int_{\xi-(\Delta\xi/2)}^{\xi+(\Delta\xi/2)} 1 - ae^{-\xi'} d\xi'$$

$$= \frac{\Delta\theta}{2} \left[\Delta\xi - a\left(e^{-\xi-(\Delta\xi/2)} - e^{-\xi+(\Delta\xi/2)}\right) \right]$$

$$P_1(\xi) = f \sin\frac{\Delta\theta}{2} \int_{\xi-(\Delta\xi/2)}^{\xi+(\Delta\xi/2)} (e^{\xi'} - a)^{-1} d\xi'$$

$$= f \sin\frac{\Delta\theta}{2} \frac{1}{a} \left[-\frac{\Delta\xi}{2} + \log\left(\frac{e^{\xi+(\Delta\xi/2)} - a}{e^{\xi-(\Delta\xi/2)} - a} \right) \right]$$

References

Albright, T. (1989). Centrifugal directionality bias in the middle temporal visual area (MT) of the macaque. *Visual Neuroscience*, 177–188.

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

Ball, K., & Sekuler, R. (1980). Human vision favors centrifugal motion. *Perception, 9*, 317–325.

Banks, M., Ehrlich, S., Backus, B., & Crowell, J. (1996). Estimating heading during real and simulated eye movements. *Vision Research, 36*, 431–443.

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

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

Braccini, C., Gambardella, G., Sandini, G., & Tagliasco, V. (1982). A model of the early stages of the human visual system: functional and topological transformations performed in the peripheral visual field. *Biological Cybernetics, 44*, 47–58.

Burr, D., Badcock, D., & Ross, J. (2001). Cardinal axes for radial and circular motion, revealed by summation and by masking. *Vision Research, 41*, 473–481.

Chaikin, G., & Weiman, F. (1980). Conformal computational geometry for machine vision. In *Proceedings of the 5th international conference on pattern recognition*.

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

Cutting, J., Springer, K., Braren, P., & Johnson, S. (1992). Wayfinding on foot from information in retinal, not optical, flow. *Journal of Experimental Psychology: General, 121*, 41–72.

Dahmen, H.-J., Franz, M., & Krapp, H. (2001). Extracting ego-motion from optic flow: limits of accuracy and neural matched filters. In J. Zanker, & J. Zeil (Eds.), *Motion vision* (pp. 143–168). Berlin: Springer.

Duffy, C., & Wurtz, R. (1991). Sensitivity of MST neurons to optic flow stimuli I. A continuum of response selectivity to large-field stimuli. *Journal of Neurophysiology, 65*(6), 1329–1345.

Duffy, C., & Wurtz, R. (1995). Response of monkey MST neurons to optic flow stimuli with shifted centers of motion. *Journal of Neuroscience, 15*, 5192–5208.

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

Duffy, J., & Wurtz, R. (1997b). Medial superior temporal area neurons respond to speed patterns in optic flow. *Journal of Neuroscience, 17*, 2839–2851.

Dyre, B., & Andersen, G. (1997). Image velocity magnitudes and perception of heading. *Journal of Experimental Psychology: Human Perception and Performance, 23*, 546–565.

Fischer, B. (1973). Overlap of receptive field centers and representation of the visual field in the cat's optic tract. *Visres, 13*, 2113–2120.

Geesaman, B., & Andersen, R. (1996). The analysis of complex motion patterns by form/cue invariant MSTd neurons. *Journal of Neuroscience, 16*, 4716–4732.

Gibson, J. (1950). *The perception of the visual world*. Boston: Houghton Mifflin.

Gibson, J. (1966). *The senses considered as perceptual system*. Boston: Houghton Mifflin.

Graziano, M., Andersen, R., & Snowden, R. (1994). Tuning of MST neurons to spiral motions. *Journal of Neuroscience, 14*, 54–67.

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

Hanada, M., & Ejima, Y. (2000). Method for recovery of heading from motion. *Journal of Optical Society of America, 17*, 966–973.

Heeger, D., & Jepson, A. (1990). Visual perception of three-dimensional motion. *Neural Computation, 2*, 129–137.

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

Jain, R., Bartlett, S., & O'Brien, N. (1987). Motion stereo using ego-motion complex logarithmic mapping. *IEEE Transactions of Pattern Analysis and Machine Intellegence, 3*, 356–369.

Johnston, A., & Wright, M. (1983). Visual motion and cortical velocity. *Nature, 304*, 436–438.

Johnston, A., & Wright, M. (1986). Matching velocity in central and peripheral vision. *Vision Research, 26*, 1099–1109.

Koenderink, J. (1986). Optic flow. *Vision Research, 26*(1), 161–179.

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

Lagae, L., Raiguel, S., & Orban, A. (1993). Speed and direction selectivity of macaque middle temporal neurons. *Journal of Neurophysiology, 69*(1), 19–39.

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

Lappe, M., Bremmer, F., & van den Berg, A. (1999). Perception of self-motion from visual flow. *Trends in Cognitive Sciences, 3*, 329–336.

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

LonguetHiggins, H., & Prazdny, K. (1980). The interpretation of a moving retinal image. *Proceedings of Royal Society of London B, 208*, 385–397.

Marcar, V., Xiao, D., Raiguel, S., Maes, H., & Orban, G. (1995). Processing of kinetically defined boundaries in the cortical motion area MT of the macaque monkey. *Journal of Neurophysiology, 74*(3), 1258–1270.

Marr, D., & Ullman, S. (1981). Directional selectivity and its use in early visual processing. *Proceedings of Royal Society of London B, 211*, 151–180.

Maunsell, J., & Van Essen, C. (1983). Functional properties of neurons in middle temporal visual area of the macaque monkey I. Selectivity for stimulus direction, speed, and orientation. *Journal of Neurophysiology, 49*(5), 1127–1147.

Morrone, M., Tosetti, M., Montanaro, D., Fiorentini, A., Cioni, G., & Burr, D. (2000). A cortical area that responds specifically to optic flow, revealed by fmri. *Natural Neuroscience, 3*, 1322–1328.

Orban, G. (1992). The analysis of motion signal and the nature of processing in the primate system. In: *Artificial and biological vision system*. ESPRIT Basic Research Series (pp. 24–56).

Orban, G., Lagae, L., Raiguel, S., Xiao, D., & Maes, H. (1995). The speed tuning of medial superior temporal (MST) cell responses to optic-flow components. *Perception, 24*, 269–285.

Orban, G., Lagae, L., Verri, A., Raiguel, S., Xiao, D., Maes, H., & Torre, V. (1992). First-order analysis of optical flow in monkey brain. *Proceedings of National Academy of Science, 89*, 2595–2599.

Perrone, J. (1986). Anisotropic responses to motion toward and away from the eye. *Perception of Psychophysics, 39*, 1–8.

Perrone, J. (1992). Model for the computation of self-motion in biological systems. *Journal of Optical Society of America, 9*(2), 177–194.

Perrone, J., & Stone, L. (1994). A model of self-motion estimation within primate extrastriate visual cortex. *Vision Research, 34*(21), 2917–2938.

Poggio, T., Verri, A., Torre, V., 1991. Green theorems and qualitative properties of the optical flow. Technical Report A.I. Memo 1289, MIT.

Raiguel, S., Hulle, M., Xiao, D., Marcar, V., Lagae, L., & Orban, G. (1997). Size and shape of receptive fields in the medial superior temporal area (MST) of the macaque. *NeuroReport, 8*(12), 2803–2807.

Raiguel, S., Hulle, M., Xiao, D., Marcar, V., & Orban, G. (1995). Shape and spatial distribution of receptive fields and antagonist motion surrounds in the middle temporal area (v5) of the macaque. *European Journal of Neuroscience, 7*, 2064–2082.

Rauschecker, J., Grünau, M., & Poulin, C. (1987). Centrifugal organization of direction preferences in the cat's lateral suprasylvian visual cortex and its relation to flow field processing. *Journal of Neuroscience, 7*, 943–958.

Raymond, J. (1994). Directional anisotropy of motion sensitivity across the visual field. *Vision Research, 34*, 1029–1037.

Richards, W. (1975). Visual space perception. In E. Carterette, & M. Friedman (Eds.), *Handbook of perception: Vol. 5 seeing* (pp. 351–386). New York: Academic Press.

Rieger, J., & Lawton, D. (1985). Processing differential image motion. *Journal of Optical Society of America, 2*(2), 354–360.

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

Roy, J.-P., Komatsu, H., & Wurtz, R. (1992). Disparity sensitivity of neurons in monkey extrastriate area MST. *Journal of Neuroscience, 12*, 2478–2492.

Roy, J.-P., & Wurtz, R. (1990). The role of disparity-sensitive cortical neurons in signalling the direction of self-motion. *Nature, 348*, 160–162.

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

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

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

Sandini, G., & Tistarelli, M. (1992). Vision and space-variant sensing. In H. Wechsler (Ed.), *Neural networks for perception I* (pp. 398–425). New York: Academic Press.

Schwartz, E. (1977). Spatial mapping in primate sensory projection: Analytical structure and relevance to perception. *Biological Cybernetics, 25*, 181–194.

Schwartz, E. (1980). Computational anatomy and functional architecture of striate cortex: A spatial mapping approach to perceptual coding. *Vision Research, 20*, 645–669.

Tanaka, K., & Saito, H. (1989). Analysis of motion of the visual field by direction, expansion/contraction, and rotation cells clustered in the dorsal part of the medial superior temporal area of the macaque monkey. *Journal of Neurophysiology, 62*(3), 626–641.

Telford, L., & Howard, I. (1996). Role of optical flow field asymmetry in the perception of heading during linear motion. *Perception of Psychophysics, 58*, 283–288.

Tistarelli, M., & Sandini, G. (1993). On the advantages of polar and log-polar mapping for direct estimation of time-to-impact from optical flow. *IEEE Transactions of Pattern Analysis of Machine Intelligence, 15*, 401–410.

Verri, A., Straforini, M., & Torre, V. (1992). Computational aspects of motion perception in natural and artificial systems. *Philosophical Transactions of Royal Society of London B, 337*, 429–443.

Wang, R. (1996). A network model for the optic flow computation of the MST neurons. *Neural Networks, 9*(3), 411–426.

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

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

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

Wechsler, H. (1992). Multiscale and distributed visual representations and mappings for invariant low-level perception. In H. Wechsler (Ed.), *Neural Networks for Perception I* (pp. 462–476). Academic Press.

Xiao, D., Raiguel, S., Marcar, V., Koenderink, J., & Orban, G. (1995). Spatial heterogeneity of inhibitory surrounds in the middle temporal visual area. *Proceedings of National Academy of Science, 92*, 11303–11306.

Xiao, D., Raiguel, S., Marcar, V., & Orban, G. (1997). The spatial distribution of the antagonistic surround of MT/v5 neurons. *Cerebral Cortex, 7*(7), 662–677.

Yeshurun, Y., & Schwartz, E. (1989). Shape description with space-variant sensor: Algorithms for scan-path, fusion and convergence of multiple scans. *IEEE Transactions on Pattern Analysis on Machine Intelligence, 11*, 1217–1222.

Zetzsche, C., Krieger, G., & Wegmann, D. (1999). The atoms of vision: Cartesian or polar? *Journal of Optical Society of America: A, 16*, 1554–1565.

Zhang, K., Sereno, M., & Sereno, M. (1993). Emergence of position-independent detectors of rotation and dilation with hebbian learning: an analysis. *Neural Computation, 5*, 597–612.