

Perception of self-motion from visual flow

Markus Lappe, Frank Bremmer and A.V. van den Berg

Accurate and efficient control of self-motion is an important requirement for our daily behavior. Visual feedback about self-motion is provided by optic flow. Optic flow can be used to estimate the direction of self-motion ('heading') rapidly and efficiently. Analysis of oculomotor behavior reveals that eye movements usually accompany self-motion. Such eye movements introduce additional retinal image motion so that the flow pattern on the retina usually consists of a combination of self-movement and eye movement components. The question of whether this 'retinal flow' alone allows the brain to estimate heading, or whether an additional 'extraretinal' eye movement signal is needed, has been controversial. This article reviews recent studies that suggest that heading can be estimated visually but extraretinal signals are used to disambiguate problematic situations. The dorsal stream of primate cortex contains motion processing areas that are selective for optic flow and self-motion. Models that link the properties of neurons in these areas to the properties of heading perception suggest possible underlying mechanisms of the visual perception of self-motion.

Vision provides a major source of information for the control of self-movement. It is very difficult to walk to a goal with eyes closed. The visual motion we experience as a result of walking, running, or driving – the optic flow¹ – is a powerful signal to control the parameters of our own movement. Its value becomes apparent when optic flow is not matched to the true self-motion. For instance, when the walls of a surrounding room are set in motion, toddlers that have just learned to walk fall over². Adults modify their walking speed depending on optic flow³. In stationary subjects optic flow induces the illusory feeling of self-movement² and causes motion sickness after prolonged exposure. The speed of the optic flow can be used for collision detection² (but see Ref. 4) and, with restrictions, for the estimation of travelled distance⁵. This review focuses on the contribution of optic flow to the control of the direction of self-motion, or heading. The importance of optic flow for the control of heading and visual navigation was first recognized by Gibson¹. He noted that the visual motion in the 'optic array surrounding a moving observer' radially expands out of a singular point along the direction of heading. Hence he termed this point the 'focus of expansion' (FOE) and suggested that heading is estimated by localizing this point.

The problem is more difficult than Gibson's analysis suggests, though. The analysis of motion in the optic array is complicated by the fact that the sensors of the visual system, the retinae of the eyes, can move with respect to the body. Many eye movements normally occur during self-motion (Box 1). These eye movements (and head movements, for that matter) are superimposed on body movements. Eye rotation induces coherent visual motion on the retina. This

motion is superimposed on the optic flow. The result is a retinal motion pattern that is composed of translational (body movement) and rotational (eye movement) components. One therefore has to distinguish retinal flow from optic flow, and note that the visual system has to use retinal, not optic flow as the basis of self-motion estimation⁶⁻⁸.

Retinal flow is often very different from the simple expansion pattern of optic flow. Examples of typical cases that have been used in studies of heading detection are illustrated in Fig. 1. They include different combinations of translation and eye rotation as well as different visual environments, because the structure of retinal flow also depends on the distances of the visible objects from the observer. A main problem of estimating heading from retinal flow is to separate translational and rotational components (Box 2).

Heading detection during eye rotation

Experimental investigations of visual self-motion perception have benefited tremendously from the availability of specialized 3-D graphics workstations that can simulate movement through virtual environments in real time. The most basic experiments use linear movement in simple random-dot environments devoid of recognizable image features (Fig. 1B). The resulting visual motion is presented on a large screen in front of the subject that covers a substantial part of the visual field. Heading judgements are determined either as just-noticeable-differences (JNDs) with respect to a reference target or by a pointing response.

For simple linear movement without eye rotation, the FOE can be used as an indicator of heading. JNDs for the estimation of the FOE are 1–2 degrees of visual angle⁹. This

M. Lappe and F. Bremmer are at the Department of Zoology and Neurobiology, Ruhr-University Bochum, 44780 Bochum, Germany.

tel: +49 234 700 4369
fax: +49 234 709 4278
e-mail: lappe@neurobiologie.ruhr-uni-bochum.de, bremmer@neurobiologie.ruhr-uni-bochum.de

A.V. van den Berg is at the Medical Faculty, Erasmus University Rotterdam, 3000 DR Rotterdam, The Netherlands.

tel: +31 10 408 7585
fax: +31 10 408 9457
e-mail: vandenber@fys.fgg.eur.nl

Box 1. Gaze during self-motion

Most natural behaviors are accompanied by eye movements. Eye movements during self-motion serve two functions: they direct gaze to objects of interest and they help to maintain stable vision.

During walking gaze is directed towards obstacles along the future path or towards future landing positions of the feet (Refs a,b). Gaze is typically one or two steps ahead of the current body position. Frequent gaze shifts also occur during car driving, for instance when approaching and passing an intersection (Ref. c). A specific gaze strategy has been observed when car drivers negotiate a curve. In this case, gaze is continuously directed towards a specific point at the inner side of the curve, the 'tangent point' (Ref. d), which is the point where the retinal image of the edge of the road reverses direction. For straight road segments gaze is directed towards the road ahead, several meters in front of the car. Depending on the demands, gaze might also divert from the road to look at road signs or other peripheral objects of interest. More often than not, therefore, the direction of heading falls in the retinal periphery.

Each saccadic gaze shift results in abrupt changes of the retinal flow. Saccades segment the visual input into discontinuous samples, each of a few hundred milliseconds duration. Between two saccades slow eye movements occur for the purpose of stabilization of the retinal image. Like any visual motion, self-motion-induced retinal flow generates problems for visual recognition. Because of this, several oculomotor reflexes attempt to keep the visual image stable on the retina (Ref. e). In the case of optic flow only part of the image can be stabilized, because the velocities in the flow field are so different. In walking monkeys, gaze is approximately stable in space; that is, rotational eye and head movements compensate for the translation of the head (Ref. f). Stabilizing eye movements during translation are induced by the vestibulo-ocular reflex (Ref. g) but also directly by optic flow. When humans or monkeys

view optic flow fields their eye movements involuntarily track the current motion in the direction of gaze (Refs h,i). This tracking is updated with every saccade that changes the direction of motion on the fovea. In addition, radial optic flow also induces reflex vergence responses (Ref. j). The normal oculomotor pattern thus consists of phases of tracking eye movements that last several hundred milliseconds separated by saccades that direct gaze to a new target. Therefore, one must assume that the eye is often in motion and that eye-movement-induced visual motion is superimposed on the optic flow.

References

- a Hollands, M.A. et al. (1995) Human eye movements during visually guided stepping *J. Motor Behav.* 27, 155–163
- b Patla, A.E. and Vickers, J.N. (1997) Where and when do we look as we approach and step over an obstacle in the travel path? *NeuroReport* 8, 3661–3665
- c Land, M.F. (1992) Predictable eye-head coordination during driving *Nature* 359, 318–320
- d Land, M.F. and Lee, D.N. (1994) Where we look when we steer *Nature* 369, 742–744
- e Miles, F.A. (1998) The neural processing of 3-D visual information: evidence from eye movements *Eur. J. Neurosci.* 10, 811–822
- f Solomon, D. and Cohen, B. (1992) Stabilization of gaze during circular locomotion in light: I. Compensatory head and eye nystagmus in the running monkey *J. Neurophysiol.* 67, 1146–1157
- g Paige, G.D. and Tomko, D.L. (1991) Eye movement responses to linear head motion in the squirrel monkey: I. Basic characteristics *J. Neurophysiol.* 65, 1170–1183
- h Lappe, M., Pekel, M. and Hoffmann, K-P. (1998) Optokinetic eye movements elicited by radial optic flow in the macaque monkey *J. Neurophysiol.* 79, 1461–1480
- i Niemann, T. et al. (1999) Ocular responses to radial optic flow and single accelerated targets in humans *Vis. Res.* 39, 1359–1371
- j Busetini, C., Masson, G.S. and Miles, F.A. (1997) Radial optic flow induces vergence eye movements with ultra-short latencies *Nature* 390, 512–515

is within the range necessary for successful control of self-motion and avoidance of obstacles⁸. Accuracy is largely independent of the 3-D layout and density of the random dots⁹. JNDs vary somewhat with the retinal eccentricity of the FOE (Refs 10–12). Larger errors occur when the FOE is outside the visible area of the screen¹³. The estimation of the FOE is mostly based on the pattern of directions of the individual dot movements, less on their speeds¹⁴.

Visual versus extraretinal mechanisms

Eye rotation, or combined eye-head rotation, induces additional retinal image motion which modifies the retinal-flow pattern and uncouples retinal from optic flow (Fig. 1D–I). In particular, this upsets the use of the FOE as an indicator of heading. Hence, a different strategy must be used. Two alternatives have been proposed. First, eye or eye-head rotations are usually accompanied by non-visual, 'extraretinal' signals. These encompass proprioceptive or vestibular signals or an internal copy of the motor command ('efference copy'). The first hypothesis therefore assumes that extraretinal signals are used to compensate for the rotational component of the retinal flow and to reconstruct the focus of expansion. On the other hand, retinal flow itself often carries enough information to separate translational and rotational components (see Box 2). The second hypothesis therefore proposes that retinal flow is used directly to recover retinal heading by a purely visual mechanism.

Studies of the relative contribution of visual and extraretinal signals usually involve the paradigm of simulated eye movements^{15,16}. The idea is to present retinal flow normally

experienced during combined translation and eye rotation to the stationary eye, that is, to a subject fixating a stable point on the screen. In this case, it is argued, only visual mechanisms of heading detection can be used, because the extraretinal signals available during real eye movement are absent. The possibility that extraretinal signals might also indicate fixation and thus create a conflict between visual and extraretinal cues is seldom discussed in these studies¹⁷. Comparing this paradigm with the case when the subject actually performs the same eye movement during presentation of only translational optic flow is thought to reveal the contribution of extraretinal mechanisms.

In the latter case, the non-conflict situation, heading errors are small (2–4 degrees) (Refs 7,17–20). This is true for eye rotation and for active head rotation²¹. However, this result does not by itself prove that extraretinal signals are required because both extraretinal and visual signals are available (and congruent) in this situation. In the simulated eye movement paradigm, the conflict case, results are more variable. The initial study by Warren and Hannon showed small errors, comparable to those observed during real eye movements, provided that the simulated scene contained large depth variations⁷. Warren and Hannon modelled their stimuli after normal locomotor behavior. They simulated rather slow (up to 1.5 deg/s) eye movements that stabilized gaze on an environmental target (compare Box 1 and Fig. 1D,E). Banks and colleagues^{17,19} used higher rotation rates and simulated pursuit of an independently moving target (Fig. 1G–I). They found much larger errors, up to 15 degrees for rotation rates of 5 deg/s. They suggested that direct visual estimation of

heading can only be performed for very slow eye movements. At rotation rates greater than 1 deg/s extraretinal information would be required. Others, however, found low errors with simulated rotation rates up to 16 deg/s (Refs 15,18,22,23), instead supporting the hypothesis of direct visual estimation of heading.

The true heading performance and the prediction for an observer that ignores the rotation and simply estimates the FOE in the retinal flow can be compared (Fig. 2). The former would correspond to pure visual heading detection, the latter to complete reliance on extraretinal input. It is apparent from Fig. 2 that neither hypothesis captures all of the data. Moreover, there are large variations between studies, which indicate that additional factors (rate of eye rotation and the simulated environment) must influence heading judgements.

Path perception

The above descriptions assume linear motion and a rotational component induced by eye movement. However, a combination of translational and rotational self-motion also arises during movement along a curved path. In this case the rotation axis is not in the eye but at the center of the motion curve. This creates a further problem for heading detection from retinal flow because the flow field cannot specify the location of the rotation axis and hence the origin of the rotational component. The retinal flow is ambiguous in that respect. Decomposition of the retinal flow in rotational and translational components only specifies the momentary retinal heading. To relate this to a path in the visual scene requires additional transformations which can give rise to additional errors²³. In principle, curved movements can be distinguished from straight movements²⁴. Observers can discriminate whether an object is on their future curved path with similar precision as during linear movement²⁵. Sometimes, however, subjects erroneously perceive curved motion paths where linear motion is presented. When pure expansion patterns differ with respect to the average speed of motion in the left and right hemifields, a curved self-motion towards the side with the lower speed is perceived²⁶.

Because of the above ambiguity, erroneous curved path percepts also often occur during combinations of straight translation with simulated eye movements^{20,22,27,28}. Simulated eye rotations might be falsely interpreted as path curvatures. Royden proposed that the extraretinal signal functions to differentiate linear path plus eye rotation from curved movements²⁷. An alternative explanation suggests that different types of visual-heading estimation are carried out in parallel for different axes of rotation²⁰. A possible mechanism is to impose various constraints on the rotation, estimate heading for each case, and combine the results. This has been suggested for different kinds of eye movements^{29,30}. The approach might be applied to different rotation axes as well. An extraretinal signal could then provide a bias towards a rotation axis in the eye.

In summary, heading errors during simulated rotation and translation may be caused by errors in decomposition, by errors in path extrapolation, or both. Taken together, the present evidence^{20,23,28} suggests that visual decomposition is possible even for high rotation rates but path extrapolation is a source of error.

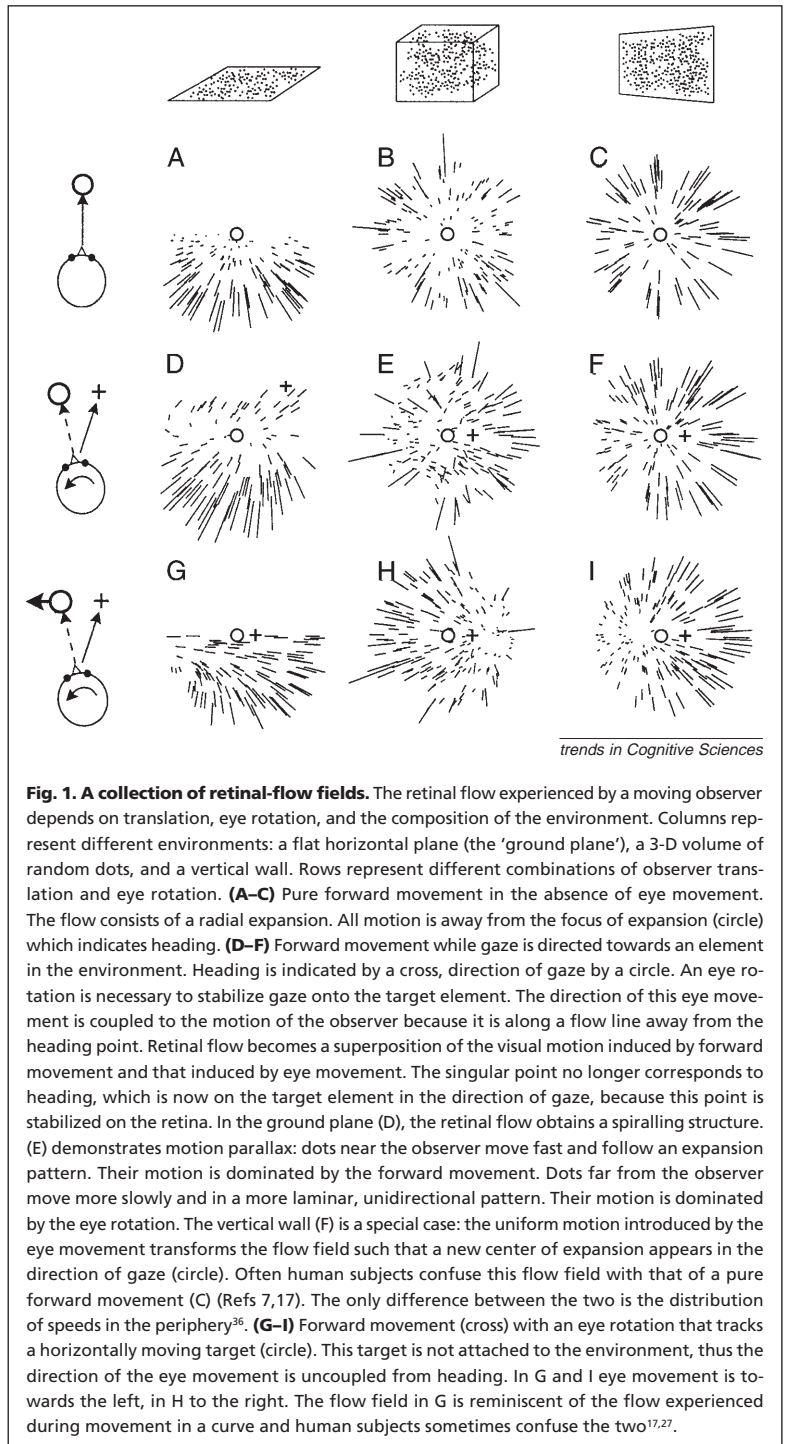


Fig. 1. A collection of retinal-flow fields. The retinal flow experienced by a moving observer depends on translation, eye rotation, and the composition of the environment. Columns represent different environments: a flat horizontal plane (the 'ground plane'), a 3-D volume of random dots, and a vertical wall. Rows represent different combinations of observer translation and eye rotation. (A–C) Pure forward movement in the absence of eye movement. The flow consists of a radial expansion. All motion is away from the focus of expansion (circle) which indicates heading. (D–F) Forward movement while gaze is directed towards an element in the environment. Heading is indicated by a cross, direction of gaze by a circle. An eye rotation is necessary to stabilize gaze onto the target element. The direction of this eye movement is coupled to the motion of the observer because it is along a flow line away from the heading point. Retinal flow becomes a superposition of the visual motion induced by forward movement and that induced by eye movement. The singular point no longer corresponds to heading, which is now on the target element in the direction of gaze, because this point is stabilized on the retina. In the ground plane (D), the retinal flow obtains a spiralling structure. (E) demonstrates motion parallax: dots near the observer move fast and follow an expansion pattern. Their motion is dominated by the forward movement. Dots far from the observer move more slowly and in a more laminar, unidirectional pattern. The motion is dominated by the eye rotation. The vertical wall (F) is a special case: the uniform motion introduced by the eye movement transforms the flow field such that a new center of expansion appears in the direction of gaze (circle). Often human subjects confuse this flow field with that of a pure forward movement (C) (Refs 7,17). The only difference between the two is the distribution of speeds in the periphery³⁶. (G–I) Forward movement (cross) with an eye rotation that tracks a horizontally moving target (circle). This target is not attached to the environment, thus the direction of the eye movement is uncoupled from heading. In G and I eye movement is towards the left, in H to the right. The flow field in G is reminiscent of the flow experienced during movement in a curve and human subjects sometimes confuse the two^{17,27}.

Combining retinal flow with information about the environment

Another factor that could influence heading judgements is information about 3-D scene layout. Knowledge of the depth structure of the scene could aid the separation of translation and rotation, because the motion of objects in the flow depends on their distance from the observer. The motion of distant points can be used to estimate rotation while the motion of near points is more useful to obtain translational information (see Fig. 1E). Independent knowledge about the depth structure of the scene is normally available through binocular vision. Stereoscopic depth improves noise tolerance for movements in a random, noisy 3-D environment^{31,32}.

Box 2. Mathematical considerations

The pattern of motion seen by the eye of a moving observer (or by a moving camera for that matter) is determined by the parameters of the movement and the layout of the environment (Ref. a). At any instance, the motion of the eye, like any rigid body motion, can be described by translation and rotation. Each has in principle three degrees of freedom. The image motion of an element of the environment depends on these parameters and on its distance from the eye. For translation, induced visual speed of each element is inversely proportional to distance. This is known as 'motion parallax'. In contrast, a rotation induces equal angular speed in all image points, independent of distance. Hence motion parallax is an important cue to segregate translational from rotational motion.

Heading estimation requires the determination of the direction of translation. Mathematically this is a problem with many unknown parameters. These are the six degrees of freedom of self-motion plus the distances of all visible points from the eye. Accurate measurement of the retinal flow provides information to solve this problem, namely the direction and speed of every moving point. This allows the mathematical decomposition of the flow into translational

and rotational components and the estimation of heading once more than six moving points are registered (Ref. b). Usually many more points are available but their measurements are noisy. In this case, redundant information provided by more than six points can be used (Refs c,d). Limiting factors for heading detection from a mathematical point of view are small fields-of-view, high rotation rates, and limited depth variations in the visual field which lead to reduced motion parallax (Ref. d).

References

- a Gordon, D.A. (1965) Static and dynamic visual fields in human space perception *J. Opt. Soc. Am.* 55, 1296–1303
- b Longuet-Higgins, H.C. and Prazdny, K. (1980) The interpretation of a moving retinal image *Proc. R. Soc. London Ser. B* 208, 385–397
- c Bruss, A.R. and Horn, B.K.P. (1983) Passive navigation *Comput. Vis. Graph. Image Process.* 21, 3–20
- d Koenderink, J.J. and van Doorn, A.J. (1987) Facts on optic flow *Biol. Cybern.* 56, 247–254

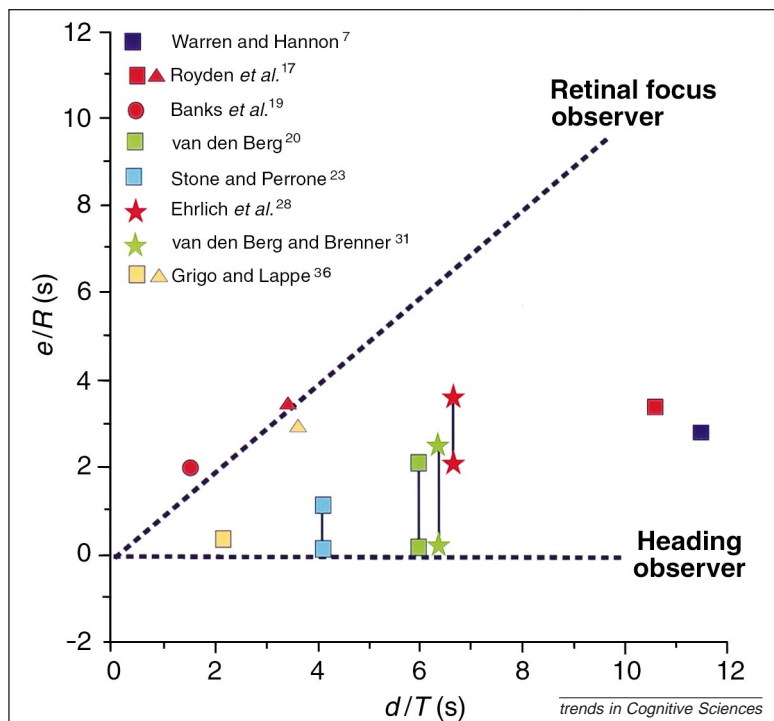


Fig. 2. Perceived heading during eye movements. Data obtained in several studies are compared with predictions of two opposite theories of heading perception during eye movements. The first is perfect visual decomposition which would find the true heading (heading observer). The second is complete reliance on extraretinal eye movement signals and the focus of expansion (retinal focus observer). Because extraretinal signals were absent in all of these studies, the retinal focus observer theory would predict errors to the extent that the retinal focus of expansion is distorted by eye movements. Predictions for the error were derived from a model tuned to purely expanding flow⁶⁷. The predicted error depends on the rate of eye rotation and on the layout of the environment. For a frontoparallel plane, eye rotation shifts the focus of expansion (see Fig. 1) by an amount $(d/T)R$, where d is the distance of the plane from the observer and T and R are observer translation and rotation, respectively. We therefore converted absolute error values obtained in the various studies to heading error per unit eye rotation (e/R) and used (d/T) as a measure of the influence of the environment. For a frontoparallel plane (triangular symbols), d is given by the distance of the plane from the eye. For studies using clouds of dots (Refs 17,19,20,23,28,36), d is the average dot distance from the eye. Studies with ground planes have not been included as these contain additional depth cues that might influence the performance. Symbols connected with vertical lines indicate ranges of responses for different subjects in one study. Because most studies (except Ref. 23) asked subjects to indicate heading with reference to the visual scene the errors observed might be due to errors in path perception rather than errors in decomposition. In that sense, the values show upper limits for decomposition errors.

Stereoscopic depth also influences an illusory transformation of optic flow (Fig. 3), consistent with the hypothesis that near and far points contribute differently to the separation of translational and rotational components³³. Moreover, monocular depth cues are also used to improve the visual estimate of heading in the presence of noise³⁴.

Dynamic properties and saccadic eye movements

Saccadic gaze shifts disrupt the retinal flow and change the retinal projection of the direction of heading on average twice per second (Box 1). Heading judgements are possible for presentation times as short as 228–400 ms, that is, within the time available between two saccades^{12,23}. Yet, visual search for the heading direction is only rarely accomplished in a single saccade, indicating that heading direction is usually processed across successive saccadic intervals³⁵.

Short presentation times might even enhance the ability to estimate heading from retinal flow. Grigo and Lappe have investigated self-movement towards a vertical plane in combination with a simulated eye movement (flow field in Fig. 1F). Because the retinal flow in this situation closely resembles that of a pure forward movement (Fig. 1C), subjects often confuse the two and erroneously report a straight translation^{7,17}. These errors were reduced, however, when the presentation duration was *decreased* from 3.0 s to 400 ms (Ref. 36). This might reflect different temporal dynamics of visual and extraretinal contributions to heading perception. Systematic errors for longer durations must result in part from the conflict between visual and extraretinal signals, because they do not occur during real eye movements^{7,17}. Grigo and Lappe suggested that heading detection in the typical time interval between two saccades uses visual mechanisms and that extraretinal inputs become important only at a later time or during longer fixations or eye pursuits³⁶.

Mechanisms of heading detection

Electrophysiology in macaque monkeys has shown several areas in the posterior parietal cortex involved in optic-flow processing³⁷. Most research has focussed on the medial superior temporal (MST) area, because this is the first area in

the cortical motion pathway with genuine optic flow selectivity^{38–40}. Many MST cells also respond to real movement of the animal, even in darkness^{41,42}. Cells in MST are selective for the location of the focus of expansion^{43,44}. MST responses during combined optic flow and eye movements suggest that the neuronal population can solve the problem of rotations^{45,46}. Area MST's main input originates from the middle temporal (MT) area, which contains neurons sensitive to local motion, that is, to the individual motion vectors of the flow field.

Population heading-map model

Current computational models of heading perception attempt to reproduce the human psychophysical findings using the properties of MT and MST neurons (see Ref. 47 for a detailed review). These models typically consist of two layers of neuron-like elements which represent the retinal flow as input (MT layer) and the computed heading as output (MST layer). The computation of heading and the properties of the neurons in the second (MST) layer mainly depend on the setup of synaptic connections with the first layer. In the population heading-map model introduced by Lappe and Rauschecker^{29,48} the connections are derived from computer vision algorithms^{49,50} that estimate heading in an optimum way. Mathematically, this is achieved by finding from all possible flow fields the one that minimizes the mean-squared difference between the measured flow field (the first layer/MT activities) and all possible flow fields constructed from any combination of observer motion (translation and rotation) and scene structure. The combination that minimizes this difference is equivalent to the actual self-motion. This is a high-dimensional computational problem (see Box 2), but it can be reduced to the search for translational heading in a low-dimensional subspace⁵⁰. In the model, populations of optic-flow processing neurons compute the mean-squared differences for a large number of possible headings in parallel. Each population estimates the current likelihood of a specific heading. This results in a heading map, the elements of which are populations of neurons. The most likely heading direction is equated with the peak of activity in this map. Extraretinal eye movement signals are easily combined with this approach to provide better estimates of heading when direct visual estimation is difficult⁵¹.

The population heading-map model reproduces many basic properties of human heading detection such as the dependence on scene structure^{48,51,52}, dot density⁴⁸, eccentricity of the focus of expansion⁵², as well as illusory optic flow perception (see Fig. 3). Moreover, the model made predictions for the properties of optic-flow processing neurons which were subsequently confirmed in recordings from monkey area MST (Refs 44,53). As in model neurons, individual MST cells cannot unambiguously specify the focus. A population code based on actual responses of MST neurons can locate the focus with an accuracy near that obtained by human observers⁴⁴.

Templates for specific flow patterns

A different approach to solving the heading task proposes the construction of templates for specific flow patterns^{54,55}. The response of an individual neuron in a template model depends on the match between the input flow field and the

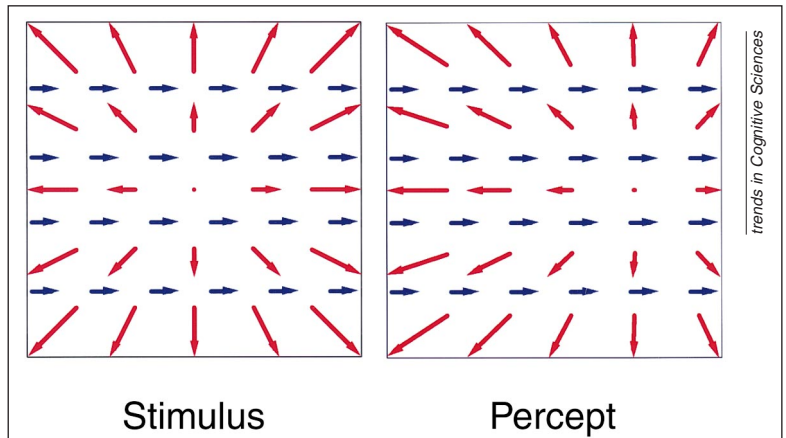


Fig. 3. An illusion of optic flow. The focus of expansion (FOE) is often thought to be synonymous with heading. This is usually incorrect, because eye movements distort the retinal flow. The percept above illustrates an illusion in which the focus of expansion appears to be shifted up to 20 degrees away from its true location. This illusion occurs when an expanding dot pattern is transparently overlapped by unidirectional motion⁴⁸. The shift cannot be explained by a simple vector averaging of the two patterns, as the FOE then must shift in the opposite direction, but has been explained by mechanisms of heading perception during combined translation and rotation³⁰. The overlapping unidirectional motion indicates a rotational component. Compensatory heading detection mechanisms shift the perceived heading against this rotation, that is, in the direction of the overlapping motion. As a result, the focus of expansion is actually 'seen' in the direction of heading even though presented at a different location on the screen. The illusory shift is influenced by binocular depth perception: when unidirectional motion is presented in front of the expansion, motion parallax and binocular disparity give conflicting information, because motion parallax implies that distant points move in a largely uniform pattern while near points follow an expanding motion. The illusory shift is strongly reduced in this case³³. The illusory stimulus has also been used to study the neural analysis of optic flow in macaque area MST. The responses of most single neurons appeared inconsistent with the illusory shift⁴⁹. This inconsistency is resolved, however, when one assumes a population code for heading as proposed by the population map model³³.

template of that neuron. Sensitivity for the direction of heading is obtained by building templates for all flow fields that could possibly occur for any given heading. For instance, the simplest expansion template, a radial arrangement of motion detectors, would respond maximally to a radial expansion which, in turn, would correspond to pure forward movement⁵⁴. This approach requires a very large number of templates⁵⁵ because an infinite number of flow patterns could arise from a single heading depending on eye movements and variations of scene structure (see Box 2 and Fig. 1). Perrone and Stone proposed to reduce the number of templates by considering only gaze-stabilization eye movements⁵⁵. This is the most prominent natural oculomotor behavior (Box 1). Constraining the eye movements in this way is an effective method of reducing the number of degrees of freedom for flow field analysis. The use of such a constraint was first proposed by Lappe and Rauschecker⁴⁸. It is one among several hypotheses about ongoing eye movements that are evaluated in parallel in the population heading-map model^{29,30}. In that model, different constraints are evaluated by different sub-populations and the whole population response is governed by the best-fitting hypothesis. However, sole use of the gaze-stabilization restriction (in Refs 55,56) appears to be inconsistent with human psychophysical data⁵⁷.

As with the population heading-map model, the behavior of the model of Perrone and Stone has been extensively compared to response properties of MST neurons, and matches many of the findings from that area⁵⁶. As this comparison was carried out only for the restricted version of the model (that

Box 3. Other visual cues for heading and the guidance of self-motion

The pattern of flow is only one among several visual cues for heading. Cutting and colleagues have described a variety of local relative motion cues that may be used for heading estimation (Refs a–c). These local cues refer to the relative motion of objects in the visual field during combined translation and gaze stabilization onto a particular object. As such, they are part of the retinal flow pattern and contribute to the information in the retinal flow, but they can also be used more directly in a somewhat heuristic manner. For instance, objects in a small area along the line of sight actually undergo inward motion in the retinal flow, that is, motion towards the point of gaze. This area is located on either side of the direction of gaze, dependent on the distance of the object. For objects closer than the fixation target, inward motion occurs on the side where heading is located. For more distant objects, inward motion occurs on the side opposite to heading. Hence, when an estimate of the depth structure of the scene is available it is possible to infer heading with respect to the center of gaze (a nominal judgement) from an analysis of inward motion. In similar manner, information about heading is provided by the distribution of acceleration and deceleration of object motions in the visual field (Ref. a), by the motion direction of the object nearest to the observer (Refs a,b), or by the analysis of the motions of pairs of objects (Ref. c). To use these cues, however, independent knowledge of the depth layout is required.

The idea to use object motion for heading estimation is actually quite old. Llewellyn suggested that guidance of self-motion towards a target could simply be achieved by continuously adjusting the travel path to cancel possible drift motion of the target (Ref. d). This strategy becomes difficult, however, when the target is stabilized by eye movement. In this case, parameters of the eye movement might be used instead. Because the motion of the target directly results from the motion of the observer, speed and direction of the eye movement are linked to the parameters of self-motion (see Box 1). This constrains the resulting retinal flow and the axis of rotation (Ref. e). In combination with other constraints – given, for instance, by the horizon when self-movement is parallel to the ground – this could allow the observer to estimate heading (Refs f,g).

The perceived location of the goal is obviously an important factor for the control of locomotion. Rushton *et al.* studied locomotor paths in subjects who wore prism glasses to deflect their perceived location of objects in the visual field (Ref. h). They found consistent path errors corresponding to the

deflection angle. Because such prisms influence the pattern of flow and the perceived object positions alike, they argue that locomotor control by maintaining the target and the retinal heading aligned could not result in such errors. Hence, they suggest that perceived target location is an independent control parameter that is used instead of flow analysis. Similarly, visual landmarks (Ref. i) and mental maps (Ref. j) are important for visual navigation. Car driving is a particular situation in which static cues for heading are available by the road edges and markings. The position and orientation of the road edges are used to estimate the current position of the car in the lane and to control steering (Refs k–m).

To summarize, visual navigation uses a variety of motion and non-motion cues. The interplay between these cues remains an open question.

References

- a Cutting, J.E. (1996) Wayfinding from multiple sources of local information in retinal flow *J. Exp. Psychol. Hum. Percept. Perform.* 22, 1299–1313
- b Cutting, J.E. *et al.* (1999) Human heading judgements and object-based motion information *Vis. Res.* 39, 1079–1105
- c Wang, R.F. and Cutting, J.E. (1999) Where we go with a little good information *Psychol. Sci.* 10, 71–75
- d Llewellyn, K.R. (1971) Visual guidance of locomotion *J. Exp. Psychol.* 91, 245–261
- e Lappe, M. and Rauschecker, J.P. (1995) Motion anisotropies and heading detection *Biol. Cybern.* 72, 261–277
- f Banks, M.S. *et al.* (1996) Estimating heading during real and simulated eye movements *Vis. Res.* 36, 431–443
- g van den Berg, A.V. and Brenner, E. (1994) Humans combine the optic flow with static depth cues for robust perception of heading *Vis. Res.* 34, 2153–2167
- h Rushton, S. *et al.* (1998) Guidance of locomotion on foot uses perceived target location rather than optic flow *Curr. Biol.* 8, 1191–1194
- i Gillner, S. and Mallot, H.A. (1998) Navigation and acquisition of spatial knowledge in a virtual maze *J. Cogn. Neurosci.* 10, 445–463
- j Vishton, P.M. and Cutting, J.E. (1995) Wayfinding, displacements and mental maps: velocity fields are not typically used to determine one's aimpoint *J. Exp. Psychol. Hum. Percept. Perform.* 21, 978–995
- k Beall, A.C. and Loomis, J.M. (1996) Visual control of steering without course information *Perception* 25, 481–494
- l Land, M.F. and Lee, D.N. (1994) Where we look when we steer *Nature* 369, 742–744
- m Land, M.F. and Horwood, J. (1995) Which parts of the road guide steering *Nature* 377, 339–340

of Ref. 55) it is difficult to judge how the comparison would appear for the unconstrained version. In the population heading-map model, the use of several parallel and non-exclusive assumptions about eye movements leads to corresponding differences in the behavior of the associated neurons^{29,44}. Some model neurons become completely immune against any type of rotation in the flow field. These neurons respond only to expansion and contraction, as do a few MST neurons^{38,39}. Other neurons in the model respond to expansion/contraction and to rotation stimuli. Their response to one of these patterns changes when another pattern is added to the stimulus. This behavior is found in most MST neurons^{39,40,58,59} and is also the dominant behavior in the Perrone and Stone model.

A different way to avoid many templates was suggested by Beintema and van den Berg^{60,61}. Their model combines two types of templates: (1) a template tuned to pure observer translation; and (2) a template that represents the derivative of the first template to rotation. This is formally equivalent to a Taylor expansion of the type-1 template's activity to the rotational flow. The activity of the derivative template is used to compensate for changes in activity of the type-1 template when the eye rotates. Such a combination of templates is tuned to the head-centric flow because it prefers the same

observer translation irrespective of the eye rotation. However, because the compensation is done to the first-order only, it is inevitably inadequate for high rotation rates. This limitation can be relaxed if an oculomotor signal is used to modulate the activity of the derivative template extending the effective range of rotations for which compensation is successful. In this way the model can account for the differences observed between real and simulated eye movements in human subjects.

Zemel and Sejnowski have trained a neural network to develop a neural representation of flow fields⁶². This network consisted of three layers of neurons: input, hidden, and output layers. It learned to reproduce a set of flow fields that contained several independent moving objects. The hidden layer neurons developed response properties that were similar to several properties of MST neurons. Based on these responses, it was possible to train a second network to estimate heading as well as the motion of individual objects. Because the simulations regarding heading in this study were scarce, however, it is difficult to evaluate how well this model predicts the properties of human heading estimation. But, importantly, it demonstrates that the recovery of heading and the recovery of the motions of individual objects might use similar mechanisms.

Comparison of local image motions as a unifying concept

Visual-heading estimation from retinal flow requires the analysis of the 2-D pattern of image motion. The central question is how this analysis is performed. In terms of neural computation the question becomes how the local 2-D motion sensitivities must be organized in the receptive field of optic-flow processing neurons. Current models might offer an answer to this question.

Let us first consider the simple idea which formed the starting point of early template models⁵⁴. Selectivity for expansion during pure translation could be simply constructed by arranging 2-D motion sensitivities in a radially expanding template pattern. However, neurophysiological experiments have clearly disproved this arrangement. Duffy and Wurtz, who at the time called it the 'direction mosaic' hypothesis, tested it directly by comparing optic flow selectivity to the selectivity for small 2-D motion in different parts of the receptive field⁶³. Clearly, for true expansion-selective cells the 2-D motion selectivities in subparts of the receptive field did not match the hypothesis.

Interestingly, several other models independently arrived at a different mechanism: the use of differences between flow vectors. Differences between flow vectors are useful because of the properties of motion parallax (Box 2). As the rotational component of all flow vectors is identical, the difference between any two flow vectors depends only on the translational component. For example, differences between neighboring flow vectors can be used to compute a rotation-independent, local-motion-parallax field, which reconstructs the focus of expansion⁶⁴. Such a computation could be implemented by motion-selective neurons with center-surround, opponent-motion selectivity⁶⁵, which are quite common in cortical area MT. Sensitivity for local opponent motions is also an important mechanism in the model of Beintema and van den Berg⁶¹. The neurons that represent the derivative templates, that is, the neurons that subserve the decomposition of translation and rotation, actually compute local motion parallax⁶¹.

The usefulness of differences between flow vectors is not restricted to local regions, however. One might argue that parallax information from widely separated regions of the visual field could often be even more useful because most local image regions contain only small depth variations, that is, limited motion parallax. An analysis of the connection structure in the population heading-map model reveals that its computations, too, comprise comparisons between small groups of 2–5 flow vectors^{52,66}. If these flow vectors are near each other then the comparison is equivalent to an opponent-motion detector. If they are further apart the computation becomes more complex, consisting of a comparison of both the speed and direction of motion. Local opponent motion can therefore be regarded as a special case of a more global flow analysis in this model.

In summary, several current models contain similar operators despite their initially different computational approaches. This suggests that these operators reflect a common principle of optic-flow processing. It will be interesting to see whether evidence for this can be found in neurophysiological properties of real optic-flow processing neurons.

Outstanding questions

- The future motion path is often more important for the control of self-motion than the current instantaneous heading. How is path information obtained from retinal flow and extraretinal signals and how is the path predicted?
- Humans can use many other cues besides optic flow for visual navigation (Box 3). How is optic flow combined with other navigational strategies?
- Most previous studies have used passive judgements of heading. Normally, however, heading judgements are required during active locomotor behavior. What role does optic flow play during active behavior? How are eye movements actively used to support heading perception and flow analysis?
- What are the computations performed by optic-flow-selective neurons in the brain? What is the structure of their receptive fields? How are the signals from different neurons combined? What is the structure of the heading map in area MST?

Conclusion

Goal-directed spatial behavior relies heavily on vision. Retinal flow provides visual input to monitor self-motion, navigate and guide future movements, and avoid obstacles. This article has reviewed the large body of knowledge about how humans analyse retinal flow that has accumulated in psychophysical studies. Humans can in principle use retinal flow for the determination of heading, in addition to several other visual cues (see Box 3). To solve the problem of eye rotations robustly, the visual system combines retinal-flow analysis with a multitude of other sensory signals including efference copies of motor commands, proprioceptive signals, and monocular as well as binocular depth cues. Neurophysiological studies have investigated the neuronal mechanisms of optic-flow processing in primate cortex. Computational models based on physiological and psychophysical data have developed unifying concepts of how the brain solves the complex computational problems inherent in retinal-flow analysis.

Acknowledgements

We thank Antje Grigo and Bart Krekelberg for comments on the manuscript. This review and much of the cited work was made possible by financial support from the Human Frontier Science Program (RG-71/96B and RG 34/96B), the German Science Foundation (SFB 509), and the Dutch Research Council (SLW-NWO 805-33.171-P).

References

- 1 Gibson, J.J. (1950) *The Perception of the Visual World*, Houghton Mifflin
- 2 Lee, D.N. (1980) The optic flow field: the foundation of vision *Philos. Trans. R. Soc. London Ser. B* 290, 169–179
- 3 Prokop, T., Schubert, M. and Berger, W. (1997) Visual influence on human locomotion: modulation to changes in optic flow *Exp. Brain Res.* 114, 63–70
- 4 Tresilian, J.R. (1991) Empirical and theoretical issues in the perception of time to contact *J. Exp. Psychol. Hum. Percept. Perform.* 17, 865–876
- 5 Bremmer, F. and Lappe, M. (1999) The use of optical velocities for distance discrimination and reproduction during visually simulated self-motion *Exp. Brain Res.* 127, 33–42
- 6 Regan, D. and Beverley, K.I. (1982) How do we avoid confounding the direction we are looking and the direction we are moving? *Science* 215, 194–196
- 7 Warren, Jr, W.H. and Hannon, D.J. (1990) Eye movements and optical flow *J. Opt. Soc. Am. Ser. A* 7, 160–169
- 8 Cutting, J.E. et al. (1992) Wayfinding on foot from information in retinal, not optical, flow *J. Exp. Psych. Gen.* 121, 41–72
- 9 Warren, Jr, W.H., Morris, M.W. and Kalish, M. (1988) Perception of

- translational heading from optical flow *J. Exp. Psychol. Hum. Percept. Perform.* 14, 646–660
- 10** Warren, Jr, W.H. and Kurtz, K.J. (1992) The role of central and peripheral vision in perceiving the direction of self-motion *Percept. Psychophys.* 51, 443–454
- 11** D'Avossa, G. and Kersten, D. (1996) Evidence in human subjects for independent coding of azimuth and elevation for direction of heading from optical flow *Vis. Res.* 36, 2915–2924
- 12** te Pas, S., Kappers, A. and Koenderink, J. (1998) Locating the singular point in first-order optical flow fields *J. Exp. Psychol. Hum. Percept. Perform.* 24, 1415–1430
- 13** Crowell, J.A. and Banks, M.S. (1993) Perceiving heading with different retinal regions and types of optic flow *Percept. Psychophys.* 53, 325–337
- 14** Warren, Jr, W.H. et al. (1991). On the sufficiency of the velocity field for perception of heading *Biol. Cybern.* 65, 311–320
- 15** Rieger, J.H. and Toet, L. (1985) Human visual navigation in the presence of 3-D rotations *Biol. Cybern.* 52, 377–381
- 16** Cutting, J. (1986) *Perception With an Eye for Motion*, MIT Press
- 17** Royden, C.S., Crowell, J.A. and Banks, M.S. (1994) Estimating heading during eye movements *Vis. Res.* 34, 3197–3214
- 18** van den Berg, A.V. (1992) Robustness of perception of heading from optic flow *Vis. Res.* 32, 1285–1296
- 19** Banks, M.S. et al. (1996) Estimating heading during real and simulated eye movements *Vis. Res.* 36, 431–443
- 20** van den Berg, A.V. (1996) Judgements of heading *Vis. Res.* 36, 2337–2350
- 21** Crowell, J.A. et al. (1998) Visual self-motion perception during head turns *Nat. Neurosci.* 1, 732–737
- 22** Cutting, J.E. et al. (1997) Heading and path information from retinal flow in naturalistic environments *Percept. Psychophys.* 59, 426–441
- 23** Stone, L. and Perrone, J. (1997) Human heading estimation during visually simulated curvilinear motion *Vis. Res.* 37, 573–590
- 24** Turano, K. and Wang, X. (1994) Visual discrimination between a curved path and straight path of self motion: effects of forward speed *Vis. Res.* 34, 107–114
- 25** Warren, Jr, W.H. et al. (1991) Perception of circular heading from optical flow *J. Exp. Psychol. Hum. Percept. Perform.* 17, 28–43
- 26** Dyre, B.P. and Andersen, G.J. (1997) Image velocity magnitudes and perception of heading *J. Exp. Psychol. Hum. Percept. Perform.* 23, 546–565
- 27** Royden, C.S. (1994) Analysis of misperceived observer motion during simulated eye rotations *Vis. Res.* 34, 3215–3222
- 28** Ehrlich, S.M. et al. (1998) Depth information and perceived self-motion during simulated gaze rotations *Vis. Res.* 38, 3129–3145
- 29** Lappe, M. and Rauschecker, J.P. (1993) Computation of heading direction from optic flow in visual cortex, in *Advances in Neural Information Processing Systems* (Vol. 5) (Giles, C.L., Hanson, S.J. and Cowan, J.D., eds), pp. 433–440, Morgan Kaufmann
- 30** Lappe, M. and Rauschecker, J.P. (1995) An illusory transformation in a model of optic flow processing *Vis. Res.* 35, 1619–1631
- 31** van den Berg, A.V. and Brenner, E. (1994) Why two eyes are better than one for judgements of heading *Nature* 371, 700–702
- 32** Lappe, M. (1996) Functional consequences of an integration of motion and stereopsis in area MT of monkey extrastriate visual cortex *Neural Comput.* 8, 1449–1461
- 33** Grigo, A. and Lappe, M. (1998) Interaction of stereo vision and optic flow processing revealed by an illusory stimulus *Vis. Res.* 38, 281–290
- 34** van den Berg, A.V. and Brenner, E. (1994) Humans combine the optic flow with static depth cues for robust perception of heading *Vis. Res.* 34, 2153–2167
- 35** Hooge, I.T.C., Beintema, J.A. and van den Berg, A.V. Visual search of the heading direction *Exp. Brain Res.* (in press)
- 36** Grigo, A. and Lappe, M. (1998) An analysis of heading towards a wall, in *Vision and Action* (Harris, L.R. and Jenkin, M., eds), pp. 215–230, Cambridge University Press
- 37** Lappe, M., ed. *Neuronal Processing of Optic Flow*, Academic Press (in press)
- 38** Tanaka, K. et al. (1986) Analysis of local and wide-field movements in the superior temporal visual areas of the macaque monkey *J. Neurosci.* 6, 134–144
- 39** Duffy, C.J. and Wurtz, R.H. (1991) Sensitivity of MST neurons to optic flow stimuli: I. A continuum of response selectivity to large-field stimuli *J. Neurophysiol.* 65, 1329–1345
- 40** Lagae, L. et al. (1994) Responses of macaque STS neurons to optic flow components: a comparison of areas MT and MST *J. Neurophysiol.* 71, 1597–1626
- 41** Bremmer, F. et al. (1999) Linear vestibular self-motion signals in monkey medial superior temporal area *Ann. New York Acad. Sci.* 871, 272–281
- 42** Duffy, C.J. (1998) MST neurons respond to optic flow and translational movement *J. Neurophysiol.* 80, 1816–1827
- 43** Duffy, C.J. and Wurtz, R.H. (1995) Response of monkey MST neurons to optic flow stimuli with shifted centers of motion *J. Neurosci.* 15, 5192–5208
- 44** Lappe, M. et al. (1996) Optic flow processing in monkey STS: a theoretical and experimental approach *J. Neurosci.* 16, 6265–6285
- 45** Bradley, D. et al. (1996) Mechanisms of heading perception in primate visual cortex *Science* 273, 1544–1547
- 46** Page, W.K. and Duffy, C.J. (1999) MST neuronal responses to heading direction during pursuit eye movements *J. Neurophysiol.* 81, 596–610
- 47** Lappe, M. Computational mechanisms for optic flow analysis in primate cortex, in *Neuronal Processing of Optic Flow* (Lappe, M., ed.), Academic Press (in press)
- 48** Lappe, M. and Rauschecker, J.P. (1993) A neural network for the processing of optic flow from ego-motion in man and higher mammals *Neural Comput.* 5, 374–391
- 49** Bruss, A.R. and Horn, B.K.P. (1983) Passive navigation *Comput. Vis. Graph. Image Process.* 21, 3–20
- 50** Heeger, D.J. and Jepson, A. (1992) Subspace methods for recovering rigid motion: I. Algorithm and implementation *Int. J. Comput. Vis.* 7, 95–117
- 51** Lappe, M. (1998) A model of the combination of optic flow and extraretinal eye movement signals in primate extrastriate visual cortex *Neural Netw.* 11, 397–414
- 52** Lappe, M. and Rauschecker, J.P. (1995) Motion anisotropies and heading detection *Biol. Cybern.* 72, 261–277
- 53** Lappe, M. and Duffy, C. (1999) Optic flow illusion and single neuron behavior reconciled by a population model *Eur. J. Neurosci.* 7, 2323–2331
- 54** Perrone, J.A. (1992) Model for the computation of self-motion in biological systems *J. Opt. Soc. Am.* 9, 177–194
- 55** Perrone, J.A. and Stone, L.S. (1994) A model of self-motion estimation within primate extrastriate visual cortex *Vis. Res.* 34, 2917–2938
- 56** Perrone, J.A. and Stone, L.S. (1998) Emulating the visual receptive field properties of MST neurons with a template model of heading estimation *J. Neurosci.* 18, 5958–5975
- 57** Crowell, J. (1997) Testing the Perrone and Stone (1994) model of heading estimation *Vis. Res.* 37, 1653–1671
- 58** Orban, G.A. et al. (1992) First-order analysis of optical flow in monkey brain *Proc. Natl. Acad. Sci. U. S. A.* 89, 2595–2599
- 59** Graziano, M.S.A., Andersen, R.A. and Snowden, R. (1994) Tuning of MST neurons to spiral motions *J. Neurosci.* 14, 54–67
- 60** van den Berg, A.V. and Beintema, J. (1997) Motion templates with eye velocity gain fields for transformation of retinal to head-centric flow *NeuroReport* 8, 835–840
- 61** Beintema, J. and van den Berg, A.V. (1998) Heading detection using motion templates and eye velocity gain fields *Vis. Res.* 38, 2155–2179
- 62** Zemel, R.S. and Sejnowski, T.J. (1998) A model for encoding multiple object motions and self-motion in area MST of primate visual cortex *J. Neurosci.* 18, 531–547
- 63** Duffy, C.J. and Wurtz, R.H. (1991) Sensitivity of MST neurons to optic flow stimuli: II. Mechanisms of response selectivity revealed by small-field stimuli *J. Neurophysiol.* 65, 1346–1359
- 64** Rieger, J.H. and Lawton, D.T. (1985) Processing differential image motion *J. Opt. Soc. Am. Ser. A* 2, 354–360
- 65** Royden, C.S. (1997) Mathematical analysis of motion-opponent mechanisms used in the determination of heading and depth *J. Opt. Soc. Am. Ser. A* 14, 2128–2143
- 66** Heeger, D.J. and Jepson, A. (1992) Recovering observer translation with center-surround motion-opponent mechanisms *Invest. Ophthalmol. Vis. Sci. (Suppl.)* 32, 823
- 67** van den Berg, A.V. Human heading perception, in *Neuronal Processing of Optic Flow* (Lappe, M., ed.), Academic Press (in press)
- 68** Duffy, C.J. and Wurtz, R.H. (1993) An illusory transformation of optic flow fields *Vis. Res.* 33, 1481–1490
- 69** Duffy, C.J. and Wurtz, R.H. (1997) Planar directional contributions to optic flow responses in MST neurons *J. Neurophysiol.* 77, 782–796