# Self-motion Perception from Optic Flow and Rotation Signals

Waarneming van Zelfbeweging
uit het Visuele Stroomveld en Rotatiesignalen

Proefschrift
ter verkrijging van de graad van doctor
aan de Erasmus Universiteit Rotterdam
op gezag van de Rector Magnificus,
Prof. Dr. P.W.C. Akkermans M.A.
en volgens besluit van het College voor Promoties.
De openbare verdediging zal plaatsvinden op
woensdag 15 maart 2000 om 15.45 uur

door

Jakob Andries Beintema

Geboren te Hoogezand-Sappemeer

Promotiecommissie

| | | |
|---|---|---|
| Promotor | : | Prof. Dr. H. Collewijn |
| Co-promotor | : | Dr. A.V. van den Berg |
| Overige leden | : | Prof. Dr. W.A. van de Grind (UU) |
| | | Dr. J. van Opstal (KUN) |
| | | Prof. Dr. C.I. de Zeeuw (EUR) |

Omslag illustratie: vulkaanuitbarsting gefotografeerd door Michiel G. Peters, bewerkt door de auteur.

ISBN 90-9013567-7

Dit proefschrift werd mede mogelijk gemaakt door financiële steun van de Nederlandse Organisatie voor Wetenschappelijk Onderzoek (NWO), Gebied Aard- en Levenswetenschappen (ALW) en het Human Frontier Science Program (HFSP).

# Contents


**1 General introduction** **5**
- 1.1 The rotation problem . . . 6
- 1.2 Physiology . . . 8
- 1.3 Modelling . . . 11
- 1.4 Psychophysics . . . 11

**2 Heading detection using motion templates and eye velocity gain fields** **13**
- 2.1 Introduction . . . 14
- 2.2 Receptive field shifts . . . 16
  - 2.2.1 Gain fields . . . 16
  - 2.2.2 Transformation of retinal to head-centric position . . . 16
  - 2.2.3 Error analysis . . . 19
  - 2.2.4 Physiology . . . 21
- 2.3 Transformation of retinal to head-centric flow . . . 22
  - 2.3.1 Retinal flow . . . 22
  - 2.3.2 Retinal flow templates . . . 24
  - 2.3.3 The analogy between transformation of retinal position and retinal flow . . . 25
  - 2.3.4 Transformation of retinal flow . . . 26
  - 2.3.5 Defining templates . . . 28
- 2.4 Retinal flow template simulations . . . 31
  - 2.4.1 Tuning properties . . . 31
  - 2.4.2 Role of motion parallax . . . 34
- 2.5 Head-centric flow template simulations . . . 36
  - 2.5.1 Oculo-motor signals . . . 37
  - 2.5.2 Visual estimation of rotation . . . 40
  - 2.5.3 Combining visual and extra-retinal signals . . . 41
- 2.6 Discussion . . . 43
  - 2.6.1 Neurophysiological evidence . . . 45
  - 2.6.2 Suggestions for neurophysiological research . . . 49
  - 2.6.3 Other physiological models . . . 51
  - 2.6.4 Conclusions . . . 52

**3 Combining extra-retinal and visual estimates of eye rotation to perceive heading** **53**
3.1 Introduction . . . 54
3.2 Two extra-retinal models . . . 55
3.3 Experiment . . . 61
3.3.1 Methods . . . 61
3.3.2 Results . . . 62
3.4 Discussion . . . 65
3.5 Appendix . . . 68

**4 Pursuit affects precision of perceived heading for small viewing apertures** **71**
4.1 Introduction . . . 72
4.2 General Methods . . . 76
4.3 Experiment I: Perceived heading and its variability . . . 78
4.3.1 Procedure . . . 78
4.3.2 Results . . . 80
4.4 Experiment II: Perceived head-centric direction and its variability . . . 87
4.4.1 Procedure . . . 87
4.4.2 Results . . . 89
4.5 General discussion . . . 91
4.5.1 Conclusion . . . 94
4.6 Acknowledgements . . . 95
4.7 Appendix . . . 95

**5 Perceived heading during simulated torsional eye movements** **99**
5.1 Introduction . . . 100
5.2 Materials and Methods . . . 103
5.3 Experiment I: Perceived heading during simulated torsion . . . 104
5.3.1 Stimuli . . . 105
5.3.2 Results . . . 105
5.4 Experiment II: Effect of depth during simulated torsion . . . 108
5.4.1 Results . . . 109
5.5 Experiment III: Effect of translation speed during simulated torsion . . . 110
5.5.1 Results . . . 111
5.6 Experiment IV: Simulated torsion during simulated horizontal rotation 111
5.6.1 Stimuli . . . 112
5.6.2 Results . . . 115
5.7 Discussion . . . 120
5.7.1 Summary . . . 120
5.7.2 Offset errors: comparison with earlier studies . . . 120
5.7.3 Performance difference between horizontal and torsional rotation 123
5.7.4 Scale errors . . . 125

5.7.5 Contribution of extra-retinal signals . . . . . 128
5.7.6 Conclusions . . . . . 128
5.8 Appendix . . . . . 129

**6 Summary and conclusions** **131**

**References** **139**

**Samenvatting** **147**

**Dankwoord** **153**

**Curriculum Vitae** **155**

**Publications** **157**

# Chapter 1

# General introduction

The value of optic flow for retrieving movement direction was recognised already two centuries ago by astronomers, searching the sky for meteorite showers. The point from which the shower appeared to emanate they termed the radiant, knowing it indicated the direction along which the meteorites were heading towards the observer. As Gibson (1966) pointed out, humans may use the shower of light that hits the eye ball when the eye moves relative to the stable world, to retrieve their direction of movement relative to the world (*heading*). Motion pictures nowadays offer opportunities to a larger audience to appreciate the value of optic flow in isolation. The ongoing display of a diverging cluster of stars, trademark of a well-known movie maker and popular among the screensavers, can bring about a strong sensation of moving forward while in fact one remains seated.

Optic flow is often accompanied by other signals on self-motion, such as the static perspective of the road edge along which one walks. Sensory input from other modalities may also contribute to our perception of self-motion, such as vestibular organs that sense linear acceleration, auditory sensors (think of Doppler's effect in sound) or somatic sensors (receptors in the skin and joints). Nevertheless, continuous motion forms an important source of information to perceive the relative movement between an object and the observer. Illustrative for this is Zihl's (1983) famous patient LM, who suffers "motion-blindness" due to a rare lesion in a motion-sensitive brain area. She reported to have difficulty in crossing roads as she could not judge the position of cars - 'When I'm looking at the car first, it seems far away. But then, when I want to cross the road, suddenly the car is very near'. Although humans probably use whatever information is available, optic flow forms an important source of feedback on one's self-movement that may be used in a variety of tasks, like keeping posture, walking along a corridor or driving a car.

Not only translation, but also rotation of the eye, head or body relative to the world is part of our self-motion. Rotation moves the projection of the world across our eye ball. Nevertheless, such self-motion does not change a subject's percept of body orientation or movement with respect to the world. Traditionally, two theories

have been proposed on the kin of information that is used by the brain to compensate for the visual effect of eye rotation. According to 'inferential' theory, dating from late 1900's, our percept is a cognitive evaluation of all sensory impressions. These sensory impressions include signals from non-retinal origin called *extra-retinal signals*. The contribution of *efference copy*, commands to the eye muscles, to our perception of space is usually demonstrated by pushing the eye aside with the finger which, unlike active eye rotation, does cause a shift of the perceived world. Gibson (1966), on the other hand, attributed no role to extra-retinal signals. With his 'direct perception' theory, he introduced the notion that people pick up so called *invariants* of their environment contained in the visual image, such as the radiant in optic flow which conveys the direction of heading. Neither view on the role of extra-retinal signals has survived the last decade of empirical research (see Wertheim 1994 for a review). The issue appears to be not whether, but how the extra-retinal and visual streams of information combine in the human perception of space and self-motion. In this thesis, we analysed the role of visual and extra-retinal information on eye rotation, specifically in the perception of heading from optic flow.

Optic flow has been well studied over the last decades by modellers, psychophysicists and electrophysiologists. Profiting from the detailed mathematical analyses of optic flow and its invariant properties, psychophysicists tested various proposed relations between human heading perception and optic flow (for review see Warren 1995). This led to more functional models. Electrophysiological studies in search of brain modules that pick up optic flow invariants shed new light on these functional models, which eventually resulted in models that were more physiologically oriented. Although the above suggests a linear development, it is in fact a cycle of empirical and modelling efforts. This continuous cycle of formulating and validating new hypotheses is crucial to a better understanding of how the brain works. We have seeked to follow that approach. In the following, I will explain the general problem. Then, I will provide some physiological background. Finally, I will give an overview of the chapters on the model and psychophysical experiments.

## 1.1 The rotation problem

The possible occurrence of an eye rotation during observer translation introduces an uncertainty in retrieving the heading direction from the motion pattern. This '*rotation problem*' is made intuitive by snapshots of the flow on the retina of a translating observer. In a situation without rotation of the eye, e.g. with an observer fixating a point on the horizon, the flow emanates from the direction of heading (Fig. 1.1a). Historically, the term optic flow has become associated with radial flow only. The term *retinal flow* is more general as it also includes the flow that occurs during eye rotation. The retinal flow may differ from the optic flow. This is demonstrated in Fig. 1.1b for the same observer, who now pursues a nearby road sign (Fig. 1.1b). If the brain does not *compensate* for the visual effect of eye rotation, the observer would

wrongly perceive the direction of heading more towards the right.

To investigate how humans might solve the rotation problem, it has been instructive to analyse the available information content. Retinal flow is a potential source of information on the 3D structure, orientation and movement of objects relative to each other and the observer's eye. Several features in the retinal flow allow distinguishing its rotational component (Fig. 1.1c) from its translational component (Fig 1.1a) (Koenderink 1986). One such feature is so called *motion parallax*, the difference in retinal velocity of points lying at different distance. Motion parallax is caused only by translation, not by rotation. A more global feature that can be used to visual compensation is the pattern of the flow field, which is different for eye rotation and eye translation. But, to compensate for the visual effect of eye rotation also extra-retinal signals on the eye's rotation velocity and axis are available (e.g. efference copy, vestibular or neck muscle signals).

The rotation problem in optic flow received interest from basic sciences, but also from applied sciences such as robotics or from car and aircraft industry. For references to earlier models, I refer to a review on physiological models by Lappe (1998). One issue raised by earlier models is the extent to which solving the rotation problem is done separately from solving the heading direction. Longuet-Higgins and Prazdny (1980), for instance, proposed a two-stage model. At a first stage, eye rotation is compensated to retrieve a representation of the translational flow. At a second stage heading is retrieved from the translational flow. Others proposed simultaneous solutions (e.g. Heeger & Jepson 1992). Another issue is how visual information is used. To remove the rotation flow component Longuet-Higgins and Prazdny (1980) and others (e.g. Rieger & Lawton 1985; Hildreth 1992) suggested to subtract locally adjacent retinal motion vectors, since the resulting 'differential motion parallax' field yields the translational component of retinal flow. Another two-stage scheme suggested is that the rotational flow is removed by subtracting an estimate of the rotational flow based on the most distant features in the scene (van den Berg & Brenner 1994a). A third issue is to what extent visual and extra-retinal signals play a role.

In an approach initiated by Warren and Hannon (1988), psychophysicists analysed the role of the extra-retinal eye velocity signal using the simulated and real rotation paradigm. In these experiments, subjects are to make a real eye movement while viewing flow on a screen as in Fig. 1.1a (real rotation), or to fixate a stationary point while viewing flow on a screen as in Fig. 1.1b (simulated eye rotation). While the retinal flow is the same in both conditions, the extra-retinal signal is informative on the rotational component of flow only during real eye movement. Various laboratories reported that in the simulated eye rotation condition humans quite accurately estimate their heading, although the accuracy depended on layout, depth in the scene and field of view (e.g. Warren & Hannon 1988; van den Berg 1992, see van den Berg 1999 for a review). Others repeatedly reported large errors in perceived heading (Royden et al. 1992, 1994; Banks et al. 1996). Results for real eye rotation condition are more clear, as perceived heading is largely unaffected during real eye rotation for all types

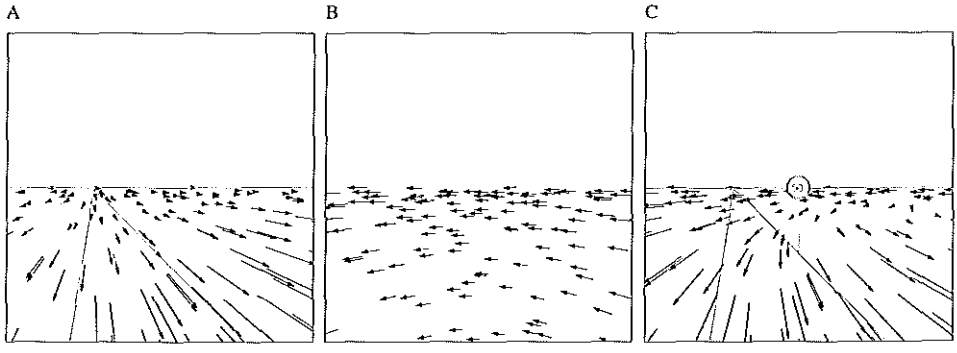



**Figure 1.1** *Flow generated on the retina during (a) horizontal observer translation along a groundplane (1 m/s), (b) observer rotation about the eye's vertical axis (4 °/s), or (c) the combination of both. As a visual aid, road edges, a horizon and road sign have been drawn (grey lines).*

of scene, which shows that extra-retinal signals play an important role as well. The challenging question that is dealt with in this thesis is how the rotation problem is solved on purely visual basis, while allowing for compensation by extra-retinal signals as well.

## 1.2 Physiology

Electrophysiological studies of monkey brain, and more recently neuro-imaging studies like PET (Positron Emission Tomography) and fMRI (functional Magnetic Resonance Imaging) have greatly enhanced our understanding of compensation mechanisms involved in self-motion perception. Electrophysiologists can record the electrical activity at the level of single neurons by inserting electrodes into the brain. Neuro-imaging studies reveal the mass activity in 3D, although at a much coarser temporal and spatial resolution.

Visual processing in the brain takes place massively parallel channels, showing a high degree of functional and hierarchical organisation. Light falling onto the retina is sensed by millions of small rods and cones. Visual information from each eye's left half of the visual field is passed on to the right half of the brain, and vice versa. Via the LGN (lateral geniculate nuclei) visual information reaches primary visual cortex (V1) at the back of the head. The visual cortex takes up a large portion of the cerebral cortex. Two major cortical pathways have been distinguished that are believed to process visual information differently (Fig. 1.2, right). A more ventral stream leading to the temporal areas is believed to be involved in recognising colour and form. A more dorsal stream leading to the parietal areas is thought to be involved in the perception

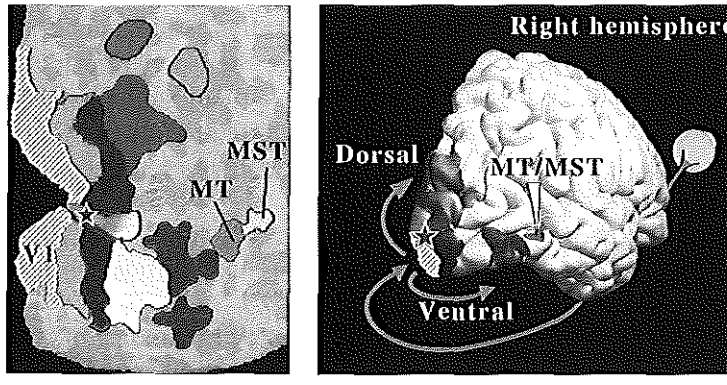



**Figure 1.2** *Human cerebral cortex overlaid with different gray levels indicating visual areas identified with fMRI. Left panel shows the cortex of the brain when flattened out. It also shows area V5a, identified as the human homologue to monkey brain MST. Adapted from Tootell et al. (1996)*

of motion and spatial relations.

From primary visual cortex (V1), the motion pathway continues to run via the middle temporal area (MT, V5 in humans) to the middle superior temporal area (MST, V5a in human) (Fig. 1.2 left). Area MST also has connections with other parietal areas, such as area 7a and the ventral intra-parietal area (VIP). Various cells in the monkey brain show sensitivity to retinal motion, including the sensors on the retina. But, it is only at brain area MT, that cells become selective for motion with a specific speed and direction. Neurons in area MT are sensitive to flow only within a few degree of their preferred retinal position. At the dorsal side of area MST (MSTd) cells are selective to the pattern of flow. Neurons in area MST are sensitive to flow covering more than a quadrant of the visual field. This makes MST suitable to process large patterns of flow such as occur during ego-motion relative to the world. Moreover, they are broadly tuned to retinal location of the expanding flow (Duffy & Wurtz 1995). Direct evidence that area MST is involved in the perception of heading has been provided recently by Britten and Van Wezel (1998). They showed that electrical stimulation of MST neurons caused systematic shifts of the response of monkeys that were trained to indicate their perceived heading. Sensitivity to flow, but not the involvement in heading tasks, was shown for other cortical areas as well, e.g. area 7a (Collum & Siegel 1996) and VIP (Colby et al. 1993). Area 7a is thought to be involved in position transformations, while area VIP is thought to be encoding nearby space.

Functional organisation is also observed within the cortical areas. Each layer in V1, for instance, is divided into columns oriented perpendicular to the cortical layer. The position of each column corresponds to the preferred retinal stimulus position of

cells, while preferred orientation or spatial frequency change systematically within the column. A retino-topic map with columnar organisation is also found in MT, where cells within the same column systematically change their preferred motion direction (Albright et al. 1984). MST is less clearly organised, although functional clustering has been reported into columns that share the same preferred retinal flow pattern, such as expansion or rotation or a combination thereof (Geesaman et al. 1997; Britten 1998). Generally, it seems that along the visual pathway, visual properties are represented by a multidimensional map of cells or activities, each cell or activity having its own specific 'line label', such as preferred retinal position, orientation, or in case of MST neurons the preferred flow pattern.

A potential location for compensation is revealed by the finding of areas that show extra-retinal activity related to eye velocity. In the lateral part of MST 'pursuit' or 'visual tracking' neurons have been found (Newsome et al. 1988; Wurtz et al. 1990). These cells seem to have a preferred axis of rotation, and fire roughly proportional to eye velocity (Thier & Erickson 1992). More recent data also showed pursuit-related activity in the dorsal part of MST (Bradley et al. 1996; Page & Duffy 1999). Pursuit-related signals have also been found in area 7a (Bremmer & Hoffmann 1993), but not in areas hierarchically situated before area MST. Neither area MT (Newsome et al. 1988), nor V1 (Ilg & Thier 1996) shows maintained activity during pursuit when the retinal stimulus is blanked out. Visual compensation might take place in MST, but also MT seems a suitable candidate. MT cells have center-surround receptive fields for motion that respond to a given motion speed and direction in the center of its receptive field, and are inhibited by motion in their surround (Allman et al. 1985). This property might be suitable for exploiting differential motion parallax (Rieger & Lawton 1985).

The findings of eye velocity signals at the level of MST encouraged us to look for a model that could take extra-retinal compensation into account at a level at which neurons are already tuned to a pattern of flow. An important observation in this respect is that extra-retinal signals are 'rate-coded' since they fire linearly with increasing rotation rate, while visual signals are 'labelled-line' coded. Because the visual and extra-retinal signals are of a different format, merely subtracting extra-retinal activity to compensate for the visual effect of eye rotation would not be adequate. The general solution that the brain might employ to compensate for visual change in the flow during eye rotation is to gain-modulate the visual signal by the extra-retinal signal. Gain modulation of labelled line encoded activity by a change in the eye position signal have been found abundantly throughout the brain. Other areas in the parietal cortex associated with compensation for changes in eye position, such as visual position maps (area 7a), auditory maps (lateral intra-parietal area, LIP) show modulation by eye position signals (see Andersen et al. 1997 for review). This suggest a common principle underlying the integration of extra-retinal signals with sensory information from another modality.

## 1.3 Modelling

The findings of a hierarchical relation between MT and MST neurons and the nature of their response properties have led to three types of physiological models. The *template* model by Perrone & Stone (1994) consists of a set of MST-like flow templates, each having its preferred specific self-motion (heading direction and rotation). The template activity reflects how well the template's preferred flow matches the retinal flow, so that the set of template activities represents the heading direction. The *population* model by Lappe and Rauschecker (1993, 1995), based on an algorithm by Heeger and Jepson (1992), assumes more intermediate layers before cells encode the heading direction. A recent model by Royden (1997) exploits differential motion parallax as proposed by Longuet-Higgins & Pradzny (1980) to visually compensate for eye rotation at the level of MT. It is based on MT cells that have center-surround receptive fields for motion, i.e. they respond to a given motion speed and direction in the center of its receptive field and are inhibited by motion in their surround (Allman et al. 1985). All three models are able to visually compensate for eye rotation. Until very recently, however, none of these physiological models was able to account for psychophysical evidence that extra-retinal signals are used.

**Chapter II** decribes a physiological model that does take extra-retinal signals into account. The aim was to make a physiologically plausible model that was compatible with existing psychophysical data and could predict particular response properties of cells in MST. The starting point was the description of a general principle that uses the above mentioned gain-modulation. The principle was demonstrated in a model for compensating the change in visual perceived position when the eye position is changed. Then, the principle was applied in an adapted version of the template model (Perrone & Stone 1994) as to compensate for the change in retinal flow during eye rotation.

## 1.4 Psychophysics

This model, called the '*velocity gain field model*', formed inspiration for the three psychophysical experiments explained subsequently. When the extra-retinal signal is explicitly used, the velocity gain model has a number of features that distinguishes it from alternative models that use extra-retinal signals. Royden et al. (1994) suggested that the extra-retinal signal is used to subtract rotational flow from the retinal flow, on the assumption that the rotational flow can be determined accurately from the extra-retinal signal, after which heading can be derived from the remaining translational flow component. This '*vector subtraction scheme*' clearly is not meant to describe the ability of humans, as others reported, to visually compensate for rotation in simulated rotation experiments. We wondered whether that model does account for perceived heading during real eye rotation. Whereas the vector subtraction requires no visual evidence for eye rotation to use the extra-retinal signal, our gain-modulation model

does require such visual evidence. **Chapter III** presents an experiment that tests these opposite assumptions.

A more physiological model that incorporated the use of extra-retinal signals was most recently proposed by Lappe (1998), who adapted the population model (Lappe & Rauschecker 1995) to this end. Essentially, his scheme also subtracts the rotational flow from the retinal flow by using the extra-retinal signal. In contrast to the vector subtraction scheme (Royden et al. 1994), such a population model does allow a visual compensation for the remaining rotational flow. Although this seems more compatible with human's ability to visually compensate for rotation, Lappe (1998) as well as Royden et al. (1994) assume that the extra-retinal signal is subtracted at a level of local motion signals. This means that the extra-retinal signal is compensated before a global estimate of visual rotation is available. Our gain-modulation model, in contrast, assumes that the interaction with the extra-retinal signal occurs after a global estimate of the rotation has become available. **Chapter IV** presents an experiment that tests these opposite assumptions.

Our model simulations indicate a possible effect of rotation about the line of sight on the compensation for rotation about the vertical axis. We refrained from further modelling this, as it would require additional assumptions on template organisation and template tuning properties. The model result did inspire us though to investigate to what extent findings for simulated rotation about the vertical axis apply to simulated rotation about the line of sight (**Chapter V**).

# Chapter 2

# Heading detection using motion templates and eye velocity gain fields


## Abstract

Eye or head rotation would influence perceived heading direction if it were coded by cells tuned only to retinal flow patterns that correspond to linear self-movement. We propose a model for heading detection based on motion templates that are also Gaussian-tuned to the amount of rotational flow. Such retinal flow templates allow explicit use of extra-retinal signals to create templates tuned to head-centric flow as seen by the stationary eye. Our model predicts an intermediate layer of 'eye velocity gain fields' in which 'rate-coded' eye velocity is multiplied with responses of templates sensitive to specific retinal flow patterns. By combination of the activities of one retinal flow template and many units with an eye velocity gain field, a new type of unit appears: its preferred retinal flow changes dynamically in accordance with the eye rotation velocity. This units activity becomes thereby approximately invariant to the amount of eye rotation. The units with eye velocity gain fields form the motion-analogue of the units with eye position gain fields found in area 7a, which according to our general approach, are needed to transform position from retino-centric to head-centric coordinates. The rotation tuned templates can also provide rate-coded visual estimates of eye rotation to allow a pure visual compensation for rotational flow. Our model is consistent with psychophysical data that indicate a role for extra-retinal as well as visual rotation signals in the correct perception of heading.



Adapted from: Beintema and van den Berg (1998), Vision Research, 38(14):2155-2179.

## 2.1 Introduction

A moving observer can perceive his direction of heading from the flow of the visual environment. The movement of points relative to a moving vantage point is called the optic flow. Because the appearing point, from which all the optic flow lines originate, coincides with the heading direction, Gibson (1966) proposed that it be could used by the human visual system to infer the direction of heading. The retinal flow, however, is defined as the flow of points relative to a possibly rotating eye, and thus contains a part due to ego-translation, and a part due to ego-rotation. When eye movements are made, the focus of outflow due to the linear movement is displaced by a rotational component. This raises the question whether humans can disregard the rotational component purely visually, or whether extra-retinal information, such as proprioceptive or efference copy signals, need to be used.

When simulated ego-translation and eye rotation is presented to a stationary eye, humans can resolve their heading direction accurately, suggesting that the visual system does not always rely on extra-retinal signals (Warren & Hannon 1988, 1990; van den Berg 1992, 1993, 1994a, 1994b). However, in conditions such as movement towards a fronto-parallel plane or at low signal to noise ratios, performance is better when real eye movements are made. Royden et al. (1992, 1994) found that significant heading errors occur for simulated rotation velocities higher than 1.5 °/s, but not when real eye movements are made. Although the latter limit is disputed, these data suggest at least that both visual and extra-retinal signals contribute to the perception of heading.

Early computational studies solve the heading direction from the optic flow by using differential motion parallax to remove the rotational component of flow (Longuet-Higgins & Prazdny 1980; Rieger & Lawton 1985), by observing the maximum of divergence in the flow (Koenderink & van Doorn 1981) or by an optimization method to find the best matching ego-motion parameters to the actual flow field (Heeger & Jepson 1992). Only recently, physiological data are taken into account (Perrone 1992; Perrone & Stone 1994; Lappe & Rauschecker 1993, 1995). These physiological models have been inspired by the discovery of motion sensitive cells in MST of the macaque. The dorsal part of MST (MSTd) contains cells that respond to large expanding/contracting, rotating or shifting flow patterns (Tanaka & Saito, 1989; Tanaka et al. 1989; Duffy & Wurtz 1991; Orban et al. 1992). In later studies, it was found that a large proportion of these cells also respond to combinations of flow patterns, such as spiral motion (Graziano et al. 1994), and that these cells may be tuned to a continuum of flow patterns, rather than to one particular flow pattern. Perrone (1992) used local flow detectors, modelled after neurons in area MT, as the input to a set of motion templates with properties similar to those of cells in MST. The model of Lappe and Rauschecker (1993) consists of a neural network, based on the subspace algorithm by Heeger and Jepson (1992), whose output neurons also showed MST-like properties.

Common to these models is that they rely on motion parallax cues to distinguish

a rotational component in the flow from a change of heading direction. Such models fail without depth differences in the scene, in contrast to humans during a smooth pursuit eye movement. We present a new model that uses the eye-velocity information explicitly to compensate for eye rotations. It is based on a theory by Koenderink (1988) that characterizes the image's local geometry by receptive field assemblies. Briefly, such an assembly allows one to compute the image's luminance profile some distance away from the current position by combining suitably the activities in the assembly. One can use such an assembly to 'shift' the activity relative to the image using eye position information. This is equivalent to a single receptive field that dynamically changes its retinal position depending on the eye position; in other words, it transforms visual position from the retinal to the head-centric or another reference frame. Note, eye position refers to the orientation of the eye, as opposed to eye rotation, which refers to the eye's rotation velocity. By analogy to the receptive field assembly in the position domain, we propose an assembly of flow templates that, in combination with eye rotation velocity signals or a visual estimate of eye rotation, can transform the retinal flow pattern into a head-centric flow pattern. We will present a model to compensate for the effects of eye rotation. However, a similar structure can be used to compensate for the effects of head-rotation using vestibular signals. Our paper concentrates on the assets and limitations of the principle of transformation by dynamic receptive fields. Although our models are inspired by physiological findings, we do not claim to provide a detailed account of cell properties.

Unfortunately, different nomenclature has been used in neurophysiological, oculomotor and visual psychophysics literature to indicate the different types of self-motion. Here, we will use the phrase 'translational' flow for any flow pattern that arises from pure translation of the eye through the environment (expansion flow in the neurophysiological literature). The phrase 'rotational' flow in our hands means a flow pattern that corresponds to rotation of the eye about any axis through the eye's center of rotation. This lumps the flow patterns that in the neurophysiological literature have been called 'rotation' and 'translational' flow. Instead, we will use the phrases 'torsional', 'horizontal' or 'vertical' flow when we need to differentiate between different axes of rotation. Furthermore, we will use the phrase 'head-centric flow' to denote the retinal motion pattern received by an eye that is stationary in the head, in contrast to 'retinal flow' for the motion pattern received by an eye that is rotating in the head.

In section II, we will demonstrate the general concept by transforming visual position from a retinal to a head-centric reference frame. We start with a summary of what is known of this problem in the parietal cortex. Then the mathematical expression for the transformation of retinal position to head-centric position will be derived. Receptive fields with gain field properties as found abundantly in the parietal cortex follow naturally from this analysis. A neural implementation and the limitations to such a model will be shown. In section III, we construct flow templates that code heading direction and form the analogy to the luminance receptive fields in the position domain. To adjust their activities for a rotational component in the flow, we

then transform retinal flow templates to head-centric flow templates. In section IV we test the properties of the retinal flow templates. In section V the performance of the heading model is compared with physiological and psychophysical data. The results are discussed in section VI.

## 2.2 Receptive field shifts

### 2.2.1 Gain fields

To represent retinal information in another frame of reference, such as a head-centric or body coordinate system, the visual input needs somehow be combined with motor signals. The posterior parietal cortex has received much attention, because here, many visual and motor pathways converge (for a review, see Andersen 1989). In area 7a of the macaque, where visual and eye position information are integrated, three types of cells have been distinguished (Andersen et al. 1985; Andersen & Zipser 1988). The first type responds only to visual stimuli. The Gaussian-shaped receptive fields of these visual cells are rather large, subtending about 40° visual angle. The second kind of cell is not visually responsive. Instead, its activity varies linearly with the eye's position in the head. Interestingly, also a third class of cells was found, that integrated visual and eye position signals. A shift of fixation, during constant retinotopic stimulation, yielded a modulation of the amplitude of the response, while leaving the retinal receptive fields unchanged. The response seemed to be the product of eye position cell activity with visual position cell activity. Referring to this multiplicative property, Andersen et al. (1985) introduced the term gain fields. The majority of the gain field cells were formed by 'planar' gain fields, i.e. cells whose activity varied linearly with horizontal eye position, vertical eye position, or both.

### 2.2.2 Transformation of retinal to head-centric position

Andersen and Zipser (1988) demonstrated with a neural network, that the three cell types, found in area 7a, form sufficient input to encode the visual position in head-centric coordinates. When the network was trained to generate head-centric responses, using visual and eye position cells as input, cells in an intermediate layer of the neural network revealed gain field like properties. However, the need for planar gain fields, to code the visual position in head-centric coordinates, can be understood differently and more explicitly by reformulating the problem in terms of 'dynamic shifts' on a map of Gaussian receptive fields.

Our model consists of an array of identical cells with retinal receptive fields that ultimately connect to an output layer of cells with head-centric receptive fields (Fig. 2.1a). In the first layer, the cell activity $L(x)$ represents the image contrast at retinal location $x$. We assume such activity is adequately described by the convolution (denoted by an asterisk) of the retinal luminance distribution $I(x)$ with a Gaussian filter $g(x;\sigma)$ centered at retinal position $x$ and having width $\sigma$:

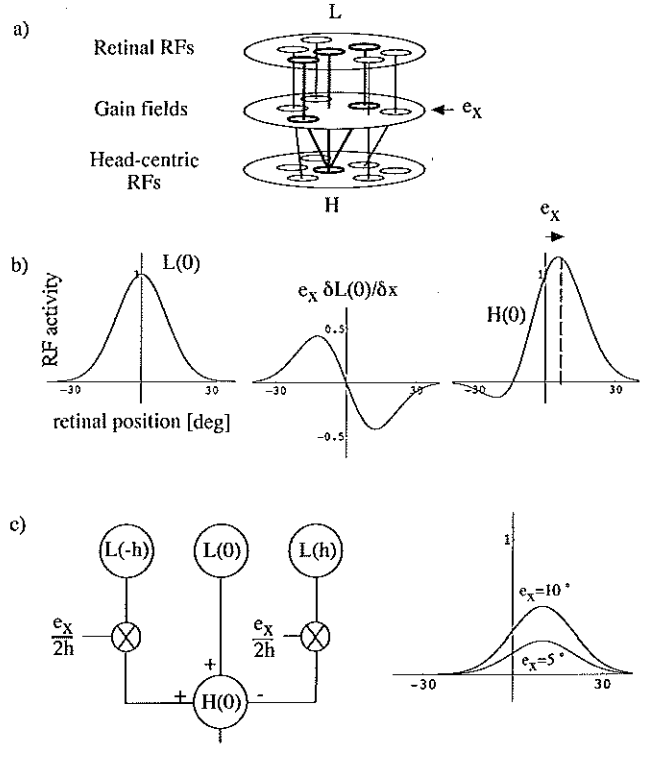



**Figure 2.1** *Scheme for coding horizontal head-centric position. a) Input and output layer of cells with respectively retinal and head-centric receptive fields. The solid lines show the projection of one set of retinal cells via an intermediate layer to one head-centric cell. The intermediate layer contains gain fields; retinal RFs which are multiplied by eye position signal $e_x$. b) Subtraction of the derivative multiplied by $e_x$ from the retinal RF activity. This creates a head-centric RF of which the preferred location is shifted dynamically relative to that of the retinal RF. This shift is somewhat less than 10°, because the magnitude of $e_x$ (-10°) is close to the RF size ($\sigma = 15°$). c) For one head-centric cell ($H(0)$) at location $x = 0°$ the interaction between retinal RFs and eye position signals is shown. Eye position cells (not shown) that provide the signal $e_x$ fire linearly with eye position. Retinal cells $L(-h)$, $L(0)$ and $L(h)$ have broadly tuned Gaussian receptive fields for visual position. The subtraction of the two RFs $L(-h)$ and $L(h)$, each having a fixed oppositely shifted center, yields the derivative $\partial L(0)/\partial x$ when scaled appropriately by a constant factor $h^{-1}/2$. Although this scaling may occur anywhere along the visual pathway, here it is carried out by scaling the eye position $e_x$. Multiplication of the retinal RFs by scaled eye position ex produces gain fields: the height of the retinal tuning curve varies linearly with eye position $e_x$.*

$$L(x) = g(x;\sigma) * I(x), \quad \text{with} \quad g(x;\sigma) = e^{-(x/\sigma)^2}$$

Now, suppose the RF of a cell in the output layer has the same properties, but $x$ now refers to the cell's preferred head-centric location. Then, the response $H(x)$ of a head-centric cell and that of a retino-centric cell are identical only when directed at the same head-centric region of the environment. Denoting the eye position relative to the head by '$e$', this leads to the following set of equations:

$$\begin{cases} \text{eye in the resting position } (e = 0): & H(x) = L(x) \\ \text{eye in an eccentric position } (e \neq 0): & H(x) = L(x - e) \end{cases} \tag{2.1}$$

The last equation suggests that, depending on the eye position, the head-centric cell should connect to a different retino-centric RF. We will show, however, that a dynamic shift of the RF center can be obtained with only a small set of fixed connections. It requires eye position in a format resembling gain fields. To satisfy the above set of equations, we use Taylor's expansion to approximate the activity of a neighbouring receptive field by the sum of the retinal cell activity at location $x$ and its first, second and higher order spatial derivatives (Fig. 2.1b).

$$H(x) = L(x - e) = L(x) - e\frac{\partial L(x)}{\partial x} + \text{higher order terms of e} \tag{2.2}$$

Taking advantage of the Gaussian kernel's property that the derivative of the 'blurred image' equals the 'blurred derivative' of the image, see Koenderink (1988), the first order derivative from Eq. 2.2 becomes:

$$\frac{\partial L(x)}{\partial x} = \frac{\partial g(x;\sigma)}{\partial x} * I(x),$$

so that the head-centric response can be written as:

$$H(x) = L(x - e) \simeq \left\{ g(x;\sigma) - e\frac{\partial g(x;\sigma)}{\partial x} \right\} * I(x) \tag{2.3}$$

Thus, the head-centric response $H(x)$ is derived from purely local convolutions of the retinal image $I(x)$ with $g(x;\sigma)$ and $\partial g(x;\sigma)/\partial x$, taking into account eye position. If the eye is in its resting position ($e = 0$), the contribution of the derivative term vanishes, in which case $H(x)$ and $L(x)$ become equal (Eq. 2.1)

Eq. 2.2 and 2.3 reveal already a multiplication of eye position with visual RF activities. Yet, it remains unclear whether first or higher order derivative cells are present in area 7a. However, the derivative term of Eq. 2.2 can be decomposed into activities of cells that do resemble known gain fields. To this end, we approximate the derivative RF at location '$x$' by the difference of two Gaussian receptive fields,

whose preferred centers are slightly offset in opposite retinal directions relative to the location '$x$' by an arbitrary small constant '$h$':

$$\frac{\partial g(x;\sigma)}{\partial x} \simeq \frac{g(x+h;\sigma) - g(x-h;\sigma)}{2h} + \text{2nd and higher order terms of h} \quad (2.4)$$

Note, this equation arises from two Taylor series of $g(x-h;\sigma)$ and $g(x+h;\sigma)$, and that the accuracy with which the derivative is approximated depends on $h$. Substitution into Eq. 2.3, and a further extension to shifts in two dimensions, now gives:

$$\begin{aligned} H(x,y) \simeq & \left\{ g(x,y;\sigma) - \frac{e_x}{2h} g(x+h,y;\sigma) + \frac{e_x}{2h} g(x-h,y;\sigma) \right. \\ - & \left. \frac{e_y}{2h} g(x,y+h;\sigma) + \frac{e_y}{2h} g(x,y-h;\sigma) \right\} * I(x,y) \end{aligned} \quad (2.5)$$

The above equation holds for all head-centric RFs, each having its own preferred head-centric location $(x,y)$ and receiving input from a different set of retinal RFs with corresponding preferred retinal locations. Each such set is composed of a pure retinal RF and four neighbouring RFs. The latter four are retinally displaced by a constant h and multiplied by either the horizontal ($e_x$) or vertical ($e_y$) eye position. These reflect the multiplicative properties of planar gain fields, because their responses vary linearly with eye position, given the retinal stimulus remains constant.

Fig. 2.1c schematically shows the interactions between neurons involved in the retinal to head-centric transformation, as proposed by Eq. 2.5. For simplicity, we have omitted the vertical dimension. Note, the terms involving the eye position signal can attain both negative and positive values, whereas physiologically, it is not possible for neuron activities to change sign. However, negative values can be avoided by raising the level of activity, followed by inhibition further along the visual pathway. For instance, eye position $e_x$ might be represented by the difference of two eye position signals $(a + e_x/2)$ and $(a - e_x/2)$, that each do remain positive throughout the range of possible eye positions. Following multiplication by the RF activity, the resulting signals should then be subtracted at the level of the head-centric cell.

### 2.2.3 Error analysis

Eq. 2.2 holds only for a limited range of eye positions around the eye's resting position ($|e| < \sigma$), whereas Eq. 2.4 is only justified as long as the RFs of the two neighbouring retinal cells overlap sufficiently ($h < \sigma$). Approaching either limit affects the form of the head-centric cell's tuning to the retinal position of a point stimulus, and causes an incomplete shift of its peak response (Fig. 2.1b). To gain insight into the effect of varying $h$ and $\sigma$, we calculated the error in the shift as a function of eye eccentricity $e$. As expected, we find the error remains less than 1° for small offsets ($h < \sigma/3$)

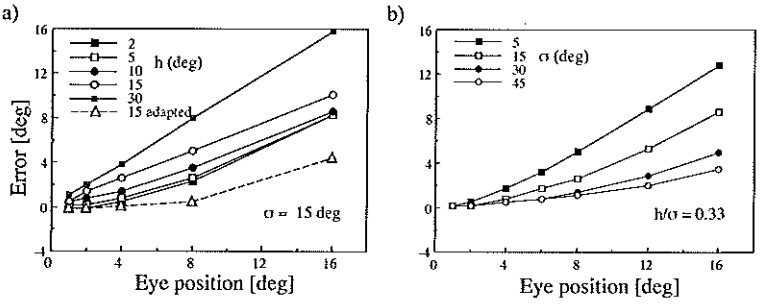



**Figure 2.2** *Shift of the preferred head-centric location of the head-centric template as a function of the eye position. Positive error means that the preferred location of the head-centric detector shifts in the direction of the eye movement. a) Smooth lines represent the error for different values of h at constant RF width ($\sigma = 15°$). Dotted line shows the error for $h = 15°$, when the derivative term is scaled with a factor '$\exp(h^2/\sigma^2)$'. b) Error curves for different Gaussian widths. No scale factor was used to improve the approximation to the derivative.*

and small eccentricities ($e < \sigma/3$) and increases for larger offsets or eccentricities (Fig. 2.2a). For large h, however, one can obtain a better approximation of the first order derivative RF by scaling it with a constant factor: $\exp(h^2/\sigma^2)$. This significantly decreases the error for $h = \sigma$ (Fig. 2.2a, dashed line). Furthermore, Fig. 2.2b shows that the range of eye positions for which the errors remain small is proportional to the RF width $\sigma$, given a fixed ratio of $h$ to $\sigma$.

More accurate approximations of the head-centric response can be obtained using a higher order Taylor expansion. This requires second or higher order derivatives of the Gaussian receptive field together with signals that are proportional to second or higher powers of the eye shift. Any higher order derivative can be approximated by suitably combining receptive fields at neighbouring retinal locations. Furthermore, given a set of eye position cells, that fire linearly with eye position, any non-linear higher power function of eye position can be approximated. Fig. 2.3 demonstrates how a quadratic function can be achieved from the sum of linear units that have different slopes and intercepts. A similar approach was used in the neural network model by Andersen and Zipser (1988). Recently, Squatrito and Maioli (1996) have found eye position cells that indeed show a linear increase of the activity as a function of the eye position, and that start responding at various eye positions.

Higher order differentiations may increase the sensitivity to noise. However, the effect of noise in the image is strongly reduced because the RF integrates the spatial derivative over a large area. Thus, one may expect higher order derivatives to provide meaningful responses only for large receptive fields. Since the higher order derivatives are constructed by means of subtracting RF responses, noise at the level of the cell's output, just prior to the subtraction operation will seriously limit the usefulness of

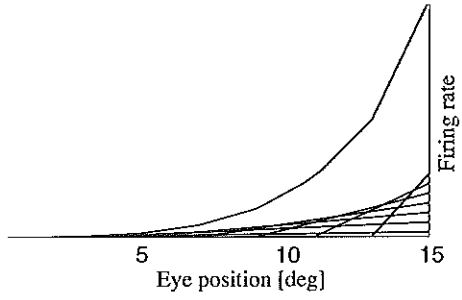



**Figure 2.3** *Shown are the responses of model units that fire in proportion to the eye position with different onset eccentricities. The quadratic curve at the top shows the sum of the linear activities.*

higher order differentiations.

### 2.2.4 Physiology

Physiological evidence for retino-centric to head-centric transformation of visual signals was found in area 7a. This area contains three type of cells (Andersen 1989) of which the response properties look remarkably similar to properties of units in the scheme of Fig. 2.1c. First of all, the Gaussian-tuning profile to retinal position of the visual cells is an essential property to enable a dynamic shift of a RF. Secondly, our model indicates that large receptive fields are necessary to fully compensate for large changes in eye position; the wide RFs as found in area 7a are therefore well fit to sustain the transformation throughout the oculo-motor range. Thirdly, cells with gain fields whose activity varies linearly with eye position (Andersen 1989) may reflect the most elementary components of the intermediate layer. We do not know of any visual cells with first order derivative receptive fields, but these may be constructed from Gaussian receptive fields with linear gain fields. Head-centric receptive fields[1], on the other hand, have not been found in area 7a, but these cells may be found in other areas along neural pathways leading from area 7a, such as area V6 (Galetti et al. 1993). Alternatively, the visual system may very well supply other areas with head-centric visual information in the distributed form of the gain fields.

In the model we have used only extra-retinal signals that code eye positions along the vertical or horizontal meridian. But, position cells in area 7a are tuned to other meridians as well (Andersen et al. 1990; Squatrito & Maioli 1996). The proposed mechanism is easily extended to include eye signals with intermediate axes of rotation. Each eye signal should then combine with a pair of receptive fields, whose shifted

[1] Recently, such head-centric receptive fields were found in area VIP of the monkey (Duhamel et al. 1997)

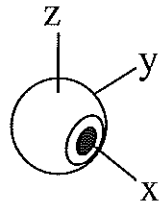



**Figure 2.4** *Orientation of positive retinal coordinate axes relative to the eye.*

centers lie on the meridian through which the preferred axis of the eye position cell passes. Our model does not deal with the special class of torsional eye positions. In that case, compensation could involve a combination of oriented receptive fields (Koenderink 1988) with cells coding rotation about the line of sight.

## 2.3 Transformation of retinal to head-centric flow

This section describes how one can derive a representation of heading direction from the retinal flow, that is invariant to the rotation velocity of the eye. First, we give a mathematical description of the retinal flow. We introduce retinal flow templates, which form the units in the first layer of the heading model, and are the equivalent of retinal position cells in the position model. Then, we show how to arrive at head-centric flow templates. At the end we construct the motion templates that collect motion signals, just as MST cells are believed to collect MT inputs, and that to some extent show similarity with MST responses found. But, for templates to code heading direction and to connect them with the compensation scheme, we must assume additional cell properties that have not yet been found or simply have not been searched for in MST.

### 2.3.1 Retinal flow

We will now show how the retinal flow is mathematically expressed as the sum of a component related to the eye rotation, and a component related to the direction of heading. To describe the visual environment, the ego-motion and the resulting flow field we use an oculo-centric system, centered on the nodal point of the eye. The $x$-axis points forward along the line of sight, while the $y$ and $z$-axis lie perpendicular to the $x$-axis, along the horizontal and vertical retinal meridian respectively (Fig. 2.4).

The retinal flow is characterized by a collection of motion vectors, that each denote the retinal projection of the 3D movement of a fiducial point in the environment. This assumes that the correspondence problem has been solved. This description only partially captures the information in the optic flow, for it does not describe the higher

order temporal derivatives, or local spatial derivatives of the flow. However, several psychophysical studies argue that this motion pattern is sufficient for the perception of heading (Warren et al. 1988, 1990).

Since the visual system measures only visual directions and angular speeds, we shall use spherical coordinates. Thus, a point at a 3D position $\mathbf{d}_i$ is written as lying at a distance $d_i$ in a visual direction denoted by the unit vector $\hat{\mathbf{d}}_i$. As a convention, we use a fat symbol to indicate a 3D vector, and a capped symbol to refer to a unit 3D vector. To each point is associated a flow vector $\mathbf{p}_i$ which always lies in the plane perpendicular to the visual direction. Although $\mathbf{p}_i$ has only two degrees of freedom, we maintain the three-dimensional notation for computational convenience. We can define the flow as the following set:

$$\text{flow} = \{\mathbf{p}_i, \hat{\mathbf{d}}_i\}$$

Any movement of the eye can be decomposed into a 3D translation vector $\mathbf{T}$ and a 3D rotation vector $\mathbf{R}$ whose rotation axis passes through the eye. The direction of a 3D rotation vector lies along the rotation axis, and its length determines the rotation speed. $\mathbf{R}$ can have any direction relative to $\mathbf{T}$ or the eye's viewing direction. Thus, it may be torsional, horizontal or vertical or any combination thereof. If we ignore the distance between the eye's nodal point and its center of rotation, the displacement $\Delta\mathbf{d}_i$ of a point in an infinitesimal time $\Delta t$ can be approximated by:

$$\Delta\mathbf{d}_i = -(\mathbf{T} + \mathbf{R} \times \mathbf{d}_i)\Delta t$$

While only movement perpendicular to the visual direction gives rise to the retinal flow of a point, we take the time derivative of the visual direction to find the motion parallax $\mathbf{p}_i$. This flow vector can be split into a part related to the pure ego-rotation $\mathbf{R}$, and a part related to the 'apparent rotation' $\mathbf{A}_i$ caused by ego-translation (Koenderink & van Doorn 1987):

$$\mathbf{p}_i = \frac{\partial\hat{\mathbf{d}}_i}{\partial t} = -(\mathbf{T} - (\mathbf{T}\cdot\hat{\mathbf{d}}_i)\hat{\mathbf{d}}_i)/d_i - \mathbf{R}\times\hat{\mathbf{d}}_i = -(\mathbf{A}_i + \mathbf{R})\times\hat{\mathbf{d}}_i, \qquad (2.6)$$

with

$$\mathbf{A}_i = \frac{T}{d_i}(\hat{\mathbf{d}}_i \times \hat{\mathbf{T}})$$

Note, that in the last equation the translation $\mathbf{T}$ has been replaced by the product of translation speed $T$ and the unit vector in the heading direction $\hat{\mathbf{T}} = \mathbf{T}/T$. Eq. 2.6 allows one to solve the 2D heading direction and the 3D ego-rotation, given a minimum of five flow vectors and visual directions. Also the reciprocal of 'time to contact' $T/d_i$ can be reconstructed, but the ego-speed (or depth of points) can not be determined without knowing the actual distance of a point (or speed). With this general description of the flow we can now derive the model's general structure.

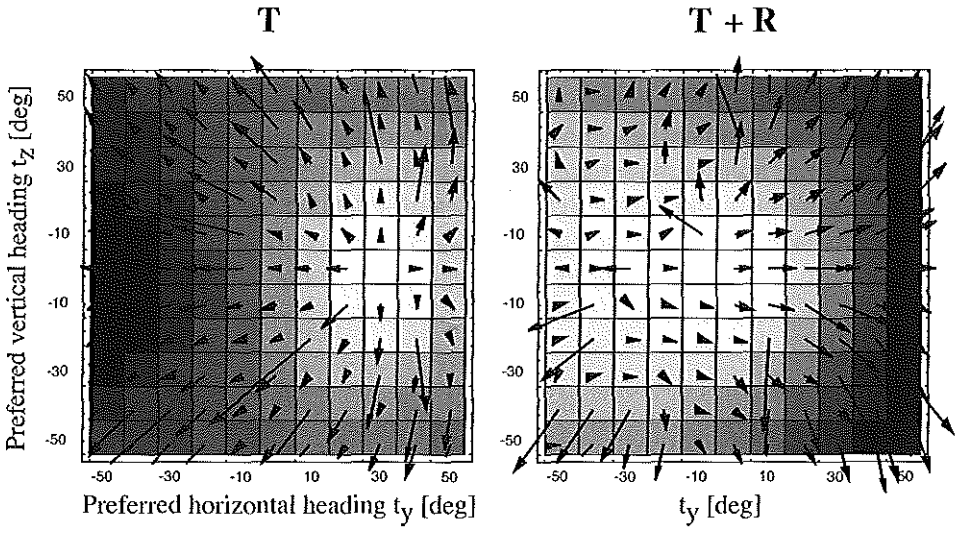



**Figure 2.5** *Response maps of detectors tuned to pure translation in the direction ($t_y$, $t_z$). Shown are the responses for flow belonging to pure ego-translation towards the right (left panel) and for flow belonging to the same ego-translation combined with ego-rotation towards the left (right panel). The flow corresponding to motion through a cloud of dots is superimposed on the map of activities. The location of the maximum response and the center of motion shifts towards the left during ego-rotation, although the direction of self-motion is still 30° towards the right.*

### 2.3.2 Retinal flow templates

Perrone (1992) was the first to propose a template-based model for the detection of heading direction. Basically, each template evaluates the evidence that the flow pattern is consistent with the template's preferred flow field. The template receives its input from a set of local motion detectors at a range of visual directions, much in the way an MST cell is believed to receive its input from motion detectors in MT.

Given a set of templates, each tuned to flow belonging to pure translation of the eye in a specific direction, heading direction can be encoded by the preferred heading direction of the most active template. This can be visualized by a 2D map of responses, in which cell position corresponds to the template's preferred heading direction (Fig. 2.5). Each of these cells prefers retinal flow that radiates outward from the preferred heading direction. An eye rotation, however, adds a rotational component to the flow, causing the apparent focus of expansion to shift. Now, the preferred heading direction of the most active cell and the direction of ego-motion do not correspond. This problem can be overcome by using templates tuned to a combination of translation and rotation.

Sampling the flow for all possible combination of 2D heading directions ($\hat{\mathbf{T}}$) and 3D ego-rotations ($\mathbf{R}$) requires a 5D set of templates, provided the templates' responses do not depend on the translation speed and the distances of points in the scene. Our approach is to create a large set of retinal flow templates, tuned to different heading directions on a fine grid, and to sample the rotation dimensions very sparsely. Subsequently, we apply the theory of the previous paragraphs to build a collection of templates, of which the output is not affected by ego-rotation. This effectively reduces the number of templates needed to sample flow space.

### 2.3.3 The analogy between transformation of retinal position and retinal flow

There are a number of similarities between the problem of disregarding the rotational flow during an eye movement and the problem of compensation for eye shifts to achieve head-centric receptive fields.

First, both models involve a 2D map of cell responses. In the position model, a visual direction of a point is indicated by a peak response in a 2D map of receptive fields. Likewise, the visual direction of ego-translation is indicated by a peak in a 2D map of motion templates tuned to heading direction. For neither the position nor the ego-translation vector do we need to represent the third dimension, because the depth information is lost.

Secondly, both models carry out a coordinate transformation to compensate for eye rotations. To attain a head-centric representation of target position that is invariant for shifts of fixation, the retinal receptive fields are converted to head-centric receptive fields using rate-coded eye position signals. In the retinal flow, a pursuit eye movement changes the responses of retinal templates that are tuned to the center of expanding flow, because a rotational component is added to the translational flow. The head-centric flow templates, however, should be invariant to that rotation component. To compensate for the rotational flow, we define an operation on the retinal flow template that accomplishes a 'shift' in flow space along the direction of ego-rotation, using eye velocity signals.

Thirdly, some characteristic properties of cells in area 7a that support our hypothesis on a proposed transformation mechanism can also be found for cells in MST. The position scheme requires visual position cells with large receptive fields for retinal 'shifts' due to the change of fixation, together with eye position signals in rate code. In MST, spiral motion cells (Graziano et al. 1994) are selective to heading direction and Gaussian-tuned to torsional rotation. Also, pursuit cells that code eye velocity in rate code have been found in MST (Thier & Erickson 1992; Bremmer & Hoffmann 1993).

It is important to note that the above transformation aims to discount the effect of the eye rotation on the templates' activities. The transformed activities do not represent head-centric heading direction, because each template is defined relative to the retinal frame. If the eye is not in its resting position, we still need to carry out a

position transformation on the retinal heading direction to get the head-centric heading direction. The latter transformation can be accomplished by applying the position model on the transformed motion template activities using eye position signals.

### 2.3.4 Transformation of retinal flow

We will now explain the transformation from a retinal to a head-centric representation of the flow. Whenever we refer to the head-centric heading direction, we mean the heading direction that is corrected for the eye velocity, but not for the eye's deviation from the null position. All coordinates are expressed in the retinal frame, e.g. with the $x$-components $T_x$ and $R_x$ along the line of sight. In the sequel, we will use capital letters $\hat{\mathbf{T}}$ and $\mathbf{R}$ to denote the heading direction and the ego-rotation of the observer, and lowercase letters $\hat{\mathbf{t}}$ and $\mathbf{r}$ for the template's preferred heading direction and ego-rotation.

As in the position model, our aim is to construct two representations of heading maps (Fig. 2.6). A heading map is a 2D set of templates, with each grid position referring to a preferred retinal heading direction $\hat{\mathbf{t}}$. The first layer '$O$', consists of the templates tuned to retinal flow. The preferred flow at one grid position corresponds to pure ego-translation ($O_0$ map in Fig. 2.6) or to a combination of ego-translation and a rotational component of the flow, denoted by $\mathbf{r}$ ($O_z$ and $O_{-z}$ maps in Fig. 2.6). Put differently, we use for each heading direction and axis of rotation, a pair of templates with opposite directions of preferred rotation about this axis but equal preferred magnitudes of rotation, and one template tuned to zero rotation about this axis. We call such a triple of templates a 'template bundle'.

To denote the response of such templates to a particular flow, characterized by $\hat{\mathbf{T}}$ and $\mathbf{R}$, we could use a notation like $O_{\hat{\mathbf{t}},\mathbf{r}}(\hat{\mathbf{T}},\mathbf{R})$. However, because we wish to express relations between templates within the bundle when they respond to one instance of the flow, we drop the $\hat{\mathbf{T}}$ and $\mathbf{R}$ parameters and only refer to the template activities by the preferred heading direction $\hat{\mathbf{t}}$ and rotation $\mathbf{r}$: '$O(\hat{\mathbf{t}},\mathbf{r})$'.

A cell in layer '$O$' codes oculo-centric heading direction and responds purely to the retinal flow, while a cell in layer '$H$', at the same grid position, codes the same retinal heading direction, but with its activity corrected for the eye velocity. That is, it responds to the head-centric flow, by combining the activity of $O$-templates with eye velocity signals. As in the position model, we concentrate initially on eye rotation $\epsilon$ about the $z$-axis (horizontal rotation). We require:

$$\begin{cases} \text{stationary eye: } (\epsilon = 0): & H(\hat{\mathbf{t}}) = O(\hat{\mathbf{t}}, 0) \\ \text{rotating eye: } (\epsilon \neq 0): & H(\hat{\mathbf{t}}) = O(\hat{\mathbf{t}}, \epsilon) \end{cases}$$

Thus, when the eye is rotating ($\epsilon \neq 0$) the head-centric flow template $H(\hat{\mathbf{t}})$ should respond identically as a retinal flow template $O(\hat{\mathbf{t}}, \epsilon)$ with the same preferred heading

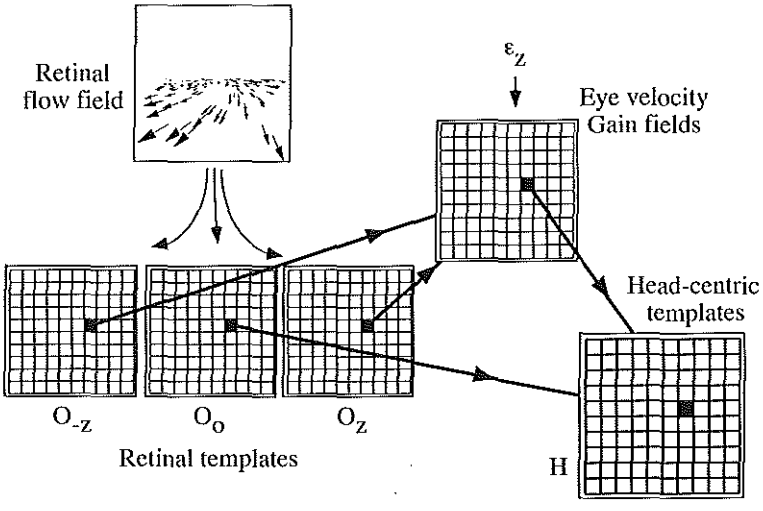



**Figure 2.6** *Construction of head-centric heading map 'H' from three retinal heading maps. Within a map, each unit represents a template with a preferred heading direction that corresponds to its location in the map. The retinal heading map '$O_0$' consists of pure expansion templates (preferred rotation $r_z = 0$ °/s), while the other two retinal heading maps are tuned to a combination of ego-translation and either positive ($O_z$) or negative ($O_{-z}$) rotation about the z-axis. To arrive at a 'H' template of which the activity remains invariant under rotation, the activities of the corresponding retinal template bundle are combined with an eye velocity signal that codes eye rotation about the same axis as to which the retinal flow templates are tuned. Top: retinal flow field input.*

direction $\hat{\mathbf{t}}$ and a preferred rotation about the $z$-axis that corresponds to the eye rotation $\epsilon$.

By analogy to the two approximations in the retinal to head-centric position transformation, we express the head-centric flow response in terms of retinal flow responses and eye rotation velocities. The change of activity of a retinal flow template due to an eye rotation is equal to the derivative of the activity $O(\hat{\mathbf{t}}, 0)$, multiplied by the eye rotation velocity $\epsilon$. Thus, for the head-centric flow template's response to flow corresponding with a particular heading direction and eye velocity $\epsilon$, we get:

$$H(\hat{\mathbf{t}}) = O(\hat{\mathbf{t}}, \epsilon) \simeq O(\hat{\mathbf{t}}, 0) - \epsilon \frac{\partial O(\hat{\mathbf{t}}, 0)}{\partial R} \tag{2.7}$$

Note the similarity with Eq. 2.2. As in Eq. 2.4, the derivative part can be approximated by two retinal flow templates that have opposite amounts of preferred rotation rate ($\omega$):

$$\frac{\partial O(\hat{\mathbf{t}},0)}{\partial R} = -\frac{\partial O(\hat{\mathbf{t}},0)}{\partial r} \simeq \frac{1}{2\omega}[O(\hat{\mathbf{t}},-\omega) - O(\hat{\mathbf{t}},\omega)] \tag{2.8}$$

Now, Eq. 2.7 can be rewritten as

$$H(\hat{\mathbf{t}}) \simeq O(\hat{\mathbf{t}},0) - \frac{\epsilon}{2\omega}[O(\hat{\mathbf{t}},-\omega) - O(\hat{\mathbf{t}},\omega)] \tag{2.9}$$

Clearly, we have the same construction as in the position model (Eq. 2.4). The eye-velocity signal $\epsilon$ provides the gain term, and is multiplied by the response of the templates tuned to the heading direction $\hat{\mathbf{t}}$ and to rotation about the $z$-axis. Thus, we merely need a bundle of three oculo-centric templates with fixed preferred rotational velocities ($-\omega$, 0 and $\omega$) and an eye velocity signal $\epsilon$ to create a template $H(\hat{\mathbf{t}})$ whose response will not depend on the rotation velocity of the eye.

There are several ways to extend the scheme to different directions of rotation. One method involves a decomposition of eye velocity in the three cardinal directions ($\epsilon_x$, $\epsilon_y$, $\epsilon_z$). To simplify the notations, we define '$O_0$' as the response of a pure translation template. In addition, we need six more templates, tuned to rotation velocity of constant magnitude $\omega$, only varying in their preferred axis and sign of rotation, denoted by a subscript. For example, '$O_{-y}$' indicates the response of the template with preferred rotation about the negative $y$-axis, and the same preferred heading direction as '$O_0$'. The head-centric flow response then becomes:

$$\begin{aligned} H(\hat{\mathbf{t}}) = O_0(\hat{\mathbf{t}}) - \frac{\epsilon_x}{2\omega}[O_{-x}(\hat{\mathbf{t}}) - O_x(\hat{\mathbf{t}})] - \frac{\epsilon_y}{2\omega}[O_{-y}(\hat{\mathbf{t}}) - O_y(\hat{\mathbf{t}})] \\ - \frac{\epsilon_z}{2\omega}[O_{-z}(\hat{\mathbf{t}}) - O_z(\hat{\mathbf{t}})] \end{aligned} \tag{2.10}$$

An alternative method would involve a larger collection of rotational axes, that are not necessarily perpendicular, each with their corresponding pair of templates and eye velocity signal. In such a scheme, several different '$H$' templates, each corresponding to a particular rotation axis, need to be combined to arrive at a template sensitive to head-centric flow for the general case.

### 2.3.5 Defining templates

Until now, we have avoided a precise definition of the relation between a retinal flow template and its local motion detectors. We will now construct a retinal flow template that prefers a specific combination of $\hat{\mathbf{t}}$ and $\mathbf{r}$.

A retinal flow template '$O$' evaluates the evidence that the local flow vectors are consistent with its preferred ego-motion. It is a geometrical fact, however, that the precise structure of the flow field depends on the structure of the environment in relation to the ego-speed: the translational component of the flow is related to the distance of the fiducial point divided by the ego-speed (Eq. 2.6). Thus, for the

same preferred ego-motion there exists a set of different flow fields corresponding to different environments. To construct templates that are independent of the distances in the environment and the preferred ego-speed, first note that the translational component of the preferred flow lies along radial lines emanating from the template's preferred retinal direction of heading. To become insensitive to the magnitude of the translational component of flow, the template's response should depend only on the component of the local motion perpendicular to these radial lines. For a visual direction $\hat{\mathbf{d}}_i$, we denote this local perpendicular direction by the symbol $\hat{\mathbf{a}}_i$:

$$\hat{\mathbf{a}}_i = \hat{\mathbf{d}}_i \times \hat{\mathbf{t}} / |\hat{\mathbf{d}}_i \times \hat{\mathbf{t}}|$$

In the following, we exploit this constraint to define templates that are not sensitive to changes of depth.

The template is assumed to collect the activities $W_i$ from $N$ different locations in the visual field:

$$O(\hat{\mathbf{t}}, \mathbf{r}) = \sum_i W_i / N$$

At each location '$i$', the retinal flow vector $\mathbf{p}_i$ contributes an amount $W_i$ to the template's activity according to its deviation from the template's preferred velocity vector ($\mathbf{q}_i$) at that location. How large is this contribution? Let $P_i$ and $Q_i$ denote the velocity components along $\hat{\mathbf{a}}_i$ of, respectively, the flow vector $\mathbf{p}_i$ and the template's preferred flow vector $\mathbf{q}_i$ at that location ($P_i = \mathbf{p}_i \cdot \hat{\mathbf{a}}_i$; $Q_i = \mathbf{q}_i \cdot \hat{\mathbf{a}}_i$). Then, we can define the weight function $W_i$ as a 'receptive field' in the local velocity space (Fig. 2.7). The local velocity space is the collection of all possible retinal flow vectors at that location. We define the receptive field in local velocity space as:

$$W_i = e^{-|P_i - Q_i|^2 / \sigma^2} \tag{2.11}$$

Eq. 2.11 corresponds to a ridge in the local velocity space at viewing direction $\hat{\mathbf{d}}_i$. The ridge is oriented parallel to the preferred translational flow, i.e. motion along the great circle connecting the preferred heading direction and the viewing direction. When the template is tuned to pure translational flow, the maximum of the ridge is centered on that line. When the template is also tuned to a component of rotation, the ridge is offset in velocity space from that line by an amount ($Q_i$) that depends on the eye rotation to which the template is tuned. To this end, we calculate the preferred flow vector $\mathbf{q}_i$ using Eq. 2.6. By definition, the translational component is perpendicular to $\hat{\mathbf{a}}_i$, so this part vanishes in the inner product ($\mathbf{q}_i \cdot \hat{\mathbf{a}}_i$). Therefore, only the rotational component contributes to $Q_i$:

$$\mathbf{q}_i = -(\mathbf{r} \times \hat{\mathbf{d}}_i)$$

The Gaussian width $\sigma_i$ for the local weight in Eq. 2.11 is defined proportional to the offset ($1.5|Q_i|$), with a lower boundary $\sigma_{min}$ that also scales with the preferred

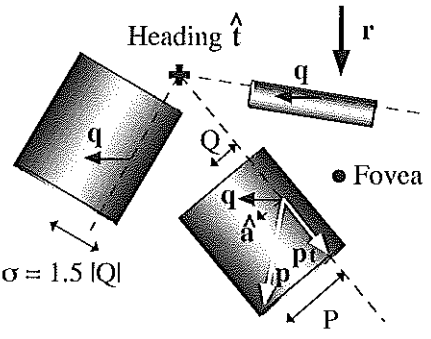



**Figure 2.7** *Construction of a template that is tuned to heading to the left and above of the fovea and an eye rotation to the right (vector* $\mathbf{r}$*). At three retinal locations the local velocity space is depicted. The meridian through the heading direction and the retinal location is one axis of this local velocity space (indicated with a dashed line). The other cardinal direction is perpendicular to that meridian. At each location the component of flow due to the preferred rotation is shown* ($\mathbf{q}$)*. At one location its projection (*$Q$*) on the line perpendicular to the local meridian (i.e. parallel to* $\hat{\mathbf{a}}$*, see text) is shown. At this location, the local flow vector* $\mathbf{p}$ *and its component of translational flow* $\mathbf{p}_t$ *are also shown. The Gaussian ridge in local velocity space represents the weight attributed to the local flow by this template* $O(\hat{\mathbf{t}}, \mathbf{r})$*. The peak of the ridge is offset from the local meridian by* $Q$*. This means that to each local flow vector* $\mathbf{p}$*, of which the component* $P$ *along* $\hat{\mathbf{a}}$ *equals* $Q$*, is attributed the maximal weight for that location. The widths of the ridges determine the width of the rotational tuning of the template, which should be the same for the pure expansion template and the rotation tuned templates. The templates' rotation tuning width* $\sigma$ *is defined by the ensemble of local widths* $\sigma_i$ *of the ridges (see definition templates). Each* $\sigma_i$ *depends on the preferred amount of rotation* $\omega$ *by* $\sigma_i = 1.5|Q_i|$*, and saturates at low values at* $\sigma_{min} = 0.85\,\omega$ *for* $O_0$*, and at* $\sigma_{min} = 0.1\,\omega$ *for* $O_z$ *and* $O_{-z}$*.*

rotation. Note, in the above equations only $\mathbf{p}_i$ is determined by the stimulus, whereas $\mathbf{q}_i$ and $\hat{\mathbf{a}}_i$ are vectors that characterize the sampling of local velocity space by the template that prefers ego-motion parameters $\hat{\mathbf{t}}$ and $\mathbf{r}$.

Templates built according to this definition possess the following properties:

- The template's response is independent of the speed/distance ratio of points in the environment, if the heading $\hat{\mathbf{T}}$ corresponds to the preferred heading $\hat{\mathbf{t}}$.
- The template is Gaussian-tuned to the magnitude of eye rotation about the preferred axis, if $\hat{\mathbf{T}} = \hat{\mathbf{t}}$.
- The template has a maximum response for preferred heading direction $\hat{\mathbf{T}} = \hat{\mathbf{t}}$.

The first property arises, because $W_i$ does not vary along the direction of preferred translational flow in local velocity space but only in the perpendicular direction, i.e. along $\hat{\mathbf{a}}_i$. The local motion component along $\hat{\mathbf{a}}_i$ only depends on the rotational component of the flow, provided the actual heading and the preferred heading correspond. Because that dependence is Gaussian, the template achieves Gaussian-tuning to rotation velocity about the preferred axis of rotation $\mathbf{r}$. Because the local Gaussian width is proportional to the preferred local flow, each flow vector will contribute equally to

the template's response. Finally, as the contribution of the translational component to the flow along $\hat{\mathbf{a}}_i$ increases with the extent to which the heading direction deviates from the preferred heading direction, we have simultaneously acquired the third property, namely tuning to heading direction.

## 2.4 Retinal flow template simulations

In the foregoing, we have derived a mathematical description of retinal flow templates that code the direction of heading, and can be transformed to head-centric flow templates of which the responses are invariant to the eye's rotation. However, to allow such a transformation, we should check whether the retinal flow templates have the desired Gaussian-tuning to the rotation velocity, and analyse to what degree the rotation tuning depends on the orientation of the rotation axis. In addition, we investigate how motion parallax can be exploited to achieve tuning to a preferred rotation, independent of the deviation of the heading eccentricity from the bundle's preferred heading direction.

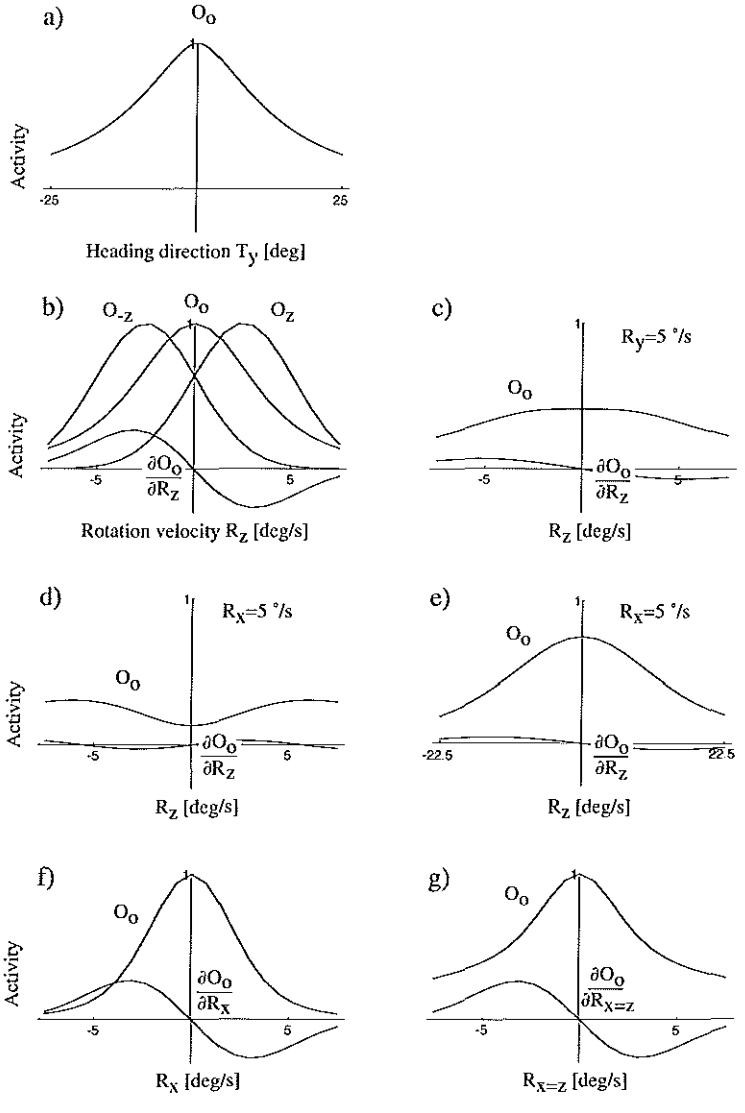
We implemented the model in Mathematica on a Power Macintosh 8500/120. We analyse in detail the response of a single bundle of retinal templates to different types of flow. This template bundle has a preferred heading direction along the retinal $x$-axis. Alternatively, one could analyse the response to a single instance of the flow by an array of template bundles with different preferred head-centric flow. These two types of analysis are equivalent provided (1) the template bundles with different preferred heading do not interact, and (2) the stimulus extends throughout the visual field, touching the receptive fields of all templates.

### 2.4.1 Tuning properties

The template and the structure of the flow is invariant under any rotation in the $y$-$z$ plane. Therefore, we analyse the tuning to heading direction along only one meridian, and tuning to rotation about only one axis in the image plane. We treat the effects of torsional rotation on the template's response separately.

The tuning to horizontal heading direction ($T_y$) of the pure expansion template ($O_0$) is Gaussian (Fig. 2.8a). The maximum is located at the template's preferred heading direction $t_y = 0°$.

Next, we analyse the tuning of retinal templates ($O_0$, $O_z$ and $O_{-z}$) to rotation about the vertical axis. Fig. 2.8b shows that the pure expansion template and the two templates tuned to expansion and opposite rotations (preferred rotation magnitude $\omega = 2.5$ °/s) about the vertical axis have similar Gaussian-shaped receptive fields. The maxima are located at the preferred rotation velocities 0, -2.5 and 2.5 °/s, and the tuning width $\sigma$ estimated from the tuning curve is about 3.8 °/s. The 'rotation+expansion' templates overlap sufficiently to allow the approximation of the derivative RF with respect to rotation about the $z$-axis (Eq. 2.8). As explained




in the error analysis of the position model, the derivative function is more closely approximated when it is multiplied with a constant factor. Substituting $h$ by $\omega = \frac{2}{3}\sigma$, this factor equals $\exp(\omega^2/\sigma^2) = \exp(0.67^2) = 1.55$. Hereafter, whenever the head-centric response is calculated to first order, we multiply the derivative approximation $(O_{-z} - O_z)/(2\omega)$ with this factor (Fig. 2.8b).

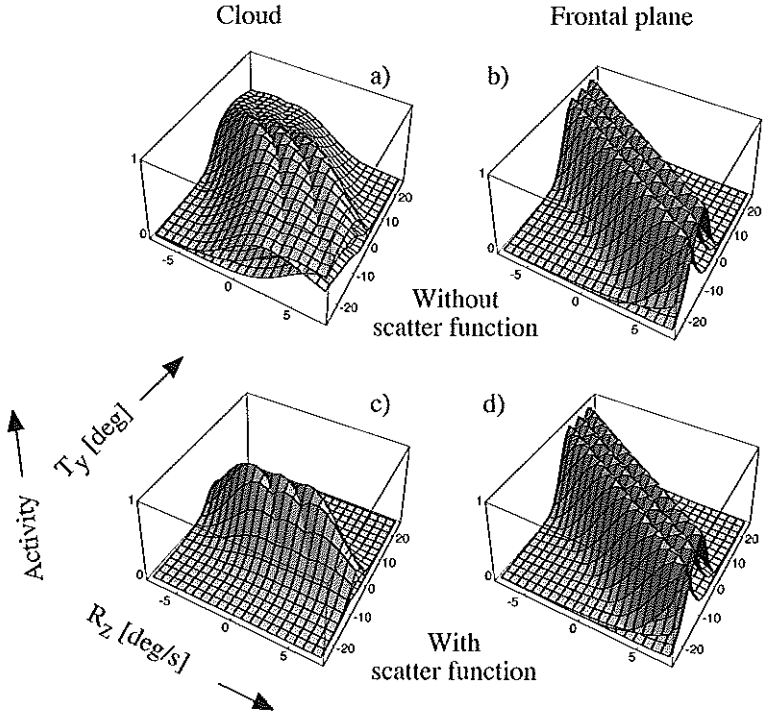
Fig. 2.8c illustrates that addition of 5.0 °/s rotation about the horizontal $y$-axis, scales the response of the expansion template and its approximated derivative equally, independent of the rotation about the vertical axis. This behaviour is desired, because

**Figure 2.8** *Response of retinal templates to flow belonging to self-motion at an ego-speed of 1 m/s. The cloud stimulus consists of 121 dots, lying within 100 × 100° visual angle at random depth between 1 and 10 m. a) Response to ego-translation of pure expansion template $O_0$ that prefers heading along the line of sight ($t_y = t_z = 0°$), for different horizontal heading directions $T_y$. b) Tuning of $O_0$ and three other templates to ego-translation along the templates' preferred heading direction in combination with varying amount of rotation $R_z$ about the vertical axis: $\mathbf{R} = (0, 0, R_z)$ °/s. $O_z$ and $O_{-z}$ have preferred horizontal rotation velocity $\omega = 2.5$ °/s, left and rightwards respectively. Their difference $1.55(O_{-z} - O_z)/(2\omega)$ yields the derivative $\partial O_0/\partial R_z$ to rotation about the vertical axis. c-d) $R_z$ tuning of $O_0$ and $\partial O_0/\partial R_z$ with addition of a constant component of (c) rotation about the horizontal axis: $\mathbf{R} = (0, 5.0, R_z)$ °/s, or (d) torsional rotation: $\mathbf{R} = (5.0, 0, R_z)$ °/s. e) Same as (d), with larger scale templates ($\omega = 7.5$ °/s). f) Rotation tuning of $O_0$ and $\partial O_0/\partial R_x$ (derivative to torsional rotation) for ego-rotation about the line of sight: $\mathbf{R} = (R_x, 0, 0)$ °/s. g) Rotation tuning of $O_0$ and $\partial O_0/\partial R_{x=z}$ (derivative to rotation about the $x = z$-axis) for ego-rotation about the $x = z$-axis: $\mathbf{R} = \frac{1}{2}\sqrt{2}(1, 0, 1)$ °/s. The response of $O_0$ does not drop to zero at high rotation velocities, because at each combination of torsional and horizontal rotation a number of points exist at which the flow measured along $\hat{\mathbf{a}}_i$ is exactly as preferred.*

uniform scaling of the responses will not impair the ability to compensate for rotation about the vertical axis. The uniform scaling is Gaussian as a function of the magnitude of rotation about the horizontal axis.

Addition of 5.0 °/s torsional flow (Fig. 2.8d), on the other hand, does not scale the responses uniformly. The pure expansion template prefers zero flow along circles concentric with the heading direction. These 'tangential' components of flow are no longer zero if the flow contains a horizontal rotation component, which reduces the activity of the $O_0$ template. However, torsional flow partly cancels the effect of horizontal rotation on the activity of the pure expansion template. Because the torsional flow is symmetrical with respect to the $x$-axis, this occurs for both positive and negative directions of torsion. This explains the occurrence of two peaks in the tuning curve. These side peaks, however, do not occur when the templates ($O_0$, $O_z$, $O_{-z}$) are tuned more broadly (Fig. 2.8e, $\omega = 7.5$ °/s). This illustrates that the template's tuning width sets an upper limit to which the transformation runs independently for orthogonal components of rotation.

There is a qualitative difference in the tuning of the rotational templates to rotation about the $z$-axis when a torsional ($R_x$) or when a horizontal ($R_y$) component of rotation is added to the flow. The compensation mechanism as proposed in Eq. 2.10 may fail if a fast torsional eye movement is made. As an alternative, we proposed the possibility of compensation about more than the three orthogonal axes. In Fig. 2.8f we show the response of a template bundle that compensates for rotation about the $x$-axis, as we vary the amount of torsional flow. In this case, the derivative is composed of two 'expansion+rotation' templates tuned to opposite amounts of torsional rota-




tion. Fig. 2.8g shows the response of a template bundle that compensates for rotation about the diagonal axis ($x = z$-axis), in which case the derivative is constructed from two 'expansion+rotation' templates tuned to opposite amounts of rotation about the diagonal axis. These two tuning curves look similar to that of the template bundle which compensates for rotation in the image plane (Fig. 2.8b). Thus, we conclude that the compensation about each axis is the same.

### 2.4.2 Role of motion parallax

For the transformation of retinal to head-centric flow templates, we desire retinal flow templates whose preferred rotation does not depend on the heading direction. In a limited aperture, however, the flow caused by ego-translation perpendicular to the observer's line of sight is very similar to the flow pattern caused by ego-rotation about an axis that is perpendicular to both the line of sight and the direction of translation (Koenderink & van Doorn 1987). For a frontal plane the recovery of the ego-motion parameters becomes inherently ambiguous (Longuet-Higgins & Prazdny 1980).

The effect of the ambiguity on the rotation tuning of the retinal flow templates is shown in Fig. 2.9a and 2.9b for movement towards a cloud and a fronto-parallel plane. In these plots, the horizontal heading direction ($T_y$) and amount of rotation about the vertical axis ($R_z$) is varied. Instead of three blobs located at the three preferred

**Figure 2.9** *Response to flow of a retinal template bundle as a function of the horizontal direction of translation (speed 1 m/s) and the amount of rotation about the $z$-axis. The retinal templates $O_0$, $O_{-z}$ and $O_z$, each have preferred heading along the line of sight ($t_y = t_z = 0°$) and preferred rotation about the $z$-axis of respectively -2.5, 0 and 2.5 °/s. For both (a) the cloud and (b) the frontal plane stimulus the heading direction at which a maximum occurs deviates from the preferred heading direction with increasing rotation rate. c) For the cloud, multiplication with the scatter function inhibits the response at heading directions that deviate from the preferred heading, but not so for the fronto-parallel plane (d). The scatter function uses as input the mean variance of $O_z$ and $O_{-z}$ ($\sigma_s = 1$ °/s). Stimuli consist of 121 dots distributed over 100 × 100° visual angle. Cloud distances: 1-10 m; frontal plane distance: 2 m.*

rotations ($r_z$ =-2.5, 0, 2.5 °/s) and the preferred heading direction ($t_y = 0°$), the pure expansion template and the 'rotation+expansion' tuned templates show oblique oriented ridges of activity.

Interestingly, the orientation of the ridges for the cloud stimulus (average distance 5.5 m) is more oblique than for the frontal plane (average distance 2 m). This effect follows from the geometry of the flow. The apparent displacement of the focus of expansion due to a rotational flow component is governed by the ratio of the rotation velocity $R$ and the distance scaled ego-speed $(T/d)$. Because the points in the cloud stimulus lie on average further away than the points in the fronto-parallel plane, the ridges of template activities are oriented less obliquely for the latter stimulus.

The diagonal ridge of activity, means that the template is equally activated, or nearly so in case of the cloud, when the heading direction and the eye rotation deviate simultaneously from the template's preferred heading direction and rotation. Consequently, heading judgement, based on these retinal pure expansion templates, will display a systematic error in the detected heading direction, proportionally to and in the same direction as the eye's rotation.

Psychophysical studies show that the removal of depth differences by presenting a frontal plane of points, results in systematic heading direction errors when eye rotation is simulated (Regan & Beverly 1982; Rieger & Toet 1985; Warren & Hannon 1990). With two planes at different depths, errors are much smaller (Rieger & Toet 1985; Royden et al. 1994). To a limited extent the templates show such an effect of depth in the stimulus, because the three templates respond less to the cloud stimulus for higher rotation rates, contrary to the response for a fronto-parallel plane. Yet, we wondered whether depth could be exploited more effectively by the template to accomplish separate tuning to heading direction and ego-rotation.

When the ego-motion is directed towards the template's preferred heading direction, each local flow vector's component along $\hat{\mathbf{a}}_i$ provides the same estimate of the amount of ego-rotation. But, if the actual heading direction differs from the template's preferred heading direction, the local flow along $\hat{\mathbf{a}}_i$ will contain a translational component that may show large variation in magnitude for points at different depths.

Thus, large variations of the local flow along $\hat{\mathbf{a}}_i$ over all points indicate that the current flow field does not correspond to the template's preferred heading direction. Suppression of responses of templates with such large variation in the local estimate of the rotation will favour the template with the correct preferred heading.

For the local estimate of rotation, we take the ratio of the components along $\hat{\mathbf{a}}_i$ of the local flow vector $\mathbf{p}_i$ and the preferred rotational flow vector $\mathbf{q}_i$, which we shall refer to as the local rotational gain:

$$g_i = P_i/Q_i$$

When $\hat{\mathbf{t}} = \hat{\mathbf{T}}$, the local rotational gain does not vary between viewing directions, even if the ego-rotation changes its direction or magnitude for this will scale all local rotational gains equally. Obviously, those points for which the denominator equals zero ($Q_i = 0$) need to be excluded from the set. As this is the case for all visual directions in case of the pure expansion template, only the rotation-tuned templates are used. The variance in the rotational gain over the set of flow vectors can then be used to inhibit template responses by multiplication with a 'scatter' function in which $\sigma_s$ determines the weight of the motion parallax cues:

$$W_s = e^{-variance/{\sigma_s}^2}$$

Inhibition by the scatter function ($\sigma_s = 1.0$ °/s) clearly has two benefits for the detection of heading. Firstly, inhibition occurs mainly at the templates' non-preferred heading direction ($t_y \neq 0°$), allowing the three ridges of activity to run parallel to the rotation axis (Fig. 2.9c, cloud stimulus). Evidently, the scatter function has no influence on the template responses during approach of a fronto-parallel plane (Fig. 2.9d). Thus, provided motion parallax cues are present, the scatter function reduces the ambiguity between a shift of preferred heading direction versus a shift in the preferred amount of ego-rotation. Secondly, the scatter function narrows the template's tuning to heading direction, which facilitates the localization of the maximum response in the heading map.

Lowering $\sigma_s$ may maximize the benefit of motion parallax cues, but also increases the template's sensitivity to variance in local rotational gain caused by noise at the level of local motion detectors. Such noise decreases the activity of all templates in the heading map, impairing the detecting the location of the maximum response in the heading map. Thus, at some point there will be a trade-off between the increased invariance to rotational flow against the loss of accuracy to locate the maximum response in the heading map.

## 2.5 Head-centric flow template simulations

Having checked the retinal flow properties, we examine the head-centric responses for a model that uses the oculo-motor signal, a visual signal or their combination as an

estimate of the eye's rotation. For simulations of an approach to a fronto-parallel plane subjects make large errors in judging their heading direction when eye rotation is also simulated, but not when they make real eye movements (Warren & Hannon 1988, 1990; Royden et al. 1994; Regan & Beverly 1982). However, performance for the simulated eye rotation condition improves considerably when depth information is present from motion parallax (Warren & Hannon 1990; van den Berg 1992; van den Berg & Brenner 1994a). The extent to which this performance increase occurs is currently disputed. Here, we do not wish to indulge into that debate, but merely ask whether a plausible mechanism can be conceived that captures both the role of eye movement or other extra-retinal signals and the contribution of motion parallax information in the perception of heading.

### 2.5.1 Oculo-motor signals

According to Eq. 2.9, the retinal flow templates can be combined with eye velocity signals to arrive at head-centric flow templates. We shall refer to this scheme as the extra-retinal signal model in which the oculo-motor signal eo represents the eye's velocity.

Fig 2.10a and b show the head-centric flow template responses for a cloud and frontal plane stimulus for different horizontal heading directions and different amounts of eye rotation about the $z$-axis without suppression by the scatter function. For rotation rates up to about 4 °/s, the response of the head-centric template has a maximum that is positioned at the preferred heading direction $T_y = 0$, in contrast to the response of the pure expansion template (Fig. 2.9a-b). For larger rotation rates, the peak activity does shift away from the preferred heading direction. Thus, the addition of the eye velocity gain field response $e_o[O_{-z} - O_z]/(2\omega)$ can compensate for eye rotations for both the cloud and frontal plane stimulus. For the frontal plane stimulus the maximum of the derivative term $[O_{-z} - O_z]/(2\omega)$ does not decline in magnitude. Consequently, the multiplication with a linearly varying oculo-motor signal causes the head-centric flow response to grow continuously with increasing rotation magnitude $|R_z|$. The detection of the heading direction, however, is not impaired by such increased response, since it does not affect the location of the ridge in the head-centric flow response map.

Fig. 2.10c shows the same head-centric flow template response (ridge in the middle, $t_y = 0°$) for a cloud stimulus when also motion parallax cues are exploited. Although the inhibition with the help of motion parallax can largely counter the effect of eye rotation, the contribution of the eye velocity gain field is evident, because the maximum activity of the head-centric flow template remains high up to 4 °/s rotation rates, compared with the rapid decrease of the response of the retinal pure expansion template in Fig. 2.9c.

We have also simulated the responses of two other head-centric flow templates tuned to heading direction $t_y = -20°$ and $t_y = 20°$, respectively (Fig. 2.10c). We observe that the peak response of these templates is simply shifted up and down

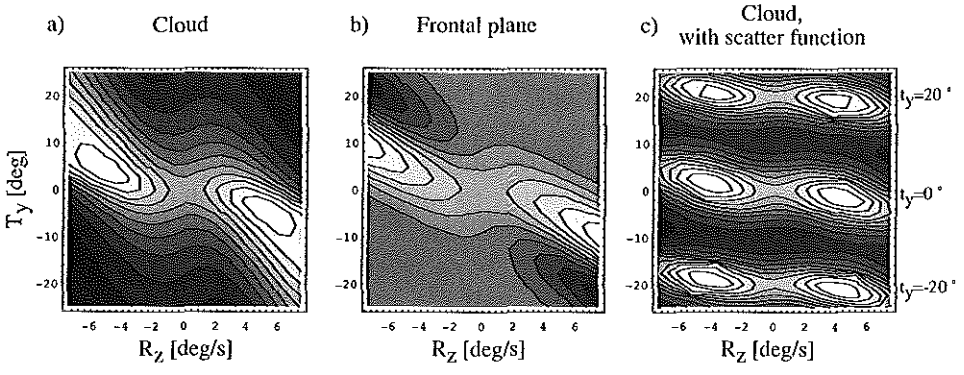



**Figure 2.10** *Head-centric template responses using an oculo-motor signal ($e_o = R_z$). Flow stimuli, see legend Fig. 2.9. a) Cloud, preferred heading $t_y = 0°$. b) Fronto-parallel plane. c) Response of three head-centric templates, superimposed. Cloud, inhibition with scatter function ($\sigma_s = 1.0$ °/s). The maximum response for each template is found at the preferred heading direction ($t_y$ =-20, 0 and 20°) up to about 4 °/s. Total activity range is divided into 10 equidistant levels, increasing from dark to light.*

along the $T_y$-axis, by an amount equal to their preferred heading direction. This implies that our assumption on the equivalence between analysing the response of a single template to flow of different heading directions, and analysing the response a population of templates to only a single flow field, is valid.

The range of eye rotations for which the head-centric flow templates detect the correct heading during motion towards a frontal plane, is small compared to human performance during real eye rotations (Royden et al. 1992). As was demonstrated in the error analysis of the position model (Fig. 2.2b), this range is limited by the width of the templates' tuning to rotation (Eq. 2.9). Fig. 2.11 shows the range is considerably increased for larger scale templates, which by our definition of templates have larger preferred rotation $\omega$ (cf. legend Fig. 2.8). For very broadly tuned retinal flow templates, the head-centric flow template tolerates eye rotations up to more than 20 °/s. However, as for the scatter function, the invariance to rotation trades off with a loss of accuracy in detecting the location of maximum activity in the population of head-centric templates.

With the second order derivative to rotation also taken into account (Eq. 2.12), the range of rotations the model can cope with is extended even more (Fig. 2.11d).

$$H(\hat{\mathbf{t}}) = O(\hat{\mathbf{t}}, 0) - \epsilon \frac{\partial O(\hat{\mathbf{t}}, 0)}{\partial R} + \frac{\epsilon^2}{2} \frac{\partial^2 O(\hat{\mathbf{t}}, 0)}{\partial R^2} \tag{2.12}$$

This second order derivative to rotation can be composed of zero order derivative RFs, according to the three-point formula (Eq. 2.13). The quadratic eye velocity signal

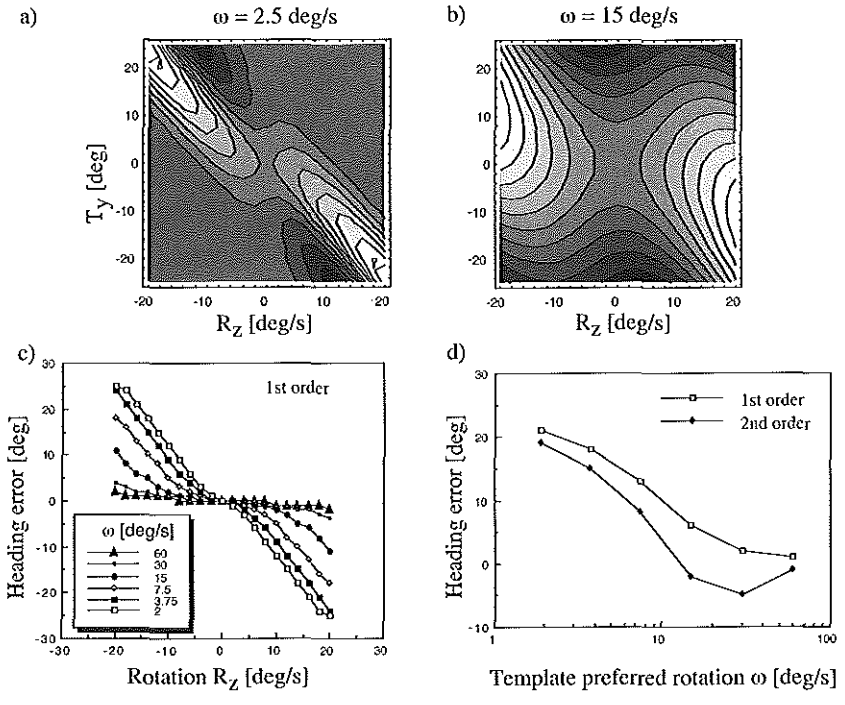



**Figure 2.11** *Effect of the scale of the template bundle on the head-centric responses. The scale is characterized by the preferred rotation $\omega$ of the retinal templates that are tuned to translational and rotational flow. Response for different heading directions ($T_y$) and amounts of real eye rotation ($e_o = R_z$) for (a) $\omega = 2.5$ °/s, and (b) $\omega = 15$ °/s. c) Heading error, defined as the difference between the heading direction at the peak response and the template's preferred heading direction, as a function of the eye rotation. d) Heading error for 16 °/s eye velocity as a function of the template's scale and order. Frontal plane, distance 0.57 m, 100 × 100° visual field, 121 dots, speed 0.38 m/s.*

($\epsilon^2$) can be constructed from linearly varying eye velocity signals with different offsets, as explained in the error analysis of the position model.

$$\frac{\partial^2 O(\hat{\mathrm{t}},0)}{\partial R^2} \simeq \frac{1}{\omega^2}[O(\hat{\mathrm{t}},-\omega) + O(\hat{\mathrm{t}},\omega) - 2O(\hat{\mathrm{t}},0)] \tag{2.13}$$

A recent report by Bradley et al. (1996) shows that certain cells in area MST do not shift their preferred center of translational flow on the screen during an eye movement compared with the preferred center when the eye is stationary. A similar effect has been found for MSTd cells during full body rotation (Shenoy et al. 1996). Interestingly, Bradley et al. (1996) found a variety of response types ranging from not shifting at

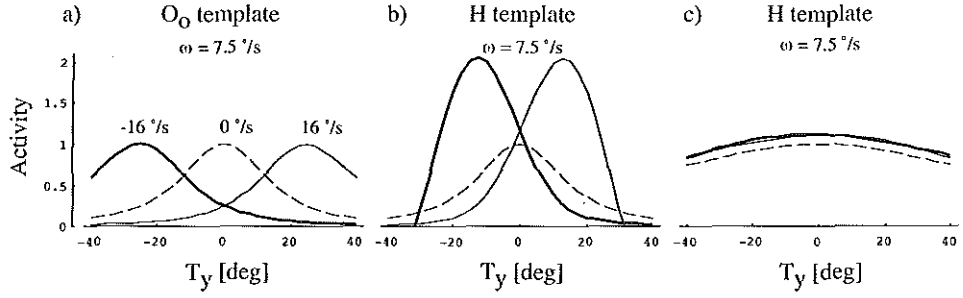



**Figure 2.12** *Each panel shows the response of a template to varying horizontal directions of ego-motion, and three different amounts of rotation about the vertical axis; $R_z = -16$ °/s (solid line), $R_z = 0$ °/s (dashed line), $R_z = 16$ °/s (thin line). a) Pure expansion template; the preferred heading direction on the screen shifts in the direction of the eye movement. b) Head-centric template for $R_z > \omega$; the peak activity shifts less on the screen than in (a). c) Head-centric template for $R_z < \omega$; there is practically no shift on the screen. For stimulus, see legend Fig. 2.11.*

all to a complete shift corresponding to a preferred retinal locus for the center of the translational flow. These response types are consistent with our head-centric and retinal flow templates, respectively (Fig. 2.12).

For other cells, the amplitude of the response to the retinally preferred locus varied depending on the direction and magnitude of the eye movement. Such results can be reproduced by our model when the output of the retinal and head-centric flow templates tuned to different amounts of rotation w are considered (cf. Fig. 2.11d; the heading error can be equated with a deficit of the required retinal shift of the preferred center of translational flow).

### 2.5.2 Visual estimation of rotation

Up to this point the model uses an eye velocity signal $\epsilon$ as input for the gain term in Eq. 2.7. This does not mean that oculomotor signals are essential to carry out the transformation to head-centric flow. The rotation tuned templates also provide a means for estimating the rotation about each axis on a pure visual basis. To create a visual signal that varies linearly with the amount of rotation, one can use the two templates of the template bundle that are tuned to opposite rotations. The difference between these Gaussian responses varies linearly with the rotation in the flow within a range that is limited by the preferred rotation of the templates (Fig. 2.13). Actually, this amounts to the same expression as the Gaussian derivative in Eq. 2.8. All left to be done is multiplication with a suitable factor; the templates' preferred rotation velocity $\omega$:

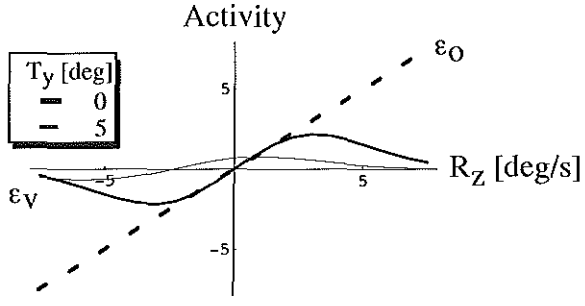



**Figure 2.13** *Visual rotation signal $e_v$ (preferred ego-motion of the templates: $t_y = 0°$, $r_z = \pm 2.5$ °/s) as a function of the simulated ego-rotation for two different heading directions (cloud, 1-10 m, speed 1 m/s). For $T_y = 0$, $e_v$ is tangent to the correct eye movement signal $\epsilon_o$ (dashed line), while for $T_y = 5°$, the response curve is decreased and shifted towards the left.*

$$\epsilon_v(\hat{\mathbf{t}}) = \omega\,[O_z(\hat{\mathbf{t}}) - O_{-z}(\hat{\mathbf{t}})] \tag{2.14}$$

Thus, substituting the visual rotation signal for the eye velocity signals in Eq. 2.9 we arrive at a purely visual model:

$$H(\hat{\mathbf{t}}) = O_0(\hat{\mathbf{t}}) + \frac{1}{2}[O_{-z}(\hat{\mathbf{t}}) - O_z(\hat{\mathbf{t}})]^2 \tag{2.15}$$

Fig. 2.13, however, shows that only the retinal flow template bundle whose preferred heading direction matches the actual direction of ego-motion is able to give a correct estimate of the true horizontal rotation component in the flow. At non-preferred heading directions ($T_y \neq 0$), the visual estimate of rotation is smaller and shifted along the rotation axis, because it suffers from the same ambiguity as the estimate of the heading direction. Therefore, the pure visual model needs motion parallax cues. Even without inhibition by the scatter function, the head-centric flow template responds at the correct heading direction for a cloud of dots for simulated eye rotations up to about 4 °/s (Fig. 2.14a). The pure visual model, however, shows large systematic errors in absence of depth differences (Fig. 2.14b).

### 2.5.3 Combining visual and extra-retinal signals

The oculo-motor model cannot explain why heading is correctly perceived during simulated eye rotations when motion parallax cues are present. On the other hand, the pure visual model cannot describe why correct heading is perceived when making real eye movements. Can we combine the visual and oculo-motor estimates of eye

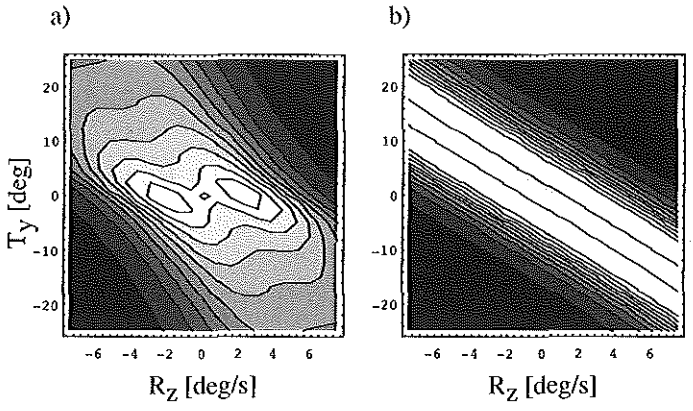



**Figure 2.14** *Head-centric response using the scatter function ($\sigma_s = 1\ °/s$) and a visual rotation signal instead of an oculo-motor signal. a) Cloud (1-10 m). b) Fronto-parallel plane (2 m).*

velocity, such that an adapted visual model can cope with simulated eye rotations during approach of a frontal plane?

As seen in Fig. 2.13, the visual estimate of rotation is underestimated when the direction of ego-motion does not correspond to the preferred heading direction of the template bundle. Hence, the difference between the extra-retinal signal, which is assumed to be correct, and the visual estimate of rotation can be used to suppress the activities of an head-centric flow template at 'wrong' heading directions, in a similar way the scatter function was used. As a measure of the amount of visuo-motor conflict we choose a Gaussian function of the error between the visual estimate of rotation $e_v$ and oculo-motor signal $e_o$:

$$W_e = e^{-[(\epsilon_o - \epsilon_v)/\sigma_{error}]^2}$$

Fig. 2.15 shows the results obtained when the head-centric flow template, that compensates with a visual signal using the scatter function ($\sigma_s = 1.0\ °/s$), is multiplied by the conflict function ($\sigma_{error} = 3.0\ °/s$). For the frontal plane, the conflict function suppresses the activity at 'wrong' heading directions during real eye rotations, while during simulated eye rotations a diagonal ridge of activity recurs. For the cloud stimulus, the visual model is already capable of detecting the correct heading direction before inhibition by the conflict function (Fig. 2.14a). This remains the same after inhibition by the conflict function, during both real and simulated eye rotation. During real eye rotation no conflict arises, whereas during simulated eye rotation inhibition has no effect on the orientation of the ridge of activity, since it already lies parallel to the rotation axis. Multiplication with the visuo-motor conflict does cause a more rapid decline of the head-centric response as a function of the rotation speed.

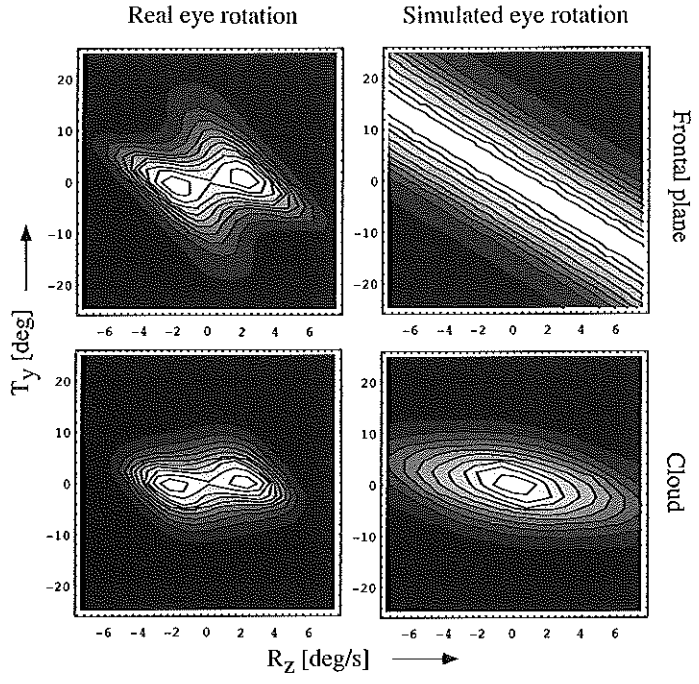



**Figure 2.15** *Head-centric response using the visual rotation signal, multiplied by the scatter function $W_s$ ($\sigma_s = 1.0$ °/s) and the conflict weight function $W_e$ ($\sigma_{error} = 3.0$ °/s). Maximum responses for the cloud stimulus (lower panels) during real ($\epsilon_o = R_z$) and simulated eye rotation ($e_o = 0$ °/s) lie at the preferred heading direction ($t_y = 0°$) for rotations up to about 3 °/s. Upper panels show that for the frontal plane the maximum only lies at the preferred heading direction during real eye rotations, but not during simulated rotations. For further details, see legend Fig. 2.9.*

## 2.6 Discussion

Our model uses motion parallax, extra-retinal signals and retinal flow to derive an estimate of the heading direction. Several psychophysical studies indicate that heading can be derived from the retinal flow under conditions where one of the other sources of information is lacking but not when both are missing, suggesting that either one contributes to heading perception. Our model (Fig. 2.16) captures this behaviour at least qualitatively.

The model uses a collection of motion templates. Basically, each template evaluates the evidence in the flow for its preferred ego-rotation and heading direction, without making assumptions on the structure of the environment. One would need a five dimensional array of such retinal flow templates to sample all possible ego-motions even when each template responds independently of the structure of the environment.

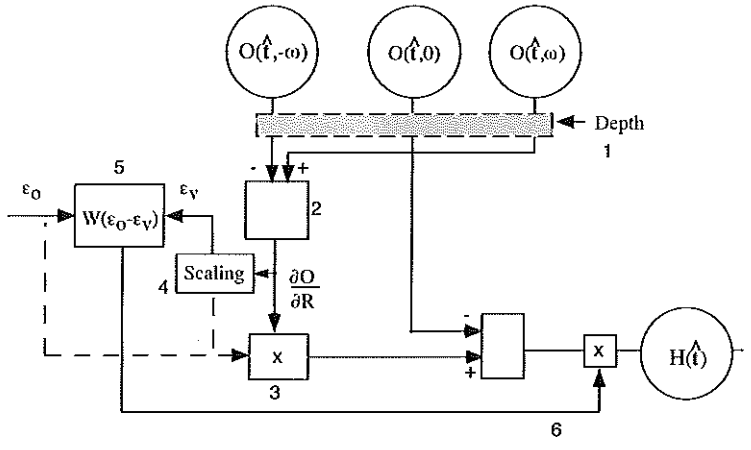



**Figure 2.16** *Diagram for transformation from retinal to head-centric flow. (1) Depth differences allow suppression of template bundle activities when their preferred heading direction deviates from the direction of ego-motion. (2) The difference between the two rotation tuned template responses provides $\partial O/\partial R_z$; a measure of how the response of the pure expansion template's activity will vary with respect to a rotational flow component. A change of the response of the pure expansion template due to rotational flow is compensated by subtracting an eye velocity gain field response, which equals $\partial O/\partial R_z$ times (3) the eye velocity signal $\epsilon_o$. Alternatively, a visual 'rate-coded' estimate of rotation $\epsilon_v$ (4) is also available after suitably scaling the derivative term. When the visual estimate of rotation and the eye velocity signal differ, the visuo-motor conflict function (5) suppresses the head-centric activity (6).*

Such an analysis of the flow by motion templates was introduced by Perrone (1992). In a later study (Perrone & Stone 1994) the number of templates was reduced by taking into account only those templates that correspond to rotations that stabilize a point of the stationary environment during forward motion.

We have taken a different approach. The templates are arranged in 'bundles'. The templates in the bundle share the same preferred heading direction but differ in the preferred ego-rotation. The bundle activities are combined with extra-retinal or visual rate-coded estimates of ego-rotation, resulting in flow sensitive cells with a gain field for eye velocity. Together, the templates in the bundle form a 'dynamic' receptive field in optic flow space, i.e. a head-centric flow cell that is tuned to different amounts of retinal rotational flow as the rate-coded eye or visual rotation signal changes. This leads to a significant reduction of the number of templates. As it is natural for the brain to use similar formats for extra-retinal and retinal estimates of self-rotation, deriving a rate code for self-rotation from the templates (Eq. 2.14) would seem appropriate. An additional asset is that estimates of self-rotation derived from different sources can be processed identically. Moreover, in this way a direct analogy

could be achieved between the retino to head-centric transformation in the position domain (area 7a, LIP) and the proposed transformation in area MST. The bundles as we present them are groups of motion templates that cooperate to carry out a certain transformation. Whether this leads to anatomical constraints is an open question to us. We merely put forward a functional scheme for the perception of heading using physiologically inspired elements. This leads us to suggest certain tuning properties of cells that have not yet been described.

An important parameter that determines to what extent the template bundle achieves tuning to head-centric (as opposed to retinal) flow is the 'scale' of the template bundle. The scale is set by the preferred rotation of the templates '$\omega$', which simultaneously determines the width of heading direction tuning and of the Gaussian-tuning to rotation. Taking into account only the first-order derivative, the head-centric flow template can tolerate rotation rates smaller than $\omega$ (Fig. 2.11). So, the amount of rotation the model can cope with can be arbitrarily set by choosing the scale of the templates. More accurate approximations can be obtained using a higher order Taylor expansion. Bradley's data suggest that in the monkey 'scales' of 15 °/s, or more, may be commonplace. For humans such data are lacking but the finding that heading perception can be accurate for pursuit eye movements as fast as 7 °/s (Royden et al. 1994; van den Berg 1996a) suggests that 'scales' of about 10 °/s are present.

The pooling of local motion signals already makes the templates sensitive to motion parallax. This holds, because the template responds equally strong for certain combinations of non preferred rotation and non-preferred heading direction as for the preferred self-motion, when no depth differences are present, but not so when depth differences occur in the scene (Fig. 2.9a and b). However, motion parallax can be exploited more effectively by the template. Each local motion signal can be used to derive a local estimate of the self-rotation. Provided the flow corresponds to the template's preferred heading direction, all these estimates are identical and the variance is minimal. By multiplying the template's output with a gain that depends inversely on the variance of the local estimates of the self-rotation, one can increase the sharpness of the tuning to the preferred self-motion parameters to any desired degree.

### 2.6.1 Neurophysiological evidence

A number of observations suggest that parts of our model may be implemented in area MST; the model relies on visual templates sensitive to pure retinal expansion, and to combinations of expansion and rotation. In area MSTd of awake monkeys cells have been found that respond to combinations of expansion and torsional rotation (Orban et al. 1992; Graziano et al. 1994; Duffy & Wurtz 1995). Earlier studies argued that MSTd cells are positional invariant (Orban et al. 1992; Graziano et al. 1994). This would render such cells unsuitable for heading detection tasks because positional invariance implies no tuning to the location of the focus of expansion in

the retinal flow pattern. However, Duffy and Wurtz (1995) and Lappe et al. (1996) showed that MST cells do respond differently when the center of an expanding motion pattern is positioned at different locations in the cell's receptive field.

Our model assumes the existence of templates with a Gaussian-tuning to the rotational flow. Most MSTd cells respond to a continuum of patterns (Duffy & Wurtz 1991; Graziano et al. 1994) consisting of combinations of expansion, torsional and horizontal rotational flow, but details on the tuning properties are scarce in the literature. Gaussian-tuning was reported by Graziano et al. (1994) for spiral motion; i.e. a stimulus that resembled addition of expansion and torsional flow. To our knowledge, no such data are reported concerning combinations of expanding flow with other directions of rotational flow.

Some cells in area MST continue to fire during blanking or retinal stabilization of a pursued point target. Pursuit cells with large visual fields were mainly found in two non-overlapping parts of MST (MSTl and MSTd) (Wurtz et al. 1990). This shows that extra-retinal input is present in MST (Wurtz et al. 1990; Thier & Erickson 1992), and possibly is being integrated with visual input. Thier and Erickson also found activity in MSTl related to head movement, which suggests MST may also integrate vestibular signals.

Recently, Lappe et al. (1996) presented data from MST that extends the findings of Duffy and Wurtz (1995) concerning the preferred location of an expanding pattern of motion. A minority of cells showed bell-shaped responses to the location of the center of expanding motion. The majority of cells , however, showed a gradient response plane as a function of the horizontal and vertical position of the focus of expansion (cf. Lappe et al. 1996, their Fig. 7 and 9). Moreover, a reversal of the gradient for contracting flow was seen. Our model can simulate these data without further modifications. These responses are reproduced by our first-order derivative cell, constructed from the difference between two broadly tuned rotation templates '$O_{-z}-O_z$' (Fig. 2.17). For templates preferring smaller rotations, the Gaussian-tuning becomes apparent as a curvature of the response surface in the direction perpendicular to the gradient. This curvedness is not visible in Lappe's figures that combine data collected from several cells with different gradients. Interestingly though, some single cell data did show saturation comparable to the simulation in Fig. 2.17a (Lappe personal communication). This suggests that the transformation to head-centric flow may take place at different scales (as defined above) simultaneously.

The gradual decline of the response when the center of expansion is shifted in one direction is in our model a property of the gradient cells ($\partial O/\partial R$) but not of the retinal motion templates themselves, which show a bell-shaped tuning to the center of expansion. If the proposed mechanism is used by the primate's visual system, Lappe's observation, that more than 80% of the cells show the gradual decline of the response when the center of expansion is shifted, implies that the majority of the MST cells compute the derivative term of Eq. 2.7. This is remarkable, because if the latter cells were always derived from two cells tuned to the same expansion but opposite directions

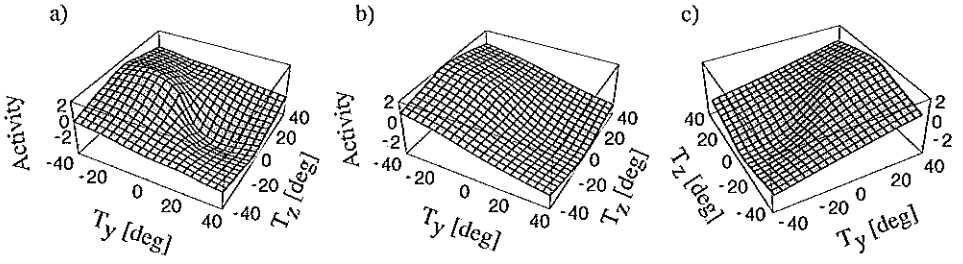



**Figure 2.17** *Response of the first-order derivative cell ($O_{-z} - O_z$) during translation in different directions. Cloud (2-40 m), 121 dots, 100 × 100°, speed 1 m/s. a) Forward motion, small scale template: $\omega = 3.8$ °/s. b) Forward motion, larger scale template: $\omega = 15$ °/s. c) Backward motion, larger scale template: $\omega = 15$ °/s.*

of rotation, one would expect a larger proportion of cells with bell-shaped tuning to heading direction. Lappe's results therefore would be compatible with the derivative term being computed directly from the local flow instead of through an intermediate step involving two templates. Because the derivative results from subtraction of two template activities, each merely based on weighted averages over the flow in different visual directions, one can carry out the subtraction also at the local level. This means that the same results are obtained when the weight functions of the two templates $W_{-z}$ and $W_z$ are combined into one weight function for the derivative template ($W_{dz}$):

$$\begin{aligned} W_{dz,i} &= (W_{-z,i} - W_{z,i}) \\ &= e^{-|P_i - Q_{-z,i}|^2/\sigma_i{}^2} - e^{-|P_i - Q_{z,i}|^2/\sigma_i{}^2} \end{aligned}$$

(Note the subscript '$i$' refers to one location in the template's receptive field; cf. Eq. 2.11). This means that the local flow component $P_i$ contributes excitatory to the derivative template when it equals $Q_{-z,i}$, and inhibitory when it equals $Q_{z,i}$. Because $Q_{-z,i} = -Q_{z,i}$, it follows that the derivative template weights opposite local velocities oppositely. Such cells have been found in area MT of the monkey (Snowden et al. 1991). These cells prefer one direction of motion in their receptive field and are inhibited by oppositely directed motion. When both motion directions are shown simultaneously in the receptive field, an intermediate response occurs. We suggest then, that these MT-cells play a role for the perception of heading in the presence of eye rotations.

Evidence for the integration of extra-retinal and flow information in MSTd comes from a recent study by Bradley et al. (1996) in which patterns of motion were presented to a rotating eye. Besides cells that shift their preferred center of expansion during eye rotation, they also found a class of cells that do not shift their preferred center, but do show a modulation of the amplitude by eye rotation. This type of cell would

correspond to our model's eye velocity gain field. Lappe et al. (1996) find that most cells in MST display a gradient response instead of a bell-shaped response. We argued that this behaviour is consistent with the computation of the derivative term $\partial O/\partial R$ from the local flow. Thus, it seems likely that the gain fields are the product of eye velocity with the derivative response $\partial O/\partial R$, rather than with the retinal $O$ templates. We suggest that part of those cells that are modulated in amplitude by extra-retinal signals, may have gradient responses to the retinal location of the center of expansion. An issue that we have virtually ignored, is the fact that the rotational component in the flow corresponds to the total rotation, i.e. the rotation of the eye relative to the environment. Oculomotor signals code the rotation velocity of the eye in the head. Strictly, our simulations apply to conditions with the head stationary. However, combination of eye-in-head and head relative-space velocity vectors, as supplied by the vestibular system, can result in extra-retinal signals that can be combined with the template signals even for a moving head. This emphasizes the importance of the finding of Thier and Erickson (1992) that certain cells in area MST integrate vestibular, oculomotor and visual velocity signals.

An additional complication arises for eccentric eye positions. If the eye velocity signals supplied by the oculomotor system are independent of the eye position, i.e. in a head-centric format, one and the same eye velocity signal will, for example, correspond to torsional flow for upward looking and to horizontal flow when looking straight ahead. Thus, eye velocity and head velocity signals need to be mapped in the retinal frame in order to be combined sensibly with the motion templates. Bremmer and Hoffmann (1993) reported pursuit related activity in MST, which was modulated by the starting position of the eye. This may indicate that eye velocity signals are coded in retinal coordinates.

We have shown the retinal flow templates respond qualitatively different to rotation about the line of sight than to rotation about an axis in the image plane. This led us to question the use of three derivatives tuned to orthogonal axes of rotation to compensate for general rotations of the eye. However, the qualitatively different behaviour is only evident when the amount of torsional flow presented to the eye is large compared with the templates' rotation tuning width. What range of torsional rotations do usually occur? During smooth pursuit of a target point, the eye rotates according to Listing's law (Tweed et al. 1992). When tracking a point in the environment, the torsional rotation is zero if the target rotates about an axis perpendicular to the eye's primary direction. Torsional rotation might occur, though, when the heading and primary direction are not aligned; for instance when making a horizontal eye movement with the head tilted down, or when a moving target is pursued such as a bird flying horizontally above the horizon. Roughly, the ratio of the torsional velocity $\omega_{torsion}$ to pursuit velocity $\omega_{eye}$ is determined by the angle $\alpha$ between the target's rotation axis and the plane perpendicular to the eye's primary direction: $\omega_{torsion} \approx \omega_{eye} \sin(\alpha/2)$ (Tweed et al. 1992; van den Berg 1996b). Under extreme conditions, ($\alpha = 40°$), torsional velocity may reach up to one third of the pursuit ve-

locity, although it remains unclear whether heading would then still be perceived well. More likely, however, torsional rotations remain an order of magnitude smaller than the pursuit velocity ($\alpha < 20°$). Because (1) the templates can compensate only for eye rotations up to about the rotational tuning velocity of the retinal flow templates (Fig. 2.11), and (2) we found irregular effects of eye torsion only when the torsional velocity exceeded the horizontal tuning velocity of the templates (Fig. 2.8e-f), we suggest that our templates will operate normally even during torsional rotations, with the exception possibly for fast torsional head movements. Yet, lacking other psychophysical and neurophysiological data, we cannot decide at this stage between the proposed alternatives to compensate for torsional flow.

The heading direction coded by our head-centric flow templates '$H$' is defined relative to the retina. To derive head-centric heading direction additional transformation is required. Perhaps the extensive projections between area MST and area 7a, which is known to contribute to a transformation of retinal to head-centric visual direction, are involved in the generation of a head-centric representation of heading direction. Recently, eye position gain fields have been found in area MST (Squatrito & Maioli 1996), while in area 7a flow sensitive cells have been reported whose response varies linearly and quadratically with eye position (Collum & Siegel 1996).

Although the presence of extra-retinal signals in MST suggests a good chance of finding velocity gain fields in MST, the integration of the extra-retinal signals and motion templates may take place at other locations along the visual and motor pathways. For example, it has recently been shown that area VIP contains cells that are sensitive to a diversity of optic flow patterns (Schaafsma & Duysens 1996) and fire during pursuit eye movements (Colby et al. 1993). These studies, however, do not allow for conclusions on the type of interaction between these signals or even whether the cells are tuned to head-centric or retino-centric flow.

### 2.6.2 Suggestions for neurophysiological research

Our model assumes that the expansion and rotation cells in MSTd are Gaussian-tuned to rotation velocity, not only about the torsional axis, but especially about axes in the fronto parallel plane. Data in support of this view are lacking currently. Cells sensitive to horizontal motion have already been found. However, relatively little is known of their speed tuning to a rotational component that is added to expanding flow. Furthermore, the sensitivity to horizontal flow should be tested with real rotation stimuli, instead of shifting patterns on a tangent screen. Especially in peripheral viewing directions, translation on the tangent screen differs from the flow belonging to a rotation about an axis in the fronto-parallel plane. Thus, it may be of interest to investigate whether MSTd cells are Gaussian-tuned to rotational velocity about axes in the fronto-parallel plane. Even more interesting would be the search for cells that are Gaussian-tuned to rotation and one direction of translation. For such cells our model predicts that if stimulated with a flow that matches the cell's preferred rotation and heading direction, the response should not vary with the simulated ego-speed,

i.e. the rate of expansion. Note, that this is not the same as saying that the cell decomposes the flow field into a translational and a rotational component, because when the preferred heading direction does differ, the simulated rate of expansion affects the response.

Eye velocity gain fields, as proposed in our model, are more likely to be found in MSTd than in MSTl, as only the former cells respond to complex optic flow patterns. According to the oculo-motor model, the template is modulated by an extra-retinal signal, which varies linearly with eye velocity. To investigate this relationship, one could present a constant retinal flow to a moving eye. According to our model, some cells' responses should vary linearly with eye velocity, but the (Gaussian) tuning to the retinal flow pattern should remain the same irrespective of the eye movement. Our model does not exclude the possibility, that the visual sensitivity of the gain field cell is only revealed during eye movement in a preferred direction, i.e. without eye movement, the cell might not appear to be visually responsive.

Another important test, is the alignment of the direction of the axis of preferred rotational flow and the axis of preferred eye rotation for MSTd neurons. If vestibular information also converges in MSTd, one could also investigate the preferred axis of vestibular rotation signals. Thier and Erickson (1992) have found cells in MSTl that are sensitive to both eye rotation and head-rotation. Their preferred direction appeared to be the same axis, and for both responses depended linearly on rotation speed.

To find the analogue of our model's output templates that code head-centric heading direction, we should differentiate between the three possible models we have proposed. The pure visual model and oculo-motor models use a visual estimation of the rotation or an extra-retinal signal as the gain factor. As a third possibility the visual rotation and oculo-motor rotation signals may be combined to suppress templates in proportion to the conflict between the two signals. A mixed model in which visual and extra-retinal estimates of the eye rotation are used in combination with different scales of the templates is another possibility that will be explored in a future study. To distinguish these different possibilities one needs to decouple the rotational component in the flow from the eye rotation. The simulated eye rotation condition in the psychophysical literature (Warren & Hannon 1990; van den Berg 1992) is one example of such a stimulus. In general, one requires that one can vary the eye velocity signal independently of the visual estimate of the ego-rotation as derived from the flow. One can then test the different models by concentrating on the head-centric flow cells or by concentrating on the modulation of cells that are tuned to retinal flow. Cells found by Bradley et al. (1996) that shift their preferred center of translational flow when the eye moves, provide an excellent means to validate the extra-retinal model. Assuming such a cell corresponds to a head-centric flow cell, measuring its tuning to the rotational flow component (and the preferred heading) for a stationary eye, immediately reveals the rotation tuning width of the retinal pure translation template. According to Eq. 2.9 and Fig. 2.11, the amount of eye rotation

the head-centric flow cell can tolerate is directly related to this rotation tuning width. This can be tested using Bradley's procedures.

Cells with activity peaking for two oppositely directed rotational components in the retinal flow (independently of the amount of expansion in the preferred heading direction and the amount of eye rotation) would provide strong evidence for the pure visual model because such behaviour corresponds to $[O(\hat{\mathbf{t}}, -\omega) - O(\hat{\mathbf{t}}, \omega)]^2$ (cf. Eq. 2.15). The extra-retinal signal model would be favoured if cells are excited by one direction of rotational flow and inhibited by the opposite direction and if a reversal of the excitatory direction occurs when the eye velocity is inverted. This type of behaviour would correspond to $\epsilon_o[O(\hat{\mathbf{t}}, -\omega) - O(\hat{\mathbf{t}}, \omega)]$ (cf. Eq. 2.9). Finally, the conflict model would be supported by cells of which the responses are modulated by the difference between the rotational flow and the eye velocity.

### 2.6.3 Other physiological models

Our model is based on templates similar to those used by Perrone (1992) and Perrone and Stone (1994). Unlike Perrone (1992), we are not concerned with modelling the motion sensors after MT cells exactly. Instead, we describe the local input to a template directly in terms of the local flow vector, the template's preferred flow vector and a local constraint line along which the flow is measured. To some extent, the local inputs to the heading detector as proposed by Perrone and Stone (1994) also achieve insensitivity to the translational component of flow, because at one retinal location a maximum response is selected from a set of motion sensors, each tuned to different ego-speeds. We think, though, their model has difficulty to account for a number of observations. First of all, to reduce the amount of templates, needed to sample all possible heading directions and ego-rotations, Perrone and Stone (1994) restrict eye rotations by assuming a point in the environment is fixated. Royden et al. (1992) find, however, heading is accurately perceived when tracking moving objects. Our scheme reduces the number of templates without posing constraints on the direction of eye rotations. Secondly, Perrone and Stone (1994) suggest extra-retinal signals can be used to emphasize the response map of templates whose preferred rotation corresponds to the eye's rotation. This would require eye velocity in labelled-line code to inhibit the response maps tuned to other rotations, while only rate-coded eye velocity in MST has been reported so far (Bremmer & Hoffmann 1993). Thirdly, it is unclear whether their templates can reproduce the recent physiological data from MST reported by Lappe et al. (1996) on gradient planes as function of the position of the focus of expansion.

Another interesting model is the neural network proposed by Lappe and Rauschecker (1993, 1995). It implements the subspace algorithm of Heeger and Jepson (1992) to set the connections between a first layer of local motion detectors and a second layer of cells that respond to large motion patterns. Both their second layer cells as well as our derivative cells $\partial O/\partial R$ can reproduce the recent data from MST (Lappe et al. 1996) in which response gradients as function of the retinal focus

of expansion were found. These similarities suggest a possible close relation between both models, for instance at the level of the sampling of the flow. In the model of Lappe and Rauschecker (1993), the second layer is organized in subpopulations of cells that share the same preferred heading direction. The total activity of a subpopulation is largest for ego-motion in the preferred direction by setting the connections between local motion sensors and second layer cells, so that a residual function is minimized. Because this residual function consists of the sum of inproducts between the selected local flow vectors and the actual flow vectors, it may be interesting to see if the selected flow vectors of second layer cells are oriented perpendicular to the translational components of flow, as we propose for our template structure. Moreover, the input connections to a second layer cell are divided into pairs (Lappe & Rauschecker 1993). It would be interesting to see if locally, such pairs of motion sensors show opposite preferred directions, as suggested for our first-order derivative cell $\partial O/\partial R$.

An important difference between our model and other models lies in its use of an extra-retinal signal, and its multiplicative interaction between the extra-retinal estimate of rotation $\epsilon_o$ with the derivative template ($\partial O/\partial R$), to create the compensating term $\epsilon_o \partial O/\partial R$. Since the compensating term scales equally with the pure expansion template $O$, it also takes into account properties of the flow, such as the number of dots, layout, contrast.

### 2.6.4 Conclusions

We have presented and tested a general model for changing the reference frame of receptive fields using extra-retinal information. Our model explains the need for broadly Gaussian-tuned receptive fields and eye position gain fields in area 7a, involved in the transformation of retinal to head-centric visual position. To retrieve the direction of self-motion from retinal flow patterns, the same method can be applied to compensate for rotational flow during eye or head rotations, but now using eye velocity signals.

Although neurophysiological studies of cells in MST seem to support our heading model, a number of model elements and assumptions have yet to be confirmed. Currently, physiological data on the retinal flow template's Gaussian-tuning to rotation, the presence of higher order derivatives and the role of motion parallax are lacking. As yet, it is not evident whether extra retinal signals, visual acquired estimates of eye velocity, or both are used for the multiplicative interactions. Qualitatively, the simulated responses of head-centric flow templates are consistent with human heading performance for simulated motion towards a cloud or fronto-parallel plane, during real or simulated eye rotations.

## Acknowledgements

This work was supported by research grant SLW 805-33.171-P from the Dutch Foundation for Scientific Research (NWO) and by the Human Frontier Science Program.

# Chapter 3

# Combining extra-retinal and visual estimates of eye rotation to perceive heading


## Abstract

An eye that translates through the environment receives an expanding motion pattern that emanates from the direction of heading Gibson (1966). When the eye also rotates, this cue is masked by the presence of a rotational component in the retinal flow: the rotation problem. Yet, humans can perceive their heading direction accurately because extra-retinal signals counter the effect of the rotational flow (i.e. Warren & Hannon 1990). We investigated whether compensation is achieved by unconditional subtraction of rotational flow, as suggested previously (Royden et al. 1994). We find, however, that the compensation occurs only when the effect of the eye rotation is also visually present in the retinal flow. This is consistent with a novel theory (van den Berg & Beintema 1997) and physiological findings (Bradley et al. 1996; Shenoy et al. 1996) that the compensation involves modulation of flow-related activities by eye velocity signals.


Adapted from: Beintema, Hooge and van den Berg, in preparation.

## 3.1 Introduction

The eye is mounted on top of a hierarchy of mobile supports, such as the legs, hips, shoulders and the head. With respect to each body part, the position and motion of the world continuously changes. To represent visually acquired information in head-centric or other coordinates, as required in navigational tasks such as perceiving the direction of heading, the visual information needs to be combined with the eye's orientation (eye position) and rotation velocity (eye velocity) relative to the frame of reference.

Translation of the eye relative to points in the environment causes their projections on the retina to flow from the direction of heading. The heading direction could therefore be retrieved by motion sensors that locate the center of flow (CF) (Gibson 1966). Rotation of the eye, however, produced for instance by pursuit of a nearby point, adds rotational flow to the retinal motion. In the changed retinal pattern, the CF is shifted in the direction of the eye movement, so the CF and the heading direction no longer project to the same retinal point. Heading would therefore be incorrectly perceived, unless a compensation for the eye's velocity takes place. In terms of reference frames, the compensation for eye velocity with respect to the head requires a transformation of retinal flow seen by the rotating eye, to head-centric flow as if seen by the eye that is stationary in the head. Note, the rotation problem does not refer to the compensation for changed eye position.

Various mechanisms have been proposed to solve the rotation problem using only the retinal flow vector field as input (Longuet-Higgins & Prazdny 1980; Koenderink & van Doorn 1981; Rieger & Lawton 1985; Heeger & Jepson 1992; Lappe & Rauschecker 1993; Perrone 1992; Beintema & van den Berg 1998a). Irrespective of the mechanism, however, the ability to separate the rotational flow from translational flow will ultimately depend on the distances of points in the scene, field of view, and the ratio of ego-speed and rotation magnitude (Koenderink & van Doorn 1987). Especially useful to differentiate rotational flow from translational flow is motion parallax, i.e. the differences in retinal velocities that occur due to the translation of points at different depths. Alternatively, the visual system could use extra-retinal eye velocity signals, i.e. efferent copies of motor commands or proprioceptive feedback, to compensate for rotational flow.

Among psychophysicists, the extent to which the compensation is done extra-retinally or purely visually is still disputed. When people view a display of expanding motion while making a pursuit eye movement, they can accurately judge their heading correctly (just notable difference of about 1.5°) up to about 10 °/s rotation, showing only a slight bias towards the direction of pursuit (Warren & Hannon 1990; van den Berg 1992; Royden et al. 1992, 1994). Without the help of extra-retinal cues, tested by presenting the same retinal flow pattern to a now stationary eye (simulated eye rotation), some scientists report large biases in the direction of simulated eye rotation at rotation rate higher than 2 °/s (Royden et al. 1993, 1994). Others, however, find observers do actually compensate for the rotational flow (Warren & Hannon 1990; van

den Berg 1992, 1993), even up to 16 °/s (Stone & Perrone 1997). Nevertheless, in contrast to the real pursuit condition, the amount of compensation found for simulated eye rotations is always lower, less robust to the presence of noise (van den Berg 1992), strongly degraded in absence of motion parallax (Warren & Hannon 1990; Regan & Beverley 1982; Rieger & Toet 1985) and more variable among subjects (van den Berg 1996a). Therefore, one may conclude that whenever a pursuit eye movement is made, the visual system seems to make use of the extra-retinal velocity signal.

We sought to test two possible schemes in which the extra-retinal eye velocity estimate is used. First, we explain these schemes. Then we derive a paradigm to discriminate between the two proposed schemes. Finally, we show the results of the heading experiment and compare them with both models's predictions.

## 3.2 Two extra-retinal models

### Vector subtraction model

To correctly perceive the direction of self-motion, the brain must have access to a representation of the flow pattern that remains invariant under changes in the eye's rotation velocity. One can attain invariance through compensation of the visual activity evoked by the rotational component of the retinal flow. Such compensation aims to null the shift of the retinal CF. One way is to reconstruct the flow field of the purely translating eye by subtracting, at each retinal location, a vector that corresponds to the eye rotation (Royden et al. 1994) (Fig. 3.1a). This requires interaction between an eye velocity signal and local retinal motion signals across the entire retinal map. Physiological studies suggest, however, that such an interaction is rare at the level of the local motion detector, but quite common at higher stages along the motion pathway like area MST in the monkey (Wurtz et al. 1990). At these stages, cells have wide fields of view and prefer different motion directions in different retinal regions. These cells resemble motion templates (Beintema & van den Berg 1998a). Vector subtraction, however, cannot work for a template, because its activity does not represent a single retinal motion vector.

### Velocity gain model

Recently, an alternative to vector subtraction was proposed in which the compensation takes place at the level of motion templates (van den Berg & Beintema 1997; Beintema & van den Berg 1998a) (Fig. 3.1b). This approach assumes that not only extra-retinal, but also retinal evidence for eye rotation is needed for compensation.

The model uses an array of retinal motion templates, each tuned to a different center of expanding flow. The activity in such templates changes when the eye rotates. This change in activity, however, is predictably related to the eye's rotation and can be countered as follows. Consider a template that is maximally activated (activity $O$), because the retinal flow matches the template's preferred expansion. If through

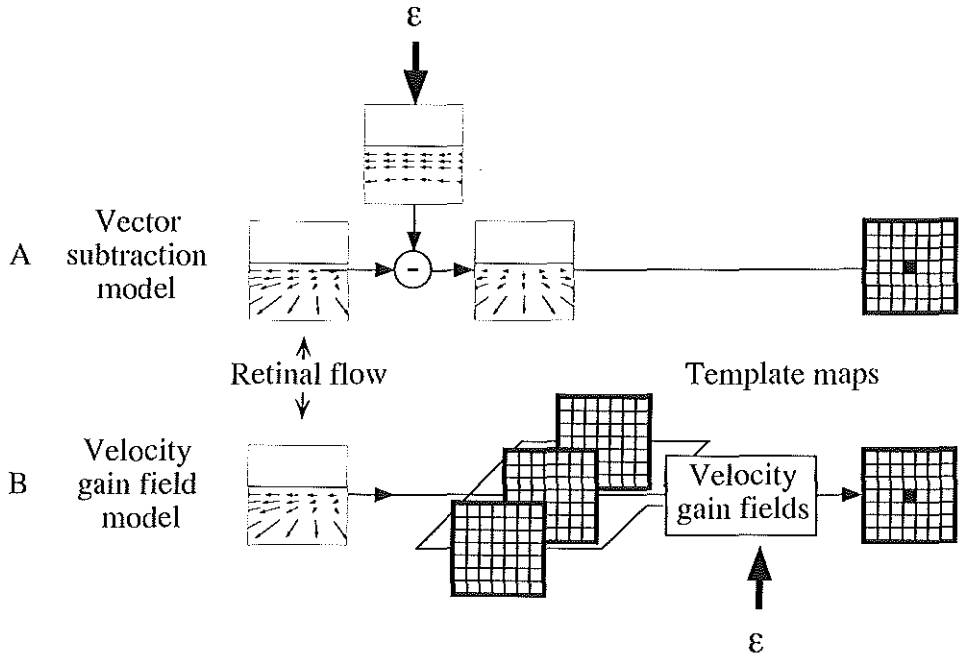



**Figure 3.1** *Two proposed ways in which an extra-retinal signal ($\epsilon$) compensates for rotational flow. Each unit in the map indicates a template that is tuned to one heading direction. The most active unit that prefers the flow corresponding to the presented heading direction is marked. To achieve a map of template activities invariant for eye rotation, either (a) the equivalent flow field of an eye rotation is subtracted from the input retinal flow, or (b) the equivalent change in template activities due to eye rotation (velocity gain fields) is subtracted from the input templates.*

an eye movement a rotational component ($R$) is added to this flow, the template's activity will diminish (Fig. 3.2). This change in activity ($\Delta O$) is described to first order by the partial derivative of $O$ to that component of rotation, multiplied by an eye velocity signal ($\epsilon$) that specifies the rotation about the same axis:

$$\Delta O = \epsilon \frac{\partial O}{\partial R}$$

The partial derivative ($\partial O/\partial R$) indicates how much the activity changes ($\delta O$) for a unit change in the rotational flow ($\delta R$). Subtracting the change ($\Delta O$) from the retinal template's activity ($O$) one arrives at a head-centric flow template ($H$), of which the activity remains invariant for a range of eye velocities. This range is mainly determined by the broadness of the template's tuning to rotational velocity (Beintema & van den Berg 1998a). The partial derivative term ($\partial O/\partial R$) is approximated by

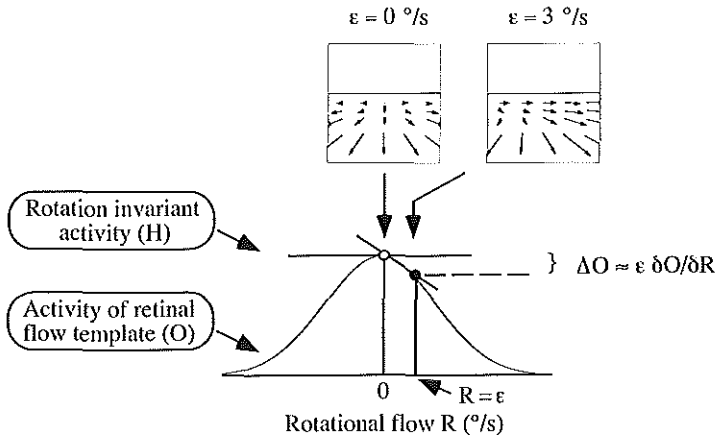



**Figure 3.2** *Response of a retinal expansion template O and a template with rotation invariant activity H as a function of the rotational flow R. The retinal flow and the magnitude e of the extra-retinal signal are shown for a stationary and a rotating eye. When the eye rotates, the activity O is reduced (solid dot). However, one can reconstruct the activity that would be seen by the stationary eye (open dot), by adding an appropriate activity (ΔO). ΔO is equal to the slope (∂O/∂R) measured at R (solid dot) times the amount of rotation R, as judged from the rotational flow R or the extra-retinal signal ϵ.*

subtracting responses of two templates that prefer opposite directions of rotational flow (van den Berg & Beintema 1997; Beintema & van den Berg 1998a) (Fig. 3.3). For rotation rates smaller than the template's tuning width, the activity of $\partial O/\partial R$ increases linearly with the magnitude of rotational flow, so we refer to the derivative term as the visual estimate of the eye's rotation. Provided the three templates respond identically in all dimensions other than the rotational flow, this arrangement provides the correct compensation independent of visual parameters like contrast or the number of elements in the display. Because a multiplicative interaction is proposed between visual and eye velocity signals, we call this scheme a velocity gain model.

### Difference between models

For compensation to occur, the velocity gain model requires both an eye rotation signal ($\epsilon \neq 0$) and visual evidence for eye rotation ($\partial O/\partial R \neq 0$) owing to their multiplicative contribution to the compensation activity ($\Delta O = \epsilon\, \partial O/\partial R$). On the other hand, for the vector subtraction model to achieve compensation, the presence of an eye rotation signal ($\epsilon \neq 0$) is sufficient. During a normal pursuit eye movement, when both visual and extra-retinal sources indicate rotation of the eye ($\Delta O = \epsilon\, \partial O/\partial R \neq 0$), either model will compensate for eye rotation. However, when during a pursuit eye movement

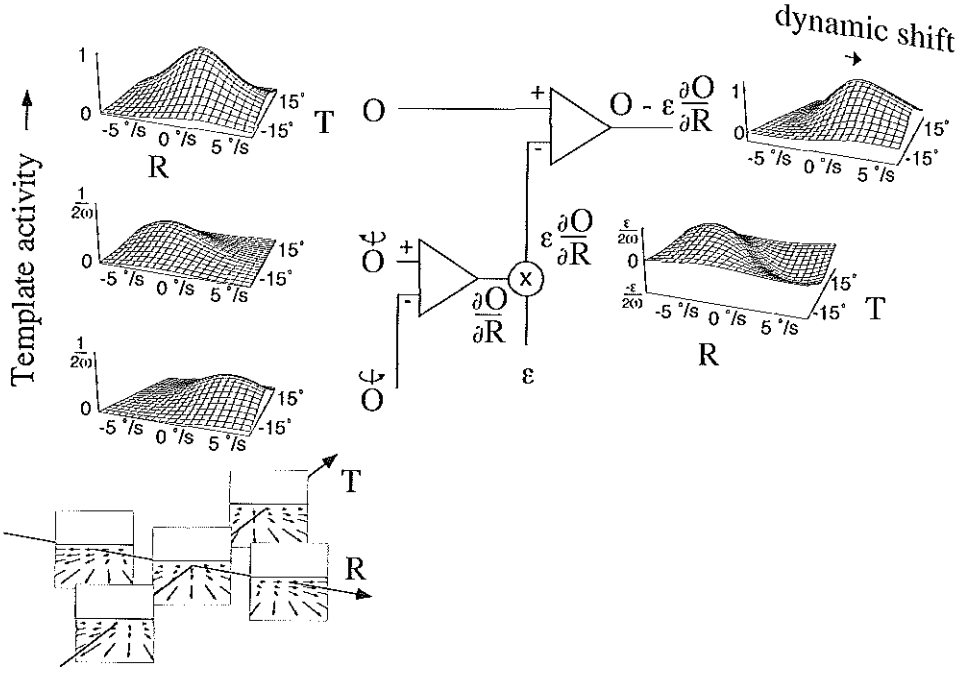



**Figure 3.3** *Velocity gain field model and responses of motion templates to retinal flow. The flow corresponds to forward motion ($T$) combined with ego-rotation about the vertical axis ($R$). To demonstrate the shift of the CF due to rotation, the flow patterns corresponding to specific combinations of $R$ and $T$ are shown below. Left: Responses of three different input units tuned to expansion and rotation. Their preferred rotational speeds ($\omega$) are 0 °/s (top), -3.75 °/s (middle) or 3.75 °/s (bottom). Lower-right: Compensation unit. Top-right: unit invariant to eye rotation, because the preferred rotational flow ($R$) shifts dynamically with the eye rotation speed ($\epsilon$).*

no visual evidence for eye rotation is present ($\partial O/\partial R = 0$), the velocity gain model predicts no compensation, despite the non-zero extra-retinal signal ($\epsilon \neq 0$). Such situation is created by pure expanding motion on the retina, as this equally activates the two templates tuned to opposite directions of rotational flow. In contrast, the vector subtraction model predicts a compensation unconditionally, even if no rotational flow needs to be compensated.

To investigate whether compensation takes place unconditionally or not, we tested perceived heading from pure expansion on the retina when the eye moves. Subjects pursued a moving fixation marker by eye, while an expanding pattern was projected on the screen, such that the CF rotated along with the fixation marker (condition rCF) about an axis through the eye (Fig. 3.4). The purely expanding pattern simulated ego-motion through a cloud of dots extending in depth, because only such stimulus

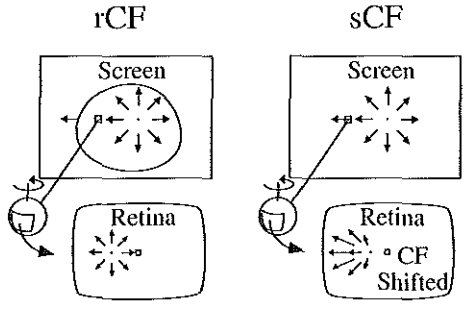



**Figure 3.4** *Test and control stimulus. In the test condition (rCF) the expanding flow pattern moves along with the fixation marker (square). The retinal CF does not shift. In the control condition (sCF) the expanding flow pattern remains stationary on the screen, so the retinal CF shifts. The line encircling the flow pattern in condition rCF used to indicate the movement of the flow pattern and fixation marker as a whole.*

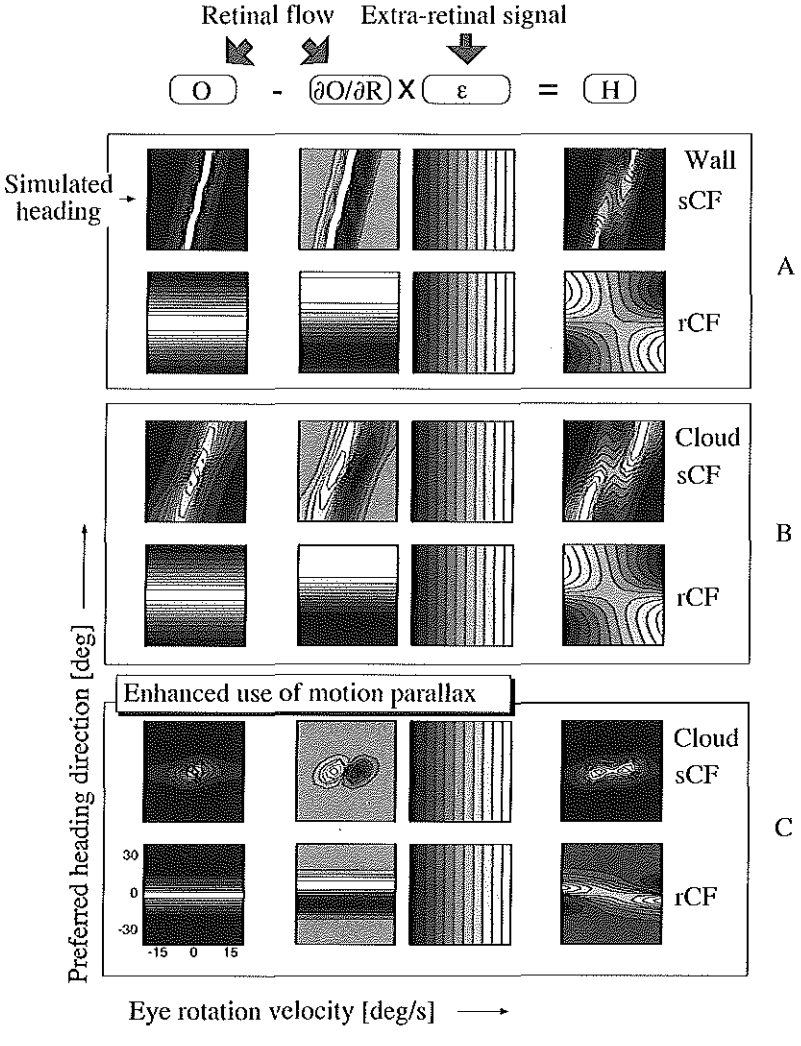
specifies unambiguously that the eye is not rotating. Because the amount of compensating shift may depend on how well eye rotation is represented by the extra-retinal activity, we included a control condition (sCF) in which the CF remained stationary on the screen. This enabled us to compare the amount of compensation between conditions in which retinal rotation was present (rCF) versus a condition without such flow (sCF).

### Motion parallax

Our experiment discriminates only between the vector subtraction and velocity gain model when depth differences are present in the scene. Without depth differences, the velocity gain model also predicts large shifts in condition rCF. Here, we explain why.

We restrict the model to compensation for rotation of the eye about the vertical axis in an array of templates that each prefer a different horizontal heading direction. How invariance for eye rotation is achieved for simulated movement in forward direction is shown in Fig. 3.5. The density plots show the activity in an array of templates, each tuned to a different heading direction. Along the horizontal axis the eye rotation is varied.

Ideally, each 'O' template is Gaussian tuned to heading, independent of its Gaussian tuning to the amount of rotational flow. Subtraction of the compensating activity ($\partial O/\partial R$) should result in rotation invariant activities ($H$). That is, in both condition sCF and rCF a sustained maximum activity (white) should be seen at 0° heading direction at each eye rotation, thus creating a horizontal ridge of activity. However, without depth differences (Fig. 3.5a, fronto-parallel plane), the independent tuning




to rotation and direction of translation breaks down, because the flow becomes too ambiguous. In conditions sCF, each rotation rate causes maximum activity at '$O$' and '$\partial O/\partial R$' templates with preferred heading direction other than simulated. This shift of the maximally activated '$O$' template depends on the rotation rate, hence diagonal ridges occur in the figure. Nevertheless, the eye velocity signal has a compensating effect, as it promotes the template with preferred rotational flow that best corresponds to the actual eye rotation. Within a certain range of eye velocities related to the templates' tuning width (Beintema & van den Berg 1998a), the output template ($H$) with preferred heading direction 0° is indeed promoted (horizontal ridge) in condition sCF. In condition rCF, however, the output template with a preferred

**Figure 3.5** *Effect of adding depth and enhancing the model's use of motion parallax on heading and rotation tuning of templates. Subsequent panels show the activation for approach (2 m/s) of (a) a wall at 10.5 m, (b) a cloud extending from 5-20 m, or (c) the same cloud with enhanced use of motion parallax. Vertically, each density plot shows an array of activities of templates, each tuned to a different horizontal heading direction, or the extra-retinal activity (e). Horizontally, the eye rotation varies about a vertical axis. With retinally stabilized flow (rCF), the activity of the retinal flow templates $O$ (left column) and $\partial O/\partial R$ (2nd column from left) does not change as eye rotation changes. For sCF conditions the eye rotation adds rotational flow and the activities do change. The simulated heading direction always corresponds to forward movement. Each density plot is rescaled, so black and white represent lowest and highest activity, respectively. The scheme at the top illustrates how vertically corresponding arrays combine.*

heading that differs from actual heading direction is promoted, resulting in errors in perceived heading. With depth in the scene (Fig. 3.5b, cloud), the presence of motion parallax reduces the ambiguity, so the output template preferring heading direction 0° dominates more. To enhance this domination requires further exploitation of motion parallax (Fig. 3.5c), which we explain in Appendix A.

## 3.3 Experiment

### 3.3.1 Methods

Five subjects monocularly viewed stimuli that were back-projected (Sony VPH 1270QM projection television) on a translucent screen (48 × 46°, distance 2 m) in a completely darkened room. To prevent head movements a chin and forehead rest were used. Stimuli were generated on a SGI workstation. The simulated world consisted of 256 white dots. Relative to the subject's viewing point, each dot was assigned a random radius (5.0 to 20 m), horizontal angle (-36 to 36°) and vertical angle (-23 to 23°).

A trial started with the presentation of the stationary world while a red fixation marker rotated separately at constant speed about the vertical axis through the subject's viewing point. Subjects pursued this fixation marker by eye until the end of the trial. After 1.5 s, when the marker passed the center of the screen, forward movement of the observer was simulated at a speed randomly chosen between 1.5 and 2.5 m/s. After 1.0 s, the dots and fixation marker stopped, upon which a red pointer appeared. Subjects were given arbitrary time to align the pointer with the final perceived direction of self-movement. In between trials a white screen was presented for 30 ms. In total 280 different trials were run: 5 rotation velocities (0, ±1.5, ±3 °/s), 4 heading directions (±5, ±10°), 2 conditions of simulation (rCF, sCF) and 7 repetitions.

We used brief (1 s) presentation of flow and relatively low pursuit eye movements to minimise the change in head-centric heading direction for the rCF condition. This

displacement was no more than 3°. We believe this has been instrumental to keep naive subjects unaware of the difference between the conditions of simulation. In fact, even the authors had difficulty to keep trials with rCF and sCF stimuli apart.

#### Eye movement recordings

Eye movement was recorded in a control session using the scleral coil technique and the same stimuli. A bite board was used to prevent head movements. During off-line data analysis the velocity of the pursuit eye movement was computed after removal of saccades. Gain is defined as the ratio of the smooth eye velocity to the velocity of the fixation marker. The pursuit remained constant throughout the whole trial, with an average gain 0.8 for 5 subjects. We found the same pursuit eye movements in both conditions rCF and sCF.

### 3.3.2 Results

In either condition the perceived heading deviates only slightly from the final simulated heading, even at the highest rotation rates (Fig. 3.6a, b). To separate a bias towards the fovea that depends on simulated heading eccentricity, from a shift independent of the simulated heading, we fitted linear functions to the data at each rotation rate. The regression slopes (Fig. 3.6c) are lower than 1.0 and indicate a bias towards the fovea, independent of rotation rate. Such bias is consistent with previous findings that subjects mostly underestimate their heading eccentricity, though it may vary largely between individuals (van den Berg 1996a). The intercepts (Fig. 3.6d) in condition sCF, reveal a small pointing bias in the direction of the eye's movement, independent of the simulated heading. This shows that compensation for the eye rotation does occur, although not completely, when the extra retinal signal and the retinal rotational flow are present. In condition rCF, the heading seems to shift in the opposite direction of eye rotation. The direction of this error is consistent with vector subtraction, but also with a velocity gain model without using motion parallax. How does the magnitude of the error compare to both models?

#### Vector subtraction model

For vector subtraction, the predicted shift of the center of flow due to rotation is difficult to estimate for points lying at varying distances. For points lying on a sphere centered on the eye (radius $d$), however, the shift of the singular point is well defined. During forward translation (speed $T = 2$ m/s) the retinal speed of a point at eccentricity $\phi$ is $p = T/d\ sin(\phi)$ rad/s. When the eye rotates with speed $\epsilon = p$, the flow at $\phi$ is nulled, displacing the focus from the fovea towards the point at eccentricity $\phi$. Hence, the shift equals $\phi = asin(\epsilon\, d/T)$. Thus, assuming each point contributes equally, we used as a substitute for the cloud extending from 3-18 m at the end of the trial, a sphere with radius equal to the average distance $((3 + 18)/2 = 10.5$ m).

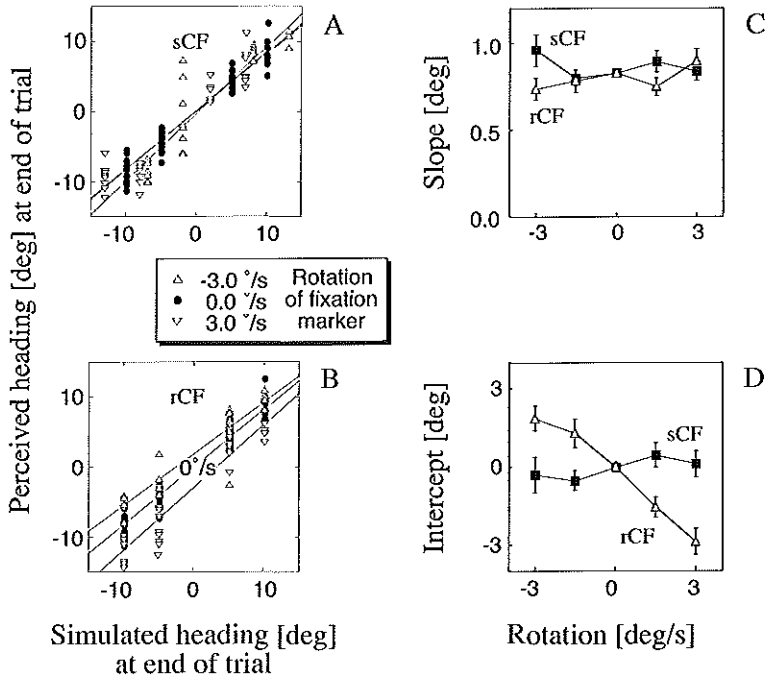



**Figure 3.6** *Data of one subject. Perceived versus simulated heading direction in angles relative to the direction of the fixation marker at the end of trial. Perceived heading at 3 rotation rates of the fixation marker are shown in condition (a) sCF and (b) rCF. The coefficients of linear regression fits at each rotation rate, split the error in perceived heading into (c) a bias towards the fixation point and (d) a constant shift relative to the fixation point (rCF: triangles; sCF: squares). Rotation rate has no effect on the bias, whereas it does cause a linear shift in the direction of pursuit in condition sCF and a somewhat larger shift in opposite direction in condition rCF.*

We modelled the incomplete compensation in condition sCF by assuming the extra-retinal signal slightly underestimates the actual eye rotation, thus being a proportion (0.93) of the eye's actual velocity. Effectively, this decreases the magnitude of the predicted error in condition rCF.

As shown in Fig. 3.7, in condition rCF the shift averaged over all subjects is much less than expected by vector subtraction (dashed lines). One could, however, argue that the observed shift might be consistent with vector subtraction on the daring assumption that subjects were judging their heading on the basis of the nearest visible points. This follows, because at close range to the eye the expansion of the dots increases relative to the rotational flow, resulting in a smaller shift of the CF as determined from the flow of the nearest dots. However, predicted shifts by the vector subtraction scheme remain too large by a factor of about 1.3, even when based only

on points at 3 m, with the assumption that the extra-retinal signal is 0.75 times the real eye velocity. Also, a different value for the pursuit gain would not decrease the discrepancy between the vector subtraction model and the experimental data, because according to vector subtraction any rotational flow accompanied by an eye rotation signal is compensated. For example, a higher pursuit gain changes the rotational flow and the extra-retinal signal in both conditions equally, so the predicted differences in perceived heading between the condition rCF and sCF are not reduced.

### Velocity gain model

For the velocity gain model, the shift of perceived heading was predicted from the preferred heading direction of the most active template in an array of output templates, as demonstrated above in Fig. 3.5. For details on the model and how each local flow vector contributes to the template's response, we refer to (Beintema & van den Berg 1998a). Essentially, there are only two free parameters; the template' broadness of tuning to rotational flow ($\omega$) and its motion parallax sensitivity ($\sigma_{var}$). The template responses and extra retinal signals were combined to obtain the rotation invariant template, using the adapted first-order approximation ($H = O - 1.55\,\epsilon\,\partial O/\partial R$). The templates were shown the simulated flow in the experiment at the end of the trial (distances 3-18 m, 112 dots visible).

In contrast to the vector subtraction model, the velocity gain model could fit both conditions sCF and rCF using $\omega = 3.75$ °/s and $\sigma_{var} = 0.45$ °/s (Fig. 3.7, solid lines).

Shifts in perceived heading opposite to the eye movement have been reported previously, in a condition similar to our rCF condition (Banks et al. 1993). Dr. Banks et al. were so kind to share some of those data with us. They used a similar range of heading directions (-5 to 5°) and ego-rotations (-4 to 4 °/s), and the same presentation time (1 s). They used a smaller ego-speed (0.5 m/s), range of distances (0-5 m), number of dots (30) and display size (20 by 20°). In the stabilised condition, they also find shifts in perceived heading of about 3° at 3 °/s rotation. To compare their results with the models, however, we must take into account two important factors. Firstly, in condition rCF they presented truly stabilised expansion on the retina by feeding back the eye's position. Secondly, the distribution of points used by Banks et al. was homogeneous, ranging from 0 to 5 m from the eye. Although the nearest point distance is now zero, on average points will lie even further away than in our radially distributed cloud, and are therefore more affected by rotational flow. The average distance of visible points[1] (3.75 m) leads to a predicted shift by vector subtraction five times higher than observed in the data (extra-retinal signal of 0.9 times the real pursuit velocity, and smooth pursuit gain of 0.8). Only a selection of points at about

[1] Given the dot density is constant, the number of visible dots is proportional to a pyramid's volume of which the top lies at the eye, and the base increases its area quadratically with distance $r$. The pyramid's volume equates to the sum of slices with thickness $dr$ and area $r^2$: $\int r^2 dr = \frac{1}{3}r^3$. The mean distance of points equates to the weighted sum of slice distances, each slice weighted according to its volume as a proportion of the whole pyramid: $\int r^3 dr / \int r^2 dr = \frac{3}{4}r$.

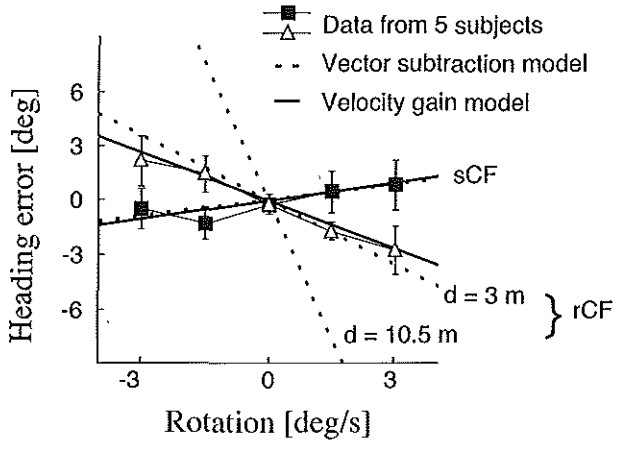



**Figure 3.7** *Modelled and measured shift of heading relative to simulated heading as a function of target rotation rate in conditions rCF and sCF. The data show the mean and error in the mean of 5 subjects (rCF: triangles; sCF: squares), individual shifts obtained from intercept as in Fig. 3.6. The remaining lines indicate the shifts predicted by the velocity gain model (solid lines) or by the vector subtraction model (dashed lines). The latter are based on the average distance of points (10.5 m) or based on the nearest points (3.0 m). Free parameters of each model were chosen to fit the sCF data.*

1.5 m or less could explain their data by vector subtraction. Given less than 2 out of the 30 dots enter that range throughout the whole trial[2], such selection seems a highly unlikely explanation. Fits to Banks et al.'s data with our velocity gain model were obtained by setting the sensitivity to motion parallax cues at $\sigma_{var} = 0.18$ °/s. The somewhat enhanced use of motion parallax in comparison to that used to fit our data ($\sigma_{var}$ =0.45 °/s) may due to other differences in the stimuli, such as the smaller number of dots or smaller display used by Banks et al.

## 3.4 Discussion

We tested whether extra-retinal signals lead to unconditional compensation in perceived heading from flow, or whether visual evidence of rotation is also taken into account. Two models were compared; vector subtraction, requiring only extra-retinal signals for compensation, and the velocity gain model, requiring both extra-retinal and visual estimates of rotation consistent with an eye rotation. To distinguish these two

[2]During the trial 0.5 m distance was covered. Points at an initial distance of 2 m or less could have moved through the plane at 1.5 m distance from the eye. In a pyramid's volume of 5 m, the fraction of dots within 2 m amounts to $(2/5)^3 = 0.064$ (see footnote 1)

models, we measured the shift in perceived heading during a pursuit eye movement, while the amount of rotational flow on the retina was varied. This paradigm differs from the 'simulated versus real rotation' paradigm, in which the contribution of the extra-retinal signal is varied, while the retinal rotational flow is kept constant. Without a rotational component in the flow (condition rCF), the compensation by the extra-retinal signal is strongly reduced, in contrast to the vector subtraction model, even under the assumption that people can select the nearest points and that the extra-retinal signal is underestimated. The data are, however, in line with the prediction of the velocity gain model.

The velocity gain model relies on a non-linear interaction between local motion detectors within the receptive field of the motion template. This enhances the effect of motion parallax to distinguish the template's response to rotational and translational flow. Note, the visual disambiguation between rotational and translational flow is a graded phenomenon that depends on the depth distribution in the scene and the speed of simulated forward motion. Therefore, the velocity gain model predicts for a given eye rotation increasingly larger errors in the same direction as the vector subtraction model, when (1) decreasing the ratio of simulated speed over distance, (2) reducing the depth range, or (3) reducing the extent to which the model exploits the motion parallax.

The velocity gain model as presented here was meant to describe how retinal flow signals can be adjusted by an eye velocity signal for the perception of heading. It does not address the issue of how heading direction can be extracted from only the retinal flow, nor how visual and extra-retinal estimates of ego-rotation might combine. A purely visual model would predict almost no error in perceived heading in condition rCF. To reduce the need for enhanced use of motion parallax, as required to fit the rCF data with an extra-retinal velocity gain model, one could also make the extra-retinal eye velocity signal $\epsilon$ depend more on the visual evidence for rotation ($\partial O/\partial R$). Not using motion parallax, though, predicts for both the purely visual model and the extra-retinal velocity gain model errors larger than found in the real eye rotation condition sCF. Therefore, a future model likely replaces the extra-retinal eye velocity term $\epsilon$ by an estimate of the ego-rotation that depends on visual and extra retinal sources and enhanced use of motion parallax.

Can the small difference in observed heading errors between condition rCF and sCF to some degree be attributed to visual or motor processing delays? We distinguish three types of delay. Firstly, in condition sCF, the retinal heading direction continuously changes its position on the retina. A delay in the update of the retinally perceived heading direction causes an error in the direction of eye rotation only in condition sCF, since in condition rCF the heading direction remains at the same retinal position throughout the trial. Therefore, such delay would only increase the difference in predicted errors between condition rCF and sCF. Secondly, a delay in the transformation from retinal to head-centric position, caused for instance by a delay or underestimation of the eye position signal, predicts an error opposite to the direction

of eye rotation. Such delay, however, affects both conditions sCF and rCF equally, therefore not reducing the difference between the two conditions. Thirdly, a delay in the processing of the head-centric heading (subjects point to the heading direction at the beginning of the trial), would predict only an error in the opposite direction of rotation in condition rCF, also increasing the difference between conditions rCF and sCF. Therefore, none of the above delays would bring a unconditional vector subtraction model closer to the data.

The velocity gain model is consistent with our results and with a number of recent neurophysiological findings. Multiplicative modulation by eye (Bradley et al. 1996) and head (Shenoy et al. 1996) velocity of MST cells, tuned to expanding motion, has been observed in the monkey, possibly reflecting a processing step towards the computation of $\Delta O$. Also cells types were found (Shenoy et al. 1996) that do not modulate in amplitude, but shift their preferred CF during eye pursuit in a range from non-compensating to fully compensating. This behaviour would be expected from cells at the in- and output stages of the transformation for eye rotation, respectively, i.e. the retinal flow templates ($O$ and $\partial O/\partial R$), and head-centric flow templates ($H$). Moreover, many cells in area MST shift their preferred location of the center of radial motion when expansion is changed into contraction (Lappe et al. 1996). Our simulations of the partial derivative $\partial O/\partial R$ also exhibit this behaviour (Beintema & van den Berg 1998a).

A mechanism to explain shifting cells during pursuit, with gain modulated cells was also proposed by Bradley et al. (1996). They showed that a shifting cell could be fitted by the sum of sinusoidal response functions of three non-shifting cells that each are differently modulated in amplitude by the pursuit. Their explanation does not address such issues as the properties of the flow or how the tuning properties of cells arise. The idea may work for a chosen set of parameters of ego-motion and stimulus, but not at other rotation rates, ego-speeds or layouts of the scene.

In our model, we do specify the tuning properties of the inputs by formulating a number of constraints, based partly on a computational transformation scheme, and partly on physiological and psychophysical data. The transformation scheme is an extension of Koenderink's (1988) theory on receptive field shifts to the domain of encoding heading direction from flow. This theory describes how a shifted Gaussian receptive field is approximated by the sum of zeroth, first and higher order derivatives of a Gaussian function, each multiplied by appropriate weights that depend on the amount shift. We linked this multiplicative interaction with the increasing evidence for gain field like interaction found in the monkey parietal cortex (Andersen 1997). This area seems to be involved in transforming labeled-line-coded sensory maps from retinal to other frames of reference using the rate-coded motor signals. The principle of compensation through visual partial derivatives multiplied by extra-retinal signals provides a novel explanation why eye position gain fields, as found in the parietal cortex of the monkey (Andersen & Zipser 1988), are an intermediate step in the transformation of visual directions from retinal to other frames (van den Berg &

Beintema 1997, Beintema & van den Berg 1998a). This suggests the brain may use similar processing steps for velocity and position transformations.

In conclusion, we have shown that a compensation for eye rotation through unconditional subtraction cannot account for the small heading errors when rotational flow on the retina is absent. In contrast, our model with eye velocity gain fields can explain the data. The model also suggests an important role for motion parallax. Future studies need to clarify to what extent an extra-retinal estimate of rotation can be replaced by a visual estimate of rotation.

## Acknowledgements

This work was supported by research grant SLW 805-33.171-P from the Dutch Foundation for Scientific Research (NWO) and by the Human Frontier Science Program (RG 34/96). We greatly acknowledge Marty Banks for sharing his data with us.

## 3.5 Appendix

In this appendix we show how templates can use motion parallax to increase the differences between the effect of an eye rotation or a change in heading direction on its activity (Beintema & van den Berg 1998a).

During linear forward translation the retinal flow expands from the direction of heading (Fig. 3.8, left). The retinal velocities vary according to their distance to the eye. The same translation combined with leftward eye rotation, gives a retinal pattern in which the CF lies more towards the left, resembling retinal flow that would occur had the heading direction been more to the left (middle patterns). Although the patterns in the middle are similar, there are differences. To bring these out more clearly, we compute at each location the difference vector between each middle-pattern and the pattern at the left. This procedure results in a flow pattern that either resembles rotational flow (lower panel) or is exactly equal to rotational flow (upper panel). The difference vectors resulting from a change in heading direction (bottom-right) are not consistent with an eye rotation, because not only do they lack the slight change in orientation at large eccentricities, but most notably they show large variation in magnitude owing to the varying distance of the points. This variation does not occur if a rotation is made.

How can a single template exploit the presence of motion parallax to enhance its tuning to heading direction? We take a template that prefers the original flow field (left). If it is presented with flow that differs from its preferred flow (middle), the activity of the template is diminished according to the difference flow field (right). These local difference vectors may be looked upon as 'local estimates of rotation'. We propose that the template is not only sensitive to the difference vectors per se (preferring these to be zero or some constant value), but also to the variation in these

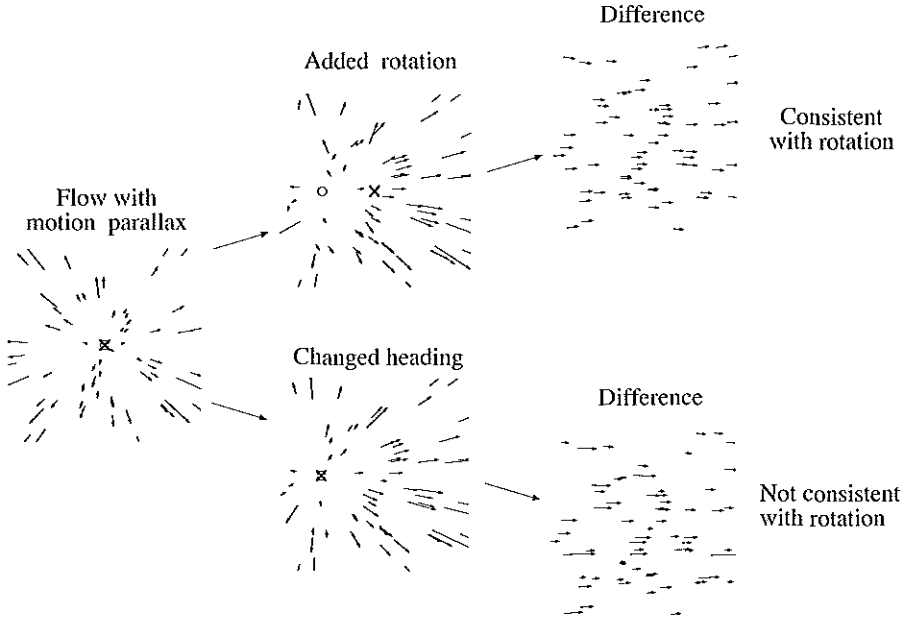



**Figure 3.8** *The effect of an eye rotation or a change in heading direction, on the retinal flow that results from ego-translation (2.0 m/s) towards a cloud of dots (3-18 m distance). Left: retinal flow corresponding to forward linear movement. The heading direction (cross) and CF (circle) both lie at the fovea. Middle: two flow patterns, each with the CF at the left, corresponding either to forward ego-translation and 3 °/s leftward rotation (top), or to ego-translation 15° to the left (bottom). Right: subtraction of the flow patterns at the middle, from retinal flow at the left side. A particular flow vector in the difference resulting from eye rotation (top) has similar orientation, but very different magnitude from a difference vector at the same retinal location in the difference resulting from a change in heading direction (bottom).*

difference vectors at different locations. According to the above, a measure of the scatter in these 'local estimates of rotation', especially the variation in magnitudes, provides a measure to the deviation from the template's preferred heading direction. Each template's response is inhibited increasingly with the amount of scatter, thus promoting the template that is tuned to the presented heading direction. Such a mechanism would enable the template to sharpen its tuning to heading direction, irrespective of the amount of eye rotation. The approach fails in the absence of depth differences because then the variation is very limited.

# Chapter 4

# Pursuit affects precision of perceived heading for small viewing apertures


## Abstract

We investigated the interaction of extra-retinal with retinal signals in order to perceive heading correctly during rotation. An important source of variability in perceived heading arises from the pattern of local motion directions and their uncertainties. For limited apertures, radial flow allows higher precision than parallel flow for the same directional uncertainty in local motion. Eye pursuit while moving towards a frontal plane transforms the pattern of retinal flow within the aperture. Without compensation for that change, perceived heading would be shifted in the direction of pursuit and its precision changed accordingly. Compensation by the extra-retinal signal potentially reduces the shift. Can it also counter the changed precision due to changes in the retinal flow pattern? It might, if at the integration stage the rotational flow has been subtracted out already. A compensation by the extra-retinal signal beyond the integration stage, however, cannot undo the change in retinal directional information, and would predict an effect of pursuit on precision. We compared perceived heading during fixation and pursuit under a variety of conditions. We find precision is affected by pursuit, whereas only little shift of perceived heading occurs. Furthermore, we find that systematic errors due to rotation decrease significantly with duration (0.5-3.0 s). A 50° increase in aperture size little increased the compensation, but did significantly increase perceived heading eccentricity. We conclude that substantial extra-retinal compensation takes place, but not before spatial integration of local motion signals.



Adapted from: Beintema and van den Berg, Vision Research, submitted.

## 4.1 Introduction

The optic flow generated by our self-motion allows us to retrieve our direction of translation with respect to the world (heading). The optic flow is the array of light projected on an imaginary sphere around the moving eye. During translation, the motion creates a pattern that expands and contracts at points where the eye's translation axis intersects the sphere. Retrieving the direction of heading from the pattern of optic flow, for instance by localising the center of flow (CF), involves an integration (pooling) over the various motion vectors. Visual motion is sampled by sensors mounted on a possibly rotating eye. So, at an early stage the visual system has access only to the representation of retinal flow. During rotation, the retinal flow may be quite different from the optic flow. Failure to compensate for this difference, i.e. the rotational flow, could lead to systematic errors in our perceived heading. We wondered at which stage, before or after pooling motion, such compensation takes place. To this end, we investigated the influence of rotational eye movement on both the accuracy (systematic errors) and the precision (random errors) with which humans perceive their direction of heading.

The brain has access to two sources of information on the eye's rotation velocity relative to the world. The first one is the retinal flow itself, which we here refer to as the retinal signal. The second one is the extra-retinal signal, a generic name for efferents to the eye muscles and afferents from proprioceptive and vestibular sensors. The role of the retinal and extra-retinal signal has been investigated extensively using the so called simulated and real eye rotation paradigm. Observers judge their heading from a display of motion, while keeping the eye fixated on a stationary point or while making a real pursuit eye movement. In either case, the displayed motion from the eye's perspective is made to look as if the observer translates through a dot scene while simultaneously rotating the eye. When rotation is simulated, people can judge the simulated direction of heading quite accurately for scenes extending in depth (e.g. Rieger & Toet 1985; Warren & Hannon 1988; van den Berg 1993), but not when approach to a fronto-parallel plane is simulated (Regan & Beverley 1982; Rieger & Toet 1985; Warren & Hannon 1988, 1990; Royden et al. 1994; Grigo & Lappe 1999; Beintema & van den Berg 2000). In the latter case, systematic errors are related in magnitude and direction to the shift of the center of retinal flow caused by the rotational flow. In contrast, during real eye pursuit, little or no systematic errors occur (Warren & Hannon 1990; Royden et al. 1992, 1994; Banks et al. 1996), even without depth in the scene (Warren & Hannon 1990; Royden et al. 1994). These findings show that the retinal signal itself can be adequate to compensate for the rotational flow, given sufficient depth in the scene, and that the extra-retinal signal also plays an important role.

Although the retinal and extra-retinal representation of eye rotation may be used in similar way (Beintema & van den Berg 1998a), in this study we investigate the compensation on the basis of the extra-retinal signal only, as to simplify the analysis. By simulating approach towards a scene without depth, we eliminate a possible solution

based on a retinal signal. We must point out, though, that depth in the scene is not strictly essential for compensation on the basis of the retinal signal. Recently, Grigo & Lappe (1999) reported heading is accurately perceived during simulated rotation given the field of view is sufficiently large (90 × 90°) and the flow briefly (0.5 s) presented. Moreover, we recently found that the compensation for simulated rotation about the line of sight is virtually complete even for a scene that lacks depth (Beintema & van den Berg 2000). As we here restrict rotation to the vertical axis, and use a very limited field of view, the conditions in our experiment do not allow a compensation on the basis of the retinal signal.

A number of suggestions have been put forward on how the extra-retinal signal might compensate for the rotational flow (Royden et al. 1994; Perrone & Stone 1994; Beintema & van den Berg 1998a; Lappe 1999). They have in common a stage at which local motion signals from different retinal locations are pooled to obtain a signal that responds selective to the pattern of flow directions. Such hierarchical architecture reflects the physiology of the brain (see Maunsell & Newsome 1987 for a review). Monkey brain studies show that signals from cells which primarily respond to local motion (area MT) are collected by cells that respond to flow covering larger parts of the visual field (e.g. area MST). Nevertheless, an important difference exists between models regarding the level at which the interaction between the extra-retinal and the retinal signals takes place.

Royden et al. (1994) suggested that the extra-retinal signal is used to subtract out the rotational flow already at the level where local motion is represented (Fig. 4.1, left). Since what is to be compensated (subtracted) at each retinal location is the representation of a motion vector, we refer to this scheme as the vector subtraction model. Alternatively, the extra-retinal signal interacts after local motion has been pooled (Fig. 4.1, right). What is to be compensated (subtracted) then is not a representation of a retinal vector but of the pattern of rotational flow. We have proposed this would require subtraction of a retinal signal that is gain-modulated by an extra-retinal eye velocity signal, hence the term 'velocity gain field' (van den Berg & Beintema 1997; Beintema & van den Berg 1998a). More specifically, imagine a visual cell that responds to a specific retinal flow pattern, such as expansion for a certain direction of heading. Its activity ($O$) will be changed by an amount ($\delta O$) when a small amount of rotational flow is added ($\delta R$). To compensate for this change, it does not suffice to subtract an extra-retinal signal $\epsilon$, because this signal represents eye velocity, not rotational flow. To obtain a compensation term that has the same dimensions and properties (e.g. preferred heading, tuning width) as the retinal flow activity $O$, it would require subtraction of an eye velocity signal $\epsilon$, multiplied by a derivative signal $\partial O/\partial R$. The latter retinal signal represents how much the visual signal ($O$) changes when a small amount of rotational flow is added to the stimulus. Thus, a vector subtraction scheme or the velocity gain field scheme both recover a representation of the optic flow, but the proposed interactions differ in being additively or multiplicatively and in taking place before or after local motion signals have been

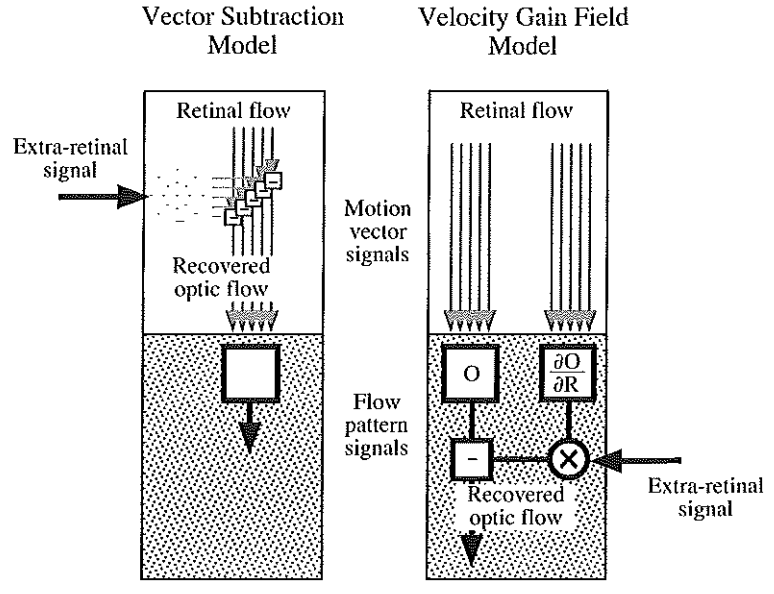



**Figure 4.1** *Two levels of interaction at the extra-retinal signal could interact with retinal motion detectors. Left: The vector subtraction model subtracts corresponding rotation flow vectors before local motion signals are pooled. Right: The velocity gain model uses the extra-retinal signal after local motion signals have been pooled.*

combined.

Various electrophysiological studies have shown an interaction between extra-retinal signals and cells that respond to retinal flow patterns, such as area MSTd (Bradley et al. 1996; Shenoy et al. 1996). Nevertheless, the hypothesis that integration has taken place at the level *before* local motion pooling, potentially at the level of local motion signals (area MT), has to our knowledge not been tested. We aimed to investigate the latter question psychophysically.

Our paradigm was inspired by findings that the precision of perceived heading depends on the pattern of flow. Thresholds for detecting a heading difference between sequentially presented trials, rose once the simulated heading was located outside the visible patch (Warren & Kurtz 1992; Crowell & Banks 1993). Moreover, Crowell and Banks (1993) showed it is the flow's pattern that determines precision, not its location on the retina per se, nor its overall magnitude. This effect is explained by the limited precision with which local retinal motion is detected (e.g. Koenderink & van Doorn 1987). To illustrate this, we plotted a fictitious curve of the variability in perceived heading as a function of the simulated heading, for a non-rotating eye (Fig. 4.2, solid line). The icons just below, along the axis of simulated heading, show the corresponding retinal flow pattern during fixation. Each vector shows a

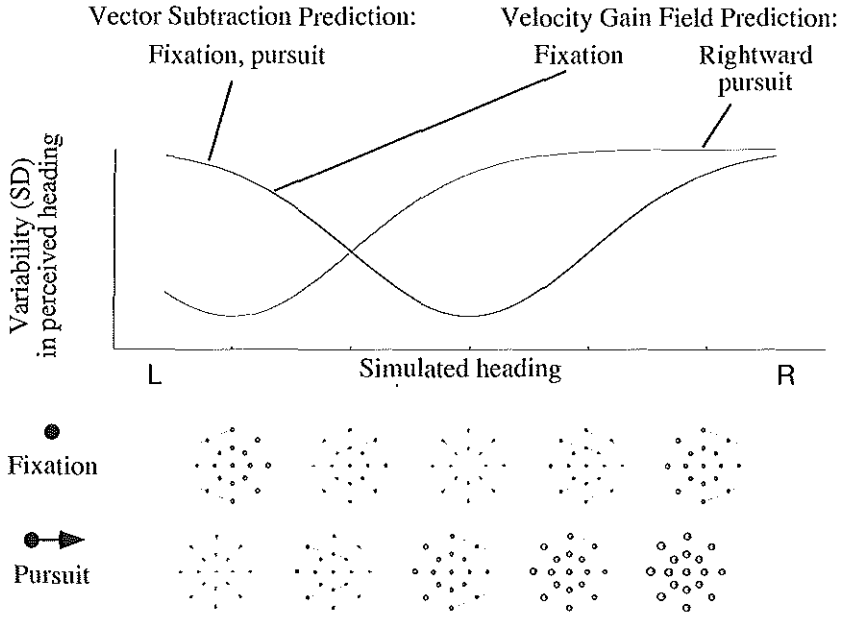



**Figure 4.2** *Predicted variability in perceived heading as function of simulated heading during fixation or rightward pursuit. The icons below show corresponding retinal flow during fixation or rightward pursuit. Also indicated is the hypothetical range of uncertainty in the represented vector direction and magnitude, indicated by a circle. The area indicates possible endpoints of the motion vector, and is a constant proportion (0.1) of the vector's length. The velocity gain field model predicts lowest variability when the center of retinal flow is located within the aperture. The vector subtraction model predicts lowest variability when the center of recovered optic flow is located within the aperture. Therefore, only the velocity gain model predicts a shift of the heading variability curve during rightward pursuit.*

hypothetical uncertainty in the retinal detected motion. The effect of uncertainty in local estimated motion directions on the recovered heading is evident when finding the point at which two motion vectors below and above the horizon intersect. For the same amount of directional uncertainty, the variability in the location of intersection increases when the flow is more parallel.

It is this geometrical relationship between a change in precision of perceived heading and a change in flow pattern that we wished to exploit. The key notion is that the change in precision takes on effect from the first stage at which cells respond to the pattern of flow and thereafter. How would pursuit affect precision? As shown by the second icon row (Fig. 4.2), rightward pursuit changes the pattern of the retinal flow. If compensation takes place after pooling (or not at all), then the cells that pool local motion will be presented with the retinal flow pattern. Therefore, we expect

the variable error curve to be offset during pursuit as much as the CF shift (velocity gain model prediction). On the other hand, if the extra-retinal signal is subtracted already at the local level, then the cells that pool local motion are presented with the recovered optic flow pattern that is the same as the the retinal flow during fixation. In that case, one would expect the variability curve to be the same during pursuit and fixation (vector subtraction model prediction).

Our main objective is to determine whether the extra-retinal signal interacts before or after the pooling local motion. We tested this in Experiment I, by measuring the precision of perceived heading during fixation and pursuit for several simulated heading directions. Pilot work revealed that not only precision but also accuracy was affected by pursuit. Recently, Grigo and Lappe (1999) showed accuracy was affected by motion duration and field of view size when rotation was simulated. This inspired us to vary duration and field of view size in Experiment I, to examine whether these parameters also affect accuracy during real eye movements. Experiment II is a control to explore an alternative explanation why pursuit affects precision.

## 4.2 General Methods

### Stimuli

Stimuli were rendered on a Silicon Graphics Onyx workstation and backprojected (Sony VPH 1270QM projection television, 815 × 611 pixels) onto a translucent 60 × 58° screen (refresh rate 60 frames/s) in an otherwise fully darkened room. Subjects were seated 1.5 m before the screen, with the head supported by a head- and chinrest. The position of the eye relative to the screen was measured for the viewing eye as to present the images in the right perspective (van den Berg 1996a).

Ego-translation was simulated by displaying image motion on the screen that, from the eye's perspective, was consistent with red dots translating along the same direction. The dots formed a fronto-parallel wall that remained parallel to the screen during the entire motion. In the heading experiment, the dot translation had a component towards the subject's eye. The horizontal angle between the simulated translation direction and the screen's normal was defined as the heading direction. In the control experiment, the dot translation was in the fronto-parallel plane.

Small aperture stimuli were created by displaying only the dots that moved within a circular boundary (10° diameter) around the fixation point. On the screen, dots were of constant size (0.2°). The motion lasted for 0.5 s or 3 s. Two types of scene were presented. In the approaching scene, the simulated distance with respect to the eye decreased in time. The scene consisted of dots distributed within a fronto-parallel plane at an initial distance to the eye of 9.0 m. The average number of visible dots was 130 initially, 110 after 0.5 s and 30 after 3 s. In the equidistant scene, the motion of dots did simulate approach, but the mean simulated distance with respect to the eye remained constant in time. The scene was created by distributing dots within a

depth volume of which only dots between the planes at 8.25 and 9.0 m distance to the eye were made visible. Thus, continuously points are moving in and out of the window of visibility, average lifetime of a dot $= 0.75/1.5 = 500$ ms and 20 dots being visible per frame on average. The 3.0-s approaching scene was also presented with large field of view ($60 \times 58°$). The number of visible dots was 440 initially and 110 dots after 3.0 s. The dimensions of the scenes were chosen so that the aperture or screen was filled with dots during the entire presentation.

The pursuit target was a red annulus with innerdiameter 0.4° and outerdiameter 0.8° at eye level. The target was not a part of the scene, its motion simulated rotation at constant speed about the vertical axis through the eye.

### Procedure

The subject was asked to follow the horizontally moving target by eye until the end of dot scene motion. The pursuit target remained visible during the whole trial. It started its motion after having been visible for 200 ms, always ending at the center of the screen. To promote stable pursuit, the target moved 750 ms before the dot scene motion was presented. Immediately after the motion stopped, the dots disappeared and a pointer appeared at the screen center. In the heading experiment, the pointer was a horizontally adjustable red dot of 0.4° diameter. In the control, the pointer was a red line, adjustable in orientation. Having adjusted the pointer according to the task, the subject pressed the mouse button to store the response and start a new trial. To facilitate the pursuit task, the aperture boundary moved along with the pursuit target, mimicking the effect of macular degeneration (tunnel vision). Pursuit was also facilitated by presented trials in blocks during which the direction and velocity of target motion were constant (Kowler & McKee 1987).

### Subjects

Seven subjects participated. Subjects BB and JB were experienced with heading stimuli and not naive towards the aim of the experiment. Subject JD and EB were experienced, but naive. Subjects ANB, JS and JG were inexperienced and naive. Subject ANB participated only in the first experiment, subject JG participated only in the first two sessions of the first experiment. Stimuli were viewed monocularly with either right (ANB, JD, JB and JG) or left eye (BB, JS and EB).

### Eye movement recordings

The 2D eye movements for subject JB were recorded using scleral coils during Experiment II. A bite board was used to prevent head movements. The velocity of the pursuit eye movement was computed after removal of saccades, and averaged over the 0.5-s period of flow presentation. The pursuit gain, defined as the ratio of the smooth eye velocity to the velocity of the fixation marker, was on average 0.95, with

SD of 0.07. This value correlates with the high pursuit gain found in a previous heading experiment (unreported data) using similar flow stimuli. The average over five individuals obtained in that experiment (0.85), we here use as the estimate of the pursuit gain.

## 4.3 Experiment I: Perceived heading and its variability

### 4.3.1 Procedure

#### Motion parameters

Within one pursuit block, eight possible heading directions ($\pm 4$, $\pm 12$, $\pm 20$ and $\pm 28°$) were presented in random order, each repeated eight times. Since the overall retinal flow magnitude increased with simulated heading eccentricity, we worried that subjects might base their judgements on perceived retinal speed, rather than on the pattern of retinal flow. Therefore, the simulated translation speed was randomised on a trial to trial basis, varying between 1.0 and 2.0 m/s (1.5 m/s on average).

#### Stimulus conditions

Data were collected in five sessions, each lasting about half an hour, run successively with no more than one day in between. A session consisted of two triplets of pursuit blocks. The first two sessions presented aperture stimuli of the approaching frontal plane. The first session consisted of a block triplet with consecutive pursuit velocities of 0, -3 and 3 °/s, duration 0.5 s, followed by a triplet with duration 3.0 s. The second session was a repetition of the first, with the order of the pursuit directions reversed. The third session presented the large field stimulus (approaching wall, duration 3.0 s) in two block triplets, balanced in pursuit direction. The fourth and fifth session were repetitions of the first two sessions, now with the 'equidistant' wall.

#### Task

Subjects were asked to indicate the direction of dot motion relative to themselves as they perceived it at the end of the trial. To familiarise subjects with the heading task, subjects were given practice trials with large field of view stimuli and short duration (0.5 s). We ran these trials prior to the first and the fourth session. A practice sequence consisted of two triplets of pursuit blocks in which each heading direction was presented only once ($2 \times 3 \times 8 = 48$ trials in total). No feedback was given during the practice, nor during the experiment. Practice trials were not analysed.

#### CF shift

We approximated the shift of the center of flow (CF) on the retina during eye rotation by the following formula:

$$\Delta H_{CF} = g_p \omega \, d/T. \tag{4.1}$$

Here, $g_p$ denotes the pursuit gain, $\omega$ the pursuit target velocity (3 °/s), $T$ the translational speed and $d$ the simulated distance of the wall with respect to the eye. Because subjects were asked to indicate their perceived heading at the end of the motion sequence, we take for $d$ the simulated distance at the end of the trial, unless stated otherwise.

### Data analysis

In experiment I, each condition (8 simulated headings × 3 rotation speeds × 5 stimuli) was presented 16 times. Mean perceived heading and its standard deviation (SD) per condition were computed for each subject. The error in the mean perceived heading and the shift of the SD-curve was retrieved from a linear and non-linear fit, respectively. The mean perceived heading ($PH$) as function of heading $h$ was least-squares fitted with the following 2-parameter function:

$$PH(h) = \Delta P + G_{bias} h. \tag{4.2}$$

Note, this function not only captures the heading error by an offset $\Delta P$ along the axis of perceived heading, but also by a slope $G_{bias}$ of the perceived versus simulated heading that may differ from unity (Fig. 4.3a). The SD-data were fitted (Levenberg-Marqhuardt) with a 4-parameter upside-down Gaussian function:

$$SD(h) = SD_{min} + A\,[1 - e^{-(h-\Delta H_{SD})^2/W^2}]. \tag{4.3}$$

This Gaussian function had horizontal width $W$ and amplitude $A$ (Fig. 4.3b). The minimum SD ($SD_{min}$) is reached for the simulated heading $h = \Delta H_{SD}$. We started the iterative fit process with parameters $A$, $W$ and $SD_{min}$ set to 5, 1 and 20°, respectively. The starting parameters for $\Delta H_{SD}$ were -15, 0 and 15° for pursuit target speeds -3, 0 and 3 °/s, respectively.

To assess the errors in the estimated parameters for the non-linear fit we employed the bootstrap method (Press et al. 1994, Efron & Tibshirani 1991). Data (either the mean perceived heading or their SD as function of the eight simulated heading directions) were pooled over five or four subjects. For reasons stated in the results, data from subject ANB and JG, as well as data from EB in the large field stimulus were not included. One dataset therefore consisted of 40 (5 × 8) datapoints for aperture stimuli, and 32 (4 × 8) datapoints for the large field stimulus. Each of the 30 datasets (3 pursuit speeds × 5 stimuli for SD-data and mean data) was bootstrapped as follows. Five hundred synthetic datasets were derived from one dataset. A synthetic dataset contained as many datapoints as the actual dataset, each datapoint randomly drawn (with replacement) from the actual dataset. Fits were computed for each synthetic dataset. The mean and standard deviation of a fit parameter taken over all synthetic

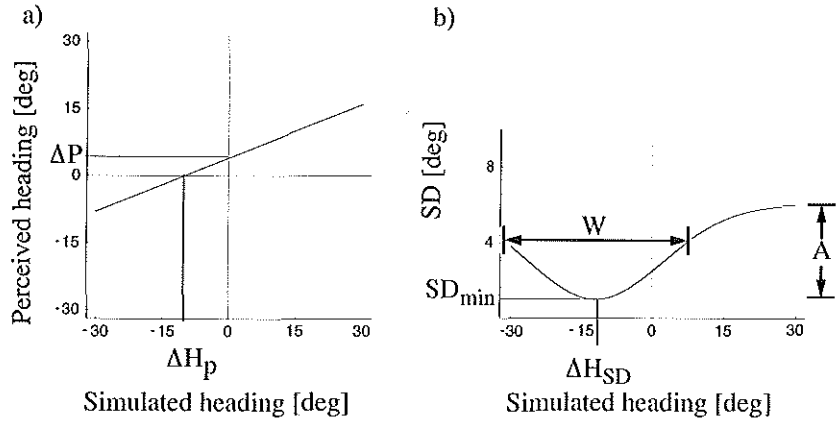



**Figure 4.3** *Functions used to fit a) the mean of the perceived heading and b) its standard deviation (SD).*

datasets provided the estimate of the fit parameter and its standard error. For each synthetic dataset, also the correlation between the fitted values and actual data value was computed. The mean correlation, taken from all synthetic datasets, was used as an estimate of the goodness of fit.

### 4.3.2 Results

#### Approaching plane, short duration

To examine whether local motion is pooled before or after the compensation by the extra-retinal signal, perceived heading was measured for three pursuit velocities (0, ±3 °/s) at eight different heading directions. First, we examine the results for simulated observer translation towards a wall (approach scene) during 0.5 s. The mean and standard deviation of perceived heading as a function of simulated heading are plotted in Fig. 4.4a and b for three subjects. Fig. 4.4b shows a clear SD-curve minimum in the condition without eye movement (solid square) for subjects JD and EB. The variable error increases for heading directions that are located outside the aperture, similar to findings by Crowell and Banks (1993). For two subjects out of seven, including subject ANB (Fig. 4.4b right), we find no clear increase of SD with heading eccentricity during fixation. Their data were discarded. For the others, we found SD-minima that were offset during pursuit (see JD and EB, triangles). As illustrated for subject JD , the offset for rightward pursuit (△: dotted line through SD-minimum) is in the direction and of a magnitude expected if precision is determined by the pattern of retinal flow (Fig. 4.2).

Regarding the accuracy of perceived heading (Fig. 4.4a), we observe that the

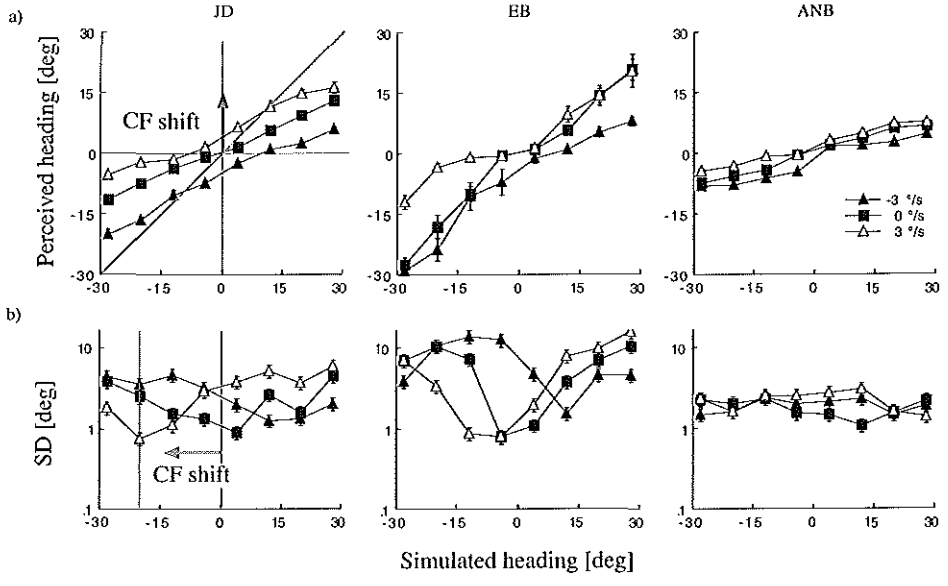



**Figure 4.4** *Results for simulated observer translation towards a wall (approach scene) during 0.5 s. Mean and standard deviation of perceived heading direction taken over 16 trials are plotted as function of the simulated heading for three subjects during fixation, leftward and rightward pursuit. a) Mean perceived heading. The errorbar indicates the standard error in the mean. The solid line with slope unity indicates veridically perceived heading. The arrow shows the offset in perceived heading expected during rightward pursuit if rotational flow is not compensated. b) Standard deviation. The errorbar indicates the error in the error, given by* $SD/(2n-2)^{1/2}$ *(Squires 1985). The arrow indicates the shift of the SD-curve expected during rightward pursuit when precision is based on the pattern of retinal flow. Given the range of simulated translation speeds and an assumed pursuit gain of 0.85, pursuit shifted the CF by 15° (3° standard deviation).*

straight line of the perceived vs. simulated heading systematically deviates from veridical (grey line with slope unity). Part of this systematic error, however, is not related to the pursuit. For all subjects except EB the slopes are less than unity during fixation. The slope is more decreased during pursuit (e.g. EB). Because this error increases with heading eccentricity, we refer to it as a scale error (Beintema & van den Berg 2000). The other part of the systematic error is directly related to CF shift. As exemplified for subject JD for rightward pursuit, the perceived heading during pursuit was offset along the perceived heading axis in the direction of the CF shift, but generally much less than the CF shift. This suggests that rotational flow is largely compensated by the extra-retinal signal.

If the slope of perceived vs. simulated heading is smaller than unity, one faces

the choice which parameter best characterises the compensation (Beintema & van den Berg 2000). If an internally represented heading is somehow scaled after compensation for rotational flow, then the offset along the perceived heading axis ($\Delta P$) overestimates the compensation by an amount inversely proportional to the slope $G_{bias}$. In that case, the offset along the simulated heading axis ($\Delta H_P = \Delta P / G_{bias}$) would be a better quantification (Fig. 4.3a). This offset is much closer to the CF shift (Fig. 4.4a). Because larger extra-retinal compensation allows a sharper distinction between extra-retinal compensation models, the same experiment was repeated at different duration. Grigo & Lappe (1999) found duration to be an important factor in the compensation for simulated rotation in simulated frontal plane approach.

### Effect of duration and simulated distance

To investigate the potential for an increase in extra-retinal compensation, we repeated the foregoing experiment with 3.0 s stimulus duration. Prolonged duration, however, also decreases the simulated distance of the approaching wall, hence decreases the CF shift. This makes it more difficult to differentiate between the models. To decouple duration and CF shift, the two durations (0.5 s and 3.0 s) were repeated for an equidistant scene that approximated a frontal plane, but remained at a constant simulated distance to the eye during the trial.

The mean and SD in perceived heading, averaged over the five subjects, are plotted in Figs. 4.5 and 4.6 as function of simulated heading. Different panels refer to different durations and types of scene. The solid lines are fits to the data; a linear regression for the perceived heading data, an up-side-down Gaussian curve for the SD-data. Duration and type of scene affected the accuracy of perceived heading (Fig. 4.5). For the approaching wall, the slopes of the regression lines are increased for longer duration. The offset, either along the perceived or simulated heading axis, is also reduced for longer duration. This implies that the extra-retinal compensation is increased at longer duration. In contrast, the fits to the SD-curves for pursuit (Fig. 4.6) appear not to be affected by duration, nor by the type of scene motion.

### Effect of aperture size

The results for 3.0-s duration (Fig. 4.5) still show systematic offsets in perceived heading during pursuit. The offset along the simulated heading axis suggests that about 40% of the CF shift has not been compensated. This seems to contrast with studies with simulated fronto-parallel plane approach, that conclude that perceived heading during real pursuit is accurate (Warren & Hannon 1990; Royden et al. 1994). In these experiments, however, the field of view was large (up to 40°) compared to our 10° aperture. To investigate whether the small aperture limited the amount of compensation, we repeated the 3.0-s approaching wall for large field of view (60 × 58°).

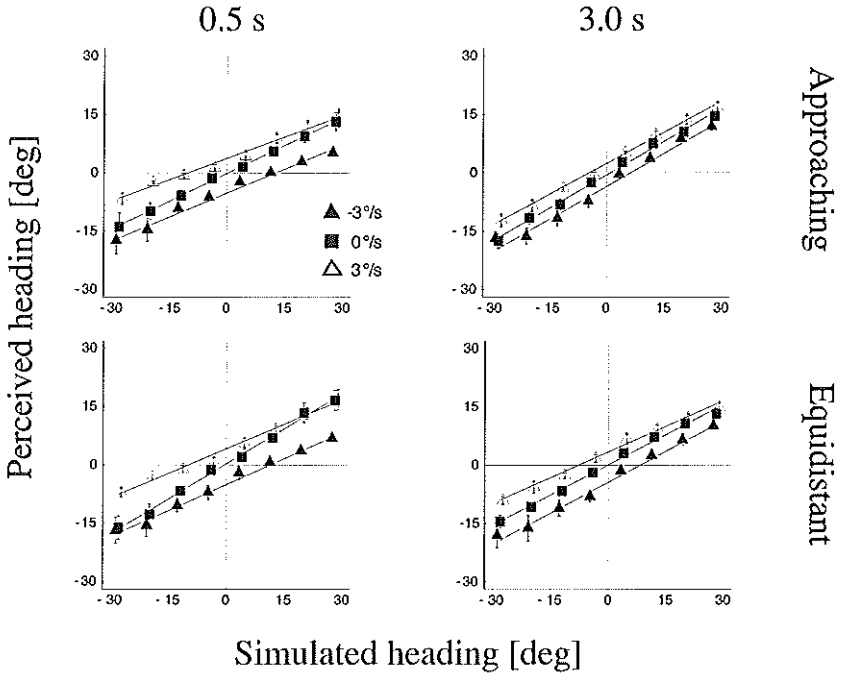



**Figure 4.5** *Perceived heading error as function of simulated heading during fixation, left and rightward pursuit, for 4 different stimulus conditions. The average perceived heading was computed for each subject, taken over 16 repetitions. Each data point represents the mean of these averages from 5 subjects. The errorbar indicates the standard error in the mean over subjects. Also drawn are the fitted linear regressions for each pursuit condition. Each panel represents data for another stimulus condition. Left and right panels differ in motion duration (0.5 s and 3.0 s, respectively). Top panels represent data for simulated translation towards a wall. Lower panels represent data for a scene that on average remains equidistant with respect to the eye.*

Fig. 4.7 shows the perceived heading and SD for large field of view, plotted as function of simulated heading, averaged over four subjects. Data from subject EB were left out because of inconsistent behaviour; a peculiar lack of heading dependency occurred during leftward pursuit, that did not occur for rightward pursuit, nor was it found in any of the other stimulus conditions or for other subjects. Contrary to our expectation, the offset in perceived heading during rotation (Fig. 4.7a: solid triangle and △) is only little decreased with respect to the small aperture condition (Fig. 4.5: 3-s approaching wall). The increase in field of view did have large effect on the perceived heading eccentricity, since the slopes for fixation and pursuit are clearly increased. The SD-data (Fig. 4.7b) show that for large field of view, precision no longer depends on heading eccentricity.

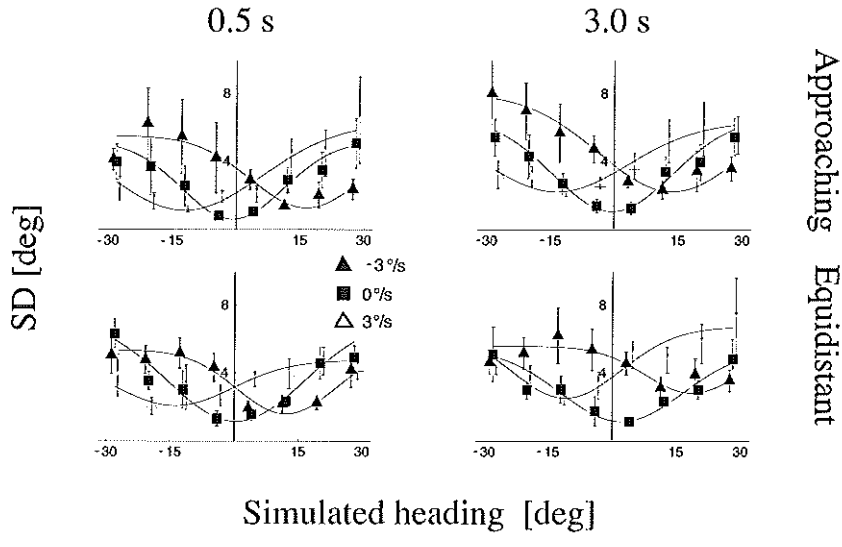



**Figure 4.6** *Variable error (SD) and fits as function of simulated heading for different stimulus conditions. For details, see legend Fig. 4.5.*

### Fit results

A more quantitative description of the data was obtained by fitting the PH-data with a line and the SD-data with an up-side down Gaussian function. Data were pooled over subjects, and standard errors were assessed using the bootstrap method. As a measure of the goodness of fit, we computed the correlation coefficient between fit and data (see methods). Correlations were high for the PH-fits ($R > 0.9$, Fig. 4.5). Although a Gaussian seems to fit the SD-data reasonably well (Fig. 4.6), correlations were much lower ($0.3 < R < 0.7$). The low correlations are not only caused by a large spread in individual SDs (compare SD-errorbars with PH-errorbars, Fig. 4.6 and 4.5), but also by the Gaussian's large degree of freedom (4 df) compared to the small number of data points (40). In random selection of datapoints for bootstrapping (see methods), left out points on the curve may then have a larger effect on the correlation. Illustrative for this is that we found a bimodal distribution for the width and amplitude of the SD-data. About 20% of the fits had large width (180°) and amplitude (140°), and about 80% had small width (20°) and amplitude (4°). The prevalence of small widths shows that a Gaussian function is a better description of the SD-data than a parabolic function. Irrespective of the width or amplitude, the minimum SD was obtained at the same simulated heading direction. Thus, the location of the minimum is a robust result in this fitting procedure. We found no significant effect of pursuit or stimulus condition on the width ($20 \pm 9°$) or amplitude ($4 \pm 2°$) of the SD-curves.

We plotted the shift of the locus of highest precision ($\Delta H_{SD}$, open symbols)

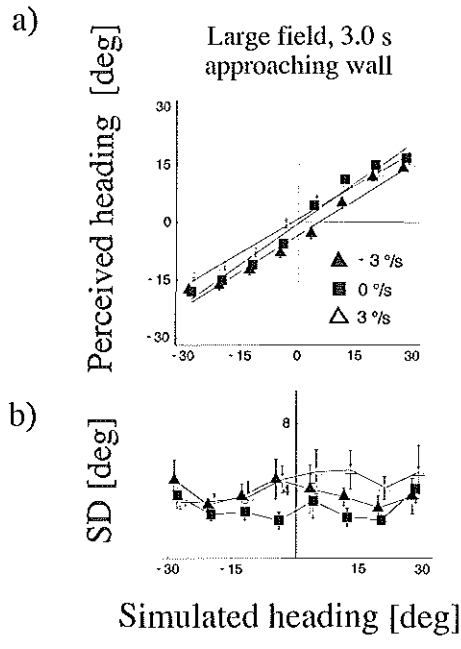



**Figure 4.7** *Results for the large field condition. a) Mean b) and standard deviation in perceived heading as function of simulated heading during fixation, left and rightward pursuit. Data are averaged over subjects. Errorbars indicate standard errors in the mean over subjects. Only for the systematic error, linear regression fits are shown.*

together with the systematic offset in perceived heading $\Delta P$ or $\Delta H_P$ in Fig. 4.8a and b, respectively, to investigate whether they are equal. Both plots show the offsets for the two types of aperture stimuli (squares: approaching; circles: equidistant) as function of duration, including the systematic error for the large field condition (triangles). Fig. 4.8a shows the SD-curve offset ($14 \pm 6°$) is much larger than the systematic offset in perceived heading ($\Delta P$) ($4 \pm 1°$). The difference is smaller, when comparing the SD-curve offset with the offset along the simulated heading axis (Fig. 4.8b). Still, we find the difference is significant for the 3.0-s aperture stimuli.

Regarding the effect of aperture size on the systematic error during pursuit, we find that the systematic error in perceived heading is smallest for the large field stimulus (Fig. 4.8a and b solid triangle), but not significantly different from the aperture condition (3-s approaching scene; solid square).

Fig. 4.8b allows a clear distinction between the effect of duration and simulated distance on the systematic errors in perceived heading during pursuit. For longer duration (3.0 s), the systematic error ($\Delta H_P$) decreases for both types of scene motion. Because the flow for the equidistant scene (solid circle) remains constant over time, we may conclude that the increased exposure time by itself reduced the systematic error. The decreased systematic error for the approaching scene (solid square) is thus an effect of the decreased simulated distance towards the end of the trial.

In Fig. 4.8b, we also plotted the predicted SD-curve offsets according to the velocity gain field model (grey lines). The SD-curve offsets follow the predictions for the equidistant scene (grey line *), but not for the approaching scene at 3.0-s duration

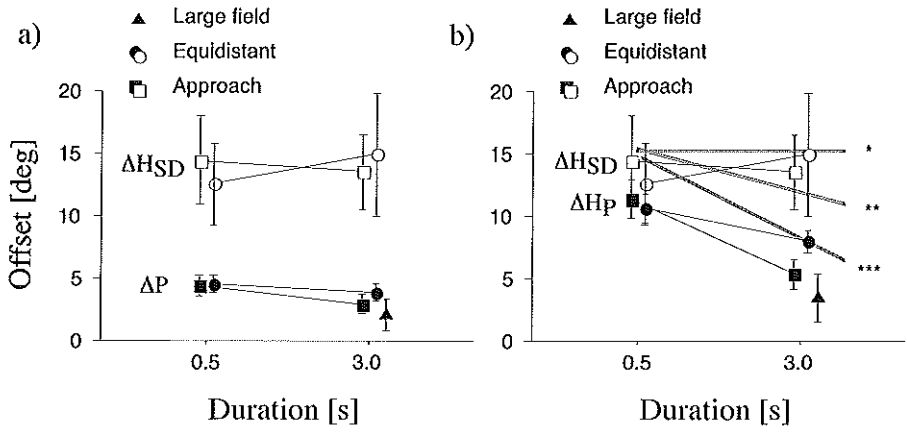



**Figure 4.8** *Offset of the SD-minimum (open symbols) and perceived heading (solid symbols) during pursuit, plotted as a function of self-motion duration. Different symbols refer to aperture stimuli (circles: equidistant; squares: approach) and large field stimulus (solid triangle: approaching scene). The offset of the SD-minimum ($\Delta H_{SD}$) along the simulated heading axis is plotted together with (a) the offset along the perceived heading axis ($\Delta P$), or (b) the offset along simulated heading axis ($\Delta H_P$). Also plotted in (b) are the grey lines indicating the CF shifts for the equidistant wall (*) and for the equidistant wall midway during trial (**) or at the end of (***) the trial. Offsets, collapsed over pursuit direction, were obtained from fits to PH-data and SD-data pooled over subjects. Errorbars indicate the standard error in the fit parameter, obtained by bootstrapping.*

if that prediction is based on the flow in the final frame (grey line ***). The velocity gain field model predictions can fit the data, on the assumption that the CF shift is based on earlier frames, such as halfway during the trial (grey line **). The SD-data would fit predictions based on the initial flow as well (grey line ***), but then we cannot explain the reduction of the systematic error in the approaching scene ($\Delta H_P$) by decreased simulated distance.

As shown in Fig. 4.9a, we find three main effects on the slope ($G_{bias}$) of the perceived vs. simulated heading. Firstly, we observe an overall decrease of the slope as a function of pursuit velocity. Such interaction between heading direction and horizontal rotation we recently reported for perceived heading during simulated horizontal rotation (Beintema & van den Berg 2000). The relation we suggested in that paper with asymmetry in the retinal flow applies to the present data as well. Secondly, we observe another effect of duration on systematic heading error. For fixation and pursuit, the slope tends to increase with longer duration (compare solid with open circles and squares). A one-tailed t-test revealed a significant effect of duration on the slope at 5% significance level, for data pooled over pursuit velocity and type of

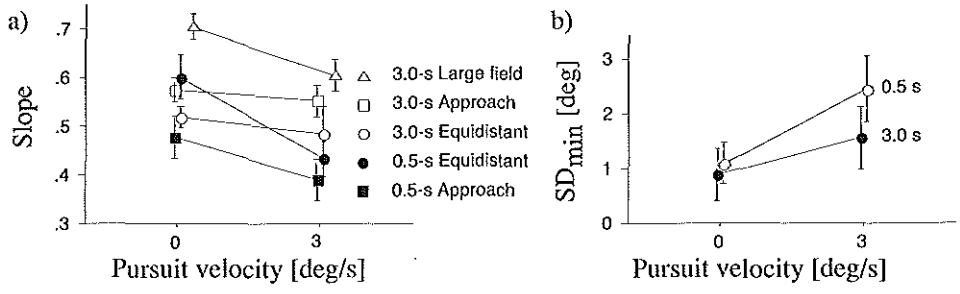



**Figure 4.9** *a) Slope of perceived vs. simulated heading as function of the absolute pursuit velocity for the approaching and equidistant scene. b) SD-minimum as function of the absolute pursuit velocity for two durations, pooled over both types of scene (small aperture data only).*

scene motion. Thirdly, contrary to the compensation for rotation, the compression of perceived heading did vary with aperture size, since the slope for the approaching wall (3 s), with data pooled over pursuit velocity, was significantly increased in the large field condition (compare △ with □).

Pursuit had only a small effect on the highest precision of perceived heading. Fig. 4.9b shows that the minimum SD during pursuit is elevated by 1° on average compared to fixation. The retinal flow is very similar (nearly radial) in all conditions of minimum SD; the extra-retinal signal is much larger. Clearly, if the extra-retinal signal combines additively with local retinal signals (vector subtraction), the limited elevation suggests that the level of noise in the extra-retinal signal is quite low. This raises the question if the variation in perceived heading is to the largest extent dependent on the retinal motion uncertainty. Is an alternative explanation of our results in terms of vector subtraction possible?

## 4.4 Experiment II: Perceived head-centric direction and its variability

### 4.4.1 Procedure

Experiment I confirmed our hypothesis that the pattern of the retinal flow - not the head-centric flow - with respect to the aperture determines precision of heading. This is in line with the velocity gain model, and argues against a vector subtraction scheme. We wondered, however, whether the data could be in line with vector subtraction scheme after all, if one assumes that pursuit affects the precision of perceived retinal velocities. When the SD minima occur, not only the center of retinal flow is located within the aperture, but also the retinal velocities are smallest (Fig. 4.2). Since

thresholds for the perceived retinal velocity follow Weber's law, i.e. decreasing with decreased retinal velocity (McKee 1981; De Bruyn & Orban 1988), local speed in our stimulus may have been detected with greater precision when pursuit shifted the center of retinal flow within the aperture. As explained in the Appendix, if after vector subtraction the speed of the recovered flow vector increases without increasing the absolute error, the relative error in the flow decreases. The variable error in perceived heading can then be small, despite the center of the recovered flow being located outside the aperture (Fig. 4.14: bottom). It would require that the increased detection of retinal velocity is not countered by noise in the extra-retinal signal. Then, for the same head-centric flow, the local recovered vector must be more precise encoded during pursuit than during fixation.

Few studies have been conducted on the precision of perceived motion during pursuit. Kowler and McKee (1987) report a striking similarity between perceptual and oculomotor measures of precision of speed, in a task of judging target velocity or pursuing a target by eye, respectively. However, oculo-motor precision would be much lower when plotted as function of the retinal instead of target velocity, so that retinal velocities alone cannot account for the oculo-motor data. Turano and Heidenrich (1996) showed that speed discrimination thresholds follow retinal speed predictions, but only when the eye moves opposite to or at lower speed than a 0.5 or 2 °/s moving background. For heading perception, however, it is not the distribution of local motion speeds, but of local motion directions that forms the essential input (Warren et al. 1991). No studies examined the effect of pursuit on perceived direction precision. With this control experiment we tested whether for the flow stimuli representative for our heading stimuli, a reduction of retinal motion during pursuit is of benefit to the precision of perceived direction, despite the interaction with the extra-retinal signal.

We repeated the aperture experiment (0.5-s approaching scene), but now simulating translation parallel to the fronto-parallel plane. The flow was presented in 12 directions (Fig. 4.10a). The direction of the flow ($\Psi_{flow}$) was defined by its angle with the horizontal ($\pm 22.5$, $\pm 45$, $\pm 67.5°$). The average number of visible dots was 130. Subjects were to judge the direction of the uniform flow within the aperture during fixation and during horizontal pursuit. Pursuit was always of the same magnitude (3 °/s) and in the same direction as the horizontal component of head-centric flow. Thus, during pursuit with gain 1.0, the retina would receive only the vertical component of motion. This way, the retinal velocity during pursuit would be decreased compared to the retinal flow during fixation. Accordingly, each retinal flow vector is detected with smaller absolute error. If the extra-retinal signal is subtracted noise free, then the direction of perceived flow should be more precise during pursuit than during fixation.

At the end of the scene motion, a line appeared as the pointer. The line was 20° of length with the 10° central segment left out (Fig. 4.10b). Subjects adjusted the direction $\Psi$ of this line to align it with the perceived direction the flow. We aided subjects in making a head-centric judgement of the flow direction, by presenting the

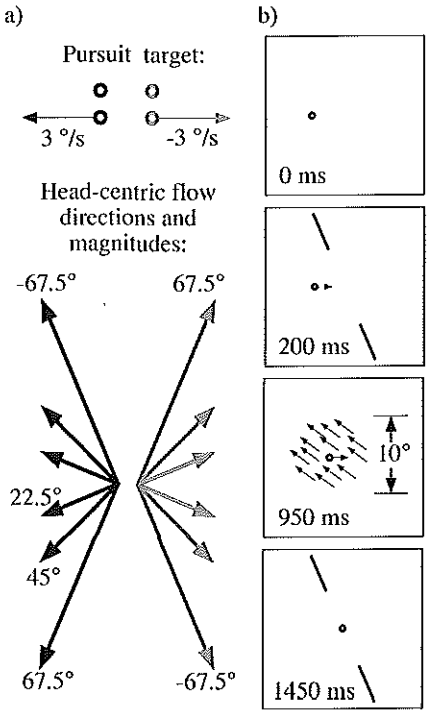



**Figure 4.10** *a) Possible directions and magnitudes of uniform flow on the screen as function of angle $\Psi_{flow}$ and simulated rotation direction. The horizontal component of the flow velocity is fixed at 3 °/s. For left and rightwards flow, a trial block with stationary or a moving pursuit target was presented (top row). Note, the pursuit velocity is of the same magnitude and in the same direction as the horizontal component of flow on the screen. b) Snapshots of stimulus (pursuit target, reference line and uniformly moving dots) within a pursuit trial, taken at the onset of the stationary target (0 ms), target motion (200 ms), uniform motion (950 ms) or at the end of the motion sequence (1450 ms).*

head-centric stationary reference line in the 950-ms interval preceding the uniform dot motion (Fig. 4.10b). To prevent the direction of the reference line from biasing the response, we randomly varied its direction $\Psi_{ref}$ from trial to trial ($\pm 22.5$, $\pm 45$, $\pm 67.5$ or $\pm 90°$).

Trials were presented in eight pursuit blocks. Each block consisted of 48 trials (6 flow directions × 8 reference directions), presented in random order. The horizontal component flow was reversed after each block. Fixation and pursuit condition were alternated in subsequent block pairs. No practice trials were given.

### 4.4.2 Results

As in the heading experiment, each condition (pursuit condition, flow direction $\Psi_{flow}$) was presented 16 times. This allowed us to evaluate the perceived flow direction (mean $\Psi$) and its uncertainty (standard deviation $SD_\Psi$) per condition for each subject. The mean error in perceived direction and its standard deviation are plotted in Fig. 4.11 as function of the absolute direction of simulated flow ($\Psi_{flow}$) for fixation and pursuit. Outliers that deviated more than 90° from the simulated direction were omitted (1.6% of all trials), because they may have been 180°out of phase; the line pointer represents only the orientation, not the direction of perceived flow. Fig. 4.11a shows that the uncertainty (SD) in the perceived direction is not decreased during pursuit.

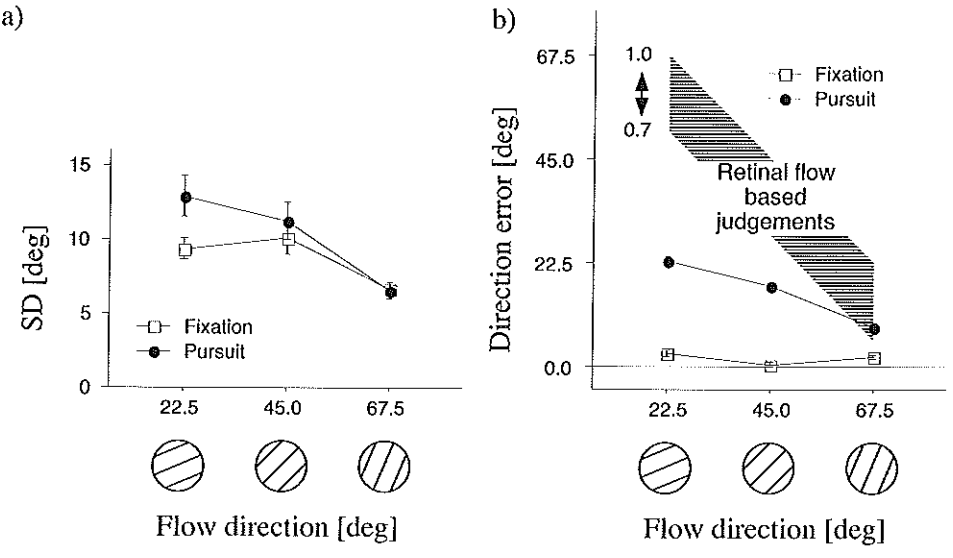



**Figure 4.11** *a) The standard deviation in the perceived direction, plotted as function of the absolute direction of flow during fixation and pursuit. b) Idem, for the mean difference between perceived and simulated direction of head-centric uniform flow. Errors for negative $\Psi_{flow}$ have been collapsed. Also plotted are the predicted errors when responding purely on the basis of retinal flow, given a pursuit gain between 0.7 to 1.0.*

A one-tailed unpaired t-test on the difference between pursuit and fixation, revealed a significant decrease at the 5% level occurred for 67.5° for one subject (EB). At 1% level no significant decrease was found for any of the five subjects.

Fig. 4.11b shows that during fixation the head-centric flow direction is perceived almost correctly, although a slight bias towards the vertical occurs for directions off the diagonal. During pursuit, on the other hand, systematic errors occurred for flow directions near the horizontal, decreasing towards the vertical. The systematic errors, especially for directions near the horizontal, are smaller than the errors expected if observers had responded to the retinal flow (Fig. 4.11b).

Thus, substantial compensation for retinal flow due to eye rotation is observed for the perceived direction of flow for all simulated directions. We find no evidence for the assumption that directional precision is increased during 3 °/s pursuit, despite the reduction in retinal velocity. This is in line with the perceptual (Turano & Heidenrich 1996) and oculo-motor data (Kowler and McKee 1985) on speed discrimination at 3 °/s. The largest extra-retinal compensation and largest elevation of the standard deviation during pursuit are found for directions near the horizontal (± 22.5°). Especially, these most parallel directions are representative for the flow within the aperture

in the heading experiment for which we wished to test the assumption.

## 4.5 General discussion

### Summary

Our main objective was to examine whether the compensation for eye rotation in the perception of heading occurs at an early stage, before pooling of local motion has taken place (vector subtraction model) or at a later stage (velocity gain model). To this end, we measured the variable and systematic errors of perceived heading during real eye movements and during fixation. We showed simulated observer translation at different angles towards a fronto-parallel plane (the plane was aligned with the screen), while only flow within a small aperture was made visible. If the rotational flow is subtracted out before the pooling of motion, variable errors in perceived heading should depend on the error in the recovered head-centric flow, whereas for compensation after pooling the variable errors should depend on the error in the retinal flow.

Under all circumstances, we found the minimum of variable error in the perceived heading was reached whenever the center of retinal flow - not the center of head-centric flow - was located within the aperture. During pursuit, the SD-curves were generally more offset than the offset in perceived heading along the perceived heading axis (Fig. 4.8a: $\Delta H_P$), indicating that the rotation was largely compensated by the extra-retinal signal. Such compensation by means of vector subtraction can not explain the larger offset of the SD-curve.

One might argue that a general compression of perceived headings has reduced the systematic errors in perceived heading, so that we have overestimated the extra-retinal compensation. But even when comparing the SD-curve offset with the systematic offset along the simulated heading axis ($\Delta H_P$), the difference is still significant for 3-s duration. Evidence for substantial compensation also follows from comparison between the CF shift (grey lines) and the systematic error in perceived heading (Fig. 4.8). The error ($\Delta H_P$) for the 3.0-s equidistant scene (Fig. 4.8b), which requires no assumptions on temporal integration to compute the CF shift, shows half of the CF shift is compensated.

A control experiment on judged direction of simulated uniform flow revealed no increased precision when pursuit reduced the retinal velocities. For directions near the horizon, about half of the rotational flow due to eye rotation was compensated. Thus, the extra-retinal signal also contributed to the compensation in the control, but did not lead to reduced directional uncertainty. The shift in the location of SD minimum in the heading experiment can therefore not be explained by the reduced retinal flow magnitude on the assumption that compensation occurs by vector subtraction. The results are compatible with a velocity gain field type of interaction.

### Physiology

Indications for extra-retinal interaction in MST has been provided by a number of studies (Bradley et al. 1996, Shenoy et al. 1996, Shenoy et al. 1999). They showed that certain cells that respond to retinal flow, maintain their tuning to retinal flow during pursuit, but are modulated in amplitude in proportion to the eye's velocity (Bradley et al. 1996) or head's velocity (Shenoy et al. 1999). Velocity gain fields also have their analogy in the spatial domain in the parietal cortex, where eye position gain fields have been found (see Andersen 1997 for a review).

### Alternative extra-retinal models

The velocity gain model as we proposed it (van den Berg & Beintema 1997; Beintema & van den Berg 1998a) uses a rate-coded representation of eye velocity. The extra-retinal signal interacts with visual units after local motion signals have been pooled. Such model is in line with the psychophysical data obtained in this study. The current data argue against a vector subtraction kind of interaction (Royden et al. 1994), as such level of interaction would predict that precision of perceived heading does not depend on the center of retinal flow with respect to the aperture. A recently proposed model by Lappe (1999) that extends a previous model by Lappe and Rauschecker (1993) to incorporate the use of extra-retinal signals, also falls into the category of the vector subtraction like interaction. The proposed interaction does not necessarily take place at the level of local motion presentation (area MT) but may take place at the input stage where motion is pooled (e.g. MST). Nevertheless, it involves a subtraction of local vector like representations of rotational flow by extra-retinal signals. These local vector representations require eye velocity encoded by a much larger population of cells, each tuned to a specific velocity and direction of eye velocity. Perrone & Stone (1994) briefly mention the possible use of extra-retinal signals by inhibition or facilitation of certain sectors of detector maps. The detectors are tuned to patterns of flow. This might be in line with our data, but it would require a more detailed description of the representation of the extra-retinal signal and the actual interaction for a fair comparison.

### Spatio-temporal aspects

The increase in extra-retinal compensation with duration (Fig. 4.8b) supports the hypothesis that the visual system relies less on the extra-retinal signal when retinal flow is only briefly presented (Grigo & Lappe, 1999). A prerequisite for demonstrating such effect is that the retinal and extra-retinal signals on eye rotation are conflicting. One such case is when simulating eye rotation, so that the retinal flow tells the eye is rotating and the extra-retinal signal tells the eye is stationary. Especially during approach of a wall, the retinal flow is ambiguous (Longuet-Higgins 1994), as such flow can hold correct and incorrect solutions. Grigo and Lappe (1999) showed that

for a large frontal plane (90 × 90°), people perceived the heading that is simulated, but only for shortly presented flow ($\leq$ 0.5 ms). For a smaller field of view (60 × 60°) the incorrect solution that ignored the simulated eye rotation was perceived for all durations. For longer duration, the incorrect solution was also perceived for the large field of view, which Grigo and Lappe explained by an increased reliance on the extra-retinal signal. At first glance, our finding during real eye rotation of decreased errors for longer duration, seems to argue against their hypothesis. Yet, like their smaller stimulus, our aperture stimulus (10 × 10°) does not allow a visual decomposition, so that the visual system by default chooses the 'incorrect' solution. Consequently, in this real pursuit case of retinal and extra-retinal signal conflict, an increased reliance on the extra-retinal signal will promote the 'correct' solution, so that we indeed expect errors to decrease with longer duration.

Pooled over subjects, we found that the size of viewing aperture had only marginal effect on the systematic error caused by the rotation (Fig. 4.8a, b). For a 60° field of view at 3-s duration, we still found small errors in perceived heading during real pursuit in the direction of the CF shift. This result complies with a recent observation by Freeman (1999) that, in contrast to earlier conclusions (Warren & Hannon 1990; Royden et al. 1992, 1994; Banks et al. 1996), the compensation for rotation during pursuit is not complete. We did find an effect of viewing aperture on the slope of the perceived vs. simulated heading (Fig. 4.9a) Examining the real pursuit data by Banks et al. (1996) for simulated translation through a cloud of dots, we observed an effect of field of view consistent with our findings. Heading errors decreased when field of view was increased from 30° to about 60° (their Fig. 3 and 6). Since they averaged heading errors across simulated heading evenly balanced around zero, their data may be consistent with our observation that the field of view increases the perceived heading eccentricity rather than reduce the rotation-related error.

### Perceived direction of flow during pursuit

In the control experiment, the perceived flow direction during horizontal pursuit systematically deviated towards the vertical. This deviation was quite large, so we were interested how these values compare to literature. Since the direction of retinal velocity during pursuit was close to vertical, the deviation may be explained as an underestimation of the extra-retinal signal (see Wertheim 1994 for a review). As more recently shown, it may also result from an overestimation of the retinal signal (Freeman & Banks 1998). The latter proposed that perceived head-centric flow is the weighted subtraction of the actual retinal velocity and the eye velocity. They find the ratio of the extra-retinal weight ($e$) and the retinal weight ($r$) ranges from 0.6 for high spatial frequencies (1 c/°) to 1.2 for low spatial ($\sim$ 0.1 c/°) frequencies. As we here wish only to compare data, not differentiate between types of extra-retinal interaction, we use like Freeman and Banks (1998) the vector subtraction model. Writing the 2D perceived head-centric velocity $\mathbf{H}$ as the sum of the 2D retinal velocity and 2D pursuit velocity, weighted with an extra-retinal gain factor $e$ and a retinal gain

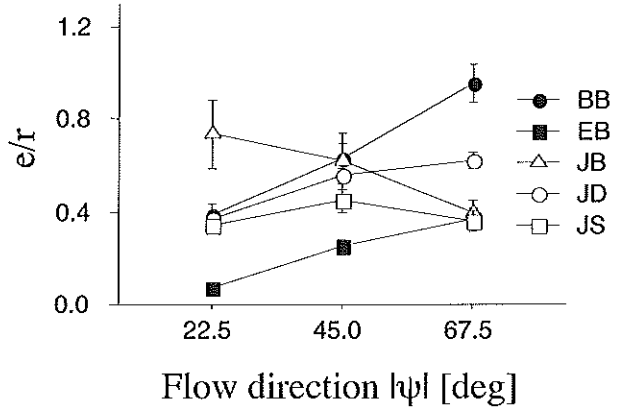



**Figure 4.12** *a) Ratio of extra-retinal gain over retinal gain as a function of the absolute flow direction ($|\Psi_{flow}|$) for each subject. Because division by the tangent of perceived directions near zero would cause large variation in the $e/r$-ratio, we excluded 0.8% of all datapoints (for which the $e/r$-ratio deviated more than three times the SD from the mean).*

factor $r$, respectively, we get for our experimental conditions:

$$\mathbf{H} = \begin{pmatrix} h \\ v \end{pmatrix} = \begin{pmatrix} r + g_p(e - r) \\ r \tan \Psi_{flow} \end{pmatrix} \omega$$

Here, $g_p$ indicates the pursuit gain (0.85) and $\omega$ (3.0 °/s) indicates the horizontal velocity of the pursuit target and of the flow on the screen. Note, the vertical contribution is purely retinal, increasing with the tangent of flow direction $\Psi_{flow}$. The ratio $e/r$ follows from the tangent of the perceived head-centric direction ($tan\Psi = v/h$):

$$\tan \Psi_{flow} / \tan \Psi = 1 + g_p(e/r - 1)$$

The $e/r$-ratio is plotted in Fig. 4.12. It varies considerably between subjects and flow directions ($SD = \pm 0.5$), and has a mean of about 0.5. Hence, it is close to the low value (0.6) found for high spatial frequencies by Freeman and Banks (1998).

### 4.5.1 Conclusion

For small viewing apertures we find that precision of perceived heading depends on pursuit. The findings confirm our hypothesis that interaction takes place after pooling of retinal motion signals (Beintema & van den Berg 1998a), and do not support an alternative hypothesis of vector subtraction (Royden et al. 1994). Furthermore, we find an influence of duration and field of view on the accuracy of perceived heading

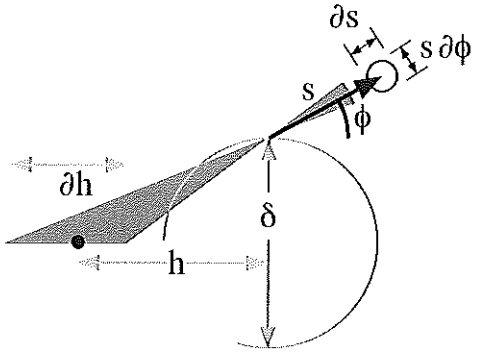



**Figure 4.13** *Geometrical relation between the direction of a motion vector ($\phi$) at the top of the aperture (diameter $\delta$) and the heading direction (solid circle at eccentricity $h$), and their uncertainties ($\partial\phi$) and ($\partial h$). Also shown is the assumed equality in directional uncertainty ($\partial\phi$) and speed uncertainty ($\partial s$) given speed $s$.*

during pursuit. The decreased accuracy during pursuit for shorter presented flow supports a recent hypothesis by Grigo & Lappe that the extra-retinal signal is less relied on for brief durations.

## 4.6 Acknowledgements

We thank Marty Banks for his critical remarks concerning the first experiment, and are grateful to Ignace Hooge for measuring pursuit eye movements. This work was supported by research grant SLW 805-33.171-P from the Dutch Foundation for Scientific Research (NWO) and by the Human Frontier Science Program.

## 4.7 Appendix

We here estimate the effect of pursuit on the heading uncertainty if rotation were compensated by vector subtraction. Specifically, we analyse the effects of changes in retinal speed and of the presence of noise in the extra-retinal signal. As heading direction is restricted to the horizontal plane, we simplify matters by computing heading ($h$) from the direction ($\phi$) of the most informative image motion vector, situated at the top or bottom of the circular aperture with diameter $\delta = 10°$ (Fig. 4.13):

$$h = \frac{\delta}{2} \tan^{-1} \phi$$

The uncertainty in this vector direction ($\partial\phi$) causes an uncertainty in heading eccentricity ($\partial h$) that equals

$$\partial h = (\frac{\delta}{2} + \frac{2}{\delta} h^2)\, \partial\phi. \tag{4.4}$$

For fixation, the retinal speed $s$ is proportional to the eccentricity of heading with respect to the location of the image motion vector, and therefore has a horizontal and vertical component:

$$s \approx \frac{T}{d}\sqrt{h^2 + (\frac{\delta}{2})^2} \tag{4.5}$$

Given the stimuli presented in the experiment ($d/T \approx 6$ s), the retinal speed for fixation thus ranges from 1 °/s ($h = 0°$) to about 5 °/s ($h = 30°$). Direction discrimination thresholds for this range of retinal speeds remain roughly constant ($\partial\phi \approx 0.1$) (Pasternak & Merigan 1984; De Bruyn & Orban 1988). The directional uncertainty and the corresponding heading uncertainty ($\partial h$) is plotted as function of heading eccentricity in Fig. 4.14a and b (fixation).

Next, we estimate heading uncertainty during pursuit predicted by vector subtraction. To this end, we must estimate the effect of changed retinal speed on the directional uncertainty of the recovered flow vector. We do this indirectly by first estimating the effect of changed retinal speed on the perceived speed uncertainty, then computing the recovered vector and its speed uncertainty and finally computing the directional uncertainty of the recovered flow vector from the speed uncertainty. For this, we assume the directional uncertainty and speed uncertainty ($\partial s$) are related by:

$$\partial\phi \approx \partial s/s. \tag{4.6}$$

This relation expresses that the directional uncertainty, in terms of a speed error orthogonal to the motion vector ($s\,\partial\phi$) equals the speed error along the motion vector ($\partial s$) (Fig. 4.13). Such proportionality at least seems to hold for the retinal speeds used in our experiment. For a much larger range of retinal speeds beyond 1°, a constancy is found for direction discrimination thresholds (Pasternak & Merigan 1984; De Bruyn & Orban 1988), as well as for speed discrimination thresholds expressed in Weber fractions (McKee 1981; De Bruyn & Orban 1988):

$$\partial s \approx 0.1\,s \tag{4.7}$$

Rightward pursuit ($R = 3$ °/s) changes the retinal speed by including a rotation component:

$$s_p \approx \frac{T}{d}\sqrt{(h + R\frac{d}{T})^2 + (\frac{\delta}{2})^2}$$

Substitution of $s_p$ into Eq. 4.7 yields the retinal speed uncertainty ($\partial s_p$) during pursuit. Under the assumption that the subtraction by the extra-retinal signal occurs noise-free, the speed uncertainty in the recovered flow vector equals the retinal speed uncertainty $\partial s_p$ during pursuit. The speed of the recovered vector, on the other hand, equals the retinal speed during fixation ($s$, Eq. 4.5). This is illustrated by Fig. 4.14 (bottom), where the recovered optic flow and its local motion error (circles) are plotted as function of simulated heading eccentricity. According to Eq. 4.6, the directional

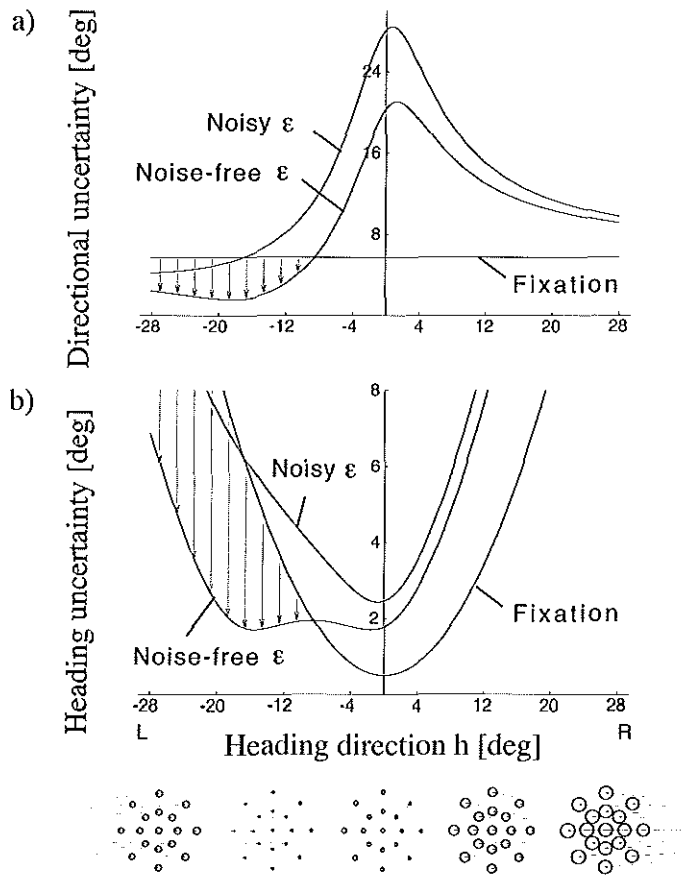



**Figure 4.14** *a) Directional uncertainty $\partial\phi$ of the recovered flow after vector subtraction as function of heading eccentricity ($h$), during fixation and pursuit. The curves for rightward pursuit are based on a noisy and noise-free extra-retinal signal ($\epsilon$). b) Corresponding heading uncertainty $\partial h$ based on directional uncertainty ($\partial\phi$). Below is depicted a representation of recovered optic flow and its local motion uncertainty during rightward pursuit after vector subtraction, on the assumption that rotational flow is noise-free subtracted.*

uncertainty equals the ratio of speed uncertainty and speed ($\partial s_p/s$). Hence, it is reduced at heading eccentricities in excess of 8° (Fig. 4.14a: compare noise-free with fixation). This increased precision of local motion coding counters the geometric effect of an increasing $\delta h$ due to more parallel flow within the aperture, up to about 20° eccentricity. This causes the heading uncertainty curve to be shifted (Fig. 4.14b: noise-free).

More likely, though, the extra-retinal signal also has a speed uncertainty ($\partial s_e$).

To illustrate its effect, we estimate this uncertainty to be in the range of perceptual thresholds for detecting 3 °/s retinal speed ($\partial s_e \approx 0.3$ °/s). If the speed uncertainty in the recovered flow vector is the vector sum of the retinal noise ($\partial s_p$) and extra-retinal noise ($\partial s_e$), then the directional uncertainty in the recovered flow vector amounts to

$$\partial\phi = \frac{\sqrt{\partial s_p{}^2 + \partial s_e{}^2}}{s}.$$

A noisy extra-retinal signal increases the directional uncertainty during pursuit (Fig. 4.14a: compare noisy with noise-free $\epsilon$). Now, the shift of the heading uncertainty curve is much reduced (Fig. 4.14b: compare noisy with noise-free $\epsilon$). This exercise shows that vector subtraction might be in accordance with a shifted SD-curve for perceived heading during pursuit, but only if the retinal speed effect is not countered by a noisy extra-retinal signal. As Fig. 4.14a shows, according to such scheme *the directional uncertainty should be lower during pursuit than during fixation* if retinal speeds are reduced by pursuit.

# Chapter 5

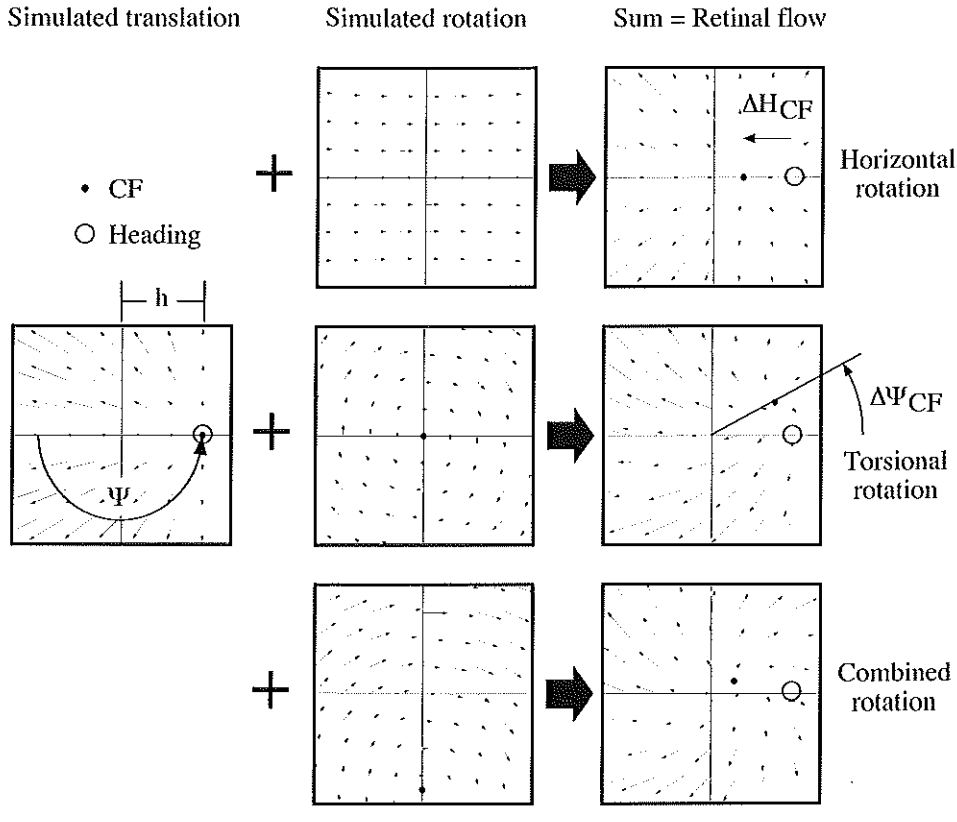
# Perceived heading during simulated torsional eye movements


## Abstract

Observer translation through the environment can be accompanied by rotation of the eye about any axis. For rotation about the vertical axis (horizontal rotation) during translation in the horizontal plane, it is known that the absence of depth in the scene and an extra-retinal signal leads to a systematic error in the observer's perceived direction of heading. This heading error is related in magnitude and direction to the shift of the center of retinal flow (CF) that occurs because of the rotation. Rotation about any axis that deviates from the heading direction results in a CF shift. So far, however, the effect of rotation about the line of sight (torsion) on perceived heading has not been investigated. We simulated observer translation towards a wall or cloud, while simultaneously simulating eye rotation about the vertical axis, the torsional axis or combinations thereof. We find only small systematic effects of torsion on the set of 2D perceived headings, regardless of the simulated horizontal rotation. In proportion to the CF shift, the systematic errors are significantly smaller for pure torsion than for pure horizontal rotation. In contrast to errors caused by horizontal rotation, the torsional errors are hardly reduced by addition of depth to the scene. We suggest the difference in behaviour reflects the difference in symmetry of the field of view relative to the axis of rotation; the higher symmetry in the case of torsion may allow for a more accurate estimation of the rotational flow. Moreover, we report a new phenomenon. Simulated horizontal rotation during simulated wall approach increases the heading-dependency of errors, causing a larger compression of perceived heading in the horizontal direction than in the vertical direction.



Adapted from: Beintema and van den Berg (2000), Vision Research, 40(5):549-566.




## 5.1 Introduction

The flow of the world's image across the moving retina (retinal flow) provides visual motion signals that can be used to infer our direction of translation relative to the scene. It is an intriguing question how the brain uses motion signals to extract the direction of relative movement between the self and the world, referred to as heading, in the presence of rotational flow that is added when the eye simultaneously rotates relative to the scene. Without sufficient information to solve this rotation problem, perceived heading may be affected, giving rise to random and systematic heading errors. Especially challenging it becomes when translating towards a fronto-parallel plane while rotating the eye about the vertical axis (horizontal rotation). In that case, the retinal flow much resembles the expanding flow that occurs during pure eye translation (translational flow), except that the center of flow (CF) is shifted horizontally, away from the heading direction (Fig. 5.1, top). Indeed, if only retinal flow information is available, systematic heading errors occur that are proportional to and in the direction of the CF shift (e.g. Royden et al. 1994). Although rotation about any axis causes the CF to be shifted when the axis of rotation and translation are

**Figure 5.1** *Shift of the center of flow (CF) away from the heading when rotation is about axes that deviate from the heading. Left panel: translation flow during pure translation (1.5 m/s) towards a fronto-parallel plane (7.5 m from the eye) has its CF (•) along the direction of heading (◯). In polar coordinates, the depicted heading direction has an eccentricity h of 12°, and a torsional angle Ψ of 180°. Central panels, top to bottom: rotational flow caused by pure rotation about the vertical axis (3 °/s), about the line of sight (8 °/s) or about the combined axis (vertical axis slanted in depth). In the resulting retinal flow (right panels) the CF is shifted horizontally by an amount $\Delta H_{CF}$ (top), torsionally by an amount $\Delta\Psi_{CF}$ (middle) or in a direction in between (bottom).*

oriented differently (Fig. 5.1, middle and bottom), psychophysical studies on heading perception so far focussed mainly on the effects of rotation about the vertical axis. We wondered whether the findings for rotation about the vertical axis also apply to rotation about an axis along the line of sight (torsional axis).

The eye not necessarily rotates about axes in the fronto-parallel plane, but may also have a considerable component of rotation about the line of sight. As formulated by Listing's law and confirmed by 3D eye movement recordings (Ferman et al. 1987), the torsional state of the eye is uniquely related to the eye's orientation with respect to the head, or more precisely, with respect to an imaginary plane through the head, called Listing's plane. From Listing's law, one can deduce that torsional velocities may occur during pursuit of points in the world when heading in a direction that is not normal to Listing's plane (van den Berg 1996b). This happens, for example, during horizontal pursuit while heading forward with the head tilted down. The torsional speed then approximately equals the pursuit velocity, multiplied by the sine of half the angle between the normal to Listing's plane and the fixation direction (Tweed et al. 1992). Limited by the eye's oculo-motor range (about 40°), torsional speeds can reach up to one third of the pursuit velocity. But, eye torsion relative to the scene can also occur without pursuit, for example when the head rolls about the line of sight and the eye does not counter-roll completely.

Several solutions to the rotation problem have been proposed on the basis of the retinal flow. Retinal image models take as input the retinal flow only (Longuet-Higgins & Prazdny 1980; Koenderink & van Doorn 1981; Rieger & Lawton 1985; Hildreth 1992; Heeger & Jepson 1992; Lappe & Rauschecker 1993; Perrone & Stone 1994; Royden 1997; Beintema & van den Berg 1998a). But, the eye's rotation velocity and its axis can also be inferred from extra-retinal information available by efference copy, proprioception or vestibular signals, which could help solve the rotation problem in several ways (Royden et al. 1994; Perrone & Stone 1994; van den Berg & Beintema 1997; Beintema & van den Berg 1998a; Lappe 1998).

Irrespective of the processing by the brain, however, how well heading can be retrieved is limited by how much the translational and rotational component in the visible retinal flow differ. This difference depends on several parameters, like the field

of view, the scene, the eye's orientation and the ego-motion. Considering that a certain amount of noise in the direction and magnitude of each flow vector is present because of 'noisy' processing in the input stage of the visual system, one can show the effect of several of these parameters (Koenderink & van Doorn 1987; Crowell & Banks 1996). For instance, limiting the field of view considerably complicates the task to differentiate rotational from translational flow. Also, the layout of the scene with respect to the eye is important. During translation, points that lie at different distances have different retinal velocities, although they might project to the same retinal area. This variation in velocity, called motion parallax, does not occur during rotation. Motion parallax is therefore a cue to help solve the rotation problem. Ego-motion parameters influence heading errors, because the rotation velocity relative to the ratio of translation speed and distance of points determines the change of the retinal flow due to rotation. But, as we will examine in this paper, also the orientation of the rotation and translation axis with respect to each other, and their orientation relative to the eye and the scene are important. Given a limited field of view, rotation about the eye's vertical axis causes flow that resembles a pure translation in the fronto-parallel plane, while rotation about the line of sight causes flow that can not be created by any translation through any environment.

The relative contribution of extra-retinal signals and visual signals has been investigated by psychophysicists using the so called simulated and real rotation paradigm. During simulated rotation, the subject fixates a stationary point to keep the extra-retinal signal non-informative of the rotational flow presented. Meanwhile, the subject views a display of motion that simulates observer translation through an environment of dots, while simultaneously rotating the eye. In the real rotation condition, only the translation is simulated, so that now the extra-retinal signal is informative of the rotational flow.

Evidence for the contribution of motion parallax in solving the rotation problem for rotation about the vertical axis has been provided in a number of studies. Simulated rotation studies suggest that humans can judge their direction of heading purely on the basis of retinal information alone (Rieger & Toet 1985; Warren & Hannon 1988, 1990; van den Berg 1992; van den Berg & Brenner 1994a; van den Berg 1996a; Stone & Perrone 1997). However, if motion parallax cues are absent because dots have been arranged in a fronto-parallel plane, people respond less correctly (Regan & Beverley 1982; Rieger & Toet 1985; Warren & Hannon 1988, 1990), making systematic errors in the direction of simulated rotation (Royden et al. 1994; Stone & Perrone 1997).

Few studies have paid attention to heading perception during rotation about the line of sight. Rieger and Toet (1985) measured heading discrimination thresholds during simulated rotation about randomly oriented axes, but did not differentiate explicitly between effects found for rotation about the line of sight and perpendicular to it. Torsional flow stimuli have been used to study the integration of motion over the visual field (Morrone et al. 1995; Ohtani et al. 1998). Others have varied the axis of rotation to study the sensitivity to full field rotation about different axes (Harris

& Lott 1995), or to study how well humans can locate the singular point in displays of simulated torsion to which is added simulated rotation about the vertical (Crowell et al. 1998b) and other axes (Te Pas et al. 1998). None of these studies, however, combined the rotation with an expanding motion component. Expanding and torsional flow have been combined in studies on the sensitivity to spiral motion (Freeman & Harris 1992; Snowden & Milne 1994) or the segregation of optic flow components (De Bruyn & Orban 1993). However, no heading task was involved, nor did a CF shift occur because translation was along the rotation axis (line of sight).

In this paper, we shall use the term compensation to quantify how well the visual system takes into account a change in retinal flow due to rotation. To this end, we compare the systematic error in perceived heading with the shift of the center of the flow (CF). We shall refer to compensation as the difference between the CF shift and the systematic error. We define performance as the ratio of compensation relative to the CF shift. Furthermore, we use the term torsional flow and horizontal flow to indicate the retinal flow induced by eye rotation about the line of sight and about the vertical axis, respectively.

In the first three experiments, we studied the compensation for simulated rotation about purely the torsional axis. Some of these data have been reported in preliminary form (Beintema & van den Berg 1998b). In experiment I, we investigate the effect of torsional flow in a stimulus without motion parallax cues. In experiment II, we analyse the effect of motion parallax by varying the depth in the stimulus. In experiment III, we study the effect of increasing the translation speed. In experiment IV, we vary the axis of simulated rotation and the amount of simulated depth. This allows us to directly compare the visual compensation for simulated torsion with the visual compensation for simulated horizontal pursuit, and to compare the contribution of motion parallax in both cases. Also, it allows us to examine the influence of torsional flow on the visual compensation for simulated horizontal pursuit.

## 5.2 Materials and Methods

Stimuli were rendered on a Silicon Graphics Onyx workstation and backprojected (Sony VPH 1270QM projection television, 815 by 611 pixels) onto a translucent 60 × 58° screen. Subjects were seated 1.5 m before the screen, with the head supported by a head- and chinrest. The position of the eye relative to the screen was measured for the viewing eye (van den Berg 1996a) in order to present the images in the right perspective.

The projected stimulus displayed the view of a camera that translates through a virtual scene of red dots and simultaneously rotates about its axis. On the display the dots were of constant size (0.2°). The dimensions of the scene were chosen such as to fill the entire field of view during the whole trial. The room was completely endarkened, apart from a dim glow of the screen that was visible in experiments I-III. In these experiments, a new trial was preceded by a 500-ms fully-illuminated screen

to prevent subjects from adapting to the dark.

During each trial, subjects had to fixate a red fixation point (0.4°) at the center of the screen. After the simulated self-motion stopped, the red dots and fixation point remained visible. Subjects indicated their heading direction relative to the scene as they perceived it at the end of the trial by moving a previously hidden red pointer (0.6°) that appeared by moving a hand-held mouse. If they felt no self-movement towards the scene, they were to indicate the direction of the moving dots relative to themselves. We explicitly told subjects that, should they perceive self-movement on a curved trajectory, not to point towards the end point of the future path, but to indicate the direction of relative movement along the tangent to the path taken at the position of the observer. We allowed 25 practice trials. No feedback was given during practice and the experiment.

In total, five subjects participated in the experiments, each having normal or corrected to normal vision. Stimuli were viewed monocularly with the left eye. Authors JB, BB and IH were experienced with heading stimuli. Subject IH was not aware of the conditions tested. Subjects ML and MF were naive towards the aim of the experiment and subject MF had no prior experience with heading stimuli.

To describe the rotations and translations, we use a coordinate system centered on the eye. The x-axis points towards the center of the screen, the z-axis upwards and the y-axis leftwards. Looking along the x-axis, positive torsion is defined as clockwise rotation of the eye, positive horizontal rotation as eye rotation to the left. The velocity and axis of rotation is denoted by a 3D vector $\mathbf{R}$, with components $R_x$ and $R_z$ referring to the torsional and horizontal rotation velocity, respectively. The torsional orientation of the eye is denoted by $\Psi$.

For two subjects, we recorded the 3D eye movements with scleral coils, using a Skalar magnetic field system (Collewijn et al. 1985).

## 5.3 Experiment I: Perceived heading during simulated torsion

Extra-retinal signals have been found to be essential for compensating horizontal (rotational) flow during approach of a fronto-parallel plane, because large errors were found for simulated rotation, but not for real eye rotation (Royden et al. 1994). Does the same hold for torsional flow? An asset of a fronto-parallel plane is that it allows a precise quantification of the shift of the center of flow (CF) because all points lie at the same distance. To calculate the CF shift when points are distributed in depth, one has to make assumptions on how each flow vector contributes because near points are more informative on the pattern of translational flow than points far away. Thus, we investigated whether the compensation for torsional flow needs extra-retinal signals by simulating eye torsion during approach of a fronto-parallel plane (wall).

### 5.3.1 Stimuli

Translation was simulated at translation speed $T$ of 1.5 m/s towards a wall consisting of dots randomly distributed in a plane parallel to the screen, perpendicular to the line of sight. Duration of the simulated self-motion was 2 s. The number of visible dots was 500 at the beginning, and 200 at the end of the trial.

Torsion was simulated at three velocities ($R_x = 0, \pm 8$ °/s). Because the CF is only shifted by rotation when heading direction and axis of rotation are not aligned, heading was chosen eccentric with respect to the line of sight, at fixed eccentricity $h$ of 15°. We presented 7 different heading directions by choosing the simulated torsional angle ($\Psi$) randomly between 0 and 360°. Because the subject's task was to indicate the final perceived heading, the CF shift was calculated for the wall distance at the end of the trial ($d = 5$ m in the last frame). For our choice of parameters, the shift of the CF is mainly in the torsional direction (Fig. 5.1, mid-right panel). Projected onto the fronto-parallel screen, the simulated heading directions and CF loci thus lie on a circle concentric with the fixation point. Because the perceived heading directions are also expected to lie on a circle, the pointer was adjustable only in torsional angle, having a fixed eccentricity (14°). Each stimulus was repeated 8 times. The 168 trials were presented in random order. Four subjects participated.

We derive the magnitude of the torsional shift from the case when the CF is initially at heading eccentricity $h$ towards the right. Clockwise rotation ($R_x > 0$) produces rotational flow that is oriented vertically at visual directions near the heading, and has angular velocity $p_R = \sin(h)R_x$. Given the radial pattern of directions in the translational flow, the rotational flow is canceled by translational flow only along the vertical meridian through the heading, at elevation $\theta$ for which $p_R$ equals the translational flow velocity ($p_T$). The latter is given by $p_T = \sin(\theta)T/d$. For small heading eccentricities ($h < 1$) and elevations ($\theta < 1$), we have a singular point in the flow at an approximate elevation of $\theta = hR_x d/T$. This quantifies the CF shift in visual angle. However, to obtain the CF shift in torsional angle, we divide by heading eccentricity $h$:

$$\Delta\Psi_{CF} = R_x\, d/T \tag{5.1}$$

Given the above parameters, we thus predict a heading error ($\Delta\Psi$) of 27° in torsional angle if none of the torsional flow is compensated. Note, in visual angle the CF shift is about 7°, given the heading eccentricity $h$ of 15°.

### 5.3.2 Results

During debriefing, subjects reported a strong sense of rotation, but a relatively weak percept of self-movement. They had no problem indicating the direction of movement of the array of dots relative to themselves.

Fig. 5.2 shows the mean torsional heading error ($\Delta\Psi$) as a function of simulated heading (torsional angle $\Psi$) for the three rotation velocities. All four subjects show

a significant bias in the direction of the shift of the singular point ($\Delta\Psi_{CF}$; indicated by arrows). Although subjects show significant variation in torsional error as function of torsional heading angle, no uniform trend across subjects is observed. Therefore, we plotted the mean error as a function of simulated torsional velocity $R_x$ for each subject (Fig. 5.3a). Two subjects (BB and JB) show a partial compensation because their errors are smaller than the shift of the singular point (indicated by the grey line). The other subjects (IH, ML) approximately pointed towards the singular point. These individual differences are also reflected in the standard deviations during simulated torsional flow, which are highest for IH and ML (Fig. 5.3b). In case of pure simulated translation ($R_x = 0$ °/s), we observe standard deviations of about 12°. We remind the reader that the torsional angle is subject to large variations when the perceived heading is close to the origin. To obtain the uncertainty in visual angle, the torsional angle must be multiplied by the tangent of the perceived heading eccentricity. However, subjects typically underestimate the simulated heading eccentricity (e.g. data by d'Avossa and Kersten (1996) show decreases in eccentricity up to 50%). On the assumption that the perceived heading eccentricity was less than simulated (15°), the uncertainty expressed in visual angle would be less than 3°. This is close to the 1-2° uncertainty reported for perceived horizontal heading eccentricity (e.g. Warren & Hannon 1988, 1990; Banks et al. 1996).

The large visual compensation found for two subjects seems to contradict the findings for simulated horizontal rotation during approach of a single plane (Rieger & Toet 1985; Warren & Hannon 1990; Royden et al. 1994). We therefore ran a control condition to investigate whether the two subjects who performed well had judged their heading on the basis of flow near the fixation point, as these motion vectors were least changed by torsional flow. To prevent the possible use of this local cue, we repeated the first experiment leaving out dots within a gap (15° radius) around the fixation point (Gap-2 s condition). A second control condition was included to investigate whether the larger errors found for the two other subjects could be due to the short (2 s) presentation time. Thus, we repeated the first experiment with increased presentation duration of 4 seconds (No gap-4 s condition).

Fig. 5.4 shows the effect of rotation for the first experiment and the two control conditions for our 4 subjects. Plotted is the slope of the linear regression (least-square) through the average heading error $\Delta\Psi$ as function of the simulated rotation $R_x$. Remarkably, elimination of the local cue (Gap-2 s) did not increase the slope with respect to the first experiment (No gap-2 s). On the contrary, as illustrated by the table of confidence intervals below the graph, leaving out the central area of visual motion even appeared to significantly improve the compensation ($p<0.05$) for 3 out of 4 subjects. Now, all subjects performed better than an observer who confuses the heading direction with the CF. Furthermore, compensation was not improved by increased presentation time (No gap-4 s). On the contrary, one subject even showed a significant decrease in compensation.

Possibly, the torsional flow stimulus elicited a torsional eye movement, which

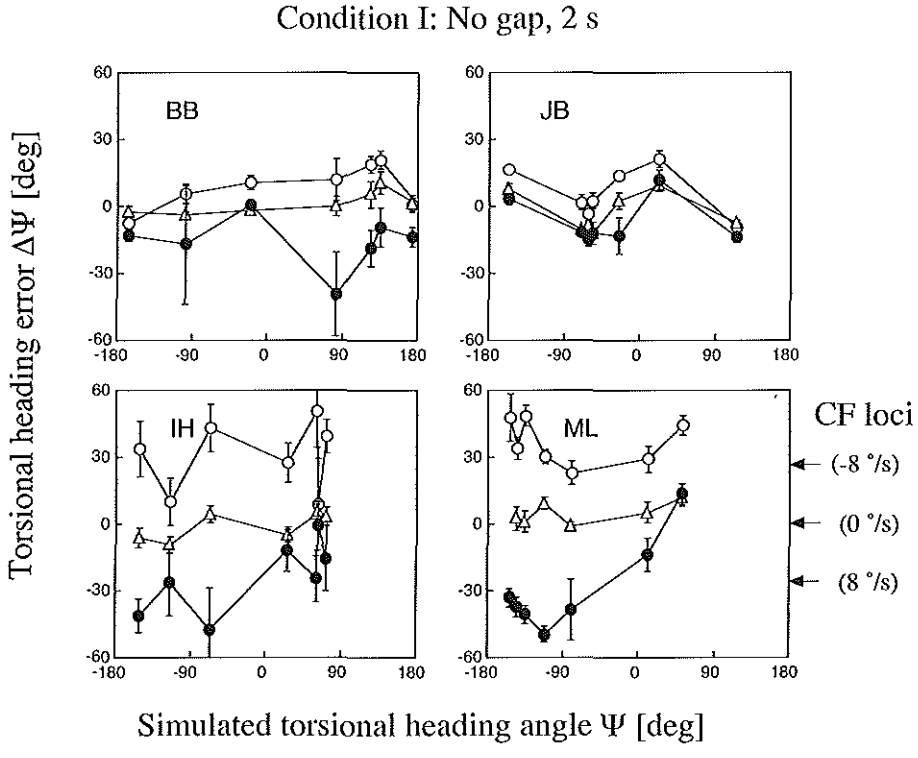



**Figure 5.2** *Difference in perceived and simulated heading as a function of simulated heading angle, at three simulated torsional velocities for all subjects. Heading angles denote torsional angles as defined in Fig. 5.1. The heading error and error bar indicate the mean and standard error in the mean from 8 trials. Simulated approach (speed 1.5 m/s; duration 2 s) towards a fronto-parallel plane (final distance 5 m), with final heading eccentricity 15°.*

would have reduced the amount of retinal torsional flow. We thus repeated the first experiment for subjects IH and BB, while measuring 3D eye movements. We found that the eye did rotate along with the stimulus, but with a gain of only 0.07 for both subjects. The perceived heading errors were the same as in the first experiment (no compensation for IH and large compensation for BB). Therefore, the gain value is too low to explain the large compensation for subject BB.

In summary, we found that torsional flow is partially compensated on a purely visual basis, without the availability of motion parallax or extra-retinal signals. The amount of compensation varied among subjects, being largest for the authors who were aware of the conditions tested. For the two naive subjects, we found no compensation in some conditions. We wondered whether the remaining error might be reduced by

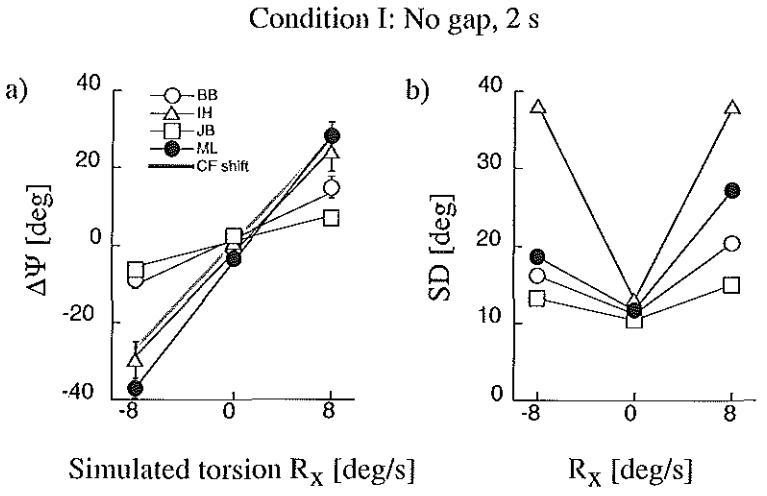



**Figure 5.3** *Heading error averaged over all simulated heading directions as a function torsional velocity for 4 subjects. Error bars indicate standard error in the mean. b) Standard deviations of the heading error as a function torsional velocity (for details, see legend Fig. 5.2).*

addition of depth to the scene, similar to what has been reported for simulated rotation about the vertical axis (e.g. Regan & Beverley 1982; Royden et al. 1994) and for simulated rotation about randomly oriented 3D axes (Rieger & Toet 1985).

## 5.4 Experiment II: Effect of depth during simulated torsion

We investigated the effect of depth by presenting simulated translation towards either a wall or a cloud. Initially, the wall was at a distance of 8 m. The cloud consisted of dots positioned in a viewing pyramid with the top at the eye, each dot having a random direction and a random radius between 6 and 10 m initially. Note, the choice of a polar distribution gives a gradient in dot density, with highest density near the eye. The number of dots is about the same in every depth plane, giving an average dot distance of about 8 m initially. Furthermore, we increased the number of different heading directions to 50, omitting any repetitions. Otherwise, parameters were the same as in the Gap-2 s condition of experiment I.

For the wall, the CF shift was computed using Eq. 5.1 (distance $d = 5$ m in the last frame). For the cloud, however, the variation in dot distance will lead to more than one singularity in the retinal flow during rotation. The shift of any such singularity lies within the range of CF shifts found for a fronto-parallel plane at the

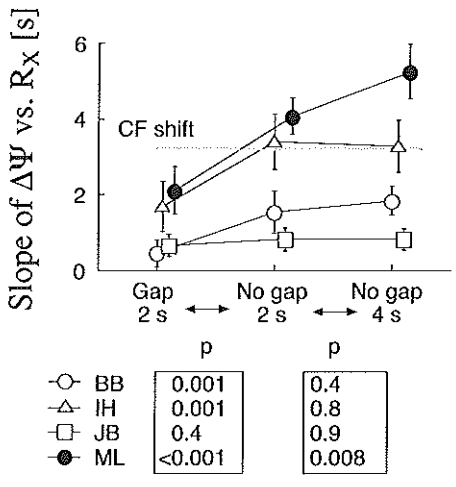



**Figure 5.4** *Regression slope of torsional heading error $\Delta\Psi$ vs. torsional velocity $R_x$, plotted for the first experiment and two control conditions. A regression slope of zero means that torsional flow is compensated fully. A slope that is equal or larger than the slope of the CF shift (grey line) means zero or negative compensation. Error bars indicate 95% confidence intervals. Below: individual significance levels (p) of difference between the slope of the first experiment and either control condition, obtained from a 2-tailed t-test (Steel & Torrie 1987, pp. 258).*

nearest and furthest dot distance. For now, we assume all dots contribute equally to the flow, and like Lappe et al. (1999), estimate the CF shift for a cloud by the CF shift for a plane at the average dot distance in the cloud ($d = 5$ m in the last frame).

### 5.4.1 Results

Fig. 5.5 shows the perceived heading angle as a function of the simulated heading angle at three torsional velocities for three subjects in the wall condition. The heading error is practically independent of the simulated heading. The intercept of the linear regression through the data gives the mean torsional heading error $\Delta\Psi$ for each rotation rate. As before, we quantify the effect of simulated torsion on this heading error by the slope of the linear regression of $\Delta\Psi$ as a function of torsional rotation $R_x$. Fig. 5.6 plots this effect as a function of depth range for each individual. Interestingly, none of the subjects showed a significant ($p<0.05$) decrease in the slope when motion parallax was added (see the table of p-values).

Thus, we find no effect of motion parallax cues on the compensation for torsional flow. This contrasts with the improved heading performance by depth in the scene that was reported for simulated horizontal rotations (Regan & Beverly 1982; Warren & Hannon 1988, 1990; Royden et al. 1994) and for simulated 3D rotations (Rieger

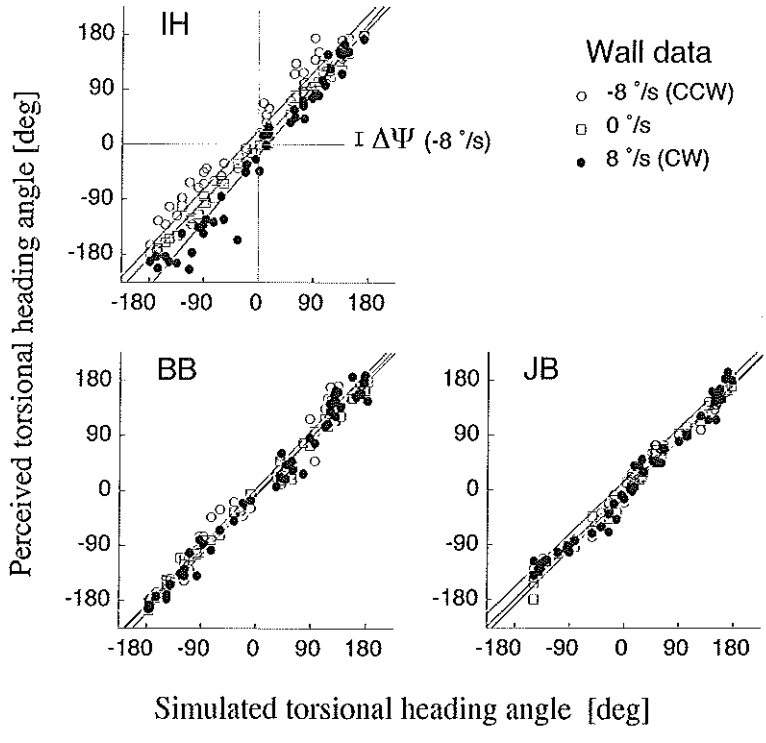



**Figure 5.5** *Heading error as a function of simulated heading angles, for 3 subjects at 3 simulated torsional velocities (0,± 8 °/s). CW and CCW refer to simulated clockwise or counterclockwise eye rotation. The average torsional heading error $\Delta\Psi$ follows from the intercept of the regression of perceived vs. simulated heading angle, as illustrated for subject IH. See Fig. 5.2 for other stimulus parameters.*

& Toet 1985). The pattern of individual differences in the amount of compensation is the same as in experiment I, with subject IH still showing largest heading errors. We wondered whether the compensation, in particular for subject IH, would be higher with increased translational flow magnitude.

## 5.5 Experiment III: Effect of translation speed during simulated torsion

We investigated the effect of translational flow magnitude by simulating motion towards a cloud at two translation speeds (T=1.5 and 3.0 m/s). If the error is related to the CF shift, we expect it to decrease since the ratio of translational over rotational flow magnitude decreases with increased speed (Eq. 5.1). A complication

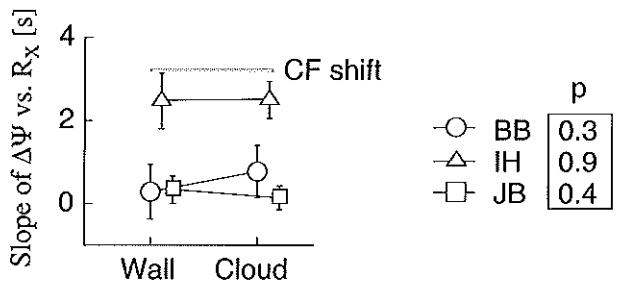



**Figure 5.6** *Effect of depth on the slope of torsional heading error $\Delta\Psi$ vs. torsional velocity $R_x$ for three subjects. Error bars indicate 95% confidence intervals. The slope for the CF shift (grey line), based on the average final dot distance, is the same for the wall (5 m finally) and the inhomogeneous cloud (3-7 m finally). Below: individual significancies (p) of difference between slopes for wall and cloud.*

arises, because increased speed also increases the travelled distance during the trial. Since decreased distance also decreases the CF shift (Eq. 5.1), we would need to make assumptions on the temporal integration of flow information as to compare the performance at two different translation speeds. We therefore decided to use clouds that remain equidistant with respect to the observer during the trial. This was accomplished by simulating translation towards dots, homogeneously distributed over a range extending beyond 18 m, of which only dots between 8 and 12 m from the observer were made visible. The number of visible dots was 160, and the mean distance of dots was 10.2 m throughout the trial. The other parameters were the same as in experiment II. Five subjects participated.

### 5.5.1 Results

Fig. 5.7 shows that for subject IH, the slope ($\Delta\Psi$ vs. $R_x$) is indeed significantly decreased for increased speed, indicating an increase in compensation. But, for most subjects, we observe no relation with the CF shift, since the slope remained the same (IH, JB, ML) or even increased (MF). In terms of performance, i.e. the proportion of compensation relative to the CF shift (grey line), it even appears that speed has a significant decremental effect on two subjects (MF, ML).

## 5.6 Experiment IV: Simulated torsion during simulated horizontal rotation

In the foregoing experiments, we found that the visual compensation for simulated torsional flow while approaching a wall is larger and less influenced by motion parallax

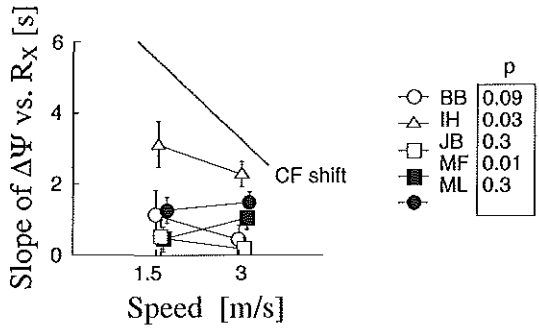



**Figure 5.7** *Effect of translation speed on the slope of torsional heading error vs. rotation rate, for five subjects. For clarity, data are slightly shifted horizontally. Error bars indicate 95% confidence intervals. In contrast to the perceived errors, the CF shift (grey lines) decreases with translation speed. Below: individual significancies (p) of difference between slopes for speed 1.5 m/s and 3.0 m/s.*

than has been reported for simulated horizontal rotation. For two reasons we wished to extend our inquiry to mixtures of horizontal and torsional rotation. First of all, we wondered why the addition of motion parallax did not improve compensation in our experiment II. Motion parallax increases with the depth in the scene. The scene that we used in experiment II is less extended in depth than the scenes that have been used in simulated horizontal rotation studies and could potentially explain the small effect. We therefore wanted to make a direct comparison between compensation for horizontal and torsional rotation in the presence of motion parallax. Secondly, we wondered to what extent the compensation for horizontal flow is influenced by torsional flow. To address these questions, we varied the orientation of the rotation axis and the amount of depth in the scene.

### 5.6.1 Stimuli

We simulated approach (1 m/s) to a flat cloud (0.5 m depth) or a deep cloud (10 m depth) at matched average distances of 7.5 m. The scene remained approximately equidistant with respect to the observer during the self-motion (1 s). This was accomplished for the flat cloud by simulating motion towards a box filled with dots, of which only dots between 7.25 and 7.75 m from the observer were made visible. For the deep cloud, dots were distributed in random polar coordinates as in experiment II, with random distances between 2.5 and 12.5 m with respect to the observer. The dots were repositioned once they came within 2.5 m from the observer or when they left the viewing pyramid. Dots were replaced each 250 ms, to ensure equal lifetime for both the flat cloud and the deep cloud. In either condition, 100 dots per frame were visible on average.

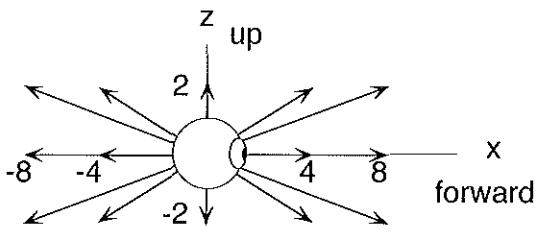



**Figure 5.8** *Possible final simulated rotation vectors. The x-axis is oriented along the subject's line of sight. The length of the vector indicates the rotational velocity, its orientation corresponds to the axis of simulated rotation.*

The orientation of the rotation axis was varied by simulating three horizontal pursuit velocities ($R_z = 0, \pm 2$ °/s) at five torsional rotation velocities ($R_x = 0, \pm 4, \pm 8$ °/s). Importantly, we simulated torsion about the simulated line of sight (Euler's definition of torsion) as opposed to a head-fixed torsional axis (Listing coordinates). Thus, we simulated a torsional axis that rotated with respect to the scene during horizontal pursuit. We made sure, however, that at the end of the trial, the simulated 3D eye orientation was always the same. Fig. 5.8 shows the 15 different axes of rotation in the last frame. When horizontal and torsional rotation are combined, the axis is along an axis slanted in depth. The torsional heading angle was varied in steps of 15°. The final heading eccentricity was fixed at 12°. In total, one session consisted of 720 trials; 15 rotations, 24 heading directions and 2 depth ranges, all presented in random order. Each of the four subjects completed three sessions. Because we also expect heading shifts in the horizontal direction, the pointer was free to move in all directions. The pointer appeared at a random location within a circle (12.6° radius) around the fixation point.

To compare the compensation between different axes of rotation, we need to quantify the CF shift. Fig. 5.1 shows how the flow pattern for heading towards the right is changed by horizontal, torsional or combined rotation. For horizontal rotation, the shift of the CF is mainly in horizontal direction by an amount $\Delta H_{CF}$. For purely torsional rotation the CF shift is mainly in the torsional direction by an amount $\Delta \Psi_{CF}$. For combined rotation, however, the shift is in both directions. We derived an exact analytical solution for the direction and magnitude of the CF given any translation $\mathbf{T}$ and rotation $\mathbf{R}$ (see Appendix). As a rule of thumb, for small angles between $\mathbf{T}$ and $\mathbf{R}$, like in our stimulus, the retinal CF is restricted to lie on a circle through $\mathbf{T}$ and $\mathbf{R}$, the circle being centered on the bisectrix of $\mathbf{T}$ and $\mathbf{R}$.

How does rotation affect the CF loci for the whole set of heading directions? We plotted the CF shifts at each simulated heading direction for different rotation axes (Fig. 5.9). Each vector starts at the final simulated heading direction and points towards the center of the retinal flow. For pure expansion (no CF shift, central panel),

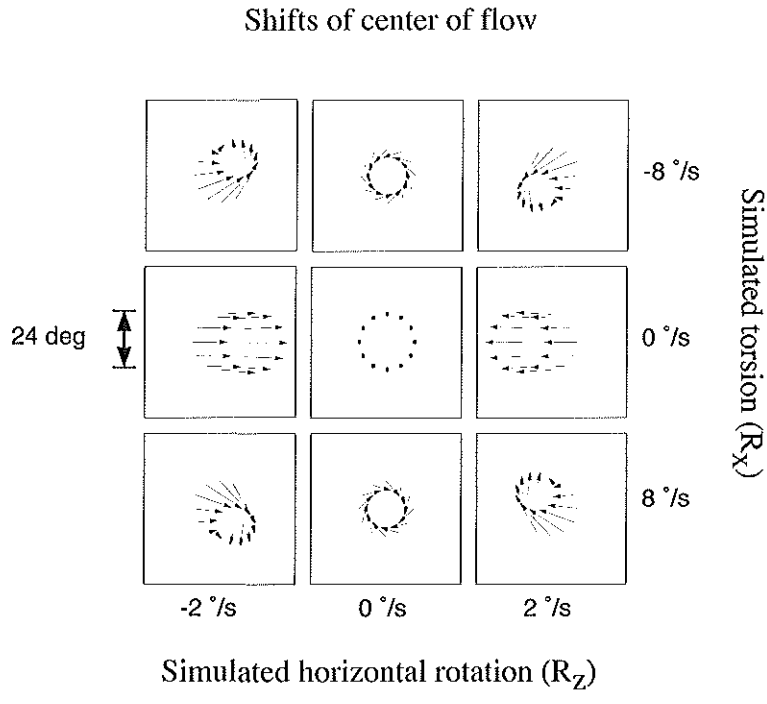



**Figure 5.9** *CF shifts for the different simulated heading directions and 9 axes of simulated rotation, as projected onto the screen. The simulated torsional and horizontal rotation velocity varies from panel to panel. Each vector starts at the simulated heading direction and points towards the corresponding CF in the last frame.*

the vectors reduce to points lying on the circle of simulated heading directions. For purely horizontal rotation (mid-horizontal panels) or pure torsion (mid-vertical panels), the endpoints are shifted mainly horizontally or torsionally, respectively. The endpoints form a circle that is slightly compressed for simulated torsion, whereas for simulated horizontal rotation the circle is shifted horizontally and slightly increased in magnitude. For combined rotation (diagonal panels) the endpoints still form a circle, which now is not only compressed and shifted horizontally, but also shifted vertically. Thus, we can quantify the effect of rotation on the circle of CF loci by 4 parameters; one for magnitude and three for the shift in horizontal, torsional and vertical direction. Because the flat and deep cloud were matched with respect to the average distance of dots, the CF shifts are the same in both conditions.

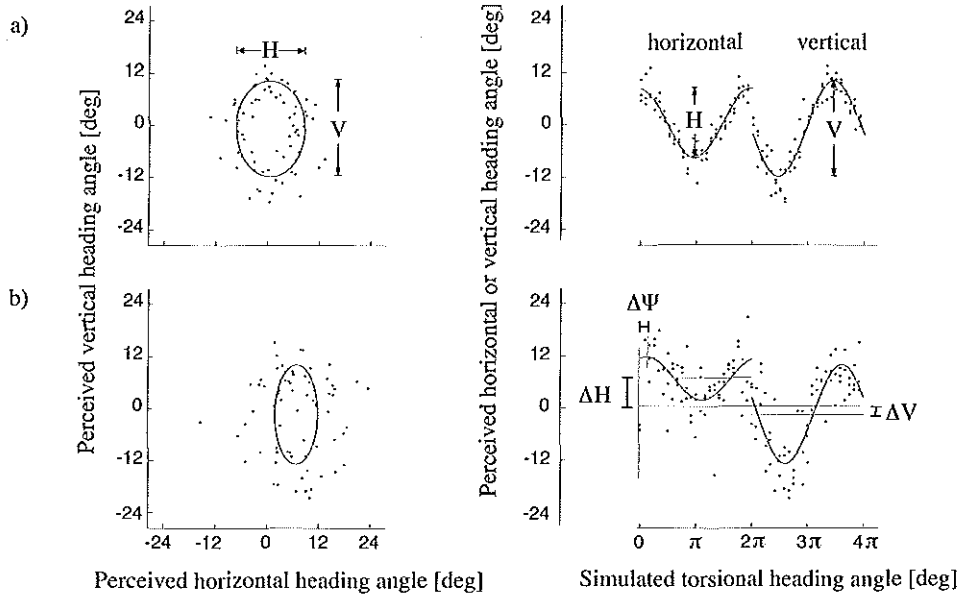



**Figure 5.10** *Perceived heading directions for subject IH for (a) the condition without rotation and approach of the deep cloud, and for (b) combined rotation ($R_x$=8 °/s and $R_z$=2 °/s) and flat cloud. Left panels show the vertical versus horizontal components of perceived heading. Right panels show the horizontal and vertical components plotted sequentially as a function of simulated torsional heading angle $\Psi$. The fits (solid lines) form ellipsoids (left panels) or discontinuous functions of $\Psi$ (right panels). The RMS errors between fit and data in a) and b) were 2.6 and 6°, respectively.*

## 5.6.2 Results

The perceived heading directions are shown for two cases (Fig. 5.10). Each data point corresponds to one trial. The solid lines are fits to the data points. In the left panels, we plotted the vertical versus horizontal direction of perceived heading. Without rotation (Fig. 5.10a) the fit to the data is an ellipsoid at the center of the screen that closely resembles the circle of simulated heading directions. For combined torsional and horizontal rotation (Fig. 5.10b), however, we see the ellipsoid is horizontally offset by an amount $\Delta H$, and vertically offset by an amount $\Delta V$. Apart from offset errors, we also observe scale errors as the ellipsoids are compressed with respect to the circle of heading direction (24° diameter). Interestingly, the ellipsoids are more compressed in horizontal than in vertical direction. This is remarkable, since anisotropic scaling does not follow from the CF shifts (Fig. 5.9). Consequently, we need both a horizontal magnitude ($H$) and a vertical magnitude ($V$) to describe the data properly.

A rotation of the ellipsoid along the line of sight, as predicted when pointing

towards the CF during pure torsion (Fig. 5.9, mid-top panel), is not revealed in the present format. Thus, another format was required to estimate the offset error in torsional angle. To this end, we sequentially plotted the horizontal and vertical perceived heading as a function of the simulated heading (torsional angle) $\Psi$ (Fig. 5.10, right panels). Veridically perceived heading would produce a cosine cycle followed by a sine cycle. The data for combined rotation (Fig. 5.10b), however, show a cosine-sine curve that has a small torsional offset ($\Delta\Psi$) along the $\Psi$-axis. Together with parameters ($\Delta H$, $\Delta V$, $H$, $V$), we can thus describe the data by the following 5-parameter function PH:

$$PH(\Psi) = \begin{cases} \Delta H + H\ \cos(\Psi + \Delta\Psi) & \text{for } \Psi < 2\pi \\ \Delta V - V\ \sin(\Psi + \Delta\Psi) & \text{for } \Psi > 2\pi \end{cases}$$

Because this function is non-linear in $\Psi$, we used the Levenberg-Marqhuardt fit procedure. Fits were run for each case of simulated rotation, depth range and subject. The root mean square (RMS) error between data and fit does not allow a quantification of $\chi^2$ or a goodness of fit, because we lack sufficient repetitions (only 3) per heading direction to estimate the uncertainty in an individual measurement (Press et al. 1989). But, even for high (6.0°) and low (2.6°) values within the total range of RMS errors (between 2 and 7°), the fits look reasonable good (Fig. 5.10, right panels). Pooled over subjects, rotation axes and heading directions, an unpaired 2-tailed t-test revealed that the RMS error was significantly smaller for the deep cloud data than for the flat cloud data ($p<0.05$). We also found that the RMS error varied for the different rotations (listed in order of magnitude: no rotation, pure torsion, pure horizontal rotation, combined rotation), but none of the differences (pooled over other conditions and subjects) were significant at 5% chance level. On the assumption that each ellipsoid fits the data well and that the uncertainty is the same for each simulated heading, the RMS error is indicative for the uncertainty in an individual measurement (Press et al. 1989). Averaged over all conditions and subjects, we find the RMS error (3.5°) is only slightly higher than the standard deviation (2-3°) in perceived heading during simulated horizontal rotations (van den Berg & Brenner, 1994a).

Fig. 5.11 shows the fitted ellipsoids for subject IH. Each panel represents a different rotation, showing the ellipsoid for the deep (thick) and flat (thin) cloud condition. Three effects are apparent. First of all, we observe a clear effect of motion parallax on the horizontal offset of the ellipsoids. The ellipsoids for flat clouds are more offset horizontally, away from the central circle of simulated headings, than the ellipsoids for the deep cloud. Secondly, we observe scale errors. Without rotation (central panel), the horizontal magnitude is smaller and less veridical than the vertical magnitude. For horizontal rotation, the horizontal compression is even increased, especially for the flat cloud condition. Thirdly, for combined horizontal and torsional rotation in the flat cloud condition, we observe a pattern of vertical offsets similar to that predicted by the CF shifts (Fig. 5.9).

To compare the compensation for simulated rotation between torsion and horizontal rotation, we take into account that each rotation affects the circle of shifted CF

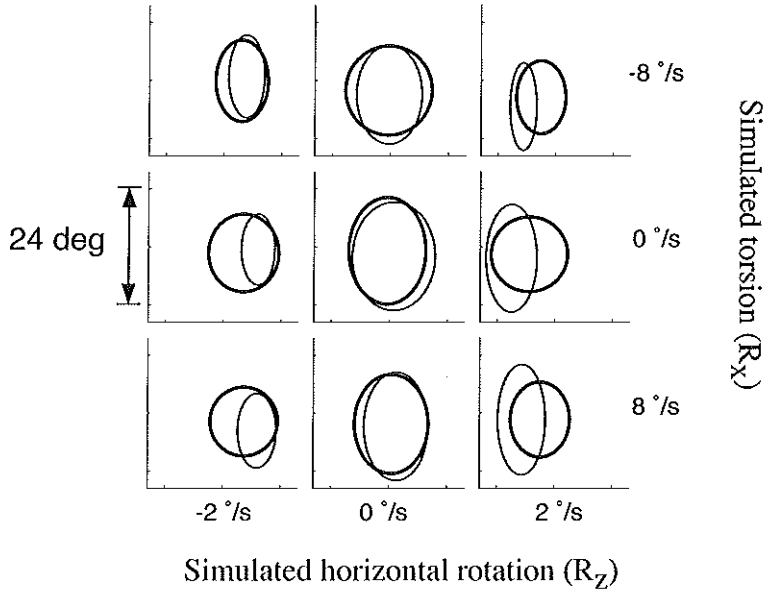



**Figure 5.11** *Fitted ellipsoids to the perceived heading data for the deep (thick line) and flat cloud (thin line). Each panel represents a different rotation axis, like in Fig. 5.9. Note that Fig. 5.9 is plotted at different scale.*

differently. We therefore compare performance, defined as the compensation as fraction of the CF shift. The main effect of pure torsion is a torsional offset (Fig. 5.12a). The flat cloud data for each individual (open symbols) show only a small effect, this being less than half of that predicted by the CF shift (solid grey line). This corresponds to a performance ranging from 0.5 up to almost unity. The addition of depth (filled symbols) does not reduce the torsional offset, apart from one subject (IH), who does show a small improvement for negative torsion velocities. The high performance and the absence of an effect of motion parallax support our previous conclusions from experiment I-II. Pure horizontal rotation, on the other hand, does have a large effect, since in the flat cloud condition (Fig. 5.12b, open symbols) the horizontal offset is larger than half of the horizontal CF shift (solid grey line). This corresponds to a performance less than 0.5. Moreover, the addition of depth (deep cloud, filled symbols) does cause a significant increase in the compensation. These results show that performance is higher during torsion than during horizontal rotation.

Next, we analysed the effect of torsional flow on all parameters that describe the ellipsoid of perceived heading. Specifically, we investigate whether torsional flow influences the compensation for simulated horizontal rotation. To this end, we plotted not only the ellipsoid's horizontal and torsional offset ($\Delta H$, $\Delta \Psi$), but also its vertical

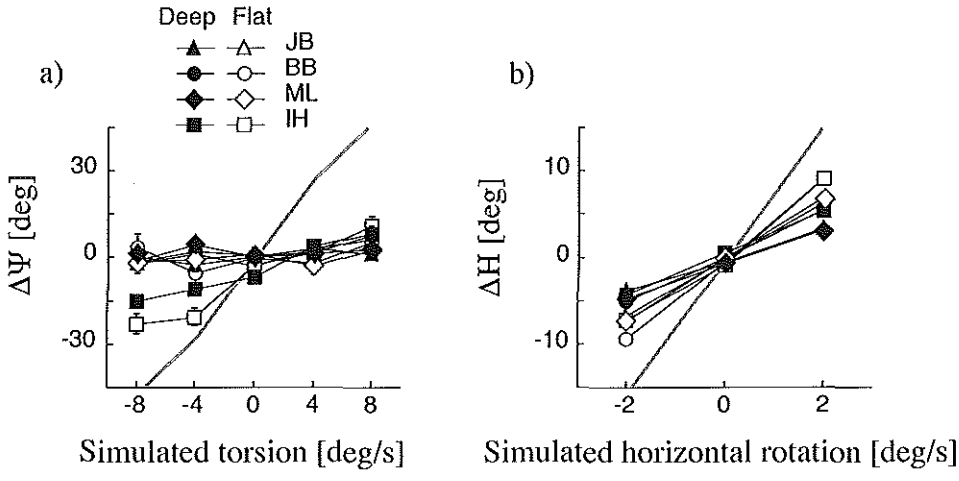



**Figure 5.12** *a) Torsional offset error as a function of simulated torsion ($R_z = 0$ °/s). b) Horizontal offset error as a function of simulated rotation purely about the vertical axis ($R_x = 0$ °/s). Open and filled symbols refer to the flat cloud and deep cloud, respectively. The ordinate values and standard error bars are obtained from the non-linear fit to the set of 2D perceived heading. Grey lines indicate the offsets predicted when pointing to the CF.*

offset ($\Delta V$) and the horizontal and vertical magnitudes ($H$, $V$), each as a function of the torsional velocity (Fig. 5.13a-e). The data show the means over 4 subjects, split by horizontal rotation and depth range (open symbols: flat cloud; filled symbols: deep cloud). Clearly, we observe no effect of torsional velocity on any of the parameters, except on the torsional offset.

More specifically, Fig. 5.13a shows that if subjects do not compensate for torsional flow during simulated horizontal rotation, and would thus point towards the CF (solid grey lines for -2 and 2 °/s), addition of torsional flow should decrease the horizontal offset. This trend is not observed in the data ($\triangle$ and $\triangledown$), since the partial compensation for simulated horizontal rotation is not changed by simulated torsion. Moreover, Fig. 5.13b shows that if subjects would point towards the shifted CF (solid grey lines -2 and 2 °/s), simulated torsion during simulated horizontal rotation should increasingly offset the ellipsoid in vertical direction, up to values as large as ±7.5°. The data show no such variation in vertical offset ($\triangle$ and $\triangledown$). Any systematic variation in vertical offset with rotation axis observed for subject IH (Fig. 5.11) apparently cancelled out by the averaging over subjects. In addition, Fig. 5.13c and d show that the horizontal and vertical magnitude is expected to decrease if the torsional flow is not compensated (solid grey lines), a trend not observed in the data either ($\triangle$ and $\bigcirc$). Finally, Fig. 5.13e shows that the predicted absence of an effect of simulated

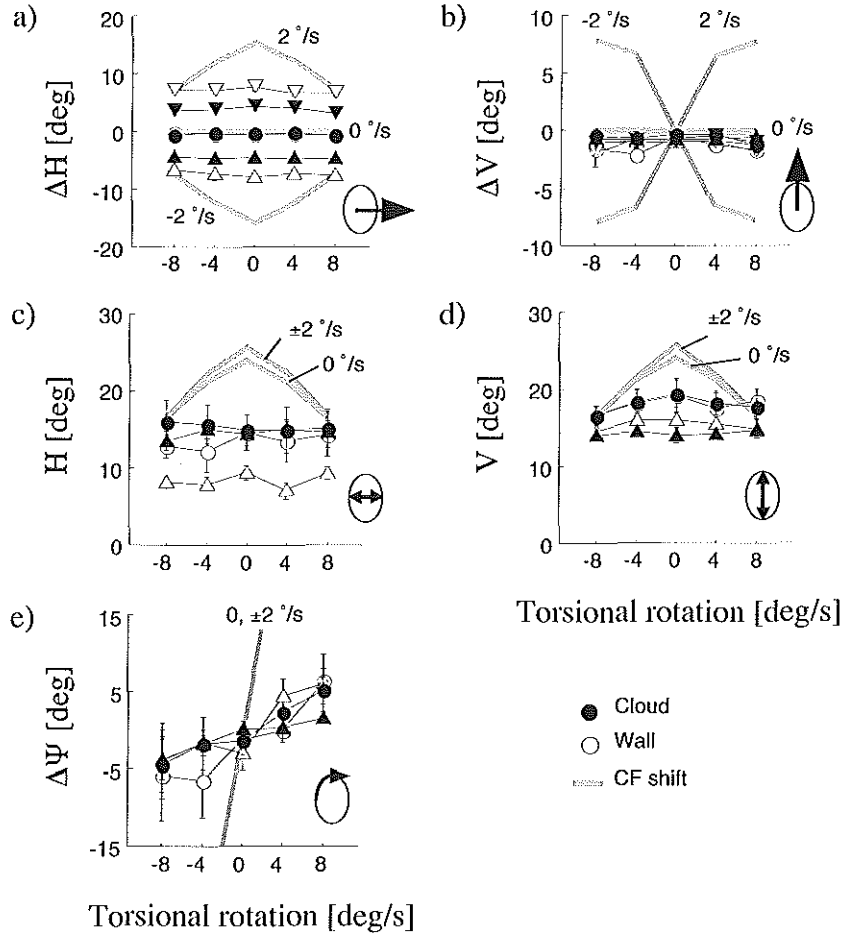
horizontal rotation on the torsional offset (solid grey lines), is indeed confirmed by the data. In summary, the compensation for simulated torsion seems virtually complete, and simulated torsion does not influence the compensation for simulated horizontal rotation.

Fig. 5.13c and d also confirm the effect of horizontal rotation already observed for subject IH in Fig. 5.11. If the deviation from veridical of each perceived heading is proportional to and in the direction of the shifted CF, we would expect a small increase in the ellipsoid's horizontal and vertical magnitude during horizontal rotation (compare solid grey lines for $\pm 2$ °/s with 0 °/s). The flat cloud data, however, show a decrease in magnitude (compare ○ with △). This decrease is large in horizontal direction, and small but still significant in vertical direction. The horizontal/vertical anisotropy disappears when depth is added to the scene (filled symbols).

## 5.7 Discussion

### 5.7.1 Summary

We investigated the systematic heading errors that occur during simulated rotation about an axis that contains the line of sight (torsional axis), while translating along a different direction. We first examined the effect of pure torsion. For translation towards a fronto-parallel plane (experiment I), we found that the heading error was small compared the shift of the center of flow (CF), which in our terminology means a large compensation (CF shift minus heading error). Longer stimulus duration did not decrease heading errors, while removal of the central part of the flow did not increase heading errors. Also, the heading error was neither decreased by increased depth in the scene (experiment II), nor by increased translational flow magnitude (experiment III). The high performance (compensation as a fraction of the CF shift) and lack of influence of depth found for simulated torsion seemed to contrast with previous reports for simulated rotation about the vertical axis. Therefore, in experiment IV, we explicitly varied the axis of simulated rotation and the depth range. We found the performance was indeed higher for torsion than for horizontal rotation. Moreover, in contrast to the compensation for torsion, the compensation for horizontal rotation did improve when depth was increased. Furthermore, we found no influence of simulated torsion on the errors caused by simulated horizontal rotation. Apart from offset errors in perceived heading, we also found scale errors which were not related to the CF shifts. In the horizontal direction perceived headings were more compressed than in the vertical direction. Moreover, for simulated horizontal rotation, we found that the horizontal compression was even further enlarged.




## 5.7.2 Offset errors: comparison with earlier studies

Experiment IV revealed a higher performance for simulated torsion than for simulated horizontal rotation. For simulated torsion, the ellipsoid of perceived heading was offset in torsional angle by about one seventh of that expected without compensation, corresponding to a performance close to 0.9 (see Fig. 5.13e). For simulated horizontal rotation, on the other hand, the ellipsoid of perceived headings was offset horizontally by one fourth (0.75 performance for deep cloud) and one half (0.5 performance for flat cloud) of the CF shift (see Fig. 5.13a). How do these findings compare to the literature?

To our knowledge, only Rieger and Toet (1985) simulated rotation about axes that included the line of sight torsional axis, as they simulated rotation about randomly oriented 3D axes. They found that thresholds for discriminating between four directions of heading, rose more quickly with increasing rotation rate for a single fronto-parallel plane than for two planes separated in depth. At first hand, their results seem to contradict our finding that the compensation for simulated torsion is not affected by depth range. Yet, we do observe an effect of depth for simulated horizontal rotation.

**Figure 5.13** *Effect of torsional rotation on a) horizontal offset error, b) vertical offset error, c) horizontal magnitude, d) vertical magnitude and e) torsional offset error. Data are the average over 4 subjects, and error bars indicate the standard error in that mean. Open and filled symbols indicate flat and deep cloud conditions, respectively. For the horizontal and vertical offset errors, data are plotted for each horizontal velocity separately, i.e. $R_z$ = -2 (▽), 0 (○) and 2 (△) °/s. For the torsional offset error, the horizontal and vertical magnitude, we show only the average over the non-zero horizontal rotation data (△). Solid grey curves show the errors expected when pointing to the CF.*

Rieger and Toet (1985) did not split out the effect of depth between rotation along and perpendicular to the line of sight. Since rotation perpendicular to the line of sight occurred in about two thirds of their trials, the effect of depth range will likely have dominated their thresholds.

Regarding simulated horizontal rotation, the large performance (0.75) found in the deep cloud condition, falls within the range of performance estimated from other studies in which cloud stimuli were used (van den Berg & Brenner 1994b; van den Berg 1996a, 1997; Stone & Perrone 1997). An overview of these performances was recently given by Lappe et al. (1999), based on heading errors as a fraction of the CF shift for a fronto-parallel plane at the average point distance in the cloud. Interestingly, their Fig. 2 revealed that the cloud data by Royden et al. (1994), which Royden et al. interpret as evidence against visual compensation, do show about half compensation for the CF shift. The performance estimated from the cloud data by Banks et al. (1996), clearly do fall below the others. One suggested explanation for this low performance is that subjects might have perceived curved paths instead of linear paths, and might have indicated a point on their future path instead of their instantaneous heading direction (Royden 1994; Ehrlich et al. 1998).

The increased compensation that we find for simulated horizontal rotation when depth in the scene is increased, is also in line with several other studies (Regan & Beverly 1982; Warren & Hannon 1988, 1990; Stone & Perrone 1997). Even the simulated rotation data by Royden et al. (1994) support an increase in performance (Lappe et al. 1999), on the assumption that all points in the cloud contribute equally. Regarding the performance in our flat cloud condition, however, we do observe a quantitative difference with other reports on systematic errors during approach of a fronto-parallel plane (Royden et al. 1994; Stone & Perrone 1997). From the data by Royden et al. (1994), we deduce that heading errors about equal the CF shift ($R\,d/T$ equals 20° at a rotation rate $R$ of 5 °/s, a translation velocity $T$ of 0.5 m/s and a distance $d$ of 2 m). The data by Stone and Perrone (1997), who simulated curvilinear motion, show heading errors that are also close the CF shift. In contrast, our flat cloud condition, which approximates a fronto-parallel plane, shows that about half of the CF shift is compensated.

The compensation for torsional rotation differed among subjects, varying from

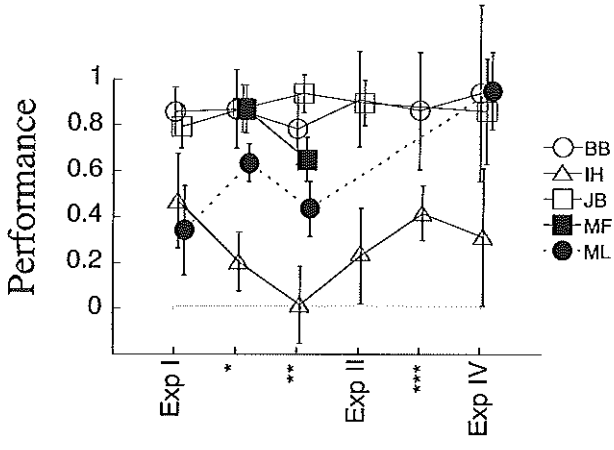



**Figure 5.14** *Performance ($= 1 - \Delta\Psi/\Delta\Psi_{CF}$) in all repetitions of experiment I (Gap-2 s), in which torsion and 1.5 m/s translation towards a fronto-parallel plane are simulated for 2 s, with the central part of the flow left out. Data are arranged in sequence of time, split by subject. The ratio ($\Delta\Psi/\Delta\Psi_{CF}$) is obtained from the ratio of regression slopes of $\Delta\Psi$ vs. $R_x$ and $\Delta\Psi_{CF}$ vs. $R_x$. Error bars indicate 95% confidence intervals. Also included are other data obtained during a repetition of experiment I Gap-2 s (*), a repetition of experiment II (**) and during eye movement recordings (***). Selected flat cloud conditions in experiment IV differed from the others in duration (0.5 s), pointer task (2D), translational velocity (1 m/s), gap condition (central part was visible).*

less than half compensation (IH, ML) to nearly complete compensation (JB, BB and MF). This pattern of individual differences in performance was roughly preserved over different conditions (Fig. 5.4a) and over repetitions of the same conditions in different sessions (Fig. 5.14). The same pattern of individual differences was reflected in the standard deviations (Fig. 5.3b), and in the compensation for horizontal rotation (Fig. 5.12b). One could wonder if errors were smallest due to experience (authors). But, large compensation was also seen for subject MF, who was naive towards the aim of the experiment and not experienced with heading stimuli. Large differences in individual performances are also reported in other studies on simulated horizontal rotation (e.g. van den Berg 1996a).

Qualitatively, the variation in perceived heading with simulated horizontal rotation, depth and subjects is in line with most other studies. We do find a quantitative difference in performance between our data and literature in the condition without depth.

### 5.7.3 Performance difference between horizontal and torsional rotation

Presented with identical simulations of observer translation, subjects performed better during simulated torsion than during simulated horizontal rotation. Moreover, whereas for simulated horizontal rotation the performance reduced when removing depth in the scene, the performance for simulated torsion remained constant, irrespective of the simulated horizontal rotation. These findings support the idea that the visual compensation is larger for torsion than for horizontal rotation. We propose this reflects a difference in the available flow information rather than a special compensation mechanism. As already mentioned in the introduction, rotation about the eye's vertical axis causes flow that for limited field of view resembles a pure translation in the fronto-parallel plane, whereas rotation about the line of sight causes flow that can not be created by any translation through any environment. We here generalise this idea to other scenes by the hand of possible information sources on rotation.

A first means to estimate the rotation is to take the integral of the flow along circles about the rotation axis. Such circle integral has the nice property that translational flow adds little to it. But note, this holds only on the assumption that the translational flow contributions are cancelled by mirror symmetry in the plane spanned by the translation and rotation axis, such as for points that are homogeneously distributed in 3D space. Because the field of view is limited to the hemisphere around the line of sight, rotation about this axis yields a more complete integral and thus a more accurate estimate of rotation than rotation about an axis perpendicular to it.

A second means of estimating the rotation, is the velocity gradient in the flow. The rotational flow velocity increases with the sine of the angle between the visual direction of the flow vector and the rotation axis. The same sine relation holds for translation flow, on the assumption that points are equidistant with respect to the eye. However, whereas translational flow is oriented along the direction of velocity gradient, the rotational flow is oriented perpendicular to the direction of the velocity gradient. Thus, measuring the magnitude of a shear velocity gradient offers a second way to estimate the rotation, without the interference of the translational flow contributions. Because the shear gradient is largest near the poles of the rotation axis, rotation about the line of sight is again potentially more accurate, since the poles of axes perpendicular to the line of sight are not visible.

Both means of estimating the rotation would predict highest performance for torsion. We must recall, though, that the use of such rotation estimates does require certain assumptions on scene layout. The circle integral assumes mirror symmetry with respect to the rotation and translation, while the shear gradient assumes equidistance with respect to the eye. These assumptions are met for motion through a homogeneous 3D cloud and for motion towards plane with its normal along the line of sight. But when, for instance, the observer translates horizontally above a ground plane, the mirror symmetry is broken (only part of the view is filled) and points are not equidistant. In that case, not only torsion, but also motion parallax creates a shear

gradient so that an accurate estimate of the rotation might fail. On the other hand, the assumptions on layout need be less strict when the heading is more along the line of sight. For if the heading is exactly along the rotation axis, the components of rotational and translational flow, locally, are perpendicular to each other throughout the whole visual field. Here again, the potential advantage of torsional flow over rotation in the fronto-parallel plane emerges; heading is most probably directed along the line of sight, not sideways.

Because the central part of the flow was visible in experiment IV, one might be tempted to explain the higher performance for torsion by the smaller rotational flow magnitude around the fixation point. However, as the control in experiment I pointed out, removal of visible dots within a 15° gap did not reduce the compensation for pure simulated torsion. Thus, we can reject the possibility that performance in experiment IV was high because the central part was included.

For our performance measure, we compared the perceived heading with the shift of the center of flow (CF). The CF was based on the location of the singularity in the flow that would arise when translating towards a plane at a distance that equaled the average distance of dots in the scene. Previous studies (Royden et al. 1994; Stone & Perrone 1997) showed such measure is a good predictor of the maximum heading error caused by horizontal rotation. Moreover, these studies show that the heading error was proportional to the simulated rotation rate, and to some degree inversely proportional to the translation speed. Therefore, we assumed that quantifying the heading error as a fraction of the CF shift would be appropriate to compare the performance for torsion and horizontal rotation.

However, some of our findings cast doubt on whether the CF shift adequately describes the effect of torsion to be expected without compensation. In Fig. 5.7, we see that a decrease in the CF shift due to an increase in translation speed is not reflected in the overall amount of compensation. Also, we note that the horizontal and vertical magnitude of the fitted ellipsoid do not show any of the change with torsional velocity expected if pointing to the CF (Fig. 5.13c and d: $R_z$=0 °/s).

To check whether for torsion the CF shift is a good predictor of the torsional heading error as well, we compared the location of the singularity with the output of a heading model based on a 2D array of pure expansion templates (Beintema & van den Berg 1998a). Each of these templates is tuned to a specific preferred 2D heading direction and evaluates the evidence that the retinal flow field resembles the template's preferred flow field. It integrates this likelihood over all flow vectors, weighting each local flow contribution such that the template's response is insensitive to the distance of points. We found that the preferred heading of the maximal responding template was not only shifted away from the true heading in the direction of the singularity shift, but also to a greater extent. This suggests that our estimation of the performance for torsional flow has even been too small rather than too large.

### 5.7.4 Scale errors

Apart from offset errors, experiment IV also revealed the occurrence of scale errors (Fig. 5.13c and d). A scale error, i.e. a compression of the ellipsoid of perceived headings, is already seen without simulated rotation, since the horizontal (12-16°) and vertical magnitude (16-20°) of the fitted ellipsoids (Fig. 5.13c and d, ○) are markedly smaller than the diameter of the circle of simulated heading directions (24°). The compression of the ellipsoid is especially large in horizontal direction and more pronounced for the flat clouds than for the deep clouds (compare open and filled symbols). Moreover, during simulated horizontal rotation, the compression is enlarged, being most pronounced in the flat cloud condition and mostly in horizontal direction (open triangles). As shown by the CF curves (solid grey lines), scale errors are to be expected when none of the rotation is compensated. However, these would not only predict an increased instead of a decreased magnitude during horizontal rotation, but also an equal effect instead of an unequal effect on the horizontal and vertical magnitude.

A bias of perceived heading towards the screen center has been reported in several studies on simulated observer translation (Llewellyn 1971; Johnston et al. 1973; Cutting et al. 1992; D'Avossa & Kersten 1996). Recently, Hooge et al. (1999) also reported the presence of a systematic bias towards the screen center for a saccadic pointing task in which subjects were asked to look towards their perceived heading. Even after 1.5 s, after a few saccades had been made, the eye's eccentricity was still less than the simulated heading eccentricity (15°).

The occurrence of both a horizontal and vertical bias towards the center of the screen in the flat cloud condition for pure translation fits with theories that predict a bias of perceived heading towards the simulated plane's normal, which in our case is the same as the screen's normal. Relying on the flow's maximum divergence (Koenderink & van Doorn 1981), a heading detection scheme proposed for its rotation invariant property, predicts a bias away from the simulated heading because the direction of maximum divergence bisects the plane's normal and the heading direction. Longuet-Higgins (1984) offers another explanation for such bias, showing that the approach of a fronto-parallel plane along a trajectory that deviates from the normal to the plane causes flow that for a brief period is indistinguishable from an approach with the normal to the plane and the heading direction interchanged. Taking the average over these possible solutions would thus also lead to a bias towards the plane's normal. Although Warren et al. (1988) found no evidence for a bias towards the plane's normal when simulating translation at oblique angles with the simulated plane's normal, a recent study by Grigo and Lappe (1999) shows that for short duration (0.5 s) such bias does occur.

The above theories, however, fail to explain why compression of the ellipsoid during pure simulated translation is larger in horizontal direction than in vertical direction. A horizontal/vertical anisotropy has been found before for simulated observer translation (D'Avossa & Kersten 1996), but contrary to our results, they report largest

compression in the vertical direction. One might, however, explain the observed horizontal/vertical anisotropy by a response bias (subjects prefer to point towards the center of the screen or fixation point if uncertain), rather than a perceptual bias (subjects actually perceive the heading closer to the center of the screen). Because all experimental conditions were represented in random order, the trials without rotation were mixed with trials with horizontal or torsional simulated rotation, but never in combination with simulated vertical rotation. Possibly, the presence of simulated horizontal rotation in part of the trials reduced the subject's certainty for the horizontal heading component in all trials, causing a response bias towards the screen's center. For two subjects we ran a session during which we recorded in which trials they found the task too hard to perform. Indeed, the trials mostly rejected were flat cloud conditions in which a horizontal component of rotation was present. Thus, a response bias might also partly explain the enlarged compression observed for simulated horizontal rotation.

The increased compression during horizontal rotation has not been reported before. One reason for this is that scale and offset errors are confounded in simulated rotation studies in which pursuit of point within the scene is simulated (e.g. van den Berg & Brenner 1994a; van den Berg 1996a; Grigo & Lappe 1999). In those studies, an offset error due to incomplete rotation compensation and a possible scale error towards the fixation point are both proportional to the heading eccentricity. Studies that did simulate rotation independently of the heading eccentricity have not reported the occurrence of increased compression during rotation (Royden et al. 1992, 1994; Banks et al. 1996; Ehrlich et al. 1998). Since their data were either averaged over simulated heading directions (Ehrlich et al. 1998), or plotted inopportune to observe scale errors, we can not deduce whether such trend was present or not.

A response bias may also to some degree explain why the offset errors observed for the flat cloud condition during purely simulated horizontal rotation were smaller than expected from the CF shifts. For, if subjects perceived their heading at the CF, but their response was compressed by about 50% due to a response bias, the resulting heading errors would match the observed horizontal offsets for pure horizontal rotation (Fig. 5.13a: $\triangle$ and $\triangledown$ for 0 °/s torsion). However, the observed amount of compression is larger (about 70%, see Fig. 5.13c: $\triangle$ for 0 °/s torsion). This would suggest that subjects' perceived heading is shifted more than the CF.

The occurrence of a scale error during simulated horizontal rotation can be looked upon differently, by plotting the horizontal component of the heading error as a function of the horizontal component of the heading direction, together with the corresponding retinal flow patterns (Fig. 5.15). Clearly, we observe that the amount of compensation varies as a function of simulated heading direction. The variation with heading direction is increased during simulated rotation. Interestingly though, smallest errors occur when heading is in the same direction as the pursuit. This effect is opposite to what one would expect from an ecological point of view. If the visual system is best trained to pursue stationary objects in the environment, one would

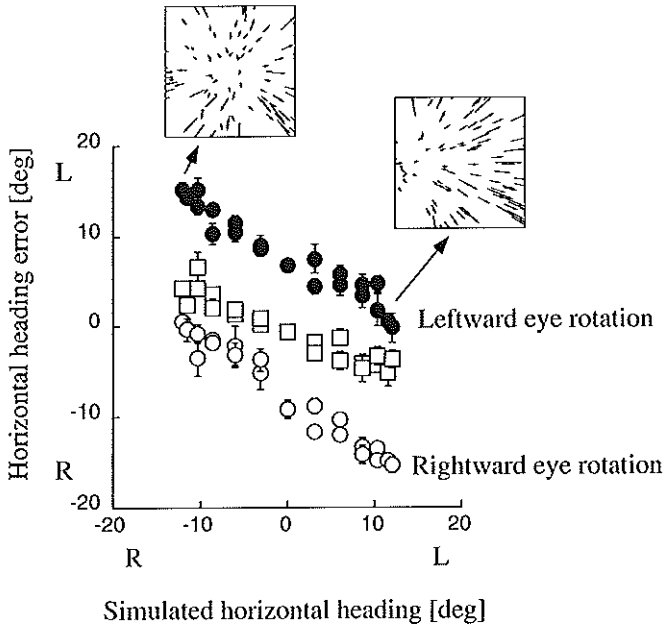



**Figure 5.15** *Horizontal component of heading error as a function of simulated heading, for the 3 horizontal rotation velocities (no torsion, flat cloud condition). The data represent the average over all subjects in experiment IV. Although the slight slope in the data without rotation already reveals an influence of heading eccentricity on the heading error, the steeper slope during simulated rotation shows a much larger dependency on simulated heading eccentricity. The flow patterns represent the flow on the screen during leftward simulated rotation for leftward (L) and rightward (R) heading. The heading error is largest when the center of flow (CF) is at the center of the screen, and smallest when the CF is more eccentric.*

expect largest compensation for eye rotation in the direction opposite to the retinal heading direction. More likely, the varying amount of compensation is related to a heading-dependent asymmetry in the retinal flow during eye rotation. Rotation shifts the CF in the direction of simulated eye rotation. This means that for leftward heading, the CF lies even more to the left, while for rightward heading the CF lies more towards the center of the screen. Thus, largest compensation occurs when the CF is more eccentric relative to the screen.

Expanding retinal flow becomes more parallel when the center of flow is more eccentric, thus resembling to larger extent flow caused by a horizontal eye rotation. We suggest such stimuli may provide more evidence to the visual system for the presence of an eye rotation than retinal flow with a central CF, and thus evoke larger compensatory action. It remains to be shown whether such behaviour would emerge

from any of the current models, or whether it requires certain properties, such as the use of compensatory units that evaluate evidence for rotational flow as proposed by Beintema and van den Berg (1998a).

### 5.7.5 Contribution of extra-retinal signals

Several studies have shown that for the compensation of horizontal rotational flow, the extra-retinal signal can play an important role (Royden et al. 1992, 1994; Banks et al. 1996; Ehrlich et al. 1998; van den Berg 1992). Royden et al. (1994) found that for simulated approach of a fronto-parallel plane, large heading errors occur during simulated rotation, but not during real eye movement. Also, for translation towards a cloud of dots, heading errors can be considerably smaller during real than during simulated rotation (Royden et al. 1992, 1994; Banks et al. 1996; Ehrlich et al. 1998). Our finding of high performance under all circumstances of simulated torsion, even for approach of a fronto-parallel plane, therefore suggests that the compensation for torsional flow need not rely on an extra-retinal signal.

### 5.7.6 Conclusions

We find higher performance for compensation of simulated torsion than for compensation of simulated horizontal rotation. Whereas the compensation for simulated horizontal rotation is small and depends on the presence of motion parallax cues, the compensation for torsional rotation is nearly complete, and little influenced by motion parallax. Moreover, simulated torsional rotation does not influence the compensation for horizontal rotation. We suggest this difference is largely due to differences in the flow information regarding the rotation. The results also imply that the considerable, but not complete, compensation observed for torsional rotation need not rely on an extra-retinal signal.

## Acknowledgements

We are grateful to Ignace Hooge for measuring torsional eye movements. We also thank Marc de Lussanet and Jeroen Smeets for their advice on fitting the data in experiment IV. This work was supported by research grant SLW 805-33.171-P from the Dutch Foundation for Scientific Research (NWO) and by the Human Frontier Science Program.

## 5.8 Appendix

Here, we derive the direction of a singularity in the retinal flow given all points lie on a sphere around the eye with radius $\rho$ (which approximates a fronto-parallel plane for large $\rho$). The translation and rotation are indicated by 3D-vectors $\mathbf{T}$ and $\mathbf{R}$. Vector lengths refer to the translation ($\tau$) and rotation velocity ($\omega$), respectively. Singularities arise for viewing directions $\mathbf{d}$ (3D unity vector) at which the translational and rotational component of the flow cancel each other. Expressing the flow perpendicular to $\mathbf{d}$ as a 3D flow vector $\mathbf{p}$ (Koenderink & van Doorn 1987), we have:

$$\mathbf{p} = -(\mathbf{R} + \mathbf{d} \times \frac{\mathbf{T}}{\rho}) \times \mathbf{d} = \mathbf{0} \tag{5.2}$$

Solving the singularity $\mathbf{d}$ in terms of $\mathbf{T}$, $\mathbf{R}$ and $\rho$ from Eq. 5.2 is yet not possible. But, we can solve the inverse problem of finding $\mathbf{T}'$ and $\mathbf{R}'$ given the singularity $\mathbf{d}'$ lies along the x-axis (an apostrophe denotes vectors in a rotated system). Then, applying a backwards rotation to the vectors $\mathbf{T}'$, $\mathbf{R}'$ and $\mathbf{d}'$ so as to align $\mathbf{T}'$ and $\mathbf{R}'$ with $\mathbf{T}$ and $\mathbf{R}$, respectively, will yield the singularity $\mathbf{d}$.

First, let us solve $\mathbf{T}'$ and $\mathbf{R}'$ in the rotated system, in which $\mathbf{d}' = (1, 0, 0)$. To simplify Eq. 5.2, we choose a translation vector $\mathbf{T}'$ that lies in the x-y plane ($T'_z = 0$):

$$\mathbf{p}' = \begin{pmatrix} 0 \\ R'_z + T'_y/\rho \\ R'_y \end{pmatrix} = \mathbf{0} \tag{5.3}$$

Equation 3 poses no restrictions on $R'_x$ and $T'_x$. But, apart from their sign, $T'_x$ and $R'_x$ are fixed by the translational ($\tau = |\mathbf{T}'|$) and rotational ($\omega = |\mathbf{R}'|$) magnitude, respectively. Using Eq. 5.3 to eliminate $R'_y$ and $R'_z$ we thus have:

$$\begin{aligned} \tau^2 &= T_x'^2 + T_y'^2 \\ \omega^2 &= R_x'^2 + T_y'^2/\rho^2 \end{aligned} \tag{5.4}$$

We can further reduce these equations by eliminating $T_y$. Denoting $\tau^2 - \rho^2\omega^2$ by the symbol $C$, we find $R'_x$ in terms of $T'_x$:

$$R'_x = \pm\frac{1}{\rho}\sqrt{T_x'^2 - C} \tag{5.5}$$

The last equation needed to solve $\mathbf{R}'$ and $\mathbf{T}'$ follows from the observation that the vector lengths and the cosine of the angle between $\mathbf{R}'$ and $\mathbf{T}'$ are invariant under rotation. Therefore, according to the cosine rule, the inproducts $(\mathbf{R} \cdot \mathbf{T})$ and $(\mathbf{R}' \cdot \mathbf{T}')$ must be equal. Substituting $T'_z = 0$ and $R'_y = 0$, we get:

$$\mathbf{T} \cdot \mathbf{R} = T'_x \cdot R'_x \tag{5.6}$$

Thus, the signs of $(\mathbf{T}\cdot\mathbf{R})$ and $T'_x$ determine the sign of $R'_x$ in Eq. 5.5. Substituting this expression of $R'_x$ into Eq. 5.6 gives a quadratic equation to the fourth power, with following solutions:

$$T'_x = \pm\left[\frac{C}{2}\left(1 \pm \sqrt{1+4\,\rho^2(\mathbf{R}\cdot\mathbf{T})^2/C^2}\,\right)\right]^{1/2} \tag{5.7}$$

Note, when taking the possible values of $(\mathbf{R}\cdot\mathbf{T})$ and $C$ into account, we find no more than two real solutions for $T'_x$.

When the rotation is about an axis perpendicular to the heading direction ($\mathbf{T}\cdot\mathbf{R} = 0$) the equations simplify greatly. Let us assume heading in the horizontal x-y plane, and rotation about the vertical. Solutions are dictated by the ratio of rotational flow relative to translational flow magnitude, i.e. $(\rho\,\omega/\tau)$. For small rotations $(\rho\,\omega/\tau < 1)$, we get 2 solutions:

$$\begin{cases} \mathbf{T}' &= (\pm\sqrt{\tau^2-\rho^2\omega^2}, \rho/\,\omega, 0) \\ \mathbf{R}' &= (0, 0, -\omega) \end{cases}$$

Thus, the two singularities lie in the horizontal x-y plane, each making an angle $\phi$ with $\mathbf{T}$ given by $\sin(\phi) = T'_y/\tau = \rho\,\omega/\tau$. Alternatively, for large rotations $(\rho\,\omega/\tau > 1)$, we get:

$$\begin{cases} \mathbf{T}' &= (0, \tau, 0) \\ \mathbf{R}' &= (\pm\sqrt{\omega^2-\tau^2/\rho^2}, 0, -\tau/\rho) \end{cases}$$

Now, the two singularities lie in a vertical plane, each making an angle $\theta$ with the vertical given by $\sin(\theta) = {R_z}'/\omega = \tau\,\rho^{-1}\,\omega^{-1}$. When the ratio equals unity $(\rho\,\omega/\tau = 1)$ only one singularity arises, having a direction perpendicular to both $\mathbf{T}$ and $\mathbf{R}$. Interestingly, when the angle between $\mathbf{T}$ and $\mathbf{R}$ decreases $(|\mathbf{T}\cdot\mathbf{R}| > 0)$, the possible retinal locations of the singularities develop into two circles intersecting $\mathbf{R}$ and $\mathbf{T}$. These circles are centered on the axis that bisects $\mathbf{T}$ and $\mathbf{R}$. This we can proof by showing that in the rotated system, the angle between the singularity direction $\mathbf{d}'$ and the bisecting vector $\mathbf{u}' = \frac{1}{2}(\mathbf{T}'/\tau + \mathbf{R}'/\omega)$ remains invariant under changes of the ratio $\rho\,\omega/\tau$. After some mathematics, we find that for small angles $\alpha$ between $\mathbf{T}$ and $\mathbf{R}$, the cosine of the angle between $\mathbf{u}'$ and $\mathbf{d}'$ is indeed constant:

$$\mathbf{d}'\cdot\mathbf{u}' = 1 - \alpha^2/4 + \text{higher order terms of } \alpha$$

Generally, given the solutions for $\mathbf{T}'$ and $\mathbf{R}'$ , we can retrieve the singularity direction $\mathbf{d}$ by using a method of projections. From $\mathbf{T}'$ and $\mathbf{R}'$ we construct a set of orthogonal basisvectors ($\mathbf{n}'_1 = \mathbf{T}'/\tau$, $\mathbf{n}'_2 = \mathbf{T}'\times\mathbf{R}'/|\mathbf{T}'\times\mathbf{R}'|$ and $\mathbf{n}'_3 = \mathbf{n}'_1\times\mathbf{n}'_2$). Similarly, a set of non-rotated basisvectors ($\mathbf{n}_1$, $\mathbf{n}_2$ and $\mathbf{n}_3$) is constructed from $\mathbf{T}$ and $\mathbf{R}$. Finally, the singularity $\mathbf{d}$ is expressed in terms of $\mathbf{T}$, $\mathbf{R}$ and $\rho$ by $\sum_{i=1}^{3}(\mathbf{d}'_i\cdot\mathbf{n}'_i)\,\mathbf{n}_i$.

# Chapter 6

# Summary and conclusions

In this thesis the rotation problem in self-motion perception from optic flow was investigated. A modelling study and three psychophysical experiments were done to gain insight into what information in the retinal flow is important for human heading perception, and how the brain integrates retinal flow signals with extra-retinal signals on eye, head or body rotation. Here, I summarise the main results, followed by some general conclusions and remarks.

## Summary

### Modelling self-motion perception during eye rotation

The main shortcoming of most physiological models for heading detection (Perrone & Stone 1994, Royden 1997) is their neglect of the extra-retinal signal. Physiologically less detailed models that do use extra-retinal signals (Royden et al. 1994; Freeman 1999) cannot cope with correct heading percepts when the rotation is visually simulated while the eye is not rotating. Neurophysiological data showed effects of various extra-retinal signals in MST neurons that encode self-motion. Our first question was whether we could model the perception of heading and could account for psychophysical and physiological data for both simulated and real eye rotation.

#### Motion templates and eye velocity gain fields (Chapter II)

We presented a model for heading detection that was inspired by the response properties of neurons in MST. The model consists of a map of units, each responding maximally to retinal flow that best fits its preferred combination of heading and rotation. Such motion templates were used before by Perrone and Stone (1994). They solved the rotation problem by including as many combinations needed to cover the full range of likely rotation velocities. Our approach was to reduce the required number of combinations by compensating the change in activity of pure heading templates

in a subsequent stage using visual and/or extra-retinal estimates of the eye rotation velocity.

The compensation for eye rotation relied on a principle that may generally underlie extra-retinal compensation for changes in 'labelled line' signals. We first demonstrated the principle by modelling the compensation for a shift in retinal position due to changed eye orientation, and then showed how the same principle could compensate for a change in template response due to a change in eye velocity. Both the visual position and the heading model assume three types of units. Units in the first map respond to the retinal signals only. Units in intermediate maps respond to retinal signals (first order and higher order derivative signals) and are modulated in amplitude by gain signals (linear, quadratic or higher powers of retinal or extra-retinal origin). The output of the first and subsequent maps are summed in a final map in which each unit can dynamically adjust its retinal tuning (i.e. preferred flow on the retina or location of retinal receptive field).

Simulations showed that the template response in the final heading map was to a large extent rotation invariant when using only the first order derivative signal, gain-modulated by a rotation signal that varies linearly with eye rotation velocity, which we refer to as eye velocity gain fields. The rotation signal can be encoded by a 'rate-coded' extra-retinal signal, but also by a visual estimate of rotation supplied by the first order derivative signal, or by mixtures of retinal and extra-retinal rotation signals. Simulations demonstrated that qualitatively we could account for previously reported heading performance for translation towards scenes with and without depth variations, in real and simulated eye rotation conditions. Despite this success, the model left several assumptions to test; concerning the role of motion parallax to narrow the templates' tuning to heading direction, the width of rotation-tuning of retinal flow templates and the existence of higher order derivatives signals.

## Discriminating between extra-retinal compensation mechanisms

Two chief elements of the model that set it apart from previous approaches to model the role of an extra-retinal eye velocity signals were experimentally investigated. The first one concerned the dependency on visual evidence for eye rotation, the second one concerned the level at which compensation takes place in the brain.

### Is visual evidence on rotation required? (Chapter III)

The dynamic tuning to retinal flow achieved by an extra-retinal rotation signal relies on gain-modulation of motion template activities (velocity gain fields). Since a gain-modulated activity is a product of a visual rotation-related signal and an extra-retinal estimate of rotation, the model predicts that extra-retinal compensation requires visual evidence for an eye rotation. Alternative models for extra-retinal compensation propose that the rotational flow is subtracted from the retinal flow using extra-retinal

signals only (Royden et al. 1994; Freeman 1999), hence do not require visual evidence for rotation.

We tested the two types of compensation using a paradigm in which the subject makes a real eye movement in conditions where visual evidence for eye rotation was present or not. This was accomplished by presenting retinal flow that was either purely expanding on the retina or contained a component of rotational flow as well. The vector subtraction model predicts that perceived heading will always be shifted away from the center of retinal flow during pursuit, even when the retinal flow contains no component of rotational flow. In contrast, our model predicts perceived heading deviates from the center of retinal flow only if visual evidence for rotation is present. We found that a shift in perceived heading occurred only when the effect of the eye rotation was also visually present in the retinal flow. This is consistent with our model's multiplicative interaction, and inconsistent with vector subtraction interaction based only on extra-retinal signals. The quantitative fits to the data also suggested an important role for motion parallax in our model.

**Does the extra-retinal compensation precede or follow local motion integration? (Chapter IV)**

Our model assumes that extra-retinal signals interact with detectors that analyse the flow pattern. Thus, local motion signals have already been integrated by such detectors. Older models assume a vector subtraction-like interaction at a level before local motion signals are integrated (Royden et al. 1994; Freeman 1999; Lappe 1998).

We tested the level of interaction by measuring precision of perceived heading (variable errors) during pursuit and fixation, while expanding flow within a small viewing aperture was presented on the screen. These different levels at which the interaction between visual motion signals and extra-retinal signals occurs lead to different predictions for the precision of perceived heading when one looks through a small aperture. For example, if the windscreen of your car is flushed by a hose of water and you can only see through right behind the wiper. In such cases, only a small part of the expanding flow is visible to the driver. Our model predicts that the eye rotation will severely reduce the driver's precision for estimating the direction of forward movement, because it will make the flow on his retina more parallel. The older models, on the other hand, would predict that eye rotation would not matter, because its effect on the retinal flow is locally subtracted out. We found that precision of perceived heading varied with pursuit velocity, being highest when the retinal flow was radial, and progressively reduced when the retinal flow became more parallel.

Through our experimental analysis we believe we have shown a glimpse of the functional architecture of the visual motion system. The interaction between non-visual motion signals on eye rotation and visual motion signals *follows* the motion integration stage that is necessary to visually represent self-motion. Thus, the extra-retinal and visual signals really interact at the level where *both* carry information on the self-motion of the organism. Currently, there are no neurophysiological or

neuro-imaging studies that have probed this question.

We also found effects of motion duration and viewing aperture on the perceived heading. Systematic errors during eye movement significantly decreased with longer duration (0.5-3.0 s). The increase in error for brief presentations of flow supports a recent hypothesis by Grigo and Lappe (1999) that the extra-retinal signal is less relied upon for brief durations. Furthermore, we found for an increase in aperture size (10-50°) little improvement of the compensation, but a significant increase of the perceived heading eccentricity.

## Visual compensation mechanisms

In the last chapter, we turned away from the issue of the role of extra-retinal signals for self-motion perception and focussed again on the visual system. The visual compensation for rotation has been investigated extensively for rotation about the vertical axis. Our eyes have three degrees of rotation, however. It seems natural to make a distinction between rotation about axes perpendicular to the line of sight and about the line of sight (torsion).

### An extension to 3D eye rotations (Chapter V)

To investigate the visual compensation for torsion, we measured systematic errors in perceived heading during simulated forward self-motion that contained also a component of rotation about the line of sight. The translational component of the simulated self-motion corresponded to heading along various horizontal and vertical directions. We found that during simulated torsion, the shift of perceived heading was small compared to the shift of the center of retinal flow, whereas during simulated rotation about the vertical axis the perceived heading shift was large compared to the shift of the center of retinal flow. Moreover, for simulated rotation about the vertical axis the errors were decreased when depth in the scene was increased. For simulated torsion the errors were little affected by depth in the scene.

These findings show that the need for extra-retinal signals is less for torsion than for horizontal rotation. We propose the difference in visual performance for vertical axis rotation and torsional rotation reflects a difference in the available information in the flow: the higher symmetry of the field of view relative to axis of rotation in the case of torsion may allow for a more accurate estimation of the rotational flow. If correct, this could mean that animals with lateral placed eyes, that sample a larger part of the visual sphere around the head, are in a better position to visually compensate for rotational flow for any axis of rotation. So far we know of no results that pertain to this hypothesis.

Another new finding is that for translation towards a fronto-parallel plane, rotation about the vertical axis caused a compression of the range of perceived heading eccentricities. A compression of perceived heading direction implies that, given the discrimination ability of heading detectors is constant, it would take a larger variation

in heading direction to detect a change in heading direction. This could mean that during landing, pilots notice a change in heading angle less well when the ground surface is devoid of features at different heights and distances, and may benefit from adding such depth variations in the scene.

## General conclusions

We have presented a model that allows for extra-retinal and visual compensation for eye rotation in self-motion perception. Studies on extra-retinal compensation mechanisms (**chapter III** and **IV**) support a multiplicative interaction of the extra-retinal signal at a stage beyond retinal flow sensitive neurons, as proposed by our model. A similar type of interaction is recently proposed to explain neurophysiological observations at the level of MST (Bradley et al. 1996; Shenoy et al. 1996; Shenoy et al. 1999).

We have tried to find explanations for two new effects that we found in the course of our studies. One effect found was a nearly complete visual compensation during simulated rotation about the line of sight when depth in the scene is absent (**chapter V**), which to us points to the importance of very wide field of view to obtain a symmetric sample of the rotational flow. The other effect was an increased compensation with longer duration during real eye rotation (**chapter IV**), which points to the possibility of a flexible use of different sources of information concerning rotation. These explanations may be interesting to explore further by modelling and testing.

### Some remarks on visual and extra-retinal mechanisms

Tangential to our modelling and experimental approach it has been investigated if visual removal of rotational flow takes place at an early, local stage. Royden (1997) and others proposed that local motion differences around depth edges emphasise the translational component of the flow and allow for removal of the rotation component on a local scale of say a few degrees. However, evidence against such local compensation mechanisms and in favour of a more global mechanism was recently given by Pack and Mingolla (1998). They studied the inducing effect of a rotational flow field carried by dots that were 'mixed' into dots that moved an expanding flow field. The addition has the effect of an 'illusory' eye rotation, shifting the perceived heading, as if the visual system compensates for the rotational flow. Pack and Mingolla report that the induced effect increases with the size of rotational flow field beyond a point where the rotational and expansion field are equal in size. This can be explained by a global visual estimate of rotation, but not by local mechanisms. Thus, several lines of evidence support a combination of global visual and extra-retinal signals on eye rotation.

**Remarks on the rotation problem in multiple reference frames**

Finally, we would like to point out that our model and results cover only a small part of the more general problem on how to compensate eye rotation for self-movement perception. In this thesis, we concentrated mainly on the compensation for eye rotation relative to the head, modelling templates that encode flow relative to the head (head-centric flow). Only when the head is stationary with respect to space does this correspond to the optic flow. In general, to obtain the optic flow representation, one also includes compensation for head and body rotations requiring other extra-retinal rotation signals. Fig. 6.1 shows how the various channels may contribute in obtaining a representation of the optic flow, starting with retinal flow. The optic flow may be obtained from eye velocity relative to space using a visual estimation of rotation or a combination of extra-retinal signals. It may also be obtained by a sequence of transformations that use eye velocity relative to the head based on efference copy signals and head velocity relative to space based on vestibular signals. Compensation for head velocity relative to space may also be achieved by using head velocity relative to the body, based on neck proprioception or efference copy, and body velocity relative to the world, based on signals from receptors in the joints and the skin. Crowell et al. (1998a) recently studied the interaction between the three extra-retinal signals related to head rotations in a series of conditions, like active head rotation, passive head or whole body-rotation. By subtracting perceived heading during either of these conditions from perceived heading during simulated rotation, they disregarded contribution of the visual rotation signal. They found that proprioceptive and vestibular signals could contribute only to compensation in the presence of the other signals, which demonstrates that the interaction between the three extra-retinal signals is non-linear. Thus, even without taking into account the contribution of the visual signal, the compensation appears to be more complex than merely a simple vector-subtraction like operation.

A final note should be made on two additional steps in the compensation for rotation velocity. At each stage in the above scheme, the representation of flow - whether retinal, head-centric, body-centric or optic - remains encoded in the retinal frame of reference. This has two important implications. First of all, the compensation for rotation velocity requires the axis of rotation to be specified in the retinal reference frame. As mentioned in Chapter II, this implies that when the eye or head orientation is changed, eye position or head position need to be taken into account. Secondly, in certain tasks it may be desirable for the brain to have access to a representation of heading direction in another frame of reference, like the head, body or space. This would for instance allow the human brain to compare a retinal representation of heading that is derived from the optic flow, with a head-centric representation of heading that is derived from head-centric signals (vestibular or auditory). The brain could perform the transformation of the retinal heading direction to head-centric heading direction by the same principle using signals on eye position relative to the head, or if necessary relative to the body or space. Whatever mechanism underlies

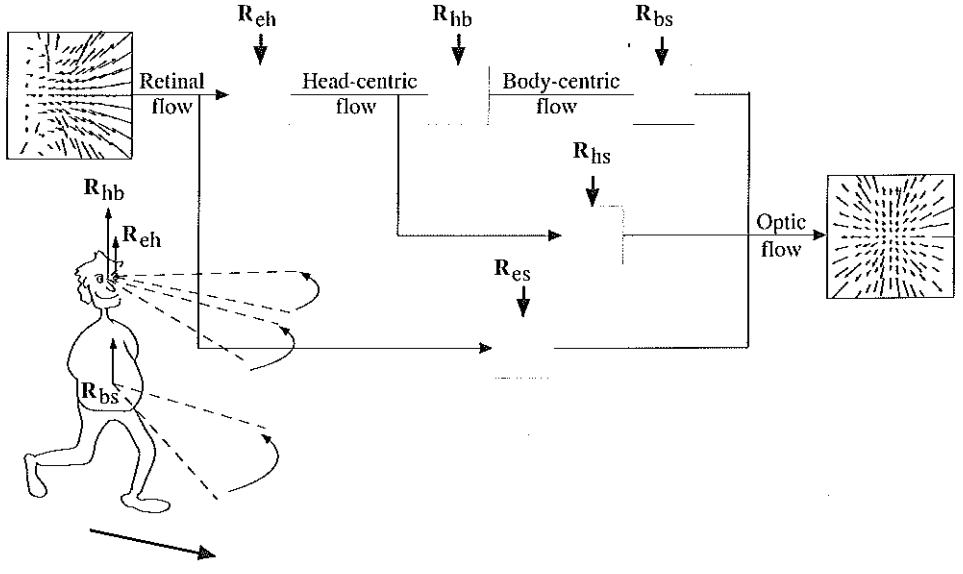



**Figure 6.1** *General scheme for obtaining a representation of optic flow from the retinal flow during eye rotation using rotation signals. The cartoon man shows how rotation of the eye relative to space ($\mathbf{R}_{es}$) is the sum of rotation of the eye relative to the head ($\mathbf{R}_{eh}$), head rotation relative to body ($\mathbf{R}_{hb}$) and body rotation relative to space ($\mathbf{R}_{bs}$). Rotation of head relative space is indicated by ($\mathbf{R}_{hs}$). Three parallel subschemes lead to compensation for rotation relative to space; using extra-retinal signals on the above rotations (bold). Eye rotation relative to space (bottom) may be given by retinal and the sum of extra-retinal signals. Not represented in this scheme are the additional inputs of eye or head position signal required to convert the extra-retinal rotation signals into the retinal reference frame or to transform the retinal representation of optic flow to other frames of reference.*

the compensation for rotation velocity, be it at local or more global level of motion analysis, these two additional position transformations would seem necessary for the perception of self-motion.

# References


Albright T. D., Desimone R. & Gross C. G. (1984). Columnar organization of directionally selective cells in visual area MT of the macaque. *Journal of Neurophysiology, 51,* 16-31.

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

Andersen R. A. & Zipser D. (1988). The role of posterior parietal cortex in coordinate transformations for visual-motor integration. *Canadian Journal of Pharmacology, 66,* 488-501.

Andersen R. A., Bracewell R. M., Barash S., Gnadt J. W. & Fogassi L. (1990). Eye position effects on visual memory, and saccade-related activity in areas LIP and 7a of macaque. *Journal of Neurophysiology, 10,* 1176-1196.

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

Andersen R. A., Snyder L. H., Bradley D. C. & Xing J. (1997). Multimodal representation of space in the posterior parietal cortex and its use in planning movements. *Annual Review of Neuroscience, 20,* 303-330.

Andersen R. A. (1989). Visual and eye movement functions of the posterior parietal cortex. *Annual Review of Neuroscience, 12,* 377-403.

Banks M. S., Ehrlich S. M., Backus B. T. & Crowell J. A. (1996). Estimating heading during real and simulated eye movements. *Vision Research, 36(3),* 431-443.

Banks M. S., Ehrlich S., Schor C. M., McCandles J. & Crowell J. A. (1993). Heading perception with misleading eye position signals. *Investigative Ophthalmology and Visual Science, 34,* 1229 (Abstract).

Beintema J. A. & van den Berg A. V. (1998a). Heading detection using motion templates and eye velocity gain fields. *Vision Research, 38(14),* 2155-2179.

Beintema J. A. & van den Berg A. V. (1998b). Effect of torsional flow on heading percept. *Investigative Ophthalmology and Visual Science, 39(4) (Suppl.),* 5020 (Abstract).

Beintema J. A. & van den Berg A. V. (2000). Perceived heading during simulated torsional eye movements *Vision Research, 40(5),* 549-566.

van den Berg A. V. & Beintema J. A. (1997). Motion templates with eye velocity gain fields for transformation of retinal to head centric flow. *Neuroreport, 8,* 835-840.

van den Berg A. V. & Brenner E. (1994a). Humans combine the optic flow with static depth cues for robust perception of heading. *Vision Research, 34*, 2153-2167.

van den Berg A. V. & Brenner E. (1994b). Why two eyes are better than one for judgements of heading. *Nature, 371*, 700-702.

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

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

van den Berg A. V. (1996a). Judgements of heading. *Vision Research, 36*, 2337-2350.

van den Berg A. V. (1996b). Optic flow and the kinematics of eye movements. *Proceedings of the Royal Society of London B, 263*, 975-981.

van den Berg A. V. (1997). Perception of heading or perception of ego-rotation? *Investigative Ophthalmology and Visual Science, 37 (Suppl.)*, 380 (Abstract).

van den Berg A. V. (1999). Human ego-motion perception In: *International Review of Neurobiology: Neuronal Processing of Optic Flow.*, edited by Lappe M., Academic Press, p. 3-25.

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

Bremmer F. & Hoffmann K. (1993). Pursuit related activity in visual cortical areas MST and LIP is modulated by eye position. *Society for Neuroscience, 19*, 1283 (Abstract).

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

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

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

Collewijn H., van der Steen J., Ferman L. & Jansen T. C. (1985). Human ocular counterroll: Assesment of static and dynamic properties from electromagnetic scleral coil recordings. *Experimental Brain Research, 59*, 185-196.

Collum J. F. & Siegel R. M. (1996). Neuronal selectivity to retino-topic centers of motion in area 7a of the behaving monkey. *Society for Neuroscience, 22*, 1619 (Abstract).

Crowell J. A. & Banks M. S. (1993). Perceiving heading with different retinal regions and types of flow. *Perception and Psychophysics, 53(3)*, 325-337.

Crowell J. A. & Banks M. S. (1996). Ideal observer for heading judgments. *Vision Research, 36(3)*, 471-490.

Crowell J. A., Banks M. S., Shenoy K. V. & Andersen R. A. (1998a). Visual self-motion perception during head turns. *Nature Neuroscience, 1(8)*, 732-737.

Crowell J. A., Maxwell M. A., Shenoy K. V. & Andersen R. A. (1998b). Retinal and extra-retinal motion signals both affect the extent of gaze-shift compensation. *Investigative Ophthalmology and Visual Science, 39(4) (Suppl.)*, 5060 (Abstract).

Cutting J. E., Springer K., Braren P. A. & Johnson S. H. (1992). Wayfinding on foot from information in retinal, not optical, flow. *Journal of Experimental Psychology Genera, 121(1)*, 41-72.

D'Avossa G. & Kersten D. (1996). Evidence in human subjects for independent coding of azimuth and elevation for direction of heading from optic flow. *Vision Research, 36(18),* 2915-2924.

De Bruyn B. & Orban G. A. (1988). Human velocity and direction discrimination measured with random dot patterns. *Vision Research, 28(12),* 1323-1335.

De Bruyn B. & Orban G. A. (1993). Segregation of spatially superimposed optic flow components. *Journal of Experimental Psychology: Human Perception and Performance, 19(5),* 1014-1027.

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

Duffy C. J. & Wurtz R. H. (1995). Response of the monkey MST neurons to optic flow stimuli with shifted centers of motion. *Journal of Neurophysiology, 15,* 5192-5206.

Efron B. & Tibshirani R. (1991). Statistical data analysis in the computer age. Second Edition. *Science, 253,* 390-395.

Ehrlich S. M., Beck D. M., Crowell J. A., Freeman T. C. A. & Banks M. S. (1998). Depth information and perceived self-motion during simulated gaze rotations. *Vision Research, 38(20),* 3129-3145.

Ferman L., Collewijn H. & van den Berg A. V. (1987). A direct test of Listing's law-I. Human ocular torsion measured in static tertiary positions. *Vision Research, 27(6),* 929-938.

Freeman T. C. A. & Banks M. A. (1998). Perceived head-centric speed is affected by both extra-retinal and retinal errors. *Vision Research, 38:7,* 941-945.

Freeman T. C. A. & Harris M. G. (1992). Human sensitivity to expanding and rotating motion: effects of complementary masking and directional structure. *Vision Research, 32(1),* 81-87.

Freeman T. C. A. (1999). Path perception and Filehne illusion compared: model and data. *Vision Research, 39,* 2659-2667.

Galetti C., Battaglini P. P. & Fattori P. (1993). Parietal neurons encoding spatial locations in craniotopic coordinates. *Experimental Brain Research, 96,* 221-229.

Geesaman B. J., Born R. T., Andersen R. A. & Tootell R. B. (1997). Maps of complex motion selectivity in the superior temporal cortex of the alert macaque monkey: a double-label 2-deoxyglucose study. *Cerebral Cortex, 7(8),* 749-57.

Gibson J. J. (1966). *The perception of the visual world.* Boston, Mass.: Houghton Mifflin.

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

Grigo A. & Lappe M. (1999). Dynamical use of different sources of information in heading judgments from retinal flow. *Journal of the Optical Society of America A, 16(9),* 2079-2091.

Harris L. R. & Lott L. A. (1995). Sensitivity to full-field visual movement compatible with head rotation: variations among axes of rotation. *Visual Neuroscience, 12,* 743-754.

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

Hildreth E. C. (1992). Recovering heading for visually-guided navigation. *Vision Research, 32(6)*, 1177-1192.

Hooge I. T. C., Beintema J. A. & van den Berg A. V. (1999). Visual search of heading direction *Experimental Brain Research, 29:4*, 615-628.

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

Johnston I. R., White G. R. & Cumming R. W. (1973). The role of optical expansion patterns in locomotor control. *American Journal of Psychology, 86(2)*, 311-324.

Koenderink J. J. & van Doorn A. J. (1981). Exterospecific component of the motion parallax field. *Journal of the Optical Society of America A, 8*, 953.957

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

Koenderink J. J. (1986). Optic flow. *Vision Research, 26*, 161-180.

Koenderink J. J. (1988). Operational significance of receptive field assemblies *Biological Cybernetics, 58*, 163-171.

Kowler E. & McKee S.P. (1987). Sensitivity of smooth eye movement to small differences in target velocity. *Vision Research, 27(6)*, 993-1015.

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

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

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

Lappe M., Bremmer F., van den Berg A. V. (1999). Perception of self-motion from visual flow. *Trends in Cognitive Sciences, 3(9)*, 329-336.

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

Llewellyn K. R. (1971). Visual guidance of locomotion. *Journal of Experimental Psychology, 91*, 245-261.

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

Longuet-Higgins H. C. (1984). The visual ambiguity of a moving plane. *Proceedings of the Royal Society of London B, 223*, 165-175.

Maunsell J. H. R. & Newsome W. T. (1987). Visual processing in monkey extra-striate cortex. *Annual Review of Neuroscience, 10*, 363-402.

McKee S.P. (1981). A local mechanism for differential velocity detection. *Vision Research, 21*, 491-500.

Morrone M. C., Burr D. C. & Vaina L. M. (1995). Two stages of visual processing for radial and circular motion. *Nature, 376*, 507-509.

Newsome W. T., Wurtz R. H. & Komatsu H. (1988). Relation fo cortical areas MT and MST to pursuit eye movements. II. Differentiation of retinal from extraretinal inputs. *Journal of Neurophysiology, 60(2)*, 604-619..

Ohtani Y., Tanigawa M. & Ejima Y. (1998). Motion assimilation for expansion/contraction and rotation and its spatial properties. *Vision Research, 38(3)*, 429-438.

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

Pack C. & Mingolla E. (1998). Global induced motion and visual stability in an optic flow illusion. *Vision Research, 38*, 3083-3093.

Page W. K. & Duffy C. J. (1999). Neuronal responses to heading direction during pursuit eye movements. *Journal of Neurophysiology, 81*, 596-610.

Pas S. F. te, Kappers A. M. L. & Koenderink J. J. (1998). Locating the singular point in first-order optical flow fields. *Journal of Experimental Psychology: Human Perception and Performance, 24(5)*, 1415-1430.

Pasternak T. & Merigan W. H. (1984). Effects of stimulus speed on direction discrimination. *Vision Research, 24(10)*, 1349-1355.

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

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

Press W. H., Teukolsky S. A., Vetterling W. T. & Flannery B. P. (1994). *Numerical recipes in C, The Art of Scientific Computing, Second Edition.* Cambridge University Press.

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

Rieger J. H. & Lawton D. T. (1985). Processing differential image motion. *Journal of the Optical Society of America A, 2*, 354-359.

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

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

Royden C. S., Crowell J. A. & Banks M. S. (1994). Estimating heading during eye movements. *Vision Research, 34(23)*, 3197-3214.

Royden C. S. (1994). Analysis of misperceived observer motion during simulated eye rotations. *Vision Research, 34(23)*, 3215-3222.

Royden C. S. (1997). Mathematical analysis of motion-opponent mechanisms used in the determination of heading and depth. *Journal of the Optical Society of America A, 14(9)*, 2128-2143.

Schaafsma S. J. & Duysens J. (1996). Neurons in the ventral intraparietal area of the awake macaque monkey closely resemble neurons in the dorsal part of the medial superior temporal area in their responses to optic flow patterns. *Journal of Neurophysiology, 76*, 4056-4068.

Shenoy K. V., Bradley D. C. & Andersen R. A. (1996). Heading computation during head movements in macaque cortical area MSTd. *Society for Neuroscience, 22*, 1692 (Abstract).

Shenoy K. V., Bradley D. C. & Andersen R. A. (1999). Influence of gaze rotation on the visual response of primate MSTd neurons. *Journal of Neurophysiology, 81*, 2764-2786.

Snowden R. J. & Milne A. B. (1996). The effect of adapting to complex motions: position invariance and tuning to spiral motions. *Journal of Cognitive Neuroscience, 8(5)*, 435-452.

Snowden R. J., Treue S., Erickson R. G. & Andersen R. A. (1991). The response of area MT and V1 neurons to transparent motion. *Journal of Neurophysiology, 11*, 2768-2785.

Squatrito S. & Maioli M. G. (1996). Gaze field properties of eye position neurones in areas MST and 7a of the macaque monkey. *Visual Neuroscience, 13*, 385-398.

Squires G. L. & Torrie J. H. (1985). *Practical physics. Third edition.* Cambridge University Press.

Steel G. D. & Torrie J. H. (1987). *Principles and procedures of statistics. A biometrical approach.* MacGraw-Hill International Editions.

Stone L. S. & Perrone J. A. (1997). Human heading estimation during visually simulated curvilinear motion. *Vision Research, 37(5)*, 573-590.

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*, 626-641.

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

Thier P. & Erickson R. G. (1992). Responses of visual-tracking neurons from cortical area MST-l to visual, eye and head motion. *European Journal of Neuroscience, 4*, 539-553.

Tootell R. B. H, Dale A. M., Sereno M. I. & Malach R. (1996). New images from human visual cortex. *Trends in Neurosciences, 19*, 481-489..

Turano K. A. & Heidenreich S.M. (1996). Speed discrimination of distal stimuli during smooth pursuit eye motion. *Vision Research, 36(21)*, 3507-3517.

Tweed D., Fetter M., Andreadaki S., Koenig E. & Dichgans J. (1992). Three-dimensional properties of human pursuit eye movements. *Vision Research, 32*, 1225-1238.

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

Warren W. H. & Hannon D. J. (1990). Eye movements and optical flow. *Journal of the Optical Society of America A, 7(1)*, 160-169.

Warren W. H. & Kurtz K. J. (1992). The role of central and peripheral vision in perceived the direction of self-motion. *Perception and Psychophysics, 51(5)*, 443-454.

Warren W. H., Blackwell A. W., Kurtz K. J., Hatsopoulus N. G. & Kalish M. L. (1991). On the sufficiency of the velocity field for perception of heading. *Biological Cybernetics, 65*, 31-320.

Warren W. H., Morris M. W. & Kalish M. (1988). Perception of translational heading from optical flow. *Journal of Experimental Psychology: Human Perception and Performance, 14(4)*, 646-660.

Warren W. H. (1995). Self-motion: visual perception and visual control. In: *Handbook of Perception and Cognition: Perception of Space and Motion*, edited by Epstein W. & Rogers S., Academic Press, p. 233-260.

Wertheim A. H. (1994). Motion perception during self-motion: the direct versus inferential controversy revisited. *Behavioral and Brain Sciences, 17*, 293-355.

Wurtz R. H., Komatsu H., Dürsteler M. R. & Yamasaki D. S. (1990). Motion to movement: cerebral cortical visual processing for pursuit eye movements. In: *Signal and Sense: Local and global order in perceptual maps*, edited by Edelman G., Gall W. E. & Cowan W. M., New York: John Wiley, p. 233-260.

Zihl J., von Cramon D. & Mai N. (1983). Selective disturbance of movement vision after bilateral brain damage. *Brain, 106*, 313-340.

# Samenvatting

## Waarneming van zelfbeweging

Dit proefschrift beschrijft de resultaten van onderzoek aan het rotatieprobleem in de waarneming van zelfbeweging. Om op koers te blijven moet men weten in welke richting het hoofd of lichaam beweegt ten opzichte van de wereld. Maar visuele informatie over de koers komt binnen via het oog dat mogelijk draait ten opzichte van hoofd of lichaam. Met behulp van rotatie-informatie (visueel of niet-visueel) moet de visuele beweging gemeten door het draaiende oog omgezet worden naar het coördinaten systeem van het hoofd of lichaam. Hoe de hersenen dit aanpakken is met een model en daaruit volgende psychofysische experimenten onderzocht.

Een ieder die wel eens in de bioscoop zit kent de sensatie zich vooruit te voelen bewegen wanneer een heelal met uiteendijende sterren getoond wordt. Het patroon van lichtbeweging over het filmdoek is vergelijkbaar met wat aan licht over het netvlies (retina) van het oog glijdt wanneer we zelf voortbewegen ten opzichte van de stilstaande wereld. Deze visuele beweging wordt 'optic flow' of het visuele stroomveld genoemd[1] en bevat informatie over de zelfbeweging van de waarnemer. Ook statische informatie kan koers aangeven, zoals de oriëntatie van strepen op het wegdek ten opzichte van de auto. Zihl's (1983) beroemde patiënte LM illustreert echter het belang van visuele beweging bij het waarnemen van relatieve beweging tussen onszelf en objecten in de omgeving. Zij is 'bewegingsblind' door een beschadiging aan een bewegingsgevoelig hersengebied. Ze rapporteert moeite te hebben met oversteken omdat ze de positie van auto's niet juist weet in te schatten - 'When I'm looking at the car first, it seems far away. But then, when I want to cross the road, suddenly the car is very near'.

De koersinformatie in optic flow wordt al een paar eeuwen gebruikt door astronomen om de baan vast te stellen waarlangs kometen de aarde naderen. Het punt aan de hemel waar de regen van kometen uit lijkt te waaieren, de radiant genoemd, ligt in de translatierichting van de kometen. Zo ook zou de mens het centrum van de optic flow kunnen gebruiken om zijn translatie richting waar te nemen. Maar wanneer tijdens de translatie het oog roteert dan is het visuele stroomveld veranderd. Door de rotatie-component in het stroomveld lijkt het stroomveld op wat het oog zou

[1] Een voorbeeld van optic flow is te zien op de omslag.

waarnemen als men in een andere richting voortbewegen zonder het oog te draaien (Fig. 1.1). Zou de mens niet compenseren voor het effect van rotatie dan zou hij zijn koers fout waarnemen. Dit is het zogeheten 'rotatieprobleem'.

Naast optic flow beschikt de mens over extra-retinale (niet-visuele) signalen informatie over zelfbeweging. Rotatie van het oog gaat bijvoorbeeld gepaard met signalen van en naar de oogspieren. Bij hoofdrotatie zijn er signalen van nekspieren en evenwichtsorganen. Hoe deze rotatiesignalen samenwerken met de optic flow informatie om het rotatieprobleem op te lossen is nog maar ten dele bekend. In zogeheten psychofysische experimenten kijkt de proefpersoon naar bewegende punten op een scherm alsof hij door een wereld van punten in een bepaalde koers voortbeweegt. De proefpersoon fixeert een vast punt of maakt een oogvolgbeweging en rapporteert zijn waargenomen richting van voortbeweging met een muis. Uit de fout in waargenomen koers heeft men afgeleid dat extra-retinale rotatiesignalen een rol spelen. Maar uit proeven waarin de oogrotatie gesimuleerd wordt (proefpersoon fixeert zodat het extra-retinale signaal niet de oogrotatie weergeeft), blijkt dat de optic flow in sommige gevallen ook voldoende kan zijn.

De wisselwerking tussen het modelleren, de psychofysica en elektrofysiologie levert vaak vruchtbare resultaten. We hebben daarom in dit onderzoek geprobeerd deze drie gebieden te combineren. Op basis van fysiologische en psychofysische data hebben we een model gemaakt voor koerswaarneming. Uit de vergelijking van dit model met andere modellen volgden twee psychofysische experimenten over extra-retinale compensatie mechanismen. Het onderzoek aan extra-retinale compensatie mechanismen geeft inzicht in hoe het brein signalen combineert die verschillend van aard zijn. Het model richt zich weliswaar op de compensatie voor oogsnelheid bij de waarneming van koers, maar het achterliggende principe zou algemener gebruikt kunnen worden voor coördinaten transformaties van sensorische informatie met behulp van motor signalen, zoals bijvoorbeeld de compensatie voor het effect van veranderde oogoriëntatie bij de visuele waarneming van object positie. Naar aanleiding van modelsimulaties volgde een derde experiment over visuele compensatie. Het onderzoek aan visuele compensatie mechanismen geeft inzicht in welke visuele bewegingsinformatie van belang voor de mens is om zijn koers waar te nemen tijdens oogrotatie. Dit soort kennis kan men mogelijk vertalen naar strategiën, automatische piloten of aanpassingen van de omgeving (autowegen, landingsbanen, e.d.) voor optimale sturing.

## Model voor koerswaarneming met gebruik van rotatiesignalen

In **hoofdstuk II** testten we een model voor de waarneming van koers die eerdere psychofysische data kan beschrijven en fysiologisch aannemelijk is. Neurofysiologisch onderzoek in hersengebied MST bij apen toont aan dat optic flow-gevoelige neuronen, die betrokken zijn bij koerswaarneming, tijdens rotatie door een extra-retinaal rotatie signaal in aktiviteit worden gemoduleerd. Een belangrijk aspect aan het model is daarom het gain-modulatie principe waarmee we de compensatie voor het effect van oogrotatie verkregen. Dit algemene principe werd eerst getest bij een model dat

compenseert voor het effect van veranderde oogoriëntatie op visueel waargenomen positie. Daarna gebruikten we hetzelfde principe bij de compensatie voor het effect van veranderde oogrotatiesnelheid op waargenomen koers uit het visuele stroomveld. De koers wordt hierbij gerepresenteerd door de aktiviteitsverdeling in een kaart van zogeheten 'templates' met de eigenschappen van MST-neuronen. Ieder template vuurt maximaal bij het retinale stroomveld dat het best past bij zijn voorkeurskoers. Om te compenseren voor een veranderde template aktiviteit tijdens rotatie wordt van iedere template een signaal afgetrokken dat bestaat uit het produkt van de oogsnelheid en een met de template corresponderende afgeleide term. Deze afgeleide term geeft aan hoeveel de template aktiviteit verandert bij een verandering in visuele rotatie. De simulaties toonden aan dat eerdere psychofysische data in zowel echte als gesimuleerde oogrotatie condities kwalitatief met het model beschreven kunnen worden, voor omgevingen met en zonder diepteverschillen. We lieten zien dat het benodigde rotatiesignaal op drie manieren gerepresenteerd kan zijn: door een extra-retinaal signaal, een visuele schatting van de rotatie afkomstig van templates die afgestemd zijn op rotatie, of door een combinatie van visuele en extra-retinale rotatiesignalen. Twee modelparameters moeten nog getest worden; het gebruik van bewegingsparallax (gevoeligheid voor diepteverschillen in de omgeving) bij de template's afstemming op de koers, en de grofheid van de template's afstemming op rotatie.

## Discriminatie tussen extra-retinale compensatie mechanismen

In het gebruik van extra-retinale signalen onderscheidt het model zich op twee punten van andere extra-retinale modellen. Beide punten zijn psychofysisch onderzocht.

In **hoofdstuk III** testten we of de extra-retinale compensatie afhangt van visuele evidentie voor een oogrotatie. In ons model wordt de compensatie verkregen door het aftrekken van aktiviteit die bestaat uit het product van het extra-retinale oogsnelheidssignaal en een visueel signaal dat bij benadering de rotatiesnelheid visueel afschat. De compenserende aktiviteit hangt daarom af van de evidentie voor een oogrotatie. Een alternatief, het zoheten vector subtractie model, stelt voor dat de rotatie component van het visuele stroomveld afgetrokken wordt op basis van alleen extra-retinale signalen (Royden et al. 1994; Freeman 1999). Dit type model gebruikt geen visuele evidentie voor rotatie. Om deze twee type modellen te testen lieten we de proefpersoon een oogbeweging maken, terwijl visuele evidentie voor een oogrotatie aanwezig was of niet. We vonden dat de verschuiving in waargenomen koers alleen optrad wanneer het effect van oogrotatie visueel aanwezig was. Dit is consistent met een gain-modulatie interactie zoals voorgesteld door ons model, en niet consistent met een vector subtractie interactie gebaseerd op alleen het extra-retinale signaal. Verder suggereren de kwantitatieve modelfits aan de data dat bewegingsparallax een belangrijke rol speelt.

In **hoofdstuk IV** testten we de tweede vraag op welk niveau compensatie in het brein plaatsvindt. In ons model nemen we aan dat extra-retinale signalen een interactie aangaan met detectoren die het patroon van het stroomveld analyseren.

Lokale bewegingssignalen zijn dus al geïntegreerd in zulke detectoren. Een ander type model, waaronder het vector subtractie model, neemt aan dat de compensatie gebeurt voordat lokale bewegingssignalen worden geïntegreerd (Royden et al. 1994; Freeman 1999; Lappe 1998). We testten het niveau van interactie door het meten van de variabele fout gedurende een oogvolgbeweging en gedurende fixatie. De stimulus op het scherm was expanderende flow binnen een kleine apertuur. Een kleine apertuur treedt bijvoorbeeld op als de voorruit van de auto door een stortvloed aan water overspoeld wordt en er alleen direct naast de ruitenwisser zicht is. In zo'n geval is alleen een klein deel van de expanderende flow zichtbaar voor de bestuurder. Ons model voorspelt dat de bestuurder's precisie (variabele fout) in waargenomen koers door een oogrotatie aanzienlijk verstoord is omdat het visuele stroomveld hierdoor meer parallel wordt. Het andere type model voorspelt echter dat de oogrotatie er niet toe doet omdat zijn effect op het visuele stroomveld al op lokaal niveau is gecompenseerd. We vonden dat de precisie in waargenomen koers afhankelijk is van oogrotatie, in overeenstemming met een interactie op het niveau na integratie van lokale visuele beweging.

Verder vonden we een effect van bewegingsduur en apertuurgrootte op de waargenomen koers. De systematische fouten gedurende oogbeweging werden significant kleiner met langere duur (0.5-3.0-s). De toename in fout voor korte aanbiedingsduur van het stroomveld ondersteunt een recente hypothese door Grigo en Lappe (1999). Zij stelden dat bij kortere duur er minder op het extra-retinale signaal vertrouwd wordt. Ook vonden we dat bij een toename van de apertuurgrootte, de compensatie weinig verbeterde, maar dat de waargenomen koerseccentriciteit wel significant toenam.

## Visuele compensatie mechanisme

In **hoofdstuk V** focusseerden we op het visuele compensatie mechanisme. De visuele compensatie is al uitvoerig onderzocht voor rotatie om de verticale as, maar niet voor rotatie om de kijkas (torsie). We vergeleken de visuele compensatie bij beide rotatie-assen in gesimuleerde oogrotatie condities. De verschuiving in waargenomen koers was klein bij gesimuleerde torsie, en groot bij gesimuleerde rotatie om de verticale as. Daarnaast gaf een toename van diepteverschillen in de gesimuleerde omgeving weinig verbetering in waargenomen koers bij torsie, maar wel bij rotatie om de verticale as. Beide bevindingen laten zien dat een extra-retinaal signaal voor het compenseren van oogrotatie minder noodzakelijk is bij torsie dan bij rotatie om de verticale as. We suggereren dat het verschil in visuele prestatie bij deze twee assen van rotatie een verschil in de beschikbare flowinformatie weergeeft. De grotere symmetrie van het zichtbare veld ten opzichte van de rotatie-as bij torsie zou een nauwkeuriger schatting van de rotatie-component in de flow mogelijk maken. Als dat zo is, zou men verwachten dat beesten met de ogen aan de zijkant - en dus een groter zichtbaar veld - beter in staat zijn om visueel te compenseren voor rotatie om een willekeurige as. Zover we weten, zijn er geen data die deze hypothese kunnen ondersteunen. Als nevenresultaat vonden

we dat voor rotatie om een verticale as, de waargenomen koerseccentriciteiten meer gecomprimeerd zijn naarmate diepteverschillen in de omgeving afnemen.

## Algemene conclusies

We hebben een model voor de waarneming van zelfbewegingsrichting gepresenteerd dat voor oogrotatie compenseren kan met behulp van extra-retinale rotatiesignalen en op basis van het visuele stroomveld alleen. Psychofysische experimenten om de extra-retinale compensatie mechanismen te onderzoeken (**hoofdstuk III** en **IV**) ondersteunen een gain-modulatie interactie tussen het extra-retinale rotatiesignaal en visuele signalen op het niveau na retinale flow detectie. Dit is in overeenstemming met ons model en is recentelijk ook elektrofysiologisch vastgesteld op het niveau van hersengebied MST bij de aap (Bradley et al. 1996; Shenoy et al. 1996; Shenoy et al. 1999).

Door onze experimentele analyse in **hoofdstuk IV** menen we een klein stukje van de functionele architectuur van het visuele bewegingssysteem te hebben laten zien. De interactie tussen niet-visuele rotatiesignalen en visuele bewegingssignalen komt na de integratie van lokale bewegingsinformatie die nodig is om visueel de zelfbeweging te representeren. De extra-retinale en visuele signalen gaan dus een interactie met elkaar aan op het niveau waarop beide informatie over de zelfbeweging van het organisme dragen. Tot op heden zijn er geen neurofysiologische of neuro-imaging studies geweest die deze vraag hebben onderzocht.

We vonden twee nieuwe effecten waarvoor we een verklaring hebben gezocht. Het eerste effect was een bijna complete visuele compensatie voor gesimuleerde rotatie om de kijkas wanneer diepte in de omgeving afwezig was (**hoofdstuk V**). Dit wijst op het belang van een groot zichtbaar veld om een symmetrische bemonstering van de rotatie-component in het stroomveld te verkrijgen. Het tweede effect was een toenemende compensatie met langere duur gedurende echte oogrotatie (**hoofdstuk IV**). Dit duidt op het mogelijk flexibel gebruik van verschillende informatiebronnen over de rotatie. We denken dan ook dat verder modelmatig en experimenteel onderzoek naar de rol van symmetrie in het visuele veld en het flexibel gebruik van verschillende informatiebronnen van belang is voor een beter begrip van compensatie mechanismen bij koerswaarneming uit optic flow.

# Dankwoord

Hierbij wil ik alle mensen bedanken die ik direct of indirect bij het maken van dit proefschrift betrokken heb, ook degene die ik hier vergeet. Twee mensen ben ik in het bijzonder erkentelijk. Bert, vier jaar lang heb je met veel interesse, ideeën, geduld en een kritische blik mijn onderzoek begeleid. Meerdere malen heb je me laten zien hoe je met optimisme en goed doordenken verder kan komen. Je stijl van onderzoek, schrijven en pragmatische houding hebben me veel geleerd. Je bleek ook een prima vent om mee samen te werken en om over andere interesses te praten. Een tweede persoon die hier een plaats verdient is Ignace Hooge. Als kamergenoot, discussie- en samenwerkpartner, mac-expert en redder in mijn last-minute acties heb je me fantastisch geholpen. Jouw komst in Rotterdam bracht voor mij veel motivatie en inspiratie mee.

Van de rest van de vakgroep in Rotterdam, speciaal dank aan de proefpersonen voor het maken van oogbewegingen en het doorklikken van mijn boeiende heading stimuli. Diverse mensen hebben mij geholpen. Maarten, behalve voor tips en uitleen van boeken om mijn horizon te verbreden kon ik ook voor echt belangrijke zaken bij je terecht. Eli met programmeertips (en ladder), Jeroen S. als raadsheer bij technische en data-analyse problemen, Jeroen G. als donateur van zijn Latex-bibliotheek, zonder jullie hulp was ik er niet uitgekomen. Han, dank voor je proefschriftcorrecties en het vertrouwen dat je in mijn onderzoek gesteld hebt. Ook de technici (Ben en anderen) en administratieve ondersteuning (Ria voor oplossen van financiële glijders, Susan voor klusjes, Yvonne voor eerste hulp bij ongelukken), bedankt voor jullie inzet. De promovendi 'voor' mij (Dennis op de zaterdag, Anne en Anil) en na mij (Marc, Jeroen S, John en Annemarie; altijd in voor discussies, afleiding, hand- en spandiensten), Michiel, Gerben en de andere onderzoekers (Editha als theeleverancier en fijne kamergenote in de laatste paar maanden, Hans, Dieke, Jos en Jan), allen bedankt voor een hele gezellige tijd.

Hierbij ook een dank aan andere collega's in het veld. Collega's uit Utrecht, bedankt voor je enthousiaste inbreng tijdens Helmholtz bijeenkomsten. Raymond en Casper, jullie waren uiteraard de basis voor mijn interesse in dit onderzoek. Ook wil ik Sonja, Maarten van der S. en de anderen in het Optic Flow programma bedanken voor de samenwerking en discussies. Colleagues in the New World (Marty, Krishna, Jim and the others in the Human Frontier program), thanks for sharing data and

ideas. Markus, fijn om met je over modellen te discussiëren, en dat je me de ruimte gaf mijn proefschriftzaken af te ronden.

Vrienden die voor de nodige ontspanning hebben gezorgd mag ik natuurlijk ook niet vergeten, waaronder Vaggers die mij de bergen of rotsen insleepten, jaarclub die me de kroeg in trok, natuurkunde vriendjes voor ruggespraak, muziekmaatjes van trio en orkest, eetclub, en oud-Kerkstraatgenoten voor brood en spelen. Mijn tweede wiel, Arjen, is een steun geweest in tijden van nood, vol ideeën, speciaal tijdens verfrissende fietstrektochten. Hem en Ignace wil ik ook bedanken voor hun inzet als paranimphen. Tot slot een woord van dank aan mijn naasten. Aan mijn zus (voor de oefenpraatjes tijdens VS-reizen). Aan mijn ouders, die me altijd hebben aangemoedigd in het volgen van mijn keuzes en veel hebben gesteund in de laatste fase van het proefschrift. En aan Lucinda, jij hebt mij de afgelopen vier jaar met heel veel geduld en relativerende woorden bijgestaan en hiermee meer voor me betekend dan in een paar woorden te vatten valt. Zonder jou was ik niet zover gekomen.

# Curriculum Vitae

Geboren 23 mei 1970 te Hoogezand-Sappemeer bracht ik mijn jeugd door in het dorp Peize (Drenthe) nabij de stad Groningen. Mijn eerste jaar middelbare school aan het Nienoord College te Leek werd onderbroken door een zeer leerzaam jaar in California, waar ik met mijn ouders en zus kennis maakte met de Amerikaanse cultuur en natuur. Terug in Nederland stapte ik na het eerste jaar over naar het Praedinius Gymnasium te Groningen. Met het VWO-diploma op zak begon ik in 1989 aan de Universiteit Utrecht met de studie Natuurkunde. Voor het afstuderen haalde Raymond van Ee mij over om onderzoek bij hem te doen op de afdeling Fysica van de Mens bij Professor Erkelens. Dit afstudeeronderzoek betrof visuele waarneming van vlakken in diepte middels twee ogen. Omdat ik tijdens mijn studie graag ervaring wilde op doen in meer toegepast onderzoek, liep ik tussentijds 3 maanden stage bij BT laboratories (Britse Telecom) te Ipswich. Daar werkte ik mee aan het ontwerpen van een virtuele 3D-omgeving voor internetdiensten. Uiteindelijk studeerde ik in 1995 af als Experimenteel Fysicus. Gefascineerd door het visuele systeem, kon ik aansluitend in Rotterdam beginnen met promotie-onderzoek bij Bert van den Berg op de afdeling Fysiologie van Professor Collewijn te Rotterdam. Daar heb ik me bezig gehouden met waarneming van visuele beweging in relatie met oogrotaties. Vanaf januari 2000 ben ik begonnen als post-doc bij Markus Lappe op de afdeling Zoologie und Neurobiologie van de Ruhr-Universität Bochum te Duitsland. Daar doe ik onderzoek aan herkenning van bewegingspatronen (zoals lopen of dansen) uit visuele beweging.

# Publications

## Articles


Beintema J. A. & van den Berg A. V. (1998). Heading detection using motion templates and eye velocity gain fields. *Vision Research, 38(14)*, 2155-2179.

Beintema J. A. & van den Berg A. V. (2000). Perceived heading during simulated torsional eye movements. *Vision Research, 40(5)*, 549-566.

Beintema J. A. & van den Berg A. V. Pursuit affects precision of perceived heading for small viewing apertures. Submitted.

Beintema J. A., Hooge I. Th. C. & van den Berg A. V. Combining extra-retinal and visual estimates of eye rotation to perceive heading. In preparation.

van den Berg A. V. & Beintema J. A. (1997). Motion templates with eye velocity gain fields for transformation of retinal to head-centric flow. *Neuroreport, 8*, 835-840.

Hooge I. T. C., Beintema J. A. & van den Berg A. V. (1999). Visual search of heading direction. *Experimental Brain Research, 29(4)*, 615-628.


## Refereed abstracts


Beintema J. A., Hooge I. Th. C. & van den Berg A. V. (1997). Interaction between eye movements and retinal flow for the perception of heading. *Investigative Ophthalmology and Visual Science, 38(4) (Suppl.)*, 381.

Beintema J. A. & van den Berg A. V. (1998). Effect of torsional flow on heading percept. *Investigative Ophthalmology and Visual Science, 39(4) (Suppl.)*, 5020.

Beintema J. A. & van den Berg A. V. (1999). Simulated roll and simulated pursuit affect perceived heading differently. *Perception, 28 (Suppl.)*, 42.