Multisensory perception of spatial orientation and self-motion

ISBN/EAN: 978-90-6464-636-2

Multisensory perception of spatial orientation and self-motion

Integratie van multisensorische informatie over ruimtelijke orientatie en zelfbeweging in de menselijke waarneming.

(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor aan de Universiteit Utrecht op gezag van de rector magnificus, prof.dr. G.J. van der Zwaan, ingevolge het besluit van het college van promoties in het openbaar te verdedigen op

woensdag 13 maart 2013 des ochtends te 10.30 uur.

door
Ksander Nikolas de Winkel
geboren op 9 oktober 1984, te Wageningen

Promotor: Prof.dr. P.J. Werkhoven
Co-promotor: Dr. E.L. Groen

What's Up, Doc?
-Bugs Bunny

Table of Contents

Chapter 1. Introduction

For many sensori-motor tasks humans require an accurate internal representation of self-motion and spatial orientation relative to the environment. For example, navigating through the environment (Dyde et al. 2006; Howard, 1982; Lackner, 1976), maintaining postural balance (Karnath, Ferber & Dichgans, 2000) and recognizing objects (Rock & Heimer, 1957) are functions that depend on a continuous sense of our orientation and motion.

Perception of spatial orientation and self-motion are achieved by the visual and inertial sensory systems. The latter comprises our inertial, or mechanical, sensors: the vestibular system of the inner ear, which reacts directly to accelerations of the body, and the somatosensory system, which consists of several types of sensory organs distributed throughout the body that are sensitive to distortions of the skin such as touch and pressure, and muscle tension.

In order to construct a percept of our spatial orientation and self-motion, our Central Nervous System (CNS) must make sense of complex patterns of information provided by the innumerable neurons that make up our visual, vestibular and somatosensory systems. Optimal sensory integration exploits all sensory information and a priori knowledge available, thereby making estimates of orientation and self-motion as accurate as possible, and enabling us to optimize our navigational effectivity.

In this thesis we investigate if and how the CNS combines sensory information on self-motion and spatial orientation to arrive at a single integrated percept.

Understanding the sensory integration mechanisms underlying effective spatial navigation will bring further the science of human perception and advance the design of effective motion simulation.

Sensory systems

Inertial systems

The vestibular system

The vestibular system is a set of tiny organs (see Figure 1.1) found in the inner ear on both sides of the head. The human vestibular system measures about one centimeter in diameter (Curthoys & Oman, 1987). Although it does not provide an experience as distinct as vision or audition, the vestibular system detects motion and orientation of the head, and directly controls reflexes that allow us to maintain postural stability and keep our eyes focused while moving about through the environment (Howard, 1982). More important

for this thesis, is that it provides us with a sense of self-motion and orientation relative to the world. The vestibular system can be divided into two main components: the semicircular canals, which are responsive to angular accelerations, and the otolith organs, which are responsive to linear accelerations.

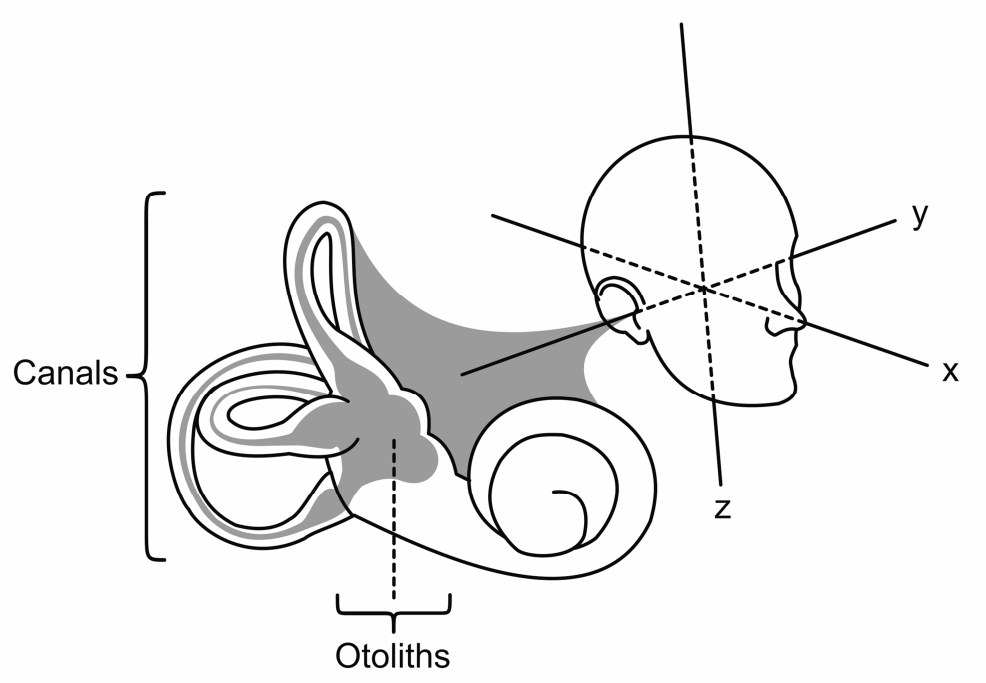

Figure 1.1. The vestibular system of the inner ear, with a set of three semicircular canals which detect angular accelerations, and two otolith organs which detect linear accelerations including gravity.

Semicircular canals

The semicircular canals are circular cavities in the cranium, filled with endolymph fluid. Each canal measures about 7 mm in diameter and has a cross-section of about 0.5mm (Curthoys & Oman, 1987). The canals are interrupted at the base, in a thicker part called the ampulla, by a viscous mass called the cupula. The cupula contains sensory hair cells known as the cilia (Figure 1.1).

To understand how the semicircular canals detect rotations of the head, consider a glass of water mounted firmly on top of a turntable. When the turntable starts rotating, the water in the glass will initially lag behind, and then gradually pick up speed due to friction between the water and the glass, until it rotates at the same angular velocity as the glass

on the turntable. Similarly, when the head rotates, the walls of the canals move relative to the endolymph fluid. This causes a deflection of the cupula. Hence, the cilia bend and send a signal to the CNS. Because the cupula has density identical to that of the endolymph fluid, it can be said to float in the endolymph fluid and is not affected by the pull of gravity.

Per ear, there are three semicircular canals: one horizontal canal, and two vertical canals – the posterior and the superior canal. The semicircular canals detect yaw, pitch, and roll rotations, which are rotations around the z-, x, and y-axis, respectively (Figure 1.1). In a naturally erect posture, the horizontal canal lies tilted backwards approximately 30° to an Earth-horizontal plane. The vertical canals are at approximately 45° angles from a vertical fronto-parallel plane. The canals are not at perfect orthogonal angles: the two vertical canals are at an angle of about 100° relative to one another, and the horizontal canal is at an angle of about 95° with the posterior canal and at an angle of 110° with the superior canal (Fischer, 1993). Moreover, the axes of rotation of the sets of canals on either side of the head are not perfectly aligned, but are at an angle of about 20-25° relative to each other. Because of these differences, any particular rotation of the head gives rise to unique pattern of stimulation of the canals, and by combining the information provided by the canals, the CNS can construct a three-dimensional representation of rotations of the head in space.

Whilst the canals are responsive to angular acceleration, the output of the system depends on motion frequency. Because of the particular diameter of the canals and the elasticity of the cupula compared to the canal's viscous resistance, cupula displacement is in phase with angular velocity for naturally occurring head-movements (i.e., approximately between 0.1 and 5 Hz). In other words: for this range of frequencies the system acts as a mathematical integrator, and the output of the system is proportional to angular velocity (Howard, 1982). At lower frequencies (< 0.1 Hz), sensitivity decreases and cupula displacement leads angular velocity. Consequently, the integration becomes imperfect and we underestimate or even fail to detect low-frequency rotations (Mergner & Rosemeier, 1998. Figure 1.2).

Otoliths

Within each ear, the otolith organs consist of two components: the horizontally placed utricle, and the vertically placed saccule (Figure 1.1). The sensory epithelium of these organs consists of hair cells that are embedded in a gelatinous matrix, covered with solid calciumcarbonate ($CaCO_3$) crystals, the otoconia. The otoconia lag behind when the head undergoes linear accelerations, causing the cilia to bend and the nerve cells to send a signal to the CNS.

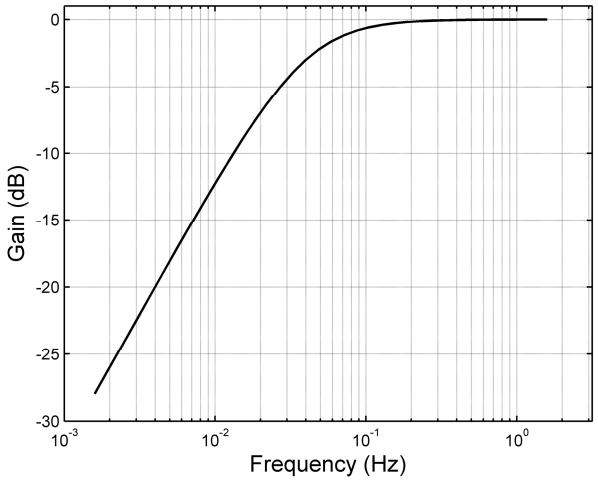

Figure 1.2. Transfer function of the semicircular canals. This plot shows the gain of the output of the canals as a function of motion frequency. As a transfer function, we used a high-pass filter with a time constant of 4s (Raphan & Cohen, 2002).

As described by Einstein's (1907) equivalence principle, a gravitational field and an accelerating frame of reference are physically equivalent. Consequently, it is impossible to distinguish gravity from any other acceleration. Consider a blindfolded observer who is placed, unknowingly, on board of a spacecraft accelerating at 9.82 m/s² in outer space where there is no pull of gravity. Because of the acceleration and the observer's inertia, the observer is forced against the spacecraft's hull. For this observer, it is impossible to determine whether he or she is accelerating in a weightless environment, or standing on the Earth's crust. The accelerations we undergo when walking or driving vehicles fuse with gravity and form a single resultant otolith stimulus. Hence, tilting the head relative to gravity and horizontal linear accelerations can cause physically equivalent stimulation, and the otoliths carry ambiguous signals (Anderson, Blanks & Precht, 1978; Dickman, Angelaki & Correia, 1991; Fernandez & Goldberg, 1976; Loe, Tomko & Werner, 1973; Si, Angelaki & Dickman, 1997; Young et al., 1984).

Because gravity does not cause an illusion of constant upward acceleration, the CNS apparently resolves this ambiguity. It has been shown that this can be achieved by taking into account visual (MacNeilage et al., 2007) and/or vestibular (Angelaki et al., 1999) information, as well as prior knowledge about intended motion (i.e., the efference copy, e.g., Bos and Bles, 2002). However, even in the absence of such additional information,

the CNS is able to distinguish tilt and acceleration. It is assumed that the CNS decomposes the ambiguous otolith signal into a component that represents gravity and a component that represents inertial accelerations by means of frequency-segregation (Mayne, 1974): because gravity is a constant acceleration (i.e., it has a 0 Hz frequency) and other accelerations are generally of short duration (i.e., of high frequency), application of a low-pass filter to the otolith signal allows the CNS to determine the magnitude and orientation of gravity, and consequently our orientation relative to gravity. This means that high-frequency distortions of the otolith signal are interpreted as linear accelerations, and low-frequency distortions as tilting of the head relative to gravity (Figure 1.3).

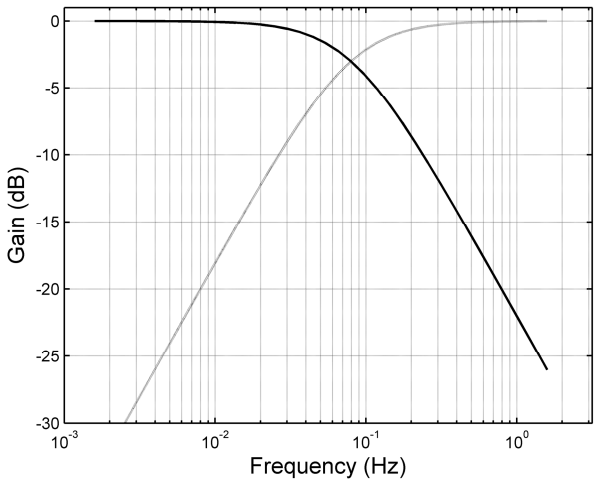

Figure 1.3. Transfer functions for otolith based tilt perception (black curve) and translation perception (gray curve). This plot shows the gain of the perceived tilt or translation, as compared to a complete interpretation of the otolith signal as either tilt or translation. As transfer functions, we used a low-pass filter for perceived tilt and high-pass filter, both with a time constant of 2s (Correia Grácio et al., 2012). Note that these functions represent the interpretation of the CNS of ambiguous otolith information, in the absence of other sources of sensory information on self-motion.

This is illustrated by illusory perceptions of tilt that occur when people are exposed to prolonged linear accelerations which generally do not occur in nature (Bos & Bles, 2002; Clément et al., 2001; Graybiel, Clark & MacCorquodale, 1947; Guedry, 1974; Holly, Vrublevskis & Carlson, 2008; Mayne, 1974; Merfeld, Zupan & Gifford, 2001; Seidman, Telford & Paige, 1998). For example, pilots often experience upward tilt (pitch) during take-off in poor visibility (the somatogravic illusion); people placed off-center in a

centrifuge experience body tilt; and periodical linear accelerations may induce a feeling of moving over a small hill (hill-top illusion).

The somatosensory system

There are several other types of sensory organs, distributed throughout the body, that also inform the CNS about self-motion and orientation. Even though most of these sensors are not stimulated directly by accelerations, they provide the CNS with information from which our orientation, the direction of gravity and other accelerations can be deduced. For example, the forces exerted by our muscles to maintain postural stability are signaled to the CNS by muscle spindles, and cutaneous pressure sensors such as Meissner's and Pacinian corpuscles inform the CNS which surfaces of our body carry our weight: a phenomenon aptly referred to as the 'seat-of-the-pants'. Another cue to orientation is provided by baroreceptors in the vascular system, which registers pressure in the arteries. As the particular distribution of blood throughout our vascular system depends on our orientation relative to gravity, these sensors inform the CNS about our orientation. For example, when we are upside down, the pressure in the head is larger than when we are in an upright position. Similarly, in zero gravity, the mass centroid of the distribution of blood in our body shifts upward relative to normal gravity. Finally, it has been argued that highly specialized graviceptor cells exist in the trunk (Mittelstaedt, 1996), which directly signal the direction of gravitational forces. Because somatosensory sensors respond to the same environmental stimuli as the vestibular system and are difficult to isolate, we will generally treat the vestibular and somatosensory systems as a single inertial system.

Visual system

The photoreceptors of our eyes are responsive to light. Light by itself does not inform us about self-motion or orientation: according to Gibson (1961), the proper stimulus for the visual system does not consist of individual photons, but of the optic array. The optic array can be defined as "the set of sources, surfaces, and media within the field of view of an observer which determines the physical features of the light entering the pupil as all the light reflected by objects in the environment" (Howard, 1982). In other words, the proper stimulus for the visual system consists of all the light that is reflected towards us by objects in our environment. Our eyes constantly sample a part of the optic array. By identifying patterns or changes in patterns in the optic array, the CNS can derive information on orientation and self-motion.

Orientation

To determine orientation from the optic array, the CNS must reference the visual information to known invariants in our environments. Hence, visual information on orientation comes from our knowledge about objects in optic array. The information can be divided into so called intrinsic and extrinsic polarity cues (Howard, 1982). Intrinsic polarity cues define "up" by their constant orientation relative to gravity (e.g., trees and buildings). Extrinsic polarity cues define "up" by their relation to intrinsic polarity cues (e.g., a book can only lie on a horizontal shelf). Manipulation of these cues in the optic array can bias the perception of orientation; for example, observers presented with a tilted line (Müller, 1916), a tilted frame (Witkin & Asch, 1948), or a tilted visual scene (Mittelsteadt, 1986), may experience illusory self-tilt. The occurrence of this illusory self-tilt has been attributed to an influence of prior knowledge on perception; as lifelong experience has taught us, it is more likely that we are tilted relative to the world, than that the world is tilted. On Earth the extent of this illusion is limited, due to the continuous presence of the gravitational reference for 'up'. However, the reliance on visual information is thought to increase in conditions of microgravity (Young et al., 1986), as illustrated by the fact that astronauts on board spacecraft often experience their percept of upright to suddenly realign with certain features of the optic array (Oman, 2007).

Self-Motion

Information on self-motion is directly present in the optic array. Given a stationary environment, its visual representation moves across the retina's of our eyes as a whole when we move. Such motion of the visual representation of environment is known as 'optic flow' (Gibson, 1950). Particular types of motion give rise to specific optic flow patterns: when we are looking in the direction of a linear motion we observe expansion, as particles of the optic array nearer to us seem to move with greater velocity than particles further away; similarly, when we are oriented orthogonally to the direction of translation, nearby objects travel over our retina's faster than distant objects, referred to as motion parallax; and rotations of the optic array give rise to circular flow patterns. Given a large enough Field-of-View and an immersive environment, optic-flow can produce a compelling sensation of self-motion, or 'vection' (Tschermak, 1931). As an example, seeing the departure of a neighboring train, when sitting in one ourselves often induces the feeling that we are moving. Visually induced perception of self-motion can typically be described by low-pass filtering (Bos, Bles & Groen, 2008: see Figure 1.4).

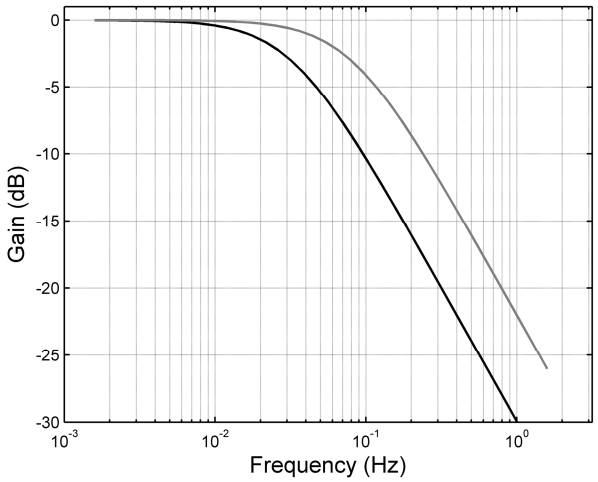

Figure 1.4. Transfer functions for visually induced perception of self-motion, or 'vection'. This plot shows the gain of the perceived circular (black curve) linear and linear motion (gray) as a function of motion frequency. As transfer functions, we used a low-pass filter with a time constant of 5s for circular vection and 2s for linear vection (Bos, Bles & Groen, 2008).

Because most of the naturally occurring motions involve both rotations and translations, optic flow generally contains complex combinations of all types of flow patterns. By identifying specific characteristics of the flow-pattern, the CNS can determine many characteristics of self-motion, such as heading, rotation, traveled distance, and velocity (e.g., Koenderink, 1986; Telford, Howard & Ohmi, 1996; Warren & Hannon, 1988; Warren, Morris & Kalish, 1988).

Multisensory perception of self-motion and orientation

Our experience of spatial orientation and self-motion is a single, unified percept. Hence, the information provided by our different sensory systems must somehow converge in the CNS. At the neural level, canal and otolith afferents have been shown to interact already at the first synapse in the brain stem and cerebellum, and interactions between visual and vestibular afferents occur throughout the visual and vestibular pathways (Angelaki & Cullen, 2008; Angelaki et al., 2011). On the perceptual level, it has been shown that perception of spatial orientation can be described as a compromise between visual information, inertial information (consisting of vestibular and somatosensoryinformation), and prior knowledge (Harm et al., 1998; Mittelstaedt, 1983, 1986, 1988; Young et al. 1986); and conceptual dynamic models have been developed that describe the

interactions between the senses that govern perception of self-motion. These models typically describe our perception of self-motion and orientation as a function of vectors that represent the information provided by the different sensory systems (Angelaki et al., 1999; Bos & Bles, 2002; Glasauer & Mittelstaedt, 1992; Merfeld, 1995; Zupan, Merfeld & Darlot 2002).

Although these models provide a conceptual description of the perception of spatial orientation and self-motion, and successfully predict perceptual illusions, they are deterministic: it is assumed that sensory observations are invariant (i.e., they do not vary for a given stimulus). However, as sensory perception is inherently noisy, there will always be some variance associated with perception (Green & Swets, 1988). This noise is introduced throughout the perceptual process: firstly, at the sensory level, the density of receptor cells affects sensor resolution (Knill & Pouget, 2004); and at the cellular level factors such as temperature, ionic channel fluctuations, and chemical reactions at the synapses affect the cell's discharge rate, also introducing noise (Serletis et al., 2011). Secondly, many neurons exhibit 'tuning curves', referring to the fact that their responsivity peaks for a certain stimulus. For example, cell-recordings in the primate and rat cortex have shown that neurons exist with a selective responsivity for the orientation of visual grating patterns (Kang, Shapley & Sompolinsky, 2004; Ringach, Shapley & Hawken, 2002), the direction of moving dots (Newsome, Britten & Movshon, 1989); the direction of reaching-behavior (Georgopoulos, Schwartz & Kettner, 2006); and spatial location with respect to a particular reference frame (O'Keefe & Dostrovsky, 1971).

Hence, given the vast number of neurons, each carrying their own noisy estimate of an environmental variable, a description of the mechanisms that underlie the perception of self-motion and orientation should consider the stochastic nature of sensory perception.

Statistically Optimal Integration

The stochastic nature of perception is addressed in Maximum-Likelihood and Bayesian models of fusion of multisensory information (Van Beers, Sittig & Dernier van der Gon, 1999; Ernst & Banks, 2002; Knill & Pouget, 2004). In these models, sensory observations of environmental properties are represented as probability distributions that have a certain mean and variance. The mean represents the most likely actual value for a given stimulus, and the variance is a measure of the size of the noise.

Observations provided by different sensors are weighted according to their reliability[1] (i.e., the inverse of their variance) and subsequently summed. As perception is the result of the combined effort of populations of neurons, 'knowledge' on sensory variance is assumed to be an inherent property of the population response (Averbeck, Latham & Pouget, 2006; Pouget, Dayan & Zemel, 2000; Zemel, Dayan, & Pouget, 1998).The result of such integration is that the integrated estimate has minimal variance. Therefore, it is considered statistically optimal. A detailed explanation of these statistically optimal models of multisensory integration is given in Chapter 2.

Conditions for Statistically Optimal Integration

Prior to any form of sensory fusion, the CNS must assess whether multisensory information originates from a common source. Hence, the CNS must evaluate whether the cues are in spatial and temporal agreement, and whether a prior coupling exists in the CNS. Such evaluation is thought to be an automatic consequence of multiple signals arriving within the constraints of a spatial and temporal 'synchrony' window (Spence & Squire, 2003). We can classify multisensory information as being either congruent (in agreement), or incongruent (in disagreement). Incongruent cues by definition violate at least one of the conditions of sematic, spatial and temporal agreement, and are generally processed separately. However, when the size of the disagreement is small, it may sometimes go unnoticed, 'fooling' the CNS. An example is the ventriloquist illusion, in which a ventriloquist creates the illusion of a talking puppet. Congruent cues are subdivided into cues that are complementary and cues that are redundant. Complementary cues are, for example, the sight and sound of a cow: this information can be combined to aid recognition, as it has been shown that their co-occurrence reduces detection times (Yuval-Greenberg & Deouell, 2009), but the information cannot be integrated. However, both the visual and auditory cue inform us of the location of the cow. Hence, the multisensory information on the cow's location is redundant. Similarly, multisensory cues that represent dimensions such as spatial location, magnitude, size, or velocity can be considered redundant. By fusing redundant information in a statistically optimal fashion, the CNS can theoretically arrive at the most reliable estimate of an environmental event given the available sensory information.

[1] In the literature, several terms are used for 'reliability'. Another common term is 'precision'. Note that both terms are defined as the inverse of a signal's variance. The variance is a measure of the size of the noise in a signal, as is the standard deviation, which is the square root of the variance.

Research questions

Thus far, a multitude of studies have provided support for the applicability of statistically optimal integration strategies to perceptual processes. For example, statistically optimal integration has been demonstrated for different visual cues on slant perception (Knill & Saunders, 2003); visual-haptic estimation of object-size (Ernst & Banks, 2002) and shape (Helbig & Ernst, 2007); visual-auditory perception of location (Alais & Burr, 2004); and multisensory information on sequences of events (e.g., Andersen, Tiipana & Sams, 2005; Shams, Ma & Beierholm, 2005).

Hence, it is tempting to assume that Optimal Integration is a general property of the CNS that also applies to perception of self-motion and orientation. Indeed, it has been shown that visual and inertial information can be fused to disambiguate otolith information (MacNeilage et al., 2007); that visual and inertial information can be fused to estimate heading of horizontal linear self-motion (e.g., Butler et al., 2010; Butler et al., 2011; Fetsch et al., 2009; Gu, Angelaki & DeAngelis, 2008; Telford et al., 1995); that people are better at indicating body rotation when both visual and inertial information are available (Jürgens & Becker, 2006); and that visual-vestibular perception of spatial orientation can be modeled using a statistically optimal approach (Vingerhoets et al., 2009).

However, these studies did not always provide conclusive evidence of the applicability of Optimal Integration to perception of self-motion and orientation for a number of reasons.

Firstly, some studies observed departures from optimality, as subjects sometimes assigned more weight to visual or inertial information than predicted (Butler et al. 2010; Fetsch et al., 2009; Wright, DiZio & Lackner, 2005), suggesting that the occurrence of integration depends on specific conditions. Secondly, while most studies investigated whether the multisensory percept is a weighted average of the constituent signals, the precise prediction on multisensory variance provided by optimal integration theory was not always assessed. Thirdly, due to difficulties in stimulating the vestibular sub-sensory systems in isolation, some parameters in state-of-the-art Optimal Integration models have not yet been verified experimentally (Laurens & Droulez, 2007; Clemens et al., 2011).

These limitations inspired us to start the research described in this thesis, with the main research question:

For which conditions can the integration of visual, semicircular canal, and otolith information on spatial orientation and self-motion be predicted by a Bayesian Integration framework based on the variances of the individual sensory signals?

We designed experiments to study the the variances of the visual system, the semicircular canals, and the otoliths in isolation, and to determine their effects on subsequent sensory integration processes. This was only possible by using the unique experimental motion platform DESDEMONA, and by carrying out an experiment in reduced gravity conditions during special parabolic flights.

Outline of this thesis

Chapter 2 provides a detailed description of the method of Maximum-Likelihood and the broader Bayesian framework to multisensory integration. As these stochastic approaches to modeling multisensory integration are the foundation of the research presented in this thesis, this description is intended as reference for the remainder of the chapters.

Chapters 3-6 describe a series of experiments in which the tenability of optimal integration theory to the perception of spatial orientation and self-motion was investigated. The general hypothesis was that multisensory perception conforms to statistically optimal integration. The studies differed in the sensory systems involved and the particular aspects of the perception of spatial orientation and self-motion that were investigated. Chapter 3 presents a study on integration of visual and otoliths information in the perception of linear motion in a horizontal plane. We exposed our subjects to visual and/or inertial horizontal linear accelerations while the heading angle was varied, and asked them to indicate their perceived heading.

Chapter 4 addresses the integration of visual and semicircular canal information on rotational motion in a horizontal plane. We applied a paradigm in which subjects discriminated stimuli based on peak angular velocity, using motion stimuli that contained visual and/or inertial information.

In Chapter 5, we investigated integration of otoliths and semicircular canal information on rotation in a vertical plane. In addition, we hypothesized that the integration of otolith and semicircular information on self-tilt would depend on stimulus frequency, since both vestibular systems have different frequency dynamics (Figures 1.2 and 1.3).

In Chapter 6, we investigated integration of visual and otoliths information in the perception of spatial orientation, and how this perception is affected in conditions of reduced gravity. This study was performed during the first Joint European Partial-G parabolic Flight Campaign, organized by the European Space Agency (ESA), Centre National d'Etudes Spatiales (CNES), and the German Aerospace Center (DLR). Subjects were lying on their side on board of an airplane that flew parabolic trajectories generating partial gravity, and simultaneously presented with visual stimuli designed to bias the perception of up. Subjects were instructed to align a visual indicator with the perceived direction of up. As the strength of otolith stimulation decreases with decreasing level of

gravity, we hypothesized that the internal weighting of the gravitational cue would also decrease with decreasing level of gravity.

Finally, in Chapter 7 we will review our findings in the light of the statistically optimal framework described in Chapter 2. We discuss whether the findings of the studies presented in Chapters 3-6 provide supporting evidence for this framework and whether alternative explanations may apply, and conclude with recommendations for future research.

Chapter 2. Bayesian Framework

A Bayesian framework can be used to construct a model that describes how the Central Nervous System (CNS) makes inferences about the world using sensory information and prior knowledge or experience (Blake, Bülthoff & Sheinberg, 1993). The framework is based on Bayes' theorem, which expresses how a degree of belief should be affected by evidence. When applied to perception, it can be used to express the probability of an event X given a sensory observation of the event, $\hat{X}$, and a priori knowledge about probability of the event:

$$\mathrm{P}(X|\hat{X}) = \frac{\mathrm{P}(\hat{X}|X)\mathrm{P}(X)}{\mathrm{P}(\hat{X})} \text{ (eq 2.1).}$$

Here $\mathrm{P}(X|\hat{X})$ is the 'posterior', which represents the probability that the environmental event is X given the observation $\hat{X}$; $\mathrm{P}(\hat{X}|X)$ is the 'likelihood', which represents the probability of observation $\hat{X}$ when the environmental event is X; $\mathrm{P}(X)$ is the 'prior', which represents a priori beliefs or knowledge about X held by the observer; and $\mathrm{P}(\hat{X})$ is the probability of observation $\hat{X}$. When we consider all observations $\hat{X}$ equally likely (i.e., $\hat{X}$ is uniformly distributed), $\mathrm{P}(\hat{X})$ acts as a normalizing constant, such that:

$$\mathrm{P}(X|\hat{X}) \propto \mathrm{P}(\hat{X}|X)\mathrm{P}(X) \text{ (eq 2.2).}$$

An estimate of X is the value that maximizes the posterior $\mathrm{P}(\hat{X}|X)$: the Maximum A Posteriori estimator, or MAP (Yuille & Bulthoff, 1996, in: Knill & Richards, 1996). If we also assume that the prior is uninformative (i.e., uniformly distributed), the posterior is proportional to the likelihood:

$$\mathrm{P}(X|\hat{X}) \propto \mathrm{P}(\hat{X}|X) \text{ (eq 2.3).}$$

In this case, the model corresponds to Maximum-Likelihood estimation (Fisher, 1922). According to the method of Maximum-Likelihood, the best estimate of X is the value that maximizes the likelihood-function. To understand the application of the Bayesian framework to the integration of multisensory information, it is useful to first introduce the method of Maximum-Likelihood.

Maximum-Likelihood Estimation

Consider a vector of observations $\hat{X} = (x_1, \dots, x_n)$ of event X from a single sensory modality. $\hat{X}$ is a normally distributed random variable with mean μ and standard deviation σ: $\hat{X} \sim \mathcal{N}(\mu, \sigma)$. We assume that perception is unbiased, such that the mean of each of the observations corresponds to the actual value of X. We can express the likelihood function L as

$$\mathrm{L}\left(\mu, \sigma \middle| \hat{X}\right) = \prod_{i=1}^{n} P(x_i|\mu, \sigma) \text{ (eq 2.4).}$$

Because we want to know the μ for which L reaches a maximum, we take the derivative of the logarithm of L, the Log-Likelihood LL, with respect to μ, and equate this derivative to zero.

$$\mathrm{LL}\left(\mu, \sigma \middle| \hat{X}\right) = \log\left[\prod_{i=1}^{n} \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(x_i-\mu)^2}{2\sigma^2}}\right] \text{(eq 2.5).}$$

As the logarithm of a product is the sum of the logarithms of the factors, and the derivative of a sum is the sum of the derivatives, we obtain:

$$\frac{\partial \mathrm{LL}}{\partial \mu} = \frac{\partial}{\partial \mu}\left[\sum_{i=1}^{n} \log\left(\frac{1}{\sigma\sqrt{2\pi}}\right)\right] + \frac{\partial}{\partial \mu}\left[\sum_{i=1}^{n} \frac{-(x_i-\mu)^2}{2\sigma^2}\right] \text{ (eq 2.6).}$$

The first factor on the right-hand side of the equation is constant with respect to μ, and hence its derivative is zero. Using the chain-rule and simplifying we find:

$$\frac{\partial \mathrm{LL}}{\partial \mu} = \frac{1}{\sigma^2}\left[\sum_{i=1}^{n} (x_i - \mu)\right] \text{ (eq 2.7).}$$

Equating the derivative to zero and solving for μ yields the Maximum-Likelihood Estimator (MLE) $\hat{\mu}$ for the mean μ, which is equal to the arithmetic mean of the vector $\hat{X}$,

$$\hat{\mu} = \frac{\sum_{i=1}^{n} x_i}{n} = \bar{x} \text{(eq 2.8).}$$

Maximum-Likelihood Integration

To understand how the method of Maximum-Likelihood can be applied to integration of multisensory information, consider two sensory systems that simultaneously provide the CNS with observations $\hat{X}_1$ and $\hat{X}_2$ of event X. We assume that both observations are normally distributed unbiased random variables with $\hat{X}_1$ and $\hat{X}_2$. Moreover, we assume that the CNS regards the information provided by these sensors as redundant, and therefore the CNS will attempt to combine $\hat{X}_1$ and $\hat{X}_2$ into an estimator $\hat{\mu}$ for X. The observations may differ in the size of their respective noises, or standard deviation σ. However, because the observations reflect the same event, there is only a single 'true' value of the event. The CNS can arrive at $\hat{\mu}$ by using the method of Maximum-Likelihood. The Likelihood function L is given by:

$$\mathrm{L}\left(\mu\middle|\hat{X}_1,\hat{X}_2;\sigma_1,\sigma_2\right)=\frac{1}{\sigma_1\sqrt{2\pi}}e^{\frac{-(\hat{X}_1-\mu)^2}{2\sigma_1^2}}\times\frac{1}{\sigma_2\sqrt{2\pi}}e^{\frac{-(\hat{X}_2-\mu)^2}{2\sigma_2^2}}\quad\text{(eq 2.9)},$$

where we treat the standard deviations σ_1 and σ_2 as known values. As in the case of unimodal sensory estimation, we want to know the value of μ for which L reaches a maximum. Hence, we follow the same steps.

First, we take the derivative of the Log-Likelihood LL with respect to μ:

$$\begin{aligned}\mathrm{LL}\left(\mu\middle|\hat{X}_1,\hat{X}_2;\sigma_1,\sigma_2\right)&\\&=\sum_{i=1}^{n}\left[\log\left(\frac{1}{\sigma_1\sqrt{2\pi}}\right)+\frac{-\left(\hat{X}_1-\mu\right)^2}{2\sigma_1^2}+\log\left(\frac{1}{\sigma_2\sqrt{2\pi}}\right)\right.\\&\quad\left.+\frac{-\left(\hat{X}_2-\mu\right)^2}{2\sigma_2^2}\right]\text{(eq 2.10)}.\end{aligned}$$

$$\frac{\partial\mathrm{LL}}{\partial\mu}=\sum_{i=1}^{n}\left[\frac{\left(\hat{X}_1-\mu\right)}{\sigma_1^2}+\frac{\left(\hat{X}_2-\mu\right)}{\sigma_2^2}\right]\text{(eq 2.11)}.$$

Equating the derivative to zero and solving for μ, it can be seen that the MLE $\hat{\mu}$ of μ is a weighted linear combination of observations $\hat{X}_1$ and $\hat{X}_2$, where the weights are proportional to their variances (i.e., σ^2):

$$\hat{\mu} = \frac{\sigma_2^2}{\sigma_1^2 + \sigma_2^2}\hat{X}_1 + \frac{\sigma_1^2}{\sigma_1^2 + \sigma_2^2}\hat{X}_2 \text{(eq 2.12).}$$

In the case of normal distributed variables as above, the MLE of μ is statistically optimal in the sense that it has the smallest possible variance. Hence, its variance is always smaller than that of either of the observations $\hat{X}_1$ or $\hat{X}_2$. The size of the reduction can be found using the general expression for the variance of a linear combination of variables:

$$\sigma^2 = \text{var}\left(\sum a_i x_i\right) = \sum_{i=1}^{n} a_i^2 \sigma_i^2 \text{ (eq 2.13),}$$

where a_i^2 represents the squared weight of variable i, and σ_i^2 its variance. Substituting the weights $a_1 = \sigma_2^2/(\sigma_1^2 + \sigma_2^2)$ and $a_2 = \sigma_1^2/(\sigma_1^2 + \sigma_2^2)$ obtained from the MLE of μ and subsequently simplifying yields:

$$\sigma^2 = \frac{\sigma_1^2 \sigma_2^2}{\sigma_1^2 + \sigma_2^2} \text{(eq 2.14).}$$

Bayesian Integration

A theoretical objection against Maximum-Likelihood Integration is that it implies that the CNS always merges redundant multisensory information. As an alternative, we consider a Bayesian framework, as proposed by Ernst (2005). This framework differs from Maximum-Likelihood Integration in two ways: firstly, it does not assume that there is a single μ for observations $\hat{X}_1$ and $\hat{X}_2$. Rather than that, $\hat{X}_1$ and $\hat{X}_2$ are considered observations from individual systems that may or may not have different expected values. Secondly, the amount by which the systems influence each other is expressed by the 'coupling prior'.

As noted at the beginning of this chapter, in the Bayesian Framework, the integrated estimates are obtained from the posterior distribution, which is proportional to the product of the likelihood and the prior. Here we represent the likelihood as a function of μ_1 and μ_2 treating $\hat{X}_1$, $\hat{X}_2$, σ_1, and σ_2 as knowns. Moreover, we assume that the observations $\hat{X}_i$ are stochastichally independent:

$$\mathrm{L}\left(\mu_1, \mu_2 \middle| \hat{X}_1, \hat{X}_2, \sigma_1, \sigma_2\right) = \frac{1}{\sigma_1\sqrt{2\pi}} e^{\frac{-(\hat{X}_1 - \mu_1)^2}{2\sigma_1^2}} \times \frac{1}{\sigma_2\sqrt{2\pi}} e^{\frac{-(\hat{X}_2 - \mu_2)^2}{2\sigma_2^2}} \text{ (eq 2.15).}$$

The coupling prior, with coupling parameter σ_p^2, is given by:

$$f_{prior}(\mu_1, \mu_2) = \frac{1}{\sigma_p\sqrt{2\pi}} e^{\frac{-(\mu_1-\mu_2)^2}{2\sigma_p^2}} \text{ (eq 2.16).}$$

Hence, the posterior is:

$$f_{posterior\_1}(\mu_1, \mu_2) \propto e^{\frac{-(\hat{X}_1-\mu_1)^2}{2\sigma_1^2}} \times e^{\frac{-(\hat{X}_2-\mu_2)^2}{2\sigma_2^2}} \times e^{\frac{-(\mu_1-\mu_2)^2}{2\sigma_p^2}} \text{ (eq 2.17).}$$

Although it is possible to derive analytic expressions for the posterior distribution of $\hat{\mu}_1$ and $\hat{\mu}_2$ by equating the derivative of $f_{posterior}$ with respect to μ_1 and μ_2 to zero and subsequently solving for these parameters (as was done for the Maximum-Likelihood example), this approach does not allow us to obtain expressions for the posterior variances and covariance. Therefore, we observe that $f_{posterior\_1}$ is quadratic in μ_1, μ_2 (i.e., it can be written as a linear form in terms of $\mu_1^2, \mu_2^2, \mu_1\mu_2, \mu_1, \mu_2$) to obtain analytic expressions for the expected value of μ_1 and μ_2 and their respective variances,. This implies that the posterior is a bivariate normal distribution with the general form (e.g., Miller & Miller, 2004):

$$f_{posterior\_2}(\mu_1, \mu_2) = \frac{\exp\left[\frac{1}{2(1-\rho^2)}\left\{\left(\frac{\mu_1-\alpha_1}{\gamma_1}\right)^2 + \left(\frac{\mu_2-\alpha_2}{\gamma_2}\right)^2 - 2\rho\left(\frac{\mu_1-\alpha_1}{\gamma_1}\right)\left(\frac{\mu_2-\alpha_2}{\gamma_2}\right)\right\}\right]}{2\pi\gamma_1\gamma_2\sqrt{(1-\rho^2)}} \text{ (eq 2.18).}$$

here, α_1, α_2 are the expected posterior values of μ_1, μ_2; γ_1^2, γ_2^2 are the posterior variances of μ_1, μ_2; and ρ is their posterior correlation.

By rewriting both the original expression for the posterior and the expression for the bivariate normal distribution in the quadratic form, the posterior means and variances (i.e., α_1, α_2 and γ_1^2, γ_2^2) can be expressed as functions of the coefficients $\mu_1^2, \mu_2^2, \mu_1\mu_2, \mu_1, \mu_2$, which are functions of the fixed values $\hat{X}_1, \hat{X}_2, \sigma_1, \sigma_2, \sigma_p$.

The original expression $f_{posterior\_1}(\mu_1, \mu_2)$ can be rewritten as:

$$f_{posterior\_1} \propto \exp\left[-\frac{1}{2}(\Omega)\right] \text{(eq 2.19).}$$

where Ω is equal to:

$$\mu_1^2\left\{\frac{1}{\sigma_1^2}+\frac{1}{\sigma_p^2}\right\}+\mu_2^2\left\{\frac{1}{\sigma_2^2}+\frac{1}{\sigma_p^2}\right\}+\mu_1\mu_2\left\{\frac{-2}{\sigma_p^2}\right\}+\mu_1\left\{\frac{-2\hat{X}_1}{\sigma_1^2}\right\}+\mu_2\left\{\frac{-2\hat{X}_2}{\sigma_2^2}\right\}$$
$$+\left\{\left(\frac{\hat{X}_1}{\sigma_1}\right)^2+\left(\frac{\hat{X}_2}{\sigma_2}\right)^2\right\} \text{(eq 2.20).}$$

and the expression $f_{posterior\_2}$ for the bivariate normal can be rewritten as:

$$f_{posterior\_2} \propto \exp\left[-\frac{1}{2}(\Psi)\right] \text{(eq 2.21).}$$

where Ψ is equal to:

$$\left(\frac{1}{1-\rho^2}\right)\left[\mu_1^2\left\{\frac{1}{\gamma_1^2}\right\}+\mu_2^2\left\{\frac{1}{\gamma_2^2}\right\}+\mu_1\mu_2\left\{\frac{-2\rho}{\gamma_1\gamma_2}\right\}+\mu_1\left\{\frac{-2\alpha_1}{\gamma_1^2}+\frac{2\rho\alpha_2}{\gamma_1\gamma_2}\right\}+\mu_2\left\{\frac{-2\alpha_2}{\gamma_2^2}+\frac{2\rho\alpha_1}{\gamma_1\gamma_2}\right\}\right.$$
$$\left.+\left\{\left(\frac{\alpha_1}{\gamma_1}\right)^2+\left(\frac{\alpha_2}{\gamma_2}\right)^2-2\rho\frac{\alpha_1\alpha_2}{\gamma_1\gamma_2}\right\}\right] \text{(eq 2.22).}$$

By equation of the corresponding coefficients from Ω and Ψ for μ_j^2, for $j = \{1,2\}$, it follows that:

$$\left\{\frac{1}{\sigma_j^2}+\frac{1}{\sigma_p^2}\right\}=\frac{1}{(1-\rho^2)\gamma_j^2} \text{(eq 2.23).}$$

and

$$\gamma_j^2=\frac{\left\{\frac{1}{\sigma_j^2}+\frac{1}{\sigma_p^2}\right\}}{(1-\rho^2)} \text{(eq 2.24).}$$

By equation of the corresponding coefficients from Ω and Ψ for $\mu_1\mu_2$ it follows that:

$$\frac{-2}{\sigma_p^2} = \left(\frac{1}{1-\rho^2}\right)\left\{\frac{-2\rho}{\gamma_1\gamma_2}\right\} \text{(eq 2.25).}$$

By filling in $\gamma_1\gamma_2$ in the above equation using the square root of expressions that we obtained for γ_1^2 and γ_2^2 and subsequently simplifying, we can express the correlation ρ as:

$$\rho = \frac{\sigma_1\sigma_2}{\sqrt{\left(\sigma_1^2+\sigma_p^2\right)\left(\sigma_2^2+\sigma_p^2\right)}} \text{(eq 2.26).}$$

By filling in the expression for ρ in the expressions for γ_1^2 and γ_2^2, obtained from the coefficients for μ_1^2, μ_2^2, and subsequently simplifying, we find that:

$$\gamma_1^2 = \frac{\sigma_1^2\left(\sigma_2^2+\sigma_p^2\right)}{\left(\sigma_1^2+\sigma_2^2+\sigma_p^2\right)} \text{(eq 2.27);}$$

$$\gamma_2^2 = \frac{\sigma_2^2\left(\sigma_1^2+\sigma_p^2\right)}{\left(\sigma_1^2+\sigma_2^2+\sigma_p^2\right)} \text{(eq 2.28).}$$

Similar to the way we derived expressions for ,$\rho, \gamma_1^2, \gamma_2^2$, solving the coefficients for μ_1 and μ_2 in α_1 and α_2, we find that α_i is a weighted sum of x_1 and ,x_2, where the weights depend on $\sigma_1, \sigma_2, \sigma_p$:

$$\mathrm{E}(\mu_1) = \alpha_1 = \omega_1\hat{X}_1 + (1-\omega_1)\hat{X}_2;\ \omega_1 = \left(\sigma_2^2+\sigma_p^2\right)/\left(\sigma_1^2+\sigma_2^2+\sigma_p^2\right) \text{(eq 2.29);}$$

$$\mathrm{E}(\mu_2) = \alpha_2 = \omega_2\hat{X}_2 + (1-\omega_2)\hat{X}_1;\ \omega_2 = \left(\sigma_1^2+\sigma_p^2\right)/\left(\sigma_1^2+\sigma_2^2+\sigma_p^2\right) \text{(eq 2.30).}$$

From these expressions, it follows that when the coupling parameter of the prior is infinite ($\sigma_p^2 \to \infty$), $\alpha_i = \hat{X}_i$, with $\gamma_i^2 = \sigma_i^2$. In this case, μ_1 and μ_2 are independent, and the CNS processes the information provided by the sensors independently. This is illustrated in the upper row of Figure 2.1. The left panel shows the prior coupling between the systems. The flat prior indicates that there is no coupling, and hence that all combinations of $\hat{X}_1, \hat{X}_2$ are equally likely. The middle panel shows the actual observations $\hat{X}_1$ and $\hat{X}_2$ and the associated variances, and the right panel shows the posterior estimates $\hat{\mu}_1$ and $\hat{\mu}_2$.

When the systems are weakly coupled, the coupling parameter has a finite value. This implies that the integrated estimate is largely determined by the observation with the smallest variance, but shifted towards the other observation. Statistically, such weak coupling provides an advantage, as in this case the integrated estimate is slightly more precise than either constituent observation (Figure 2.1, middle row).
When there is a strong coupling between the sensory systems, the value of the coupling parameter approaches, or is equal to zero, and the sensory observations are completely fused. In this case, the Bayesian model reduces to Maximum-Likelihood Integration (Figure 2.1, lower row): the posterior estimates $\hat{\mu}_1$ and $\hat{\mu}_2$ are a weighted average of observations $\hat{X}_1$ and $\hat{X}_2$, where the weights are proportional to their variances (i.e., σ^2). Hence $\hat{\mu}_1$ and $\hat{\mu}_2$ are statistically optimal in the sense that it has the smallest possible variance across all convex combinations of $\hat{X}_1$ and $\hat{X}_2$.

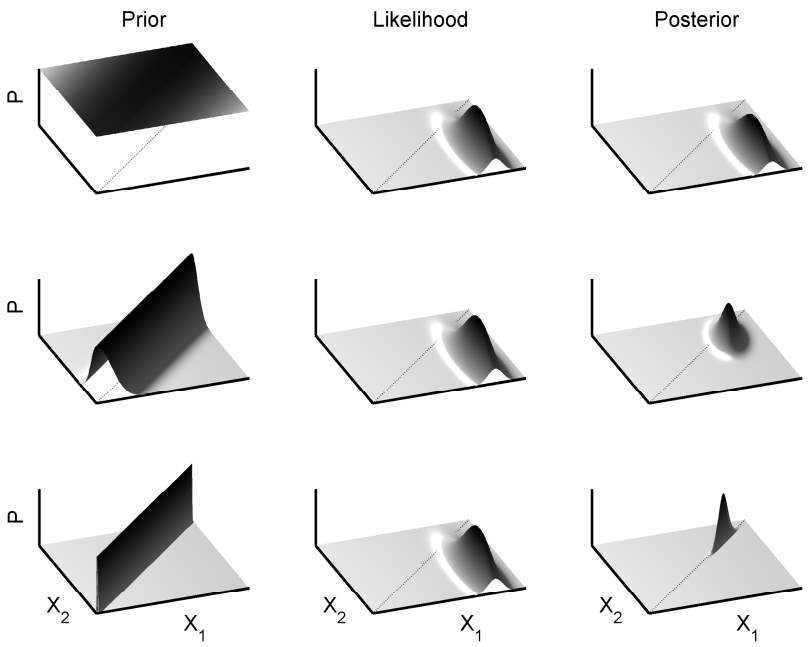

Figure 2.1. Visual illustration of the Bayesian framework including a coupling prior. The upper row shows the case of sensory fusion with a coupling prior with infinite coupling parameter; the middle row shows sensory fusion with a coupling prior with a finite coupling parameter; and the lower row shows sensory fusion with a coupling prior with a coupling parameter that approaches zero.

Chapter 3. Integration of Visual and Inertial Cues in Perceived Heading of Self-motion[1]

Summary

In the present study we investigated whether the perception of heading of linear self-motion can be explained by Maximum Likelihood Integration (MLI) of visual and inertial sensory cues. MLI predicts smaller variance for multisensory judgments compared to unisensory judgments. Nine participants were exposed to visual, inertial, or visual-inertial motion conditions in a moving-base simulator, capable of accelerating along a horizontal linear track with variable heading. Visual random-dot motion stimuli were projected on a display with a 40° horizontal × 32° vertical field-of-view (FoV). All motion profiles consisted of a raised cosine bell in velocity. Stimulus heading was varied between 0-20°. After each stimulus, participants indicated whether perceived self-motion was straight-ahead or not. We fitted cumulative normal distribution functions to the data as a psychometric model, and compared this model to a nested model in which the slope of the multisensory condition was subject to the MLI hypothesis. Based on likelihood-ratio tests, the MLI model had to be rejected. It seems that the imprecise inertial estimate was weighed relatively more than the precise visual estimate, compared to the MLI predictions. Possibly, this can be attributed to low realism of the visual stimulus. The present results concur with other findings of overweighing of inertial cues in synthetic environments.

[1] This chapter has been published as: De Winkel, K.N., Weesie, J., Werkhoven, P.J., & Groen, E.L. (2010) Integration of visual and inertial cues in perceived heading of self-motion. Journal Of Vision 10(12).

Introduction

Perception of self-motion and –orientation in the environment is based on neural integration of inputs from the visual, vestibular, kinesthetic, and tactile senses. There has been extensive research on this topic, in particular on the contribution of the vestibular and visual systems (e.g., Bischof, 1974; Guedry, 1974; Howard, 1982). The visual system provides multi-dimensional information on self-motion by means of optic flow, and on spatial orientation by means of visual frame and polarity information (Howard & Childerson, 1994). The vestibular system in the inner ear detects linear motion by means of the otolith organs. Furthermore, there is evidence for extra-vestibular sensory neurons dedicated to the perception of gravity; so called 'graviceptors' (Mittelstaedt, 1996; Zaichik, Rodchenko, Rufov, Yashin & White, 1999). The dynamics of the different sensory systems and interactions between them have been represented in mathematical models to explain fundamental psychophysical characteristics of self-motion and –orientation (e.g., Bos & Bles, 2002; Zupan, Merfeld & Darlot, 2002). For example, models have successfully described the multisensory processes in the onset of perceived self-motion (e.g., Young, Dichgans, Murphy & Brandt, 1973; Henn, Cohen & Young, 1980); the neural disambiguation of the gravito-inertial force into gravity and linear acceleration by means of visual information (MacNeilage, Banks, Berger & Bülthoff, 2007); and visual-vestibular interaction in the perception of self-tilt (Vingerhoets, Van Gisbergen & Medendorp, 2009). However, heading perception, i.e., the direction of linear motion along the naso-occipital axis, has mainly been studied as a visual task with non-moving observers and has not received much attention in the literature on multisensory perception. Furthermore, although multisensory perception has been associated with higher precision than unisensory perception for several perceptual tasks, it is not yet known whether this principle also holds for the perception of heading. In this paper, we describe an experiment in which we measured variability in visual, inertial, and combined visual-inertial perceptions of heading in order to test whether multisensory stimuli yielded more precise heading judgments than their unisensory constituents.

Visual heading perception

During self-motion, heading is specified by the point in the optic array from which the surrounding image radially expands (the Focus-Of-Expansion, FOE). Heading can be estimated by localizing this point (Gibson, 1950; Warren & Hannon, 1988). Warren, Morris and Kalish (1988) presented to participants moving horizontal random-dot planes and instructed the participants to report whether it looked as if they moved to the left or right of a target. 75%-correct thresholds were in the range of 0.65° to 1.2°, depending

somewhat on the density of the dots and stimulus speed. In a paper by Telford, Howard, and Ohmi (1995) a 75% correct detection threshold of 5.5° was reported, averaged over participants. Accuracy of heading judgments depends upon the part of the fovea stimulated, eccentricity of the FOE (Warren and Kurtz, 1992; Crowell and Banks, 1993), and the coherence of motion of particles in the optic flow field (Gu, Angelaki & DeAngelis, 2008).

Inertial heading perception

The vestibular system responds to rotational and linear accelerations of the head associated with changes of self-motion. It is less useful for the perception of the actual speed and trajectory of self-motion. For these aspects we rely on the visual system. However, the otolith organs do provide information on the direction of linear acceleration (Gu, DeAngelis & Angelaki, 2007; Gu, Angelaki & DeAngelis, 2008; Ohmi, 1996; Telford et al. 1995), as do several extra-vestibular graviceptors throughout the body (Mittelstaedt, 1996; Zaichik, Rodchenko, Rufov, Yashin & White, 1999). For humans, a 75%-correct detection threshold of non-zero heading (i.e., not straight ahead) for inertial stimuli has been reported to be 11.4°, averaged over participants (Telford et al., 1995). Thresholds as small as 1.3° have been reported for extensively trained macaque monkeys, suggesting that inertial signals can indeed contribute to precise heading judgments (Gu et al., 2007). Since extra-vestibular sensors respond to the same environmental stimuli as the vestibular system and are difficult to isolate from the vestibular system, we will treat them as a single system, here collectively designated inertial heading sensors.

Multisensory heading perception

It is likely that the brain combines estimates of heading made by multiple senses in a way that allows us to benefit from having multiple sources of information. Such a benefit can be expressed by improvements in precision (i.e., a reduction of variance) of the estimate. A well-known integration scheme is Bayesian Integration (BI). BI models have been applied to describe the influence of prior knowledge on sensory integration (e.g., Bresciani, Dammeier & Ernst, 2006; Laurens & Droulez, 2007; Jürgens & Becker, 2006; MacNeilage, Banks, Berger & Bülthoff, 2007). In its most simple form, the Bayesian scheme is essentially Maximum-Likelihood Integration (MLI) (Ernst & Banks, 2002). This is a statistically optimal strategy to combine multiple cues (“observations”). Assuming normality and independence of noises in internal representations, MLI effectively states that multisensory estimates of the same environmental property are combined as a weighted average, with weights proportional to each estimate’s reliability (i.e., inverse of its variance) (Howard, 1997). MLI yields minimal variance in the integrated estimate

among all weighted averages. Thus, MLI predicts how the (parameters of the) multisensory condition relate to the (parameters of) unisensory conditions. It has been shown that the brain acts according to MLI for several psychophysical phenomena, such as integration of visual and haptic information on an object's size (Ernst & Banks, 2002) and for integration of multisensory information on sequences of events (e.g., Andersen, Tiipana & Sams, 2005; Shams, Ma & Beierholm, 2005). In these studies, multisensory estimates were more precise than unisensory estimates. In a recent study Gu et al. (2008) investigated whether this integration strategy also holds for the perception of heading. Macaque monkeys were trained to perform a discrimination task in which they were passively moved along a linear track in the horizontal plane with a certain heading. The monkeys indicated whether the experienced motion was to the left or right of straight ahead heading. The results suggested that the monkeys were more precise in their judgments when multisensory visual-inertial cues were presented than when either cue was presented in isolation, consistent with the MLI hypothesis. Furthermore, Fetsch, Turner, DeAngelis and Angelaki (2009) recently showed that cue weighting is a dynamic process. They presented macaque monkeys and humans with multisensory heading stimuli. Reliability of the visual heading cue was manipulated between successive trials by varying coherence of the direction of moving dots in the visual field between 25 and 70%. It was observed that the weight attributed to each cue was updated on a trial-to-trial basis.

Our interest in the present study was to investigate whether multisensory presentation yields a more precise estimate compared to its unisensory constituents, when the objects in the optical array move in a completely coherent fashion. We hypothesized that the variance of the estimate would be smallest when both visual and inertial cues were available, compared to when either cue was presented in isolation. We focused on MLI theory, which gives a precise prediction of the extent to which the variance is reduced. The study was performed in a moving base simulator which allowed us to independently manipulate visual and inertial motion cues.

Methods

Participants

Nine paid volunteers took part in this experiment (six men, three women, mean age 28.7, standard deviation 6.9). All participants reported normal vestibular function and normal or corrected-to-normal vision. After receiving general instructions of the experimental goals and procedures all participants signed an informed consent.

Apparatus

The experiment was performed using the DESDEMONA simulator at the TNO institute in Soesterberg, the Netherlands (Bles & Groen, 2009). This centrifuge-based simulator features a motion platform with six degrees-of-freedom. For this study, only two degrees-of-freedom were used: rotation about the cabin's vertical yaw axis, and linear motion along the 8m horizontal track. Participants were seated on a padded seat inside the cabin, and secured by a five-point safety harness. Foam cushions were placed between a head-rest and the left and right side of the head to minimize head movements.
Inside the cabin, an out-the-window (OTW) visual stimulus was projected on a screen at about 1.5m in front of the participant. Participants wore a mask which restricted their field-of-view to 40° (horizontal) × 32° (vertical) of the OTW display, blocking stationary visual cues from the cabin interior. The mask served as a substitute for a lacking cockpit canopy and made the OTW scenery appear in the background. The latter is important, since it is known that vection (the visually induced sense of self-motion) is induced more effectively when a visual motion stimulus is presented in the perceptual background (Howard & Heckmann, 1989). Participants wore an audio-headset that allowed for constant contact with the experimenters. Responses were given verbally and were noted by the experimenters.

Stimuli

Stimuli consisted of visual, inertial, or combined visual-inertial linear horizontal motion with different headings. Heading is defined as the direction of motion with respect to the median plane of the body: A heading of 0° corresponds to linear forward motion along the participants' naso-occipital axis; a 90° heading corresponds to linear rightward motion along the inter-aural axis. We did not expect consistent differences between perception of left and rightward motion. Therefore, only rightward motion was presented, which reduced the number of trials.
Visual stimuli consisted of linear horizontal motion through a star field. Different angles of visual heading were achieved by shifting the Focus-Of-Expansion (FOE) sideways. The star field consisted of a cloud of solid white circles, placed in random positions on a three-dimensional grid in a dark surrounding environment. This stimulus was generated at random with each trial (Figure 3.1). The objects never appeared at the FOE. Displacement of the visual objects was coupled linearly with the inertial motion. Absolute velocity of the visual motion was arbitrary since participants neither had objective information about the distance between objects, nor could they determine their size. Velocity of movement through the star field was determined in a pilot study; the stimulus amplitude was chosen

such that it subjectively matched the velocity of the inertial motion. Participants never reported a feeling of discrepancy between the visual and inertial velocity.

Inertial stimuli consisted of motion along the linear horizontal track of the simulator over a total length of 7 meters. The velocity profile was a raised cosine bell with maximum velocity of 1.5 m/s. Maximum acceleration was 0.5 m/s². Each motion profile lasted 9.3 s (see additional material: HeadingI.wmv). Since the vestibular system is responsive to acceleration and not to constant velocity, a constantly changing velocity profile was used to ensure vestibular reactivity.

Participants reported whether or not they perceived a heading of 0°. Psychometric curves (cumulative normal distribution functions) of the probability that a participant perceived the stimulus as being not-straight ahead as a function of heading angle were determined based on six different fixed stimulus headings. When using fixed stimuli it is important that the data points fall as much as possible on the steepest part of the psychometric curve. Based on previous reports we knew that the detection thresholds differ between the different sensory modalities. For example, visual detection thresholds fall in the range of 0.65°-5.5° (Warren, Morris & Kalish, 1988, Telford et al., 1995, respectively) and an inertial detection threshold of 11.4° have been reported (Telford et al., 1995). The range of stimulus values for which the percentage correct responses gradually increases thus varies between sensory modalities. Therefore we used different ranges of stimulus values for each condition, which we verified in a pilot experiment (Table 3.1).

Figure 3.1. Screen-capture of the visual stimulus

Table 3.1. Heading angles used in the three conditions. V, I, & C stand for the visual, inertial and combined visual-inertial condition, respectively.

	Heading angle 'α'							
	0°	1°	3°	5°	7°	10°	15°	20°
V	x	x	x	x	x	x		
I	x		x		x	x	x	X
C	x		x	x	x	x	x	

Procedure

Stimuli were presented in separate simulator runs of about 30s. At the beginning of each run the cabin was positioned at one end of the linear track. The run started with rotation of the cabin about its yaw axis to orient the participant at the desired stimulus heading relative to the linear track. The cabin always rotated the longest distance (i.e., around 180°) with an angular velocity between 12-13.33°/s, depending on the stimulus angle. The duration of this rotation was kept constant to eliminate it as a possible cue to stimulus heading. To allow the after-effects of the response of the semi-circular canals for this yaw-rotation to wash out, a 6s pause was implemented before the actual stimulus started. After this pause the cabin was moved over 7m to the other end of the linear track in the inertial and visual-inertial conditions. In the visual-only condition, the cabin remained stationary at the end of the linear track, and only visual motion was presented. Linear motion stimuli always lasted 9.3s. Following stimulus presentation, participants gave their verbal response in a 1s pause before the next run started with the reorientation of the cabin. Participants were asked to judge the direction of perceived self-motion by means of a two-alternative forced-choice (2AFC) task. In other words, they indicated after each stimulus whether they perceived the motion as "straight ahead" or "not straight ahead". Participants were instructed to use all sensory information on self-motion, irrespective whether this was inertial, visual, or both.

The runs were presented in three separate blocks, or conditions: a condition with visual-only (V) runs, a condition with inertial (I) runs, and a condition with combined visual-inertial (C) runs. In each condition, all six stimulus angles were presented 10 times, totaling 60 runs per condition. Participants also experienced an extra set of 60 visual-inertial stimuli to answer another research question, which will be reported elsewhere. The order of stimuli within a condition was randomized. The order of conditions was randomized as much as possible using Latin squares. Each participant performed a total of 240 runs in four 30-minute blocks. After each block they had a 15-minute break outside the simulator. Including instruction, the whole experiment lasted about four hours for each participant.

Theoretical model

We assume that participants have an internal noisy but 'continuous' representation X^* of their heading angle α, with normal distributed noise, $X^* \sim N(\alpha, \sigma)$, where the standard deviation σ reflects the size of the noise in the random variables. We expect that participants experience to move "not straight ahead" and respond accordingly (i.e., binary response; X = 1) when a certain internal threshold is exceeded: $X^* > \tau$ (Long 1997: Section 3.2). When the heading angle increases, the probability of responding "not straight ahead" also increases.

Maximum-Likelihood Integration Hypothesis

MLI hypothesizes how the standard deviation parameters for the multisensory (combined) condition (σ_c) depend on the standard deviation parameters for the unisensory inertial (σ_i) and visual (σ_v) conditions. The value of an environmental property, such as an assessment of heading, can be represented by an 'internal' random variable X^*. When we have two assessments, $X_i^*$and X_v^*, of a single environmental property, as is the case with multiple senses in the combined condition, the value of that property can be estimated by a convex (i.e., coefficients sum to one) combination of the unisensory representations

$$X_c^* = w_i X_i^* + w_v X_v^*, w_i, w_v > 0, w_i + w_v = 1 \text{ (eq. 3.1)}$$

where the *w*'s are weights. Assuming unbiased unisensory noisy representations of the true heading angle α, the linear combination is also an unbiased noisy representation of α. Since a linear combination of normal variates is itself normal distributed, the noise in the combined estimate X_c^* is also normal-distributed. Assuming that the random noises are stochastically independent, the variance σ_c^2 of X_c^* is

$$\sigma_c^2 = w_i^2 \sigma_i^2 + w_v^2 \sigma_v^2 \text{ (eq 3.2)}$$

The statistical method of Maximum-Likelihood (ML) can be used to give a prediction about the weights (w_i and w_v). Since we assumed normal distributed noise, the likelihood L_j for the internal representation $X_j^*$for the sensory condition j is

$$L_j(\alpha; X_j^*, \sigma_j) = \frac{1}{\sqrt{2\pi\sigma_j^2}} \exp\left\{\frac{1}{2}\left(\frac{X_j^* - \alpha}{\sigma_j}\right)^2\right\} \text{ (eq. 3.3)}$$

In the multisensory condition, the likelihood-function of (X_i^*, X_v^*) is given by the product of the likelihoods of the unisensory variables X_j^* because we assume that the noises are independent across senses. Treating the σ_j as knowns, the maximum of this function yields the ML-estimate of the true angle α in terms of these parameters. It can be derived mathematically that the values of the expression where the maximum is attained indeed takes the linear form (3.1) where

$$w_i = \frac{\sigma_v^2}{\sigma_v^2 + \sigma_i^2}, \text{ and } w_v = \frac{\sigma_i^2}{\sigma_v^2 + \sigma_i^2} \quad \text{(eq. 3.4)}$$

Hence the variance σ_c^2 of X_c^* corresponding to these weights is

$$\sigma_c^2 = \frac{\sigma_v^2 \sigma_i^2}{\sigma_v^2 + \sigma_i^2} \quad \text{(eq. 3.5)}$$

We conclude that MLI yields a precise prediction of how the variance in the combined condition depends on the variances in the unisensory conditions. As an aside, we mention that the same prediction (3.5) can also be attained by another statistical principle, namely that the weights w_i and $w_v$are chosen so that the variance σ_c^2 is minimal across all convex combinations.

Data analysis

Since we assume normally distributed unbiased internal representations of heading, this results in a probit-regression of the binomially distributed binary response $X_{j\alpha}$ on the condition $j = \{i, v, c\}$ and angle α, and their interaction. Here i, v and c represent the inertial, visual and combined condition, respectively. More specifically,

$$\pi_{j\alpha} = \Pr\left(X_{j\alpha} = 1\right) = \Phi\left(\frac{\tau_j - \alpha}{\sigma_j}\right) \quad \text{(eq. 3.6)}$$

where Φ denotes the cumulative standard normal distribution. The model is a dichotomous analogue to the familiar ANCOVA for continuous responses, where we regress the binary dependent variable 'response' on the independent variable 'sensory modality tested', in the presence of the continuous independent variable 'heading angle'.
We estimated the parameters $\left(\tau_j, \sigma_j\right)$ of the psychometric curve for condition j by maximum-likelihood estimation (MLE, not to be confused with MLI), assuming that all observations of a participant were stochastically independent since no feedback on

performance was provided. We found considerable inter-subject differences in psychometric curve parameters. Since the number of participants was too small to warrant a random effect specification, and we had large numbers of observations per participant, we fitted the model with six parameters (three τ_j and three σ_j) separately for each participant. A Pearson's χ^2 showed satisfactory goodness-of-fit of the psychometric curves, so that we finally fitted the model (3.6) with the MLI-induced constraint (3.5). This restricted model has five parameters (three τ_j and two σ_j), as the slope in the combined condition was predicted by MLI of the parameters of unisensory conditions. Since the τ-parameter is free to vary for each condition, a comparison of the unrestricted and restricted model is only affected by the slopes (σ) of the fitted functions. A comparison of the fit of the models with and without this constraint, using likelihood ratio tests with one degree of freedom, allowed us to test whether our data supports the MLI hypothesis.
In other work on testing MLI in cue integration (e.g., Helbig & Ernst, 2007; Fetsch et al., 2009), the standard deviations σ of the underlying Gaussians for the unisensory and multisensory conditions are often derived from the slope of fitted cumulative Gaussians (3.7).

$$slope = \sigma\sqrt{2} \quad \text{(eq. 3.7)}$$

The standard deviation of the multisensory condition is subsequently compared to the value that is predicted using the standard deviations of the unisensory conditions (3.5). Our approach essentially does not differ, but is a more direct evaluation of the MLI hypothesis that is statistically more efficient.

Results

Figure 3.2 shows the results for each individual participant. In each panel the fitted probability that a stimulus was not perceived as 'straight ahead' (SA) is plotted against heading angle for each condition (visual, inertial, and combined visual-inertial). The curves plotted through the data points are cumulative normal distribution functions. The shaded areas represent pointwise 95% confidence intervals. Table 3.2 describes the two fitted models and two associated tests for each participant. More specifically, the table lists the log-likelihood of the unrestricted model, Pearson's goodness-of-fit test for the unrestricted model, the log-likelihood of the MLI-restricted model, and a likelihood ratio test for the MLI-restriction.

Table 3.2. Model log-likelihoods and parameters. Pearson X^2 goodness-of-fit test results are presented for the unrestricted model; p-values of χ^2 are based on a chi-square distribution with 12 degrees of freedom. Significant goodness-of-fit results indicate poor model fit. p-values for likelihood ratio tests are based on a chi-square distribution with 1 degree of freedom. Thus, small p-values indicate that the MLI hypothesis on the variance has to be rejected. Significant test-results are boldfaced.

	Unrestricted model			MLI	Likelihood ratio test for MLI	
		Goodness of fit				
participant	Log-likelihood	Pearson X^2	p	LL	LR	p
1	-85.69	10.84	0.542	-86.02	0.65	0.419
2	-69.87	21.77	0.040	-74.89	10.03	**0.002**
3	-69.31	44.31	0.000	-75.14	11.67	**0.001**
4	-88.61	17.11	0.146	-93.04	8.85	**0.003**
5	-82.62	7.68	0.810	-86.05	6.87	**0.009**
6	-67.18	12.97	0.371	-79.74	25.13	**0.000**
7	-51.75	12.80	0.384	-71.71	39.92	**0.000**
8	-75.76	13.57	0.330	-84.24	16.97	**0.000**
9	-71.53	9.38	0.671	-89.58	36.10	**0.000**

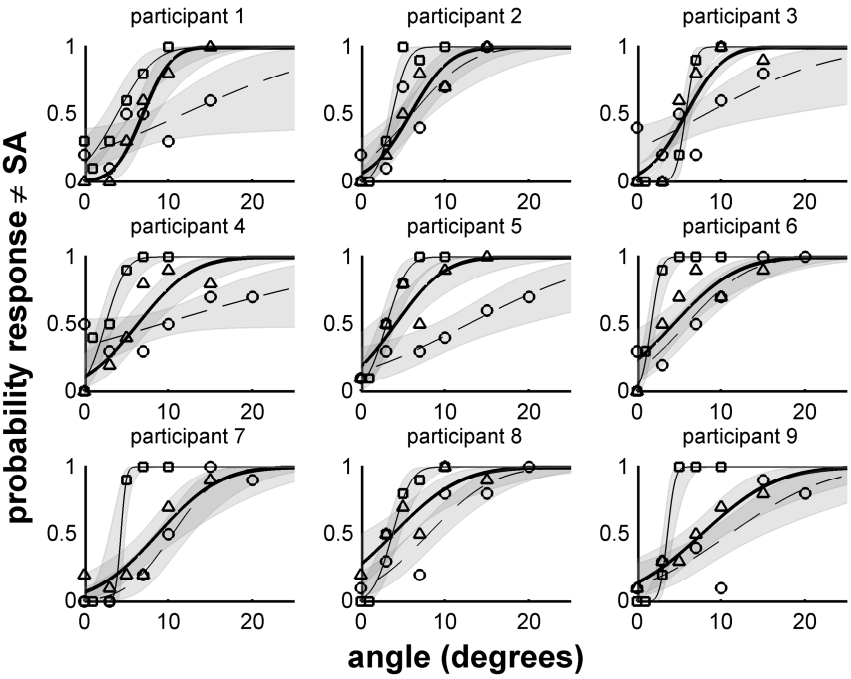

Figure 3.2. Data points and fitted psychometric functions for each participant (panels) and condition with 95% confidence intervals for the visual (squares, normal curve), inertial (circles, dashed curve), and combined (triangles, bold curve) condition.

The fit of the probit-model is based on a Pearson's X^2 test for goodness-of-fit with 180 observations equally distributed over 18 heading-angle × condition combinations. In general, the model fits the data well, and hence we may proceed below to test whether the 6 parameters of this model satisfy the MLI constraint. The violation for participant number three is caused by a poor fit of the model in the combined condition. The results reported below do not differ substantially if we include or exclude participant three.

A likelihood-ratio test showed large differences in the parameters of the fitted cumulative normal distributions (τ, σ) between participants (χ^2 (48) = 168.92, $p < 0.001$). Differences between the curves were also assessed for each participant individually using Wald χ^2 tests (Table 3.3).

Table 3.3. Statistical comparison between the parameters of the psychometric curves for the visual (V), inertial (I) and combined (C) conditions. 'ns' stands for 'not-significant'. * = p < .05.

	participant								
	1	2	3	4	5	6	7	8	9
V vs. I	*	*	*	*	*	*	*	*	ns
V vs. C	*	*	ns	*	*	*	*	*	ns
I vs. C	*	ns	*	*	*	ns	ns	ns	ns

According to the MLI hypothesis, the variance of the multisensory estimate should be lower than the variance of either unisensory estimate. We compared an unrestricted model (See: Data Analysis) to a model in which the standard deviation of the multisensory condition was constrained to the value predicted by MLI (5). The average standard deviations for the visual, inertial and combined visual-inertial conditions were: 1.07° (SD=0.68), 7.20° (SD=3.94), and 3.56° (SD=1.04), respectively. The model likelihoods and likelihood ratios, as well as a measure of the model goodness-of fit are given in Table 3.3. Associated observed and predicted standard-deviations are presented in Figure 3.3.

Combined over all participants, we have strong evidence against MLI (χ^2 (9) = 156.19, p < 0.001). For all but one participant, we have strong evidence that their heading perceptions conflict with MLI. To allow comparison with previous studies on heading perception, we also calculated the group's average 75% detection thresholds for the three conditions, using the estimated model parameters and an inverse CDF function. These values amounted to 4.6° (SD=1.3), 16.1° (SD=5.2), and 9.4° (SD=1.94) for the visual, inertial and combined visual-inertial condition, respectively. Note that these thresholds indicate the shift of a subjective judgment from 'straight ahead' to 'not straight ahead'. These thresholds should be seen as a 'Point of Subjective Straight Heading'.

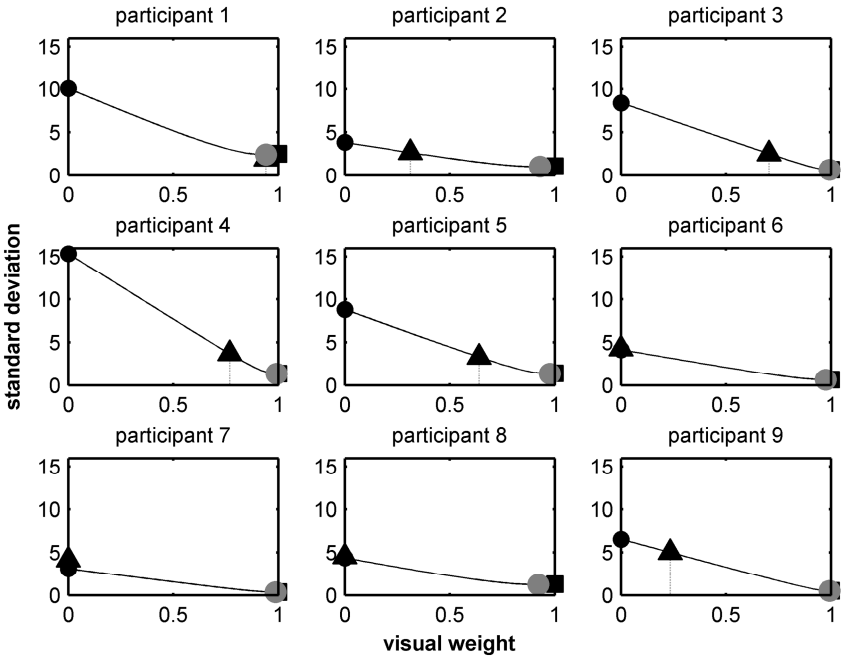

Figure 3.3. Standard deviations (SD; square root of the variance) for each participant and condition. The markers represent the observed inertial (black circle), visual (square) and combined (triangle) visual/inertial SD's, respectively. The location of the black and green dots in terms of the x-coordinate indicates the associated visual weight. Gray circles are the MLI-predicted optimal weights (x-coordinates) and corresponding SD's (y-coordinates). If MLI-theory applies, the gray and triangular marker should be close to each other.

Discussion

The results of this study showed that the 75% correct detection threshold for deviations from subjective straight ahead was much lower for visual motion than for inertial motion in all participants. The average 75% threshold was 4.6° in the visual condition and 16.1° for the inertial condition. These values correspond to earlier findings, e.g., by Telford et al. (1995). However, the present multisensory results differ from earlier observations. We found that the detection threshold in the combined condition was always larger (average 9.4°) than in the visual-only condition, while Telford and colleagues reported a near identical visual and combined visual-inertial threshold of 5.5° and 5.7°, respectively. More importantly, they reported a slight, but non-significant reduction in variance when both visual and vestibular information was available, compared to when only visual information was available. This is in line with the predictions of a Maximum-Likelihood Integration

(MLI) model. According to MLI, the variance of a combined estimate should always be smaller than the variance of the best unisensory constituent. Although Telford et al. (1995) did not explicitly test the MLI hypothesis, Gu et al. (2008) found supporting evidence for MLI in macaque monkeys. In contrast, our results indicated that the variance of the combined estimate was actually larger than that of the best unisensory estimate. Provided that our assumptions on independent and normally distributed errors in the internal representation of stimuli were met, this implies that the weight assigned to the inertial estimate was too large and that MLI has to be rejected.

Our observation does concur with other findings that inertial cues in a synthetic simulator environment are weighted more heavily relative to visual cues (Groen & Bles, 2004; Groen et al. 2001; Harris, Jenkin & Zikovitz, 2000). Furthermore, Fetsch et al. (2009) also observed that some participants tended to overweight inertial cues. Remarkably, this overweighting seemed to occur mostly for multisensory stimuli with the highest level (70%) of visual motion coherence (how many of the dots in the visual stimulus coherently moved in the same direction), In the present experiment, the dots in the visual cue moved in a fully coherent fashion (100%) and we observed overweighting of the inertial cue for every participant. Although the different experimental setup used by Fetsch et al. (2009) prohibits direct comparison of the results, it seems that the less reliable inertial information gets more weight when motion in a random dot pattern becomes more coherent.

The observed overweighting of inertial cues might be explained by a violation of the assumption that sensory estimates are unbiased. Although this assumption is standard for MLI, it has recently been shown by Todd, Christensen, and Guckes (2010) that when this assumption does not hold, a 2AFC experiment can yield biased estimates of the size of the internal noise. Using these biased estimates in an integration scheme will result in erroneous estimates of the sensory weights. Although a 2AFC paradigm, such as used in the present study, does identify sensory noise (variable error), it does not provide information on a possible bias (constant error). Therefore, we cannot validate whether sensory estimates were biased. For this purpose it would be necessary to measure heading judgments at a continuous scale.

An alternative explanation of this overweighting is that participants did not perceive the visual and inertial motion as associated. In the present study, inertial motion may have given a much stronger and compelling sensation of self-motion than the random dot pattern. It is therefore likely that the visual cue was interpreted a separate event. Although in the debriefing all participants asserted to have followed the instruction to use all sensory information available, they may instinctively have discarded the visual cue. If so, performance in the multisensory condition would correspond to performance in the

inertial condition, which was the case for five out of nine participants. As noted by Fetsch et al. (2009), this is analogous to causal inference models (Körding et al., 2007; Sato, Toyoizumi & Aihara, 2007). These integration models include a step prior to actual integration of multisensory estimates; it is first evaluated whether two cues arise from a single source or from multiple sources. In case multisensory cues are attributed to separate events, integration does not occur.

The quality of a visual stimulus itself may affect the extent to which it induces self-motion (or vection). It has been suggested that binocular visual cues are more effective in inducing vection than monocular cues, although Fetsch et al. (2009) did not find any effects of stereo vision on cue integration. However, it has been shown that a larger FoV (Allison, Howard & Zacher, 1999) and photorealistic visual cues (Trutoiu, Mohler, Schulte-Pelkum & Bülthoff, 2009) enhance vection. Compared to the stimuli used in the present experiment, the stimuli used by Gu et al. (2008), Fetsch et al. (2009) and Telford et al. (1995) were displayed with larger FoV; further, the stimulus used by Telford et al. depicted the actual surroundings, which may have increased vection. In future experiments, we plan to investigate the effects of these visual factors on cue integration.

Chapter 4. Integration of Visual and Inertial Cues in the Perception of Angular Self-Motion[1]

Summary

The brain is able to determine angular self-motion from visual, vestibular and kinesthetic information. There is compelling evidence that both humans and non-human primates integrate visual and inertial (i.e., vestibular and kinesthetic) information in a statistically optimal fashion when discriminating heading direction. In the present study, we investigated whether the brain also integrates information about angular self-motion in a similar manner. Eight participants performed a 2-Interval Forced Choice task in which they discriminated yaw-rotations (2s sinusoidal acceleration) on peak velocity. Just-noticeable differences were determined as a measure of precision in unimodal inertial-only, and visual-only trials, as well as in bimodal visual-inertial trials. The visual stimulus was a moving stripe pattern, synchronized with the inertial motion. Peak velocity of comparison stimuli was varied relative to the standard stimulus. Individual analyses showed that data of three participants showed an increase in bimodal precision, consistent with the optimal integration model; while data from the other participants did not conform to maximum likelihood integration schemes. We suggest that either the sensory cues were not perceived as congruent, that integration might be achieved with fixed weights, or that estimates of visual precision obtained from non-moving observers do not accurately reflect visual precision during self-motion.

[1] This chapter is under revision for publication as: De Winkel, K.N., Soyka, F., Barnett-Cowan, M., Bülthoff, H.H., Groen, E.L., & Werkhoven, P.J. (*under revision)* Integration of Visual and Inertial Cues in the Perception of Angular Self-Motion.

Introduction

The broad range of evidence supporting statistically optimal integration in perceptual phenomena may suggest that optimal integration is a general property of the CNS (e.g., Ernst and Bülthoff 2004; Körding 2007). Optimal integration theory (Maximum-Likelihood Integration, MLI) provides two predictions about multimodal perception. Firstly, it predicts that a multimodal percept is a weighted average of the unimodal constituent estimates, where the weight assigned to each unimodal estimate is proportional to its precision, defined as the inverse of the variability in repeated measurements under unchanged conditions. Secondly, it gives an exact prediction about the precision of a multimodal percept, given the precision of the unimodal constituent estimates.

A number of studies support the applicability of optimal integration theory to perception of self-motion. For example, it has been shown that both humans and non-human primates are able to estimate the heading of horizontal self-motion more precisely when presented with both visual and inertial (i.e., vestibular and kinesthetic) information, than when either modality is stimulated in isolation (Butler et al. 2010; Butler et al. 2011; Fetsch et al. 2009; Gu et al. 2008). Similarly, information on body tilt and acceleration are thought to be combined in a statistically optimal fashion (MacNeilage et al. 2007). Conversely, some studies report departures from optimal integration in the perception of self-motion: the integrated percept of slightly discrepant multimodal cues does not always reflect the predictions on unimodal weightings (e.g., Butler et al. 2010; Fetsch et al. 2009), and the multimodal percept is not always more precise than the best unimodal precision (De Winkel et al. 2010; Butler et al. 2011).

Jürgens and Becker (2006) investigated multimodal integration in estimates of self-rotation about an Earth-vertical yaw axis. They asked participants to actively rotate to various target angles, aided either by visual, vestibular, or kinesthetic information, or combinations of these. It was found that the targeting gain was closer to unity when more sources of sensory information were available, and that targeting precision increased. Similarly, Bakker et al. (1999) reported that targeting precision for active yaw-rotations in a virtual environment was highest when visual, vestibular, and kinesthetic information were simultaneously available.

Although the studies by Jürgens and Becker (2006) and Bakker et al. (1999) provide support for the notion that sensory information on rotatory self-motion is integrated in a statistically optimal fashion, the targeting paradigms employed in these studies are not particularly well suited to assess the *exact* prediction on improved multimodal precision provided by MLI. This is due to the fact that targeting performance is affected by factors that may introduce additional noise, such as reaction time, and voluntary body motions.

This led us to carry out a new experiment in which we directly assessed the tenability of the predictions of optimal integration theory on the precision of multimodal perceptions of rotation around an Earth-vertical yaw axis. Motion stimuli in the present study did not include constant velocities, as used in previous studies, but rather consisted of a raised cosine bell velocity profile (see Benson, 1989). This was done to comply with the fact that the semicircular canals of the vestibular system are sensitive to changes in angular velocity rather than constant velocity. We used a Two-Interval Forced-Choice (2IFC) paradigm with passive rotational stimuli. Passive rotation was used as the motor component in an active rotation paradigm may introduce additional variance in the psychophysical data and because active head movements have been shown to suppress vestibular afferent signals that drive perception (Barnett-Cowan and Harris 2011; Barnett-Cowan et al. 2012) and action (Roy & Cullen 2004). The 2IFC task was performed in three conditions: a visual-only, an inertial-only and a combined visual-inertial condition. We assessed whether precision in the combined condition increased relative to the constituent estimates, and whether the size of the increase was in accordance with the predictions of an MLI model.

Methods

Participants

Eight paid volunteers took part in this experiment (three men, five women, mean age 26.1, range: 22-36). All participants reported normal vestibular function and had normal or corrected-to-normal vision. Participants received general instructions after which they signed an informed consent form in compliance with the Declaration of Helsinki.

Apparatus

The experiment was carried out on a rotating chair at TNO in Soesterberg, the Netherlands. The chair consists of a metal framework with a seat that is mounted on a large electrical motor. As an assessment of whether the motion platform was capable of generating the maximum acceleration of 110 $°/s^2$ (needed to generate the maximum velocity motion profile used in this experiment), we compared the rotating chair's actual position over time to the computer reference signal. The motion platform was indeed capable of this acceleration. The chair was fitted with an LCD screen subtending 62° (horizontal) × 34° (vertical) visual arc, mounted 45 cm in front of the participant (for a discussion of possible effects of the use of a head-fixed display in investigations of visual-inertial integration we refer the reader to the Discussion – Reduced Visual Sensitivity section). An aperture was placed 10 cm in front of the screen, with a physical fixation cross at its center. The aperture reduced the field of view to 40° × 33° visual arc, and made

the visual stimulus appear as if it was viewed through a window. Presenting visual motion stimuli in the perceptual background increases the naturalness of visual motion (Howard and Heckmann 1989). Subjects were instructed to maintain gaze at the fixation cross. During the experiment, the apparatus was completely covered by a lightproof black cloth, and lights in the room were dimmed. This was done in order to block out any external light and view of the surroundings, and hence, to ensure that the LCD screen provided the only visual stimulation.

Rotation of the chair and playback of visual stimuli were controlled by custom made hardware and MATLAB Simulink software. The visual stimulus was synchronized to the inertial motion, i.e., the motions had equal onset, amplitude, frequency, and phase (Di Luca, 2010), and moved in the opposite direction of inertial motion. The seat was padded with foam cushions, and featured a head rest for minimizing head movements. A five point safety harness secured the participants. Participants wore headphones over which pink noise was played to mask auditory cues such as motor sounds.

Stimuli

Stimuli consisted of visual, inertial, or combined visual-inertial counterclockwise rotations around an Earth-vertical yaw axis. Motion profiles were a single period (2s, 0.5Hz) of a raised cosine bell in velocity (see Benson, 1989). Peak velocity of the motion profiles was varied between 30 and 70 °/s. Visual stimuli were video recordings of the chair's surroundings, consisting of striped wallpaper (6 × 1 m) that was suspended from the ceiling. Visual stimuli were filmed from the participant's viewpoint during playback of inertial motion profiles. Hence, visual stimuli were similar to visual motion seen during rotation of an optokinetic drum (see Figure 4.1). Despite the fact that the chair was covered by lightproof cloth, the wallpaper was kept in place during the experiment to increase immersion. This was done as participants in a previous study did not believe that visual and inertial motion originated from the same source (De Winkel et al. 2010). Videos had a resolution of 907 × 680 pixels, and were recorded at a frame rate of 25 fps. Combined visual-inertial stimuli consisted of congruent combinations of visual and inertial motion stimuli.

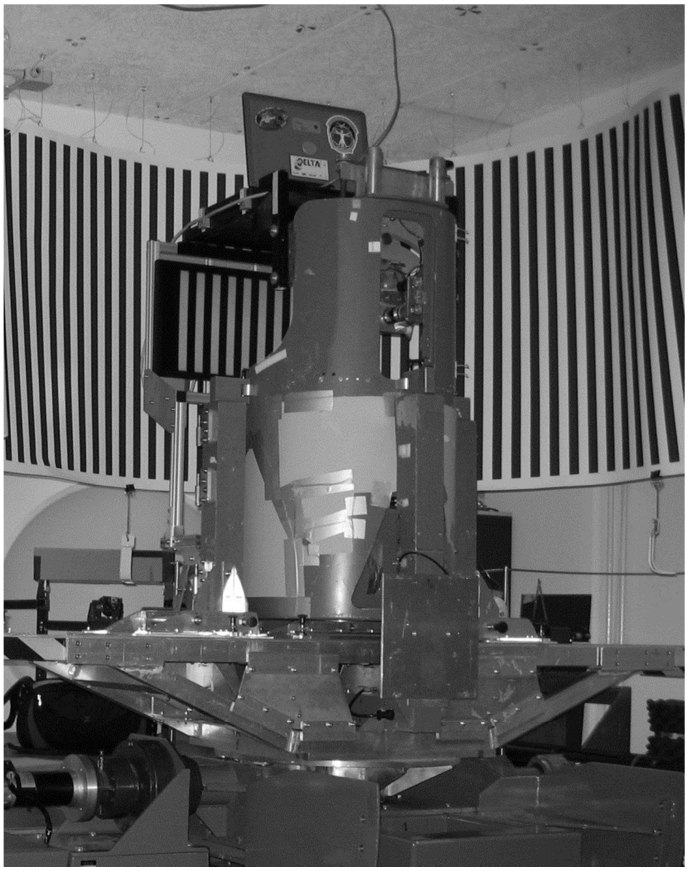

Figure 4.1. Experimental setup. Participants viewed the video stimulus on the screen through an aperture which was removed for taking this picture. During the experiment, the chair was covered with a lightproof black cloth.

According to MLI, the maximum increase of precision for multimodal stimulation is found when performance of the constituent modalities is equal (Ernst & Banks, 2002). Therefore, for each participant, performance for visual stimuli was approximately equated to performance for inertial stimuli by adjusting the visual contrast. This was done by linear interpolation of performance (JND measurements) in two visual practice runs of 50 trials each, with different contrast levels (Equation 4.1). These practice runs were run prior to the experiment proper.

$$\text{JND} = a * \text{contrast} + b \quad \text{(eq 4.1)}$$

Based on the obtained parameters a and b, a contrast value was chosen that would yield visual performance close to the performance measured in a practice run with inertial stimuli. Contrast values were manipulated by setting the contrast level in the video-player (remotely controlled VLC-player), and varied between 10 and 60%.

Procedure

Participants were instructed to maintain central fixation at all times throughout the experiment, which was intended to suppress eye movements. Participants performed a 2IFC task, in which they gave verbal judgments for which motion in a given pair had the highest peak velocity (i.e., 'first', or 'second'). Verbal responses were noted by the experimenter. A pair of stimuli (i.e., one trial) consisted of a 'standard' and a 'comparison' stimulus, presented in random order. Peak velocity of the standard stimulus was always 50 °/s, while peak velocity of the comparison stimulus was varied. There were 6 second breaks between the stimuli of a pair, as well as between the trials to prevent possible after effects (the time constants of the semi-circular canals is approx. 5 s, Fernandez & Goldberg, 1971). After each trial, participants gave their response and were moved back to the chair's initial position.

A Bayesian adaptive procedure was used to determine the next stimulus intensity based on the previously obtained data (Kontsevich & Tyler, 1999; Tanner, 2008). This was based on the assumption that the probability of a test stimulus being judged as faster than the standard stimulus follows a cumulative normal distribution (See Data Analysis). Data were gathered for three different conditions: an inertial-only condition (I), in which inertial stimuli were presented, a visual-only condition (V), in which visual stimuli were presented, and a combined condition (C), in which combined visual-inertial stimuli were presented.

For each participant, the experiment was carried out over the course of three days. On the first day, participants received instructions and performed an inertial-only practice run and two visual only practice runs. These runs were intended to familiarize them with the experiment and provided initial estimates of unimodal performance. The actual experiment was run on day two and three. Participants completed 200 trials per condition. These were divided over four blocks of 50 trials, for a total of 12 blocks. The order of blocks was arranged such that there were never two blocks of the same condition in a row. Otherwise, the order was randomized. Participants completed six blocks per day. Each block took approximately 30 minutes to complete. Participants were allowed to take breaks between blocks. Including breaks, the experiment took about 4 hours per day.

Note that although 2 IFC tasks as employed in the present experiment are to our knowledge the most suitable tasks for determination of JND's, comparisons between a standard and a comparison stimulus may not only reflect the variability of their internal

representations, but also the volatility of their memory traces. As memory capacity may be modality-specific, the particular effects of the inter-stimulus interval on the JND's may differ between modalities. However, as the interval is essential for after effects to fade out, we assume that any such effects are sufficiently countered by presenting the standard and comparison stimuli in random order.

Data Analysis

Data analysis was largely analogous to the method applied by De Winkel et al. (2010). This method assumes that people have internal estimates of angular velocity that can be represented by normally distributed, unbiased random variables. Hence, for each trial a participant has an internal estimate of the velocity of the standard as well as the comparison stimulus. The probability that a participant judges the comparison stimulus as larger than the standard stimulus increases as the maximum velocity of the comparison stimulus increases. Thus, this probability can be expressed as a monotonically increasing psychometric function of the maximum velocity of the comparison stimulus (Macmillan & Creelman 2005; Merfeld, 2011). As a psychometric function we used cumulative Gaussian distributions (Equation 4.2):

$$0.5\left[1+erf\left(\frac{x-\mu}{\sqrt{2\sigma^2}}\right)\right] \text{(eq 4.2)}$$

where erf represents the Gauss error function, x is the peak velocity of the comparison stimulus, and μ and σ are the mean and standard deviation of the internal representation of the standard stimulus, respectively. Sensory precision was defined as the σ-parameter of this function, the Just-Noticeable Difference (JND). Optimal values for the JND parameters were obtained simultaneously for all conditions by Maximum-Likelihood Estimation (MLE). In MLE, the likelihood of a model is defined as the probability of the observed data given a set of parameters (e.g., Miller and Miller, 2004).The best fitting parameters were found by minimizing the model's negative log-likelihood, using the 'fmincon' routine (MATLAB, MathWorks, MA, USA). The present procedure is known as probit regression, which is the dichotomous equivalent to the familiar ANCOVA for continuous responses: in probit regression, a dichotomous response is regressed on a continuous predictor, while in ANCOVA both the response and predictor are continuous.
Covariance matrices for the estimated parameters were obtained by taking the pseudoinverse of the Hessian matrix at the minimum negative log-likelihood (e.g., Draper & Smith, 1966). Using the covariance matrix, we can test for differences between the parameters using the 'linhyptest' routine (MATLAB, MathWorks, MA, USA). This routine

performs Wald χ^2 tests (e.g., Polit, 1996). We tested for hypothesized differences in the JNDs obtained for the three sensory modality conditions.

To assess the tenability of the MLI-model, a version of the model was fitted where the values of the slope parameter of the unimodal conditions were fixed to the values obtained from the probit regression, while the parameter of the bimodal condition was subject to the MLI induced constraint (Equation 4.3):

$$\sigma_c^2 = \frac{\sigma_v^2 \sigma_i^2}{\sigma_v^2 + \sigma_i^2} \text{ (eq. 4.3)}$$

where σ_v^2, σ_i^2, and σ_c^2 represent the visual, inertial, and combined condition respectively. We compared the fit of the unconstrained and the MLI model using Likelihood-Ratio tests.

Results

A summary of the performance in terms of JNDs is presented for each participant in Figure 4.2. Results are divided into a group of participants whose data were in agreement with the MLI model, and a group for participants whose data did not agree with model predictions.

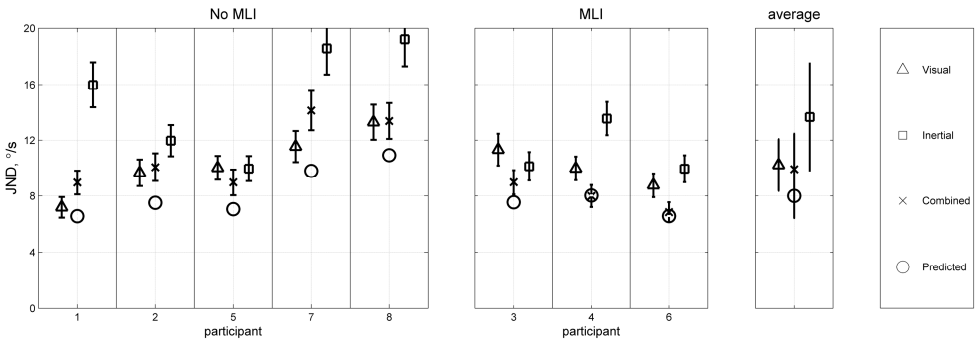

Figure 4.2. JNDs per condition, shown for each participant. The left panel shows data for participants whose results were in disagreement with the MLI model; the central panel shows data for participants whose results did agree with the model. Error bars represent standard deviation of the JNDs. The right panel shows average performances. Error bars in the right panel represent between-participant standard deviation.

As a visual illustration of the data fitting procedure, Figure 4.3 shows model fits for a participant whose data are compatible with MLI (PP 4), and for a participant whose data are not compatible with MLI (PP 7).

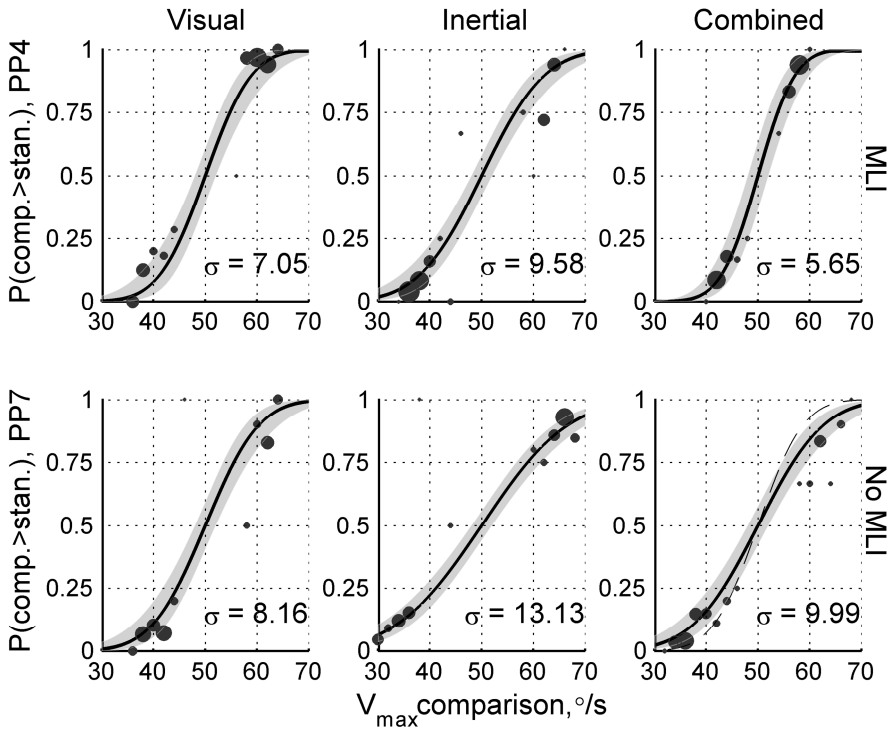

Figure 4.3. Example model fits for a participant whose data are compatible with MLI (upper row), and for a participant whose data do not agree with MLI. Note that in the former case, the fitted function has the steepest slope in the combined condition, corresponding to the smallest JND, and the predicted performance coincided with the observed performance. The dashed line in the right panel of the bottom row represents MLI predicted performance. The dots represent the individual data points, where the size of each dot is proportional to the number of repetitions of a particular stimulus. The shaded area's represent the 95% confidence intervals.

As an assessment of test-retest reliability, we tested whether there were differences between parameter estimates for the sessions completed on day 1 and day 2 of the experiment. This was done using a likelihood-ratio test between the fit of a model that allowed paramters to vary between sessions and a model that did not allow this. Although we did observe changes for three participants [1, 5, and 8], these effects were inconsistent between participants (Table 4.1). As the obtained parameters did not differ between sessions for all participants, and because the effects were inconsistent between participants for whom differences were observed, we conclude that test results did not change between sessions.

Table 4.1. Statistical comparison between JND's for the first and second session. 'LR' denotes likelihood ratio. The likelihood ratio statistic (LR) is χ^2 distributed with three degree of freedom. Boldface p-values are significant.

participant	LR	df	p
1	12.091	3	**0.007**
2	4.215	3	0.239
3	5.279	3	0.153
4	0.876	3	0.831
5	8.138	3	**0.043**
6	2.146	3	0.543
7	4.733	3	0.192
8	8.079	3	**0.044**

JNDs were compared between conditions using Wald-tests (see Methods). For five out of eight participants, the JNDs differed significantly between conditions. The participants for which conditions could be distinguished generally showed differences between the inertial JND and the other JNDs. Post-hoc tests showed that the visual JND did not differ from JND in the combined condition for any participant. Results of the individual comparisons are shown in Table 4.2.

Table 4.2. Statistical comparisons between JND for the visual (V), inertial (I), and combined (C) condition. Results are obtained by Wald-tests. Boldface p-values are significant.

	V vs. I vs. C		V vs. C		I vs. C		V vs. I	
participant	$\chi^2(2)$	p	$\chi^2(1)$	p	$\chi^2(1)$	p	$\chi^2(1)$	p
1	35.042	**0.000**	2.520	0.112	17.700	**0.000**	32.661	**0.000**
2	2.835	0.242	0.093	0.761	1.698	0.193	2.477	0.116
3	2.816	0.245	2.813	0.094	0.789	0.374	0.665	0.415
4	16.857	**0.000**	3.437	0.064	16.846	**0.000**	5.797	**0.016**
5	0.851	0.654	0.710	0.399	0.615	0.433	0.002	0.965
6	7.377	**0.025**	3.072	0.080	7.226	**0.007**	0.876	0.349
7	11.666	**0.003**	2.133	0.144	3.807	0.051	11.601	**0.001**
8	9.017	**0.011**	0.002	0.962	6.462	**0.011**	7.143	**0.008**

Applicability of the MLI model was assessed by comparing the fit of the MLI model with the fit of the unconstrained model using likelihood-ratio tests. Individual comparisons of model fit are presented in Table 4.3, showing that the fit of the MLI model did not differ from the fit of the unconstrained model for four participants (3, 4, and 6). Combined over participants the MLI model had to be rejected ($\chi^2(8)=30.859$, $p < 0.001$, See Figure 4.2).

Table 4.3. Comparisons of the fit of the unconstrained and the MLI model. The likelihood ratio statistic (LR) is χ^2 distributed with one degree of freedom. Boldface p-values are significant.

	Model Log-Likelihoods			
participant	Unconstrained	MLI	LR	p
1	-187.68	-193.98	12.595	**0.000**
2	-172.07	-177.07	10.312	**0.001**
3	-196.57	-198.38	3.613	0.057
4	-204.14	-204.14	0.003	0.960
5	-221.01	-224.25	6.480	**0.011**
6	-169.12	-169.20	0.163	0.687
7	-205.11	-213.67	17.116	**0.000**
8	-199.04	-201.50	4.902	**0.027**

Discussion

We investigated optimal integration of visual and inertial cues in the discrimination of rotational self-motion around an Earth-vertical yaw axis by comparing actual multimodal performance to the predictions of an MLI model. We found that the data agreed with the model for three out of eight participants, the responses of the other participants contradicted model predictions.

MLI during self-motion

The present study is not the first to find departures from optimal integration of visual and inertial cues to self-motion. For example, De Winkel et al. (2010) measured visual, inertial and combined visual-inertial precision for heading, and observed that the multimodal precision was in between unimodal precision while optimal integration theory predicts increased precision for multimodal perception. The findings were attributed to a lack of marked vection, i.e., visually induced sensation of self-motion. A lack of vection may cause the visually perceived motion to be interpreted as object-motion instead of self-motion. Consequently the visually and inertially perceived motions are considered incongruent information, prohibiting multimodal integration.

Other studies have reported a relative over-reliance on the inertial cue. For example, Butler et al. (2010) also investigated optimal integration theory in the perception of heading. While they did observe increased precision for multimodal stimulation, it was found that when discrepancies between the visual and inertial stimuli were introduced, the multimodal average was closer to the inertial estimate than predicted. This result suggests an overweighting of the inertial cue compared to the predicted inertial weight. Similar observations for heading estimation were reported Fetsch et al. (2009), for both humans and macaque monkeys, and Jürgens and Becker (2011) also reported a similar inertial bias for yaw-rotations. Below, we consider several alternative explanations for the deviations from optimality observed in the present study.

Visual capture

Vision has been shown to dominate multisensory perceptions on various tasks. This phenomenon is known as ‘Visual Capture’ (Rock & Victor, 1964), and might be due to the fact that visual perception is generally more precise than perception based on other senses (Ernst & Banks, 2002). Because precision in the visual condition was generally larger than in the inertial condition, and precision in the combined condition did not differ from precision in the visual condition for any of the participants (Table 4.1), one might posit visual capture as a possible explanation of the present results. However, the data for

participants three, four, and six show a trend towards an increased precision in the combined condition. Additionally, the MLI model did fit the data for these participants. For participants one and seven, precision in the combined condition was worse than in the visual-only condition. In general, there was no consistent trend for all participants. Therefore, we cannot conclude that visual capture explains the results, even if the average data suggests this explanation.

Switching strategy

One consideration concerning the variability in the applicability of MLI to self-motion perception is that, for integration of visual and inertial cues to take place, one would expect that the visual stimulus should effectively induce vection, which is a sensation of self-motion induced by motion of the visual surround, and has been shown to depend on various factors, in particular field-of-view (Dichgans & Brandt, 1978), and perceived depth (Howard & Heckmann, 1989).

Despite the fact that visual stimuli were synchronized video recordings of the actual surroundings of the chair, presenting them on a synthetic display with inherently limited spatial resolution (or field-of-view) may still have been insufficient to induce strong vection. As a consequence, participants may not have interpreted the visual and inertial stimuli as congruent, despite explicit instructions that they were. Notably, during the debriefing one participant mentioned that it was difficult to "monitor two motions at once". When cues were not considered as congruent, participants may have applied a strategy in which they relied solely on a single source of sensory information, or alternatingly based their responses on either the visual or the inertial cue on a trial by trial basis. In terms of modeling, such a 'switching' strategy states that there is a certain probability that one sensory estimate will dominate the percept on a trial-by-trial basis (Equation 4.4).

$$\sigma_C = \pi_V \sigma_V + (1 - \pi_V)\sigma_I \quad \text{(eq 4.4)}$$

where π_V represents the probability that the response is based upon the visual cue, and σ_V, σ_I, and σ_C represent the visual, inertial and combined precisions, respectively. Note that $0 \leq \pi_V \leq 1$, and $\sigma_i > 0$, for $i = \{\mathrm{V}, \mathrm{I}, \mathrm{C}\}$. This model can account for dominance of either sensory modality, as well as intermediate findings. Although the fit of the switching strategy model did not differ from the fit of the unconstrained model (see Table 4.4), it also is not able to account for the trend of increased bimodal precision observed for some participants.

Table 4.4. Comparison of the unconstrained model to a switching strategy-model.

Participant	Model Log-Likelihoods		LR	p
	Unconstrained	Switching strategy		
1	-187,.684	-187.684	0.000	1.000
2	-172.067	-172.067	0.000	1.000
3	-196.574	-196.961	0.775	0.379
4	-204.139	-205.568	2.859	0.091
5	-221.009	-221.314	0.611	0.434
6	-169.118	-170.631	3.025	0.082
7	-205.115	-205.115	0.000	1.000
8	-199.045	-199.045	0.000	1.000
overall	-1.554.750	-1.558.385	7.270	0.508

Fixed weights

Several studies investigating visual-inertial integration for self-motion perception found that the inertial cue was weighted more than expected from model predictions (Butler et al., 2010; Fetsch et al., 2009; Jürgens & Becker, 2011). The average visual weight found by Butler et al. (2010) when investigating heading motions for four different conflict levels was approximately 0.35 (see Figure 6 in Butler et al., 2010). Using the formula for the weighted combination of the variances of *n* independent random variables (Equation 4.5; c.f. Miller & Miller, 2004):

$$\sigma_c^2 = \sum_{i=1}^{n} \omega_i^2 \, \sigma_i^2 \text{ (eq 4.5)}$$

Here, ω_i refers to the weight and σ_i to the variance for modality i, with $i = \{\text{V,I}\}$. Assuming complete integration (i.e., the weights sum to unity), we can compute the expected precision (the inverse of the variance) for the combined condition from the unimodal precision estimates (Equation 4.6):

$$\sigma_C^2 = \alpha^2 \sigma_V^2 + (1-\alpha)^2 \sigma_I^2 \text{ (eq 4.6)}$$

Where α is the visual weight. Fitting the 'fixed weights' model of Equation 6, using the weights derived from the Butler et al. study yields a model fit that does not differ from the unconstrained model (Table 4.5). Thus, integration might be achieved with fixed weights, rather than optimal integration where the weights depend on the precision of the

constituent estimates. Moreover, the weights may be independent of the particular task, and vary between individuals; reflecting individual 'perceptive styles' (De Vrijer et al., 2010; Jürgens & Becker, 2011; Lambrey & Berthoz, 2003).

Table 4.5. Comparison of the unconstrained model to a model with fixed weights.

	Model Log-Likelihoods			
Participant	Unconstrained	Fixed weights	LR	p
1	-187.684	-188.552	1.736	0.188
2	-172.067	-172.979	1.825	0.401
3	-196.574	-197.385	1.624	0.654
4	-204.139	-205.061	1.843	0.765
5	-221.009	-222.322	2.626	0.757
6	-169.118	-169.176	0.115	1.000
7	-205.115	-205.422	0.616	0.999
8	-199.045	-199.045	0.001	1.000

Reduced visual precision

A final alternative explanation is based on the finding that visual sensitivity worsens during self-motion, as illustrated by the freezing illusion: when a slowly moving pattern is presented on a display that is moving itself, the pattern seems to freeze, regardless whether the observer is moving with the screen (Pavard & Berthoz 1977; Probst, Brandt & Degner, 1986; Shirai & Ichihara, 2012; Wertheim & Reymond, 2007). If visual sensitivity is indeed reduced during self-motion, this could result in reduced visual precision. Therefore, it might not be admissible to use precision estimates of the visual signal measured in stationary observers in order to calculate the MLI predictions. Accounting for the possibility of decreased visual precision during self-motion by adding a constant offset (4.89 °/s) to the measured visual-only JNDs results in predictions of the visual-vestibular precision that are in agreement with our measurements (Table 4.6). This value for the constant (4.89 °/s) was chosen such that the mean absolute error (over all participants) between measured and predicted JNDs for the combined condition was minimal (0.54°/s). The JND for the combined condition was calculated as follows (Equation 4.7):

$$\sigma_C^2 = \frac{(\sigma_V + c)^2 \sigma_I^2}{(\sigma_V + c)^2 + \sigma_I^2} \quad \text{(eq 4.7)}$$

With *c* equal to 4.89 °/s.

Changes in visual sensitivity could account for deviations from optimality that would otherwise be considered visual capture, or vestibular dominance because the calculated

weights would be based on estimates of the unimodal precision that do not reflect their actual magnitudes during self-motion.

Table 4.6. Comparison of the unconstrained model to a model with reduced visual precision during self-motion. Boldface p-values are significant.

Participant	Model Log-Likelihoods		LR	p
	Unconstrained	Red. vis. precision		
1	-187.684	-188.792	2.217	0.137
2	-172.067	-172.116	0.098	0.755
3	-196.574	-196.580	0.013	0.909
4	-204.139	-206.909	5.540	**0.019**
5	-221.009	-221.075	0.133	0.716
6	-169.118	-170.408	2.580	0.108
7	-205.115	-205.344	0.459	0.498
8	-199.045	-199.106	0.123	0.725
overall	-1.554.750	-1.560.331	11.162	0.193

Aubert-Fleischl Phenomenon

Despite the fact that our participants were explicitly instructed to maintain fixation during the course of the experiment, we did not measure eye-movements to confirm this. Although we do not believe that eye movement in itself hinders integration, as our eyes move continuously without any notable effect on the perception of self-motion, it should be acknowledged that visual stimuli of the same speed seem to be moving slower when pursued with the eyes, than when moving in front of stationary eyes; the Aubert-Fleischl phenomenon (Dichgans et al., 1975). Therefore, future experiments using eye tracking are required to assess the effects of eye-movements on visual precision.

Conclusion

The main goal of this study was to test whether multimodal precision in the perception of angular self-motion around an Earth-vertical yaw axis matched with predictions from an MLI model. It was found that MLI predictions could account for the results of three out of eight participants. Overall, the MLI model had to be rejected. As discussed above, previous research investigating multi-sensory integration between visual and inertial cues also reported deviations from the MLI predictions. Possible causes for the deviations were discussed, and it was shown that our data could be explained both by violation of the assumption that visual and inertial motion are interpreted as sharing a common cause (i.e., incongruence) or by adopting fixed weights reported in previous literature.

Additionally, our findings may be explained by a decrease in visual sensitivity during self-motion. Further research is needed to distinguish between the proposed models to better understand how visual and inertial cues are combined.

Chapter 5. Integration of Otolith and Semicircular Canal Cues in the Perception of Lateral Body Tilt.[1]

Summary

In this study we investigated whether the perception of lateral self-tilt based on sensory information from the otoliths and semi-circular canals (SCC) is consistent with statistically optimal Maximum Likelihood Integration (MLI). According to MLI, estimates of self-tilt based on information from both sensory modalities (i.e., the otoliths and the SCC) should be more precise than estimates based on only one modality. Since it is known that the tilt percept induced by stimulation of the otoliths and SCC shows more or less opposite frequency dynamics, we expected that this would also be reflected in the precision of self-tilt estimates. We hypothesized that the precision of SCC-induced estimates of self-tilt would increase with increasing stimulus frequency, whereas otolith-induced perception of self-tilt would be more precise at low-frequency motions.

Six subjects performed a two-interval forced choice task in which they compared the perceived maximum tilt angle induced by a reference stimulus to that induced by a comparison stimulus. The amplitude of the comparison stimulus was varied on a trial-by-trial basis. Motion stimuli comprised three periods of a sine function, and were presented at two frequencies: 0.075 Hz and 0.250 Hz. Isolated otolith stimulation in the subject's fronto-parallel plane (sway) was achieved by a linear track for the 0.250 Hz condition, and eccentric centrifugation for the 0.075 Hz condition, respectively. Stimulation of the SCC alone was achieved by rotating the subjects around an Earth-vertical naso-occipital (roll) axis. Simultaneous stimulation of otoliths and SCC was achieved by rotation around an Earth-horizontal roll-axis. Cumulative normal distributions were fitted to the dichotomous responses. The slope parameter of the model was used as measure of precision.

The results showed a significant interaction between sensory modality and stimulus frequency: the precision of tilt-estimates in the SCC-only condition was more precise at 0.250 Hz than at 0.075 Hz, while in the otolith-only condition, low-frequency estimates

[1] This chapter has been submitted for publication as: De Winkel, K.N., Correia Grácio, B.J., Groen, E.L., & Werkhoven, P.J. (submitted) Integration of Otolith and Semicircular Canal Cues in the Perception of Lateral Body Tilt.

were more precise. For five out of six subjects, the precision of bimodal tilt estimates agreed with the MLI model's prediction.

We conclude that the observed improvement in precision for bimodal tilt-estimates as compared to unimodal estimates supports the view that the CNS combines multisensory information on self-motion in a statistically optimal fashion. Furthermore, the opposite effects of frequency for otolith- and SCC-induced tilt-estimates suggest that the sensory dynamics should be taken into account in models that address the integration of multisensory information on self-motion.

Introduction

The vestibular system in the inner ear is important to our perception of spatial orientation and self-motion. It consists of two types of sensory organs: the semi-circular canals (SCC), which respond to angular accelerations; and the otolith organs, which are responsive to linear accelerations and gravity (i.e., the Gravito-Inertial Acceleration). By combining the information provided by these organs, the central nervous system (CNS) can construct an internal representation of our spatial orientation (e.g., Angelaki & Cullen, 2008). This is not trivial: because of the equivalence of gravity and inertial accelerations (Einstein, 1907), tilting the head relative to gravity and horizontal linear acceleration can cause physically equivalent stimulation. Since otolith primary afferents carry ambiguous signals (Loe et al. 1973; Fernandez & Goldberg, 1976; Anderson et al., 1978; Young et al., 1984; Dickman et al., 1991; Si et al., 1997), in the absence of vision the CNS discriminates between tilt and translation by combining the ambiguous otolith information with simultaneous information on rotation provided by the SCC (Guedry, 1974; Mayne, 1974; Young, 1974; Angelaki et al., 1999). However, at low frequency rotations the SCC response deteriorates, as the angular accelerations are too weak to excite the sensory neurons. Moreover, SCC information is absent during constant tilt relative to gravity (e.g., lying supine). Because stationary tilt does not induce illusions of constant linear acceleration, it is assumed that the CNS also segregates gravitational from translational accelerations signaled by the otoliths, by using knowledge on the constancy of the gravity and the transient nature of inertial accelerations. In other words: low-frequency changes in the orientation of the otolith signal are interpreted as due to changes in the orientation of the head relative to gravity, while high-frequency changes are interpreted as translation (Mayne, 1974; Paige & Tomko, 1991; Telford et al., 1997; Bos & Bles, 2002). Hence, in the absence of other sensory information, the perceived self-tilt is derived from otolith information for low-frequency rotations of the head relative to gravity, and from SCC information for high-frequency rotations.

The range of frequencies for which otolith and SCC information induce perceptions of self-tilt are not mutually exclusive, but partially overlap (De Graaf et al., 1996, 1998). Therefore, these sensors can be considered to provide redundant estimates of self-tilt for this range of frequencies. In line with findings of statistically optimal integration of redundant visual and vestibular information on self-motion (e.g., MacNeilage et al., 2007; Fetsch et al., 2009; Butler et al., 2010; Butler et al., 2011), Laurens and Droulez (2007) proposed a probabilistic model for the combination of otolith and SCC information into a unified multisensory percept, based on Bayesian inference. In this model internal representations of sensory estimates are represented as probability distributions.

Redundant sensory estimates are weighted according to their precision (i.e., inverse of the variance), summed, and subsequently multiplied by so-called 'priors', representing top-down processing of sensory information. In the model proposed by Laurens and Droulez, the priors caused a bias to perceive low-frequency linear accelerations as head tilt, and low-frequency angular rotations as an immobile head.

Although the model proposed by Laurens and Droulez successfully reproduced several perceptual effects that occur during centrifugation and off-vertical axis rotation, the model parameters were chosen by the experimenters rather than derived from actual data. Moreover, the model did not include a precision parameter for the otoliths. The purpose of the present experiment was to collect actual data on the precision of otolith- and SCC-induced perception of body tilt, and to validate whether integration of otolith and SCC information by the CNS conforms to the predictions of probabilistic integration frameworks. Assuming normally distributed internal representations of motion variables and uniform priors, the Bayesian framework is identical to Maximum-Likelihood Integration (MLI). MLI predicts that the integrated percept is more precise than the percept based on either constituent estimate in isolation. Moreover, it provides an exact prediction about the size of the improvement. We measured precision of roll-tilt estimates in response to stimulation of the otoliths and SCC in isolation, as well as for combined stimulation of both sensory modalities. We assessed whether the integrated percept was indeed more precise and whether the size of the precision agrees with the MLI model's predictions.

As described above, the dynamics of SCC- and otolith-based perception of tilt depend on motion frequency in an approximately opposite fashion (e.g., Mayne, 1974; Howard, 1982; Paige & Tomko, 1991; De Graaf et al., 1996, 1998; Seidman et al., 1998; Bos & Bles, 2002; Kaptein & Van Gisbergen, 2005). Hence, we hypothesized that their precision would also be affected by motion frequency. We measured precision of tilt-estimates in a low- and a high-frequency condition (i.e., 0.075 and 0.250 Hz). Based on the low-pass filter characteristics of the otolith-induced tilt perception (i.e., "somatogravic illusion" Clarke & Graybiel, 1949) and the high-pass filter characteristics of the SCC, we expected that the precision data of both sensory modalities would show opposite effects; we expected that the otolith estimate of roll-tilt would be less precise with increasing frequency, while the opposite was expected for the SCC.

Methods

Subjects

Seven paid volunteers participated in this study (three male, four female; mean age 32.3; range [22-55]). One subject was excluded from the analysis due to motion sickness. Another subject was only able to complete the 0.075Hz conditions. All subjects had normal or corrected-to-normal vision, and reported no history of intestinal, neurological, or vestibular illness. Three subjects (3, 5, and 6) had prior experience with the apparatus. Subjects received general instructions after which they signed an informed consent form in compliance with the Declaration of Helsinki.

Apparatus

The experiment was run in the DESDEMONA motion facility at TNO (Bles & Groen, 2009). DESDEMONA is a centrifuge-based motion simulator, which features a motion platform with six degrees of freedom. In this experiment we used the central yaw axis, the linear track, and the cabin's gimbal system.

Task

Subjects performed a 2-Interval Forced Choice task, in which they judged which stimulus of a pair induced the larger perception of peak tilt angle (i.e., the 'first', or 'second'). A pair of stimuli (i.e., a trial) consisted of a 'reference' and a 'comparison' stimulus, presented in random order. Peak tilt angle of the reference stimulus was kept constant, while peak tilt angle of the comparison stimulus was varied (see 'Stimuli' below). Responses were given using a button on the left and a button on the right-hand side of the simulator's control yoke, corresponding to the responses 'first' or 'second', respectively.

Stimuli

The task was performed in three sensory modality conditions: an SCC-only condition, in which only the SCC were stimulated; an otoliths-only condition, in which only the otolith system was stimulated; and a combined otoliths + SCC condition, in which both sensory modalities were stimulated simultaneously.

For all conditions, motion profiles consisted of three periods of a sine function. The amplitude of the first period gradually increased, while the amplitude of the last period gradually decreased (i.e., fade-in and fade-out, Figure 5.1).

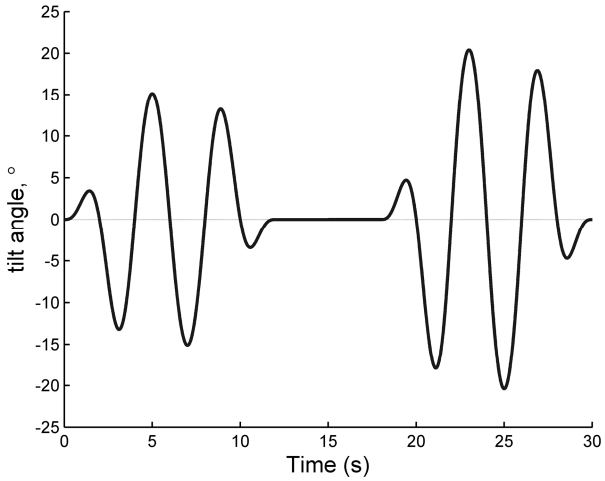

Figure 5.1. Tilt angle as a function of time for a high-frequency trial. The flat line between the two oscillations is the break between the two stimuli of a trial. In the low-frequency condition, the shape of the motion profile was equal, but was completed in about 90s.

Stimuli were presented at two different frequencies: 0.075Hz and 0.250Hz. These frequencies were chosen such that a tilt percept would be induced by isolated stimulation of both the otoliths and the canals, but that their relative sensitivities would differ (Merfeld et al., 2001; Bos & Bles, 2002; Raphan & Cohen, 2002; Correia Grácio et al., 2012). Figure 2 shows the frequency dynamics of the SCC and otoliths in terms of perceived self-tilt. Note that the otolith-induced self-tilt is based on the somatogravic illusion. The amplitude of the reference stimulus was always 15.1°. The amplitude of the comparison stimulus ranged between [9.5 – 20.4]°, in 50 equally spaced steps. Stimulus maximum angular velocity ranged between [4.5 – 9.6]°/s for the 0.075Hz conditions, and [14.9 – 32.0]°/s for the 0.250Hz conditions, which is above the detection threshold (Benson, Hutt, & Brown, 1989).

For the SCC-only conditions, the simulator cabin was tilted backwards to orient the subjects in a supine position so that the roll rotation around the subject's naso-occipital axis was aligned with the earth-vertical axis. In this way, subjects experienced roll-motion while the orientation of gravity relative to the subject was kept constant (Figure 5.3; upper right panel).

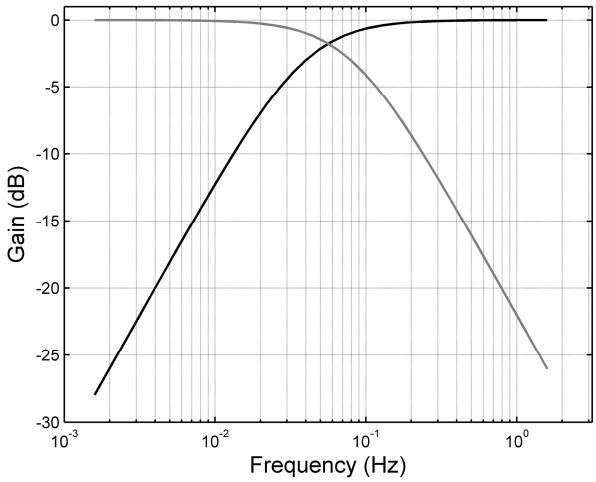

Figure 5.2. Transfer functions for SCC (black) and otolith (gray) in terms of perceived self-tilt. The SCC transfer function is a high-pass filter with a time constant of 4s (Raphan & Cohen, 2002); the otolith transfer function is a low-pass filter with a time constant of 2s (Correia Grácio et al., 2012).

In the otoliths-only conditions, we used the somatogravic illusion to induce tilt perceptions. For the two different frequencies tested, the illusory tilt was achieved in different ways. For the 0.250 Hz stimuli, the cabin was translated laterally along the simulator's linear track. The combined acceleration of the cabin and the acceleration of gravity results in a tilt of the Gravito-Inertial acceleration, which subjects interpret as body tilt (Figure 5.3; lower left panel). It was not possible to apply the same approach for the 0.075Hz stimuli, because the linear track was not long enough to accommodate the distance the cabin would need to translate. Therefore, for the 0.075 Hz stimulus, the linear acceleration was generated by centrifugation, rotating the simulator cabin and linear track around the central yaw-axis, until a constant angular velocity of 1.26 rad/s was attained. While rotating at constant velocity, the cabin was translated eccentrically along the simulator's linear track. This generates a centripetal acceleration that increases with cabin eccentricity (Figure 5.3; lower right panel). The acceleration attained in both frequency conditions was varied between 1.65 and 3.65 m/s2, corresponding to the tilt angles subjects experienced in the SCC-only and combined conditions. Although this centrifuge paradigm does generate Coriolis forces that may provide subjects with false cues, these false cues were considered negligible for the low frequency motions in a pilot study (also see: Merfeld et al., 2005). However, the paradigm was not applied in the 0.250 Hz condition because at this frequency the Coriolis false cues became noticeable.

In the bimodal condition, subjects were seated upright in the simulator cabin, and tilted such that they experienced rotations around their naso-occipital axis (Figure 5.3; upper left panel).

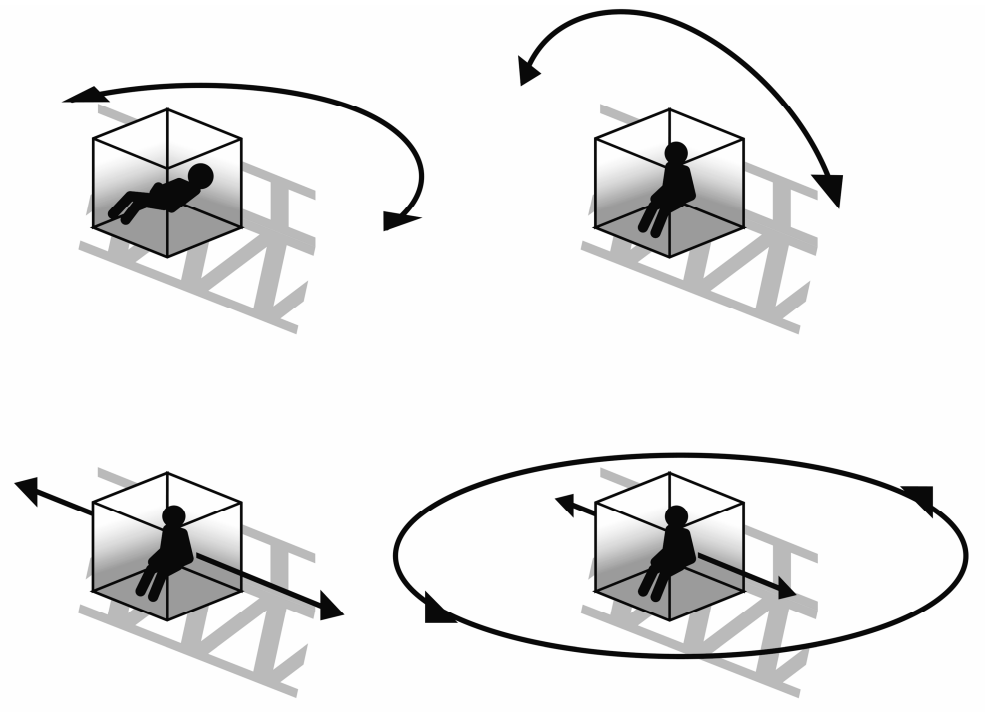

Figure 5.3. Schematic representation of simulator motion for different experimental conditions. Upper left: combined stimulation of otoliths and SCC. Upper right: isolated stimulation of the SCC. Bottom row: isolated stimulation of otoliths, for the high- (left) and low (right) frequency conditions.

Procedure

Subjects were seated inside the simulator cabin and secured by a five-point safety harness. The lights in the simulator cabin were dimmed and subjects were instructed to keep their eyes closed. Pink noise was played through the headset to mask simulator sounds. Stimuli were presented in 12 blocks of 25 trials, separated by 15 minute breaks. Per block, only stimuli of a single sensory modality (3) × motion frequency (2) condition were presented. There were 6 second breaks between the stimuli of a pair, as well as between each pair of stimuli. Subjects were notified at the end of a trial by an auditory 'beep'. They initiated the next trial themselves via button presses. Including breaks, the

experiment lasted 12 hours, divided over two days. On the first day, 'low-frequency' stimuli were presented; on the second day 'high-frequency' stimuli. Otherwise, the order of blocks and the order of stimuli within a block was randomized.

Data Analysis

Data analysis is largely analogous to the method applied in De Winkel et al. (2010). We performed a probit-regression of the binary response variable $X_{j,f,\alpha}$ on categorical predictors 'sensory modality' j, 'frequency' f, and continuous predictor 'tilt angle' α:

$$\pi_{j,f,\alpha} = Pr\left(X_{j,f,\alpha} = 1\right) = \Phi\left(\frac{\mu_{j,f} - \alpha}{\sigma_{j,f}}\right) \text{ (eq 5.1)}$$

where Φ represents the cumulative normal distribution (Equation 5.1). This means that for each sensory modality condition j and frequency f, the probability $\pi_{j,f,\alpha}$ that a comparison stimulus is judged larger than the reference stimulus ($X_{j,f,\alpha} = 1$) is given by a cumulative normal distribution with parameters $\mu_{j,f}$ and $\sigma_{j,f}$ evaluated at angle α. The slope of the cumulative normal distribution $\sigma_{j,f}$ is proportional to the standard deviation of the underlying internal tilt estimate by a factor √2 (e.g., Ernst & Banks, 2002; De Winkel et al., 2010). Therefore, we interpret the slope parameter as a measure of precision. We fitted the model for each subject individually. Note that this model is a dichotomous analogue to the ANCOVA model.

The twelve parameters of the model – six $\mu_{j,f}$ and six $\sigma_{j,f}$ (frequency (2) * sensory modality (3)) were estimated by Maximum- Likelihood Estimation (MLE), that is, minimizing the model's overall negative log-likelihoods, using the Matlab 'fmincon' routine. This model is referred to as the unrestricted model. Hypotheses on the parameters were tested using the 'linhyptest' routine in Matlab, which performs Wald χ^2-tests.

Given normally distributed, independent internal representations of tilt angle, and normal or uniform distributed priors, Optimal Integration theory states that the reciprocal of the variance of the integrated estimate is equal to the sum of the reciprocals of the constituent signals' variances. When no prior distributions are involved this is known as Maximum-Likelihood Integration (MLI) (e.g., Ernst & Banks, 2002; De Winkel et al., 2010). This yields the following constraint for the precision in the combined condition:

$$\frac{1}{\sigma^2_{scc+oto}} = \frac{1}{\sigma^2_{scc}} + \frac{1}{\sigma^2_{oto}} \text{ (eq 5.2)}$$

where ‘scc’, ‘oto’, and ‘scc+oto’ represent the SCC-only, otoliths-only and combined sensory modality conditions, respectively. To test the applicability of the MLI model, we also fitted a version of the model (Equation 5.1), in which precision in the combined condition was subject to the MLI-induced constraint (Equation 5.2). We compared the fit of these models to the unrestricted model using likelihood-ratio tests.

Results

Parameter estimates

For each subject, we estimated the precision of tilt estimates for each sensory modality condition and frequency (Table 5.1). Example model fits for a single subject are presented in Figure 5.4. For subject number 3, SCC-estimates are not available for the 0.075 Hz condition because the subject was unable to discriminate stimuli in that condition. During the debriefing this subject also noted that in the high frequency otolith-only condition he mainly felt translation, as opposed to roll-tilt, and that he had discriminated stimuli based on the perceived lateral acceleration. Because the number of subjects was small, the data was not discarded.

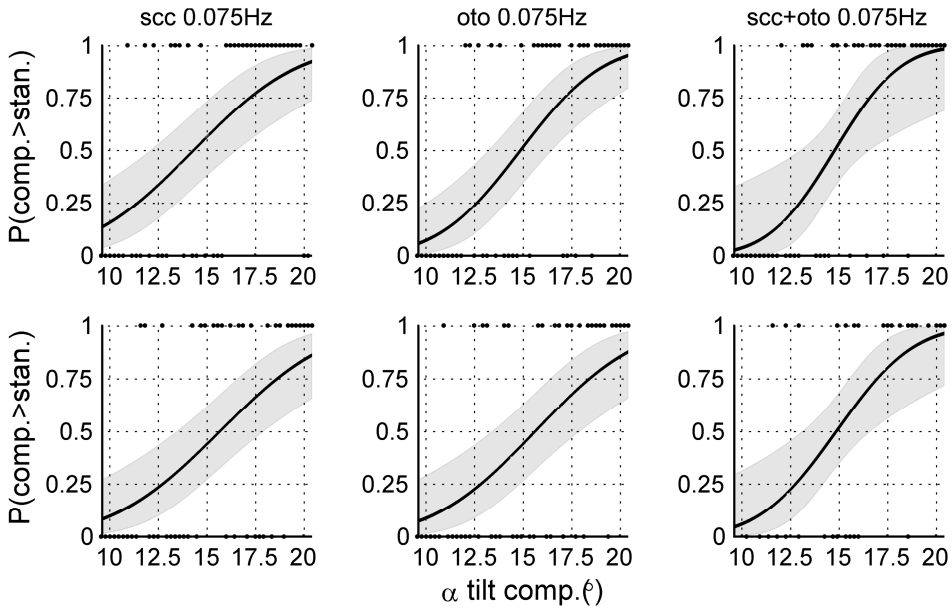

Figure 5.4. Example model fits for subject number 1. See text for explanation of acronyms. Dots represent individual data points. Shaded area marks the 95% CI.

Table 5.1. Parameter σ estimates for each condition, per subject.

	f = 0.075 Hz			f = 0.250 Hz		
Subject	scc	oto	scc+oto	scc	oto	scc+oto
1	8,739	2,627	2,808	3,891	8,696	4,033
2	9,560	4,225	5,599	6,574	3,448	3,540
3	n/a	2,883	2,584	7,304	3,019	5,631
4	18,241	6,815	6,185		n/a	
5	4,278	3,339	2,668	4,371	4,144	3,114
6	4,485	1,408	1,635	4,740	4,259	2,004
median	8,739	3,111	2,738	4,740	4,144	3,540

Effects of sensory modality and motion frequency

X^2 tests were performed to assess the main effects of sensory modality and stimulus frequency. The main effect of frequency was inconsistent among subjects, as for two subjects higher precision was observed (i.e., smaller value for slope parameter) for the low-frequency condition, while overall, precision was higher in the high frequency condition. The main effect of modality showed that the otoliths provided more precise tilt estimates than did the SCC. This effect was consistent among subjects (Table 5.2).

Table 5.2. Assessment of main effects of frequency and sensory modality on discrimination thresholds. Significant effects are boldfaced.

	frequency			modality		
Subject	χ^2	df	P	χ^2	df	p
1	8,823	3	**0,032**	0,834	2	0,659
2	1,136	3	0,768	6,321	2	**0,042**
3	5,038	3	0,169	14,425	2	**0,001**
4		n/a		0,683	2	0,711
5	0,478	3	0,924	0,756	2	0,685
6	14,820	3	**0,002**	9,052	2	**0,011**
Overall	30,296	15	**0,011**	32,071	12	**0,001**

To assess our hypothesis on the effect of sensor dynamics on the precision of tilt estimates, we tested for an interaction effect between frequency and sensory modality on the slope parameters of the 'scc' and oto condition. On individual level, the effect was significant for three out of five subjects (one subject did not complete any high-frequency condition). Analysis of the group results also revealed a significant interaction effect ($X^2(5)=59{,}538$, $p<0.001$) (Table 5.3). Post-hoc analyses did not show differences between the slope parameters for high vs. low-frequency SCC-only precisions ($X^2(5)=6.984$,

p=0.222), but did show a reduction of high-frequency otoliths-only precision ($X^2(5)=22.141$, $p<0.001$).

Table 5.3. Assessment of interaction effect of frequency and sensory modality on discrimination thresholds. Significant effects are boldfaced.

	frequency × modality		
Subject	χ^2	df	p
1	0,910	1	0,340
2	12,593	1	**0,000**
3	28,642	1	**0,000**
4		n/a	
5	0,967	1	0,325
6	16,425	1	**0,000**
overall	59,538	5	**0,000**

Optimal Integration hypothesis

To assess the tenability of the Optimal Integration model (MLI), we compared the MLI model log-likelihood to the log-likelihood of the unrestricted model, using likelihood-ratio tests. The results are shown in Table 5.4. It can be seen that the MLI hypothesis holds for five out of six subjects. A significant difference indicates a significantly worse fit of the MLI model than the unrestricted model. Hence, p-values > 0.05 can be interpreted as supportive of the MLI hypothesis.

Table 5.4. Model Log-Likelihoods (LL) for the unrestricted (unres) and the Maximum-Likelihood Integration (MLI) model, and likelihood-ratio (LR) test results. Significant differences between model-fits are boldfaced.

	LL			
Subject	unres	MLI	LR	p
1	-154,101	-154,504	0,806	0,668
2	-152,553	-153,584	2,061	0,357
3	-158,646	-161,731	6,171	**0,046**
4	-46,798	-46,801	0,006	0,997
5	-149,017	-149,038	0,042	0,979
6	-96,998	-98,770	3,545	0,170
overall	-758,113	-764,428	12,631	0,396

Note. The unrestricted model has 12 parameters, while the MLI model has 10 - therefore df=2.

Discussion

In line with reports on statistically optimal integration of visual and inertial information on self-motion, the findings of this study provide the first experimental evidence that the human brain integrates information from the otoliths and SCC in a statistically optimal fashion. According to optimal integration theory, multimodal estimates should always be more precise than each of the constituent unimodal estimates (Ernst & Banks, 2002). Our results agree with this hypothesis, as model predictions agreed with observations for five out of six subjects, and overall multimodal precision was better than unimodal precisions. For the subject for whom the model did not fit the data, the anomalies may have been caused by a failure to induce (illusory) perception of tilt for high-frequency motion in the otolith-only condition. During debriefing, this subject remarked to discriminate high-frequency otolith-only stimuli on the basis of perceived lateral translation rather than tilt.

Most current models of integration of otolith and SCC are deterministic (e.g., Bos & Bles, 2002); they do not consider sensory noise, and although theoretical models have been proposed earlier which do take sensory noise into account (Laurens & Droulez, 2006; Clemens et al. 2011), these models have not been validated with actual experimental data on precision. This may be partly due to the difficulty of collecting data on otolith-induced self-tilt. Here we used the unique motion capabilities of the Desdemona simulator to stimulate the otoliths without congruent stimulation of the canals. This allowed us to discriminate between the contribution of the otoliths and canals to the perception of tilt, and measure the precision of both systems in isolation.

The precision of otolith-induced estimates of self-tilt was about 3.1° and 4.1° for the low and high frequency conditions, respectively; SCC precisions amounted to 8.7° for the low, and 4.7° for the high frequency condition. Clemens et al. (2011) inferred the size of otolith precision using an inverse estimation approach on psychophysical data on subjective body tilt. Their average estimate of otolith precision was 2.4° for upright observers, which increased as a function of body tilt at a rate of 0.16° per degree of body tilt. Given the tilt angle of our reference stimulus (15.1°), their data suggest a precision of about 4.8°. This value is in agreement with the present observations. De Winkel et al. (2012) estimated the precision of perceived yaw rotations based on stimulation of the horizontal SCC, and found an average precision of about 14°, slightly larger than the SCC precision observed in the present study. The difference may be explained differences in stimulus frequency and maximum angular velocity.

The analysis of motion frequency seems to support our hypothesis that the sensory dynamics affect the precision of the SCC and otoliths in a different way. As expected, we found an interaction effect between sensory modality and stimulus frequency. More

specifically, the precision of otolith-based tilt perception decreased with increasing motion frequency, while the precision of SCC-based tilt perception increased with increasing motion frequency. This result implies that studies dealing with multi-sensory integration in the perception on self-motion should take the effect of frequency dynamics on the precision of the contributing systems into account.

Chapter 6. The perception of verticality in Lunar and Martian gravity conditions[1]

Summary

Although the mechanisms of neural adaptation to weightlessness and re-adaptation to Earth-gravity have received a lot of attention since the first human space flight, there is as yet little knowledge about how spatial orientation is affected by partial gravity, such as Lunar gravity of 0.16 g or Martian gravity of 0.38 g. Up to now twelve astronauts have spent a cumulated time of approximately 80 hours on the Lunar surface, but no psychophysical experiments were conducted to investigate their perception of verticality. We investigated how the Subjective Vertical (SV) was affected by reduced gravity levels during the first European Parabolic Flight Campaign of Partial Gravity. In normal and hypergravity, subjects accurately aligned their SV with the gravitational vertical. However, when gravity was below a certain threshold, subjects aligned their SV with their body longitudinal axis. The value of the threshold varied considerably between subjects, ranging from 0.03 to 0.57 g. Despite the small number of subjects, there was a significant positive correlation of the threshold with subject age, which calls for further investigation.

[1] This chapter has been published as: De Winkel, K.N., Clément, G., Groen, E.L., & Werkhoven, P.J. (2012) The perception of verticality in lunar and Martian gravity conditions. Neuroscience Letters 529(1):7-11

Introduction

Accurate spatial orientation is fundamental to a wide range of human behaviors, such as locomotion through the environment (Dyde, Jenkin & Harris, 2006; Howard, 1982; Lackner, 1976) and the recognition of objects (Lackner, 1976; Rock & Heimer, 1957). In estimating our orientation relative to the environment, the brain combines inertial, visual, and idiotropic information. The body's inertial sensors include the vestibular otolith organs in the inner ear (Howard, 1982), and somatosensory sensors such as muscle spindles and pressure sensors in the skin (Lackner & Graybiel, 1978a; 1978b), and specialized graviceptor units in the abdominal organs (Mittelstaedt, 1996). All these sensors signal the directional pull of gravity on internal organs, muscles, joints, and the skin. Visual information about our orientation consists of intrinsic and extrinsic polarity cues. Intrinsic polarity cues define "up" by their constant orientation relative to gravity, e.g., trees and buildings. Extrinsic polarity cues define "up" by their relation to intrinsic polarity cues, e.g., a book lying on a shelf (Howard, Bergström & Ohmi, 1990). In addition to these inertial and visual cues, it is assumed that the brain also uses prior knowledge, based on a lifelong experience and invariables in our orientation (Mittelstaedt, 1983). For example, the notion that our head is almost always "up" and our feet are "down" would explain that when subjects are asked to align a luminous rod in an otherwise dark room with their SV, they tend to make errors towards their longitudinal body axis when they are tilted relative to gravity (Dyde et al., 2006). This tendency has led to postulate a so-called idiotropic vector that attracts the SV (Mittelstaedt, 1996). Hence, the SV can be considered a compromise between inertial, visual, and idiotropic estimates of orientation (Groen, Jenkin & Howard, 2002).

On Earth, the pull of gravity[A] provides an omnipresent reference for verticality, and hence for spatial orientation. When this gravitational reference is absent, as in free-fall during orbital flight, humans can be spatially disoriented, to the point of nausea (Lackner, 1992; Oman, 2007; Young et al., 1984). In weightlessness, the SV is an unstable composite of visual and idiotropic information (Clément & Reschke, 2008; Young et al., 1984). Even when visual cues are available, disorientation is a common phenomenon, as illustrated by the inversion illusion, in which a subject suddenly feels upside down (Graybiel & Kellogg, 1967), and the visual reorientation illusion, in which the SV suddenly realigns with certain features of the visual surroundings (Oman, 2007). The latter illusion typically occurs when an astronaut observes a colleague floating upside-down relative to him or herself. When visual cues are also absent, some subjects completely lose their spatial orientation (Lackner, 1992; Young et al., 1984), while others may sense that the vertical is aligned with their idiotropic vector (Glasauer & Mittelstaedt, 1992). Experiments conducted in

parabolic flight have shown that participants no longer experience the inversion and visual reorientation illusions when the gravity level during the pull-out phase was between 0.1 and 0.2 g (Lackner, 1992).
In the present study, we investigated how spatial orientation is affected by the level of gravity. Subjects aligned a visual rod with their SV during different phases of parabolic flight. To disambiguate between the contributions of the body's inertial sensors and the idiotropic vector, observers were lying on their right side. We hypothesized that the error in the orientation of the SV relative to the true vertical increases with decreasing gravity level.

Material and methods

The experiment was performed during the first Joint European Partial Gravity Parabolic Flight campaign (JEPPF) organized jointly by the European Space Agency (ESA), Deutsches Zentrum für Luft- und Raumfahrt (DLR), and Centre National d'Etudes Spatiales (CNES), onboard the Novespace Airbus A300 Zero-G aircraft, in Bordeaux, France. This aircraft flew parabolic trajectories during which periods of reduced gravity were attained for periods of about 20-25 s. Six males, aged 26-55, participated in this study. One subject was excluded from analysis because no data was gathered in the 0 g phase. All subjects passed a JAR FCL3 Class II medical examination prior to be enrolled. The JAR FCL3 Class II medical certificate is developed to test whether an individual meets the health demands for a private pilot license (i.e., 6/6 visual acuity, good hearing). None of the subjects had a history of vestibular or neurological disorders. No motion sickness medication was administered. Three subjects (EG, GC, and BG) had previously experienced microgravity (0g) in parabolic flight. None of the subjects had previous experience in partial gravity (0.16-0.38g). The study was approved by the Institutional Review Board of the Principal Investigator and all subjects gave their written informed consent before participating.

Task & Stimuli

Subjects aligned a visual rod with their subjective vertical (SV). This test was based on the classic Rod-And-Frame test (Witkin & Asch, 1948). The test was performed under various gravitational magnitudes. To disambiguate between the contributions of the inertial sensors and the idiotropic vector to the SV, observers were lying on their side (Figure 6.1). Xsens™ accelerometers were used to log the gravity level. The rod was presented using an eMagin Z800 Head-Mounted Display (HMD). Resolution of the display was 800*600 pixels, subtending 40° diagonally. Stimuli were presented binocularly, non-stereo. The HMD was fixed relative to the head using a customized bite board. The rod consisted of four white dots and one orange dot indicating "down". For each trial, the rod had a random initial

orientation. Subjects adjusted the rod orientation using a trackball finger mouse until they were confident it reflected their SV.

The rod was presented against one of six visual background conditions: a city scene that was tilted relative to the gravitational upright at [-112.5, -60, 0, +60, +112.5]°, or a black background, which served as a control condition. Visual conditions were presented in random order.

A trial was confirmed via a button press, after which a new trial was automatically initiated. The HMD turned black for 0.5 s between trials.

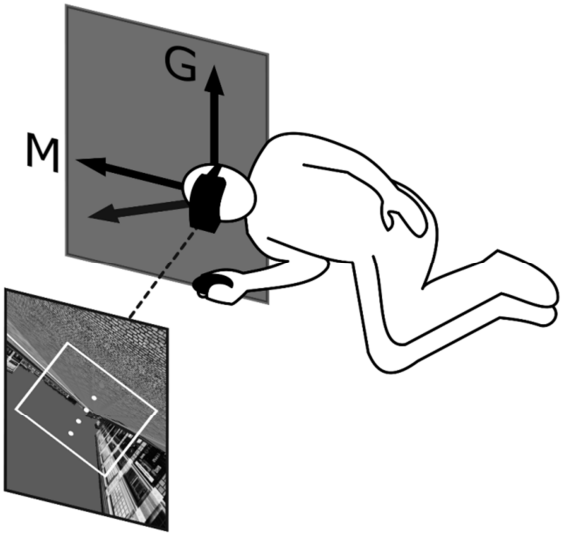

Figure 6.1. Schematic representation of the experimental setup. 'G' represents the gravity vector; 'M' the idiotropic vector (Mittelstaedt, 1983)

Procedure

Subjects were tested on three flights of 31 parabolas each. For each flight, the first 13 parabolas generated lunar gravity (approx. 0.16 g) separated by hypergravity of 1.2-1.5 g. The following 12 parabolas generated Martian gravity (approx. 0.38 g) separated by hypergravity of 1.5-1.8 g. The final six parabolas generated microgravity (approx. 0 g) separated by hypergravity of 1.8 g. In between parabolas were periods of level flight lasting 1 to 5 min. Two subjects were tested per flight. Data collection proceeded

continuously from the onset of each parabola to shortly after pull-out. Subjects were allowed breaks between parabolas. Subjects completed an average of 337.6 trials (standard deviation = 158.7), i.e., 10.9 trials per minute.

Results

The effects of the experimental manipulations on SV settings were assessed by performing an ANCOVA analysis with visual condition as a categorical predictor and the level of gravity as continuous predictor. The analysis showed that SV settings were not affected by visual condition (F(5, 30.13)=1.75, p=0.15), while there was a significant effect of gravity (F(1,1843.00=978.80, p<0.001). However, as visual inspection of Figure 2 shows, the SV settings are of a somewhat dichotomous nature: at lower levels of gravity, the data are scattered around the longitudinal body axis(response of 0°), whereas at higher levels of gravity data are scattered around the gravitational vertical (response of 90°). The observed transition of SV settings from alignment around the LBA to alignment around the gravitational vertical suggests that the LBA is used as reference for orientation at zero gravity, and that gravity must exceed a certain threshold before it is recognized as a reference to orientation. In general, thresholds can be estimated by fitting a cumulative normal distribution to dichotomous data. However, because of the observed scatter around the principal axes, we considered this approach inappropriate here. Instead, we determined the recognition threshold by applying a mixture model approach (Titterington, Smith & Makov, 1985).

Mixture model

The mixture model assumes that an SV response is either aligned with the body axis or with gravity. In the absence of gravity, observers are predisposed to use the internal representation of the Longitudinal Body Axis (LBA) as reference to orientation. In the model the LBA is represented by an angle α^*_{LBA}. Only when the magnitude m^*_F of the sensed Gravito-Inertial Force vector F^*exceeds a threshold value τ, its orientation α^*_F is interpreted as internal representation of gravity α^*_G.

Because perception is inherently noisy (Green & Swets, 1988), we model α^*_G and α^*_{LBA} as stochastic variables with a mean μ_i and normal distributed noise σ_i: $\alpha^*_G \sim \mathrm{N}\left(\mu_{\alpha^*_G}, \sigma_{\alpha^*_G}\right)$ and $\alpha^*_{LBA} \sim \mathrm{N}\left(\mu_{\alpha^*_{LBA}}, \sigma_{\alpha^*_{LBA}}\right)$, respectively. We describe the prior probability that responses are based on the representation of gravity as a cumulative normal distribution function of the magnitude of the sensed Gravito-Inertial Force m^*_F.

The above assumptions imply that the probability of a response y_i, given that it is based upon representation α^*_G or $\alpha^*_{LBA}$is

$$\Pr\left(y_i|\mu_{\alpha_G^*},\sigma_{\alpha_G^*}\right)=\varphi\left(\frac{y_i-\mu_{\alpha_G^*}}{\sigma_{\alpha_G^*}}\right),\ \text{or}\ \Pr\left(y_i|\mu_{\alpha_{LBA}^*},\sigma_{\alpha_{LBA}^*}\right)=\varphi\left(\frac{y_i-\mu_{\alpha_{LBA}^*}}{\sigma_{\alpha_{LBA}^*}}\right)\ \text{(eq 6.1)},$$

respectively, where φ represents a normal distribution function.

Furthermore, the prior probabilities that a response y_i originates from α_G^*, $\Pr(y_i\epsilon\alpha_G^*)$, or α_{LBA}^*, $\Pr(y_i\epsilon\alpha_{LBA}^*)$, given m_F^*, are given by:

$$\Pr(y_i\epsilon\alpha_G^*|\tau,\sigma_\tau)=\Phi\left(\frac{m_F^*-\tau}{\sigma_\tau}\right),\ \text{and}\ \Pr(y_i\epsilon\alpha_{LBA}^*|\tau,\sigma_\tau)=1-\Phi\left(\frac{m_F^*-\tau}{\sigma_\tau}\right)\ \text{(eq 6.2)}$$

where Φ is a cumulative normal distribution with mean τ and slope σ_τ.

We obtained optimal estimates of model parameters $\left[\tau,\sigma_\tau,\mu_{\alpha_G^*},\sigma_{\alpha_G^*},\mu_{\alpha_{LBA}^*},\sigma_{\alpha_{LBA}^*}\right]$ by maximizing the log-likelihood function LL for the set of responses $Y=y_1,\ldots,y_n$:

$$LL\left(\tau,\sigma_\tau,\mu_{\alpha_G^*},\sigma_{\alpha_G^*},\mu_{\alpha_{LBA}^*},\sigma_{\alpha_{LBA}^*}|Y\right)$$
$$=\sum_{i=1}^{n}\log\left[\Phi\left(\frac{m_F^*-\tau}{\sigma_\tau}\right)\times\varphi\left(\frac{y_i-\mu_{\alpha_G^*}}{\sigma_{\alpha_G^*}}\right)+\left(1-\Phi\left(\frac{m_F^*-\tau}{\sigma_\tau}\right)\right)\right.$$
$$\left.\times\varphi\left(\frac{y_i-\mu_{\alpha_{LBA}^*}}{\sigma_{\alpha_{LBA}^*}}\right)\right]\ \text{(eq 6.3)}$$

Because of the absence of a visual effect, data of the different visual conditions were pooled for this analysis. Individual data and model fits are presented in Figure 6.2. Estimates of all model parameters are presented in Table 6.1, where we focus on the recognition threshold τ. The threshold values vary considerably between subjects.

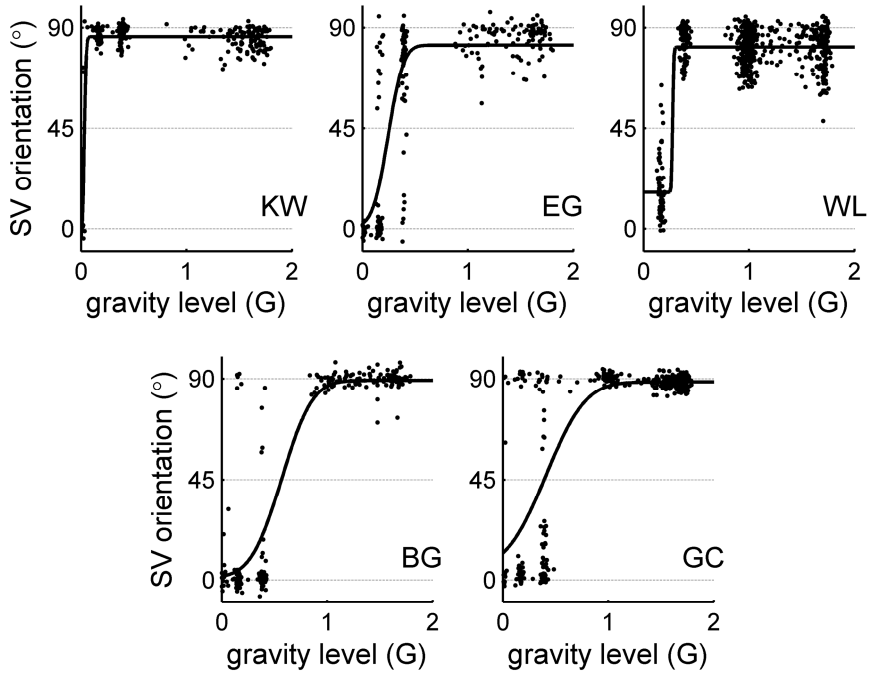

Figure 6.2. Subjective vertical (SV) settings for each subject during parabolic flight. The abscissa represents the gravity level. The ordinate represents orientation of SV estimates.

Table 6.1. Mixture model parameter estimates.

pp	τ(g)	στ(g)	μLBA(°)	σLBA(°)	μG(°)	σG(°)
KW	0,030	0,015	-6,766	7,800	86,036	4,315
EG	0,244	0,115	1,020	4,124	82,245	10,329
WL	0,276	0,009	16,384	13,871	81,400	8,255
BG	0,567	0,230	1,162	5,171	89,296	5,436
GC	0,407	0,297	4,932	6,839	88,653	4,230

As a possible explanation for the large threshold differences between subjects, we investigated whether the subject's age could play a role by means of a regression analysis. Because the uncertainty of each threshold estimate was known (error bars of each data point in Figure 6.3), we performed a random-effects meta regression instead of a regular regression. Whereas in regular regression there is a single error parameter, in meta regression the error of the model is assumed to be composed of a constant and the variance of each individual threshold estimate. As a consequence, data points with larger

uncertainty are weighted less in meta regression. Figure 3 shows the threshold estimates for each subject against the subjects' age. There was a positive correlation between the recognition threshold and age (R^2=0.65). The slope parameter of the regression model indicated an increase in threshold of 0.0114g per year ($\chi^2(1)$=5.2267, p=0.022). The function fitted to the data was: τ=-0.1731+0.0114×age.

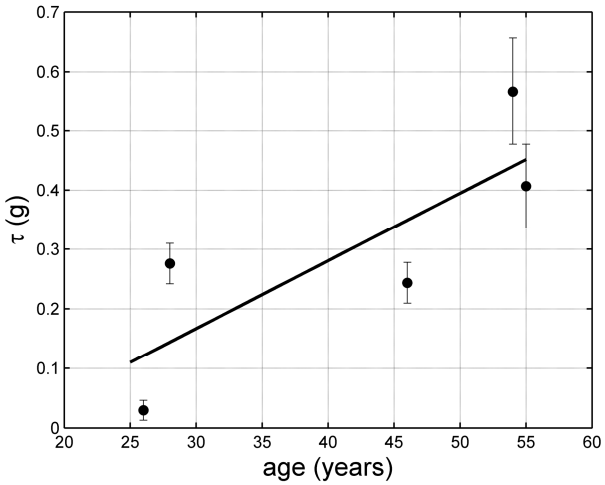

Figure 6.3. Gravity threshold estimates plotted as a function of subjects' age. Error bars represent each estimate's 95% confidence interval.

Discussion

Our results indicate that the gravity level must exceed a certain threshold to be recognized as reference for verticality. This is interesting when one realizes that a reference for verticality is of primary importance for the perception of body tilt. Three previous studies investigated the perception of body tilt during linear accelerations created by sled or short radius centrifuges in orbit (Arrott, Young & Merfeld, 1990; Benson et al., 1997; Clément et al., 2001). Subjects exposed to peak linear acceleration of 0.2 g along the interaural axis on a sled did not experience a sense of lateral tilt (Arrott et al., 1990). Eccentric rotation generating 0.22 g along the longitudinal axis at the head during 60 s of constant velocity was also insufficient to provide a vertical reference (Benson et al., 1997). However, subjects exposed to greater acceleration levels (0.5 g and 1.0 g) for up to 5 min in a centrifuge perceived lateral tilt when the resultant force was directed along the interaural axis, and inversion when the resultant force was directed towards the head (Clément et al., 2001). Thus, these studies indicated that the threshold for gravitational verticality was

somewhere between 0.22 g and 0.5 g. We found an average recognition threshold of 0.30 g, suggesting that already at the level of lunar gravity humans may lose their perception of verticality based on inertial sensory information.

Indeed, astronauts from the Apollo crews admitted to feel a little “rocky” when they emerged to walk onto the lunar surface (Paloski et al., 2008). Particularly interesting are the reports following the Apollo 14 mission, where the Lunar Module landed on a 7° slope. The crew did not feel the tilt when they were standing in the cockpit, but reported this tilt to be “very disconcerting during the sleep period” (Mitchell, Apollo 14; Godwin, 2000). They got up and looked out of the windows to check if the module had tilted. With the module powered down, they resorted to hanging a piece of string to determine the module’s tilt angle(Godwin, 2000). Also, while driving the Apollo 16 Lunar Rover, the pitch meter broke off and the crew reported that they “had no sensation of climbing a steep slope… The thing could have gradually increased to a 25° slope and I don’t think we would have realized it” (Duke, Apollo 16; Godwin, 2002). “I tell you one thing, we wouldn’t have gotten out and worked on a 20° slope. You just can’t handle it. Although, I think maybe we did when we were standing in that crater. We might have been on a 20° slope” (Young, Apollo 16; Godwin, 2002). “I think so, too” (Duke, Apollo 16). “At station 4, we had the feeling that the Rover was just going to slide off down the hill like it did on Apollo 15” (Duke, Apollo 16). “So we backed around and parked in a crater” (Young, Apollo 16).

While on Earth the threshold for the perception of lateral sway ranges from about 0.01 to 0.12 g (Fitzpatrick & McCloskey, 1994; Gundry, 1978; Janssen et al., 2011), corresponding to a body tilt of about 0.6-6.8°, one may intuitively think that in lunar gravity more body tilt is required to achieve the same gravitational stimulation in sway (3.6-48.6°). However, our results show that lunar gravity may be below the threshold for verticality, so that it does not provide any gravitational reference for body tilt, irrespective of the steepness of the slope one is on. In other words, the threshold for recognizing the gravitational acceleration as reference for verticality exceeds the sensory detection threshold. This merits the conclusion that the perception of verticality results from central processing, where the brain takes various sources of information into account. The large variance between our subjects suggests a large subjective component, or internal bias.

A recent parabolic flight study by Harris, Jenkin, and Dyde (2012) also investigated human orientation under lunar gravity. Subjects performed a discrimination task in which they judged whether an ambiguous character (a rotated ‘p’) presented against a tilted visual background was either a ‘p’ or a ‘d’ (OCHART-task, see: Dyde et al., 2006). The task was performed in lunar gravity as well as normal and hypergravity. It was found that the relative weighting of visual information in the determination of the subjective direction of “up” did not change for lunar gravity compared to normal and hypergravity, while in a

previous experiment conducted in microgravity (Dyde et al., 2009) the visual weighting was reduced. It was concluded that the threshold for verticality should therefore lie somewhere between lunar gravity and microgravity. This would imply that lunar gravity provides an adequate reference for verticality, whereas our results show that, on average, lunar gravity is below the threshold for verticality. It is interesting to speculate what may have caused this apparent difference in thresholds.

A first possible explanation is that the different thresholds are due to inter-subject variability, considering the large differences observed in the present study.

Second, the studies used different tasks. The OCHART task used by Harris, Jenkin, and Dyde was developed to assess the perceived direction of "up" under special circumstances, such as microgravity. The OCHART and SV-task have been shown to yield different results, where the OCHART estimate of "up" is more dependent on the long body axis, while the SV showed a stronger reliance on gravity (Dyde et al., 2006).

Although one may argue that the Rod and Frame test becomes useless in 0g without a physical reference for "up", we selected this test because the focus of the parabolic flight campaign was on partial gravity (consisting of 25 parabolas at partial gravity and six parabolas at 0g). We hypothesized that the partial gravity levels would still provide a useful reference for "up", so that the R&F test would be appropriate. Yet we asked subjects to continue the task in the 0g phases of flight, assuming that they still may have some internal concept of verticality, e.g., determined by the idiotropic vector. It is therefore not too remarkable that the SV settings of all subjects were aligned with the idiotropic vector at 0g. However, it is remarkable that for most subjects this was also the case at 0.16g, and for some subjects already at 0.38g.

The absence of an effect of the visual frame on SV settings was unexpected, considering the robustness of the effect in the literature. Because the effect is normally in the order of a few degrees, we think that it may have been obscured by additional noise generated by parabolic flight. A pilot study on the ground, using the same equipment and visual stimuli but upright subjects, did produce a significant average error of 2.5° at a visual angle of about 45°.

Finally, our results suggest that the threshold for verticality increases with age. In itself, such would be in line with observations of disequilibrium with aging individuals, and reductions of sensory hair cell counts for people over 30 (Richter, 1980). Although the number of subjects tested in our experiment is too small to substantiate an age effect, the result does warrant further investigation.

Chapter 7. Discussion

The main hypothesis of this thesis was that the neural integration of information on spatial orientation and self-motion as provided by the visual system, the semicircular canals, and the otoliths, can be predicted by a Bayesian framework. We assumed the purest model of optimal integration are and assumed that information provided by different sensors would be completely integrated, yielding a statistically optimal reduction of variance, consistent with the predictions of Maximum-Likelihood Integration model (MLI; Chapter 2). In our experiments we observed such statistically optimal integration in some cases, but not in others. In the following, we will review the main findings of the various experiments to determine which stimulus characteristics did and that did not yield optimal integration. Then, we will discuss the validity of our assertions.

Main findings

In Chapter 3, we investigated the integration of visual and otolith cues in the perception of linear motion in a horizontal plane. We used large field-of-view and rather abstract visual stimuli that resembled motion through a dark environment filled with white spheres (i.e., a starfield). The inertial stimuli consisted of physical translations. In this study, the data of only one subject were consistent with (optimal) integration. Five of the other eight subjects showed an alternating reliance on either visual or inertial cues, varying on a trial-to-trial basis. The remaining three subjects or showed a complete reliance on the inertial cue. Obviously, conditions for optimal integration were not fully met. We assume that this overreliance on inertial cues indicates that the visual stimulus was not sufficient to induce a sensation of self-motion.

In Chapter 4, we studied integration of visual and semicircular canal cues in the perception of rotational motion in a horizontal plane. As the results of the previous study suggested that the visual stimuli were not immersive enough, we decided to use video recordings of motion through realistic surroundings as visual stimuli in this study. The results showed a reduction of variance consistent with MLI in the multisensory condition for three subjects. For the other five subjects, it was observed that variance in the multisensory condition was either in between unisensory variances, or agreed with variance of the visual condition. For these subjects, the visual stimulus may not have been sufficient to induce a sensation of self-motion, thereby prohibiting integration. The relative overreliance on visual information may have been due to its increased realism, making the discrimination task easier to perform based on visual information.

In Chapter 5, we assessed integration of otoliths and semicircular canal cues in the perception of rotational motion in a vertical plane. During lateral body tilt, the

semicircular canals of the vestibular system contribute to our percept of rotation by signaling angular acceleration to the CNS. Simultaneously, the otoliths contribute to the percept of rotation by detecting a rotation of gravity relative to the head. As illustrated in Chapter 1, the tilt perceptions induced by the semicircular canals and otoliths have different dynamics. Therefore, we varied motion frequency. We hypothesized that the CNS integrates semicircular canal and otoliths information on tilt in a statistically optimal fashion. The results agreed with model predictions, as the variance in the multisensory condition was smaller than in either unisensory condition, for both motion frequencies tested.

In Chapter 6, we investigated integration of visual and otoliths cues in the perception of spatial orientation in a vertical plane. The level of gravity was varied as a manipulation of the realism of the inertial cue. The visual stimuli were computer-generated city scenes, presented using a head-mounted display. The results of the study showed that the perception of body-orientation was dominated by gravity, as long as its level exceeded an individual's recognition threshold. As the level of gravity dropped below this threshold value, the percept of orientation was more likely to be dominated by a-priori knowledge that the direction of gravity usually aligns with our long body axis. Integration of visual and otolith information was not observed. We assume that the absence of a visual effect was due to the small field-of-view of the head-mounted display.

Implications of findings for the tenability of the Bayesian framework

As the review of the main results of this thesis shows, we found variable support for the hypothesis that integration of visual, semicircular canal, and otolith information on spatial orientation and self-motion can be predicted using a Bayesian framework: in the cases where model predictions applied, the observed size of the reduction of variance was statistically optimal, suggesting complete integration consistent with MLI, which is a special case of the Bayesian framework (see Chapter 2). In other cases, no integration was observed at all; showing a reliance on either constituent source of information, sometimes visual, sometimes inertial. Sub-optimal reductions of variance, consistent with a partial coupling between the systems, were not evident. In the following sections we will discuss the possible factors affecting integration.

Stimulus realism

The visual stimuli used in the studies described in chapters 3, 4, and 6 varied in level of realism. Coincidently, the incidence of optimal integration also varied. The visual stimuli used in Chapter 3 were of an abstract nature, as they resembled motion through a star field, presented with a large field-of-view, and integration was observed only for a single subject. The visual stimuli used in Chapter 6 were naturalistic, as they resembled a city scene. However, they were presented with a small field of view, and integration was not observed. The visual stimuli used in Chapter 4 were realistic, as they were video recordings of motion of the actual surroundings presented on a display with field-of-view about equal to the visual stimuli used in Chapter 3, and integration was observed for three out of eight subjects. Although none of the studies provided full support for integration, it is likely that the level of realism and field-of-view may play a role (Allison, Howard & Zacher, 1999; Trutoiu, Mohler, Schulte-Pelkum & Bülthoff, 2009), that is, sensory integration requires a sufficiently large field-of-view, as well as a sufficiently realistic visual scene. In Chapter 6, we varied the realism of the gravitational input to the otoliths, that is, the strength of gravitational input relative to the gravitation experienced in natural conditions on Earth. This had a significant impact on the perception of body orientation. Hence, we conclude that realism of the gravitational stimulus also affects the incidence of integration when the otoliths are involved.

Congruence between stimuli

As noted in Chapter 1, multisensory stimuli must occur within a certain temporal and spatial window in order to be perceived as originating from the same event (e.g., Spence & Squire, 2003). With respect to perception of self-motion and body orientation, guidelines for these constraints may be obtained from studies on visual-inertial coherence zones. Coherence zones are defined as “the range of inertial motion amplitudes that, though not being a physical match to the visual cues, are still perceived by subjects as coherent.” (Van Der Steen, 1998; Valente Pais, 2010). In the perception of self-motion, the spatial constraints may be defined in terms of the degree of correspondence between the amplitude of visual and inertial acceleration, velocity or distance travelled (Valente Pais et al., 2010; 2010). Temporal constraints may be defined in terms of phase mismatch (Beckers et al., 2012; Jonik et al., 2011). With respect to the perception of body orientation, the coherence zone may be defined in terms of the degree of correspondence between orientation information provided by the visual and various inertial senses. A ‘violation’ of the coherence zone for body orientation may explain the dichotomous nature of the perceived orientation that was observed in the study presented in Chapter 6: if subjects had been positioned such that the angle of their longitudinal body axis

relative to gravity was smaller, both would have been interpreted as indicative of spatial orientation, and integration would have occurred. This is supported by a study by Groen, Jenkin, and Howard (2002), showing that the visual and body axis were weighted more strongly relative to the inertial cue in conditions where they were aligned in comparison to conditions where they were misaligned. This suggests that congruence is a necessary condition for optimal integration.

Sensory dynamics

As discussed in Chapter 1, the visual system, semicircular canals, and the otoliths have different dynamics, as their output depends differently on motion frequency. Because this factor was not considered in earlier research, we suspected that it may explain why some studies did, and other studies did not find optimal integration. We assessed whether this suspicion was correct in Chapter 5, where we investigated integration of the otoliths and semicircular canals regarding the perception of self-tilt for two different motion frequencies. The motion frequencies that were investigated (0.075 and 0.25 Hz) fall within the 'critical range', as canal-based tilt perception should improve over this range, while otoliths-based tilt perception should deteriorate. The results showed that variance of the individual systems were indeed affected by motion frequency, consistent with the hypothesized effect. Integration turned out to occur, regardless of motion frequency. Hence, we conclude that integration of otoliths and semicircular canal cues is robust for sensory dynamics.

Validity of the results

Assessment of the tenability of optimal integration in general requires measurements of variance in unisensory and multisensory conditions. Compared to studies on optimal integration in other perceptual phenomena, for example visual- haptic integration on object size, we encountered several difficulties in studying optimal integration in the perception of spatial orientation and self-motion. To obtain reliable estimates of variance using Forced-Choice paradigms (as in our experiments), it is common practice to gather up to several hundred individual measurements per condition. Since *motion* implies a change of position over time, presentation of large numbers of motion stimuli is highly time-consuming and also demanding for participants. Moreover, isolated stimulation of the vestibular (sub)systems may introduce artifacts that cannot be avoided. For example, we achieved isolated stimulation of the otoliths by using complex motion profiles in the DESDEMONA centrifuge that also produced confounding Coriolis forces. And inter-sensory conflicts may arise when measuring visually perceived motion in stationary observers, as the inertial sensors simultaneously signal stationarity. Furthermore, the extreme

conditions of parabolic flight and centrifugation may induce motion sickness. Consequently, the number of subjects that could be tested and the number of repetitions that could reasonably be obtained without introducing high drop-out rates, selection effects and maturation effects, were limited.

The limited number of repetitions that could be obtained may have implications for the validity of statistical conclusions. Specifically, the uncertainty of the variance estimates did not always allow us to discriminate between performances in different sensory conditions for individual subjects. However, judged by the goodness of fit, the obtained parameter estimates were reasonable. Furthermore our conclusions are based on consistency of findings across multiple experiments.

The results may also be affected by the passive nature of our motion stimuli; our subjects were passive observers, as they were instructed to provide judgments on motion stimuli post-hoc. It is reasonable to assume that subjects who actively control their self-motion or orientation can better anticipate, and hence better interpret, the sensory feedback on their motion. Also, active self-motion through a real environment has the highest possible immersion and realism (Van Beuzekom, Medendorp & Van Gisbergen, 2001). Because our results suggest that integration benefits from a high level of realism, integration may benefit from active tasks.

It may be obvious that despite the taken precautions, such as minimizing the number of repetitions, and allowing frequent breaks, the experimental paradigms employed were highly demanding for our subjects. Therefore, it should be noted that the extreme conditions of the present studies may affect the ecological validity of the results.

Suggestions for future research

Based on the findings of our studies, we can make recommendations for future research on the effects of realism, congruence, and an active role of the subject, on sensory integration mechanisms.

For example, as a stronger test of the effect of visual realism on the incidence of optimal integration, a study could be performed where visual stimuli are not generated using artificial displays, but where the actual surroundings of the participants are manipulated. A perfectly suited platform for such a study would for example be the ‘Villa Volta’ attraction of the Efteling theme park in the Netherlands. In this attraction, a perception of head-over-heels rotation is induced by a combination of slight inertial pitch and simultaneous full rotation of the surrounding room. If the incidence of optimal integration can be shown to be higher in such a study than the incidence reported in the studies presented in this thesis, this would be provide support for the hypothesis that visual realism affects the integration of visual and inertial cues by the CNS.

In order to assess the relation between congruence and the incidence of optimal integration, the studies performed by Valente Pais et al. (2010) provide guidelines for an experimental design. Specifically, their data provide a Point of Mean Coherence (PMC)[1], which can be thought of as the amplitudes of a pair of visual and inertial cues for which they are judged congruent in 100% of cases, and provide an estimate of the Coherence Zone Width (CZW), which shows an upper and lower threshold for which cues are judged congruent in 50% of the cases. It may be hypothesized that the likelihood of integration of visual and inertial cues is proportional to their position within the coherence zone: the highest incidence should be observed at the PMC, corresponding to complete integration, which should decrease as a function of the distance of the cues with respect to the PMC. Similarly, the studies performed by Beckers et al. (2012) and Jonik et al. (2011) provide guidelines to assess temporal constraints of optimal integration of visual and inertial cues. Hence, future research may address the integration as a function of their combined position within the coherence zone.

Beneficial effects of an active role of subjects on the incidence of integration may be assessed by using a nulling task: subjects could be presented with a 'disturbance' signal that consists of several frequencies of visual and/or inertial motion. Such motion could for example be horizontal linear surge or sway motion, or angular motion. In each condition, the subject's task is to actively counter ('null') the disturbance, using for example a joystick as a feedback device. If subjects are indeed able to integrate visual and inertial signals, their nulling performance should improve when both sources of information are available, as compared to when only visual or inertial information is available.

Conclusion

Integration of redundant (multi)sensory information yields optimal estimates of environmental variables. Hence, integration provides an evolutionary advantage, as an organism that integrates redundant information is able to interact with its environment more adequately than an organism that does not. Therefore, we assume that humans integrate visual and inertial information on self-motion and body orientation in their everyday live. Because it is the purpose of motion simulation to mimic real-life, identification of the particular conditions that yield integration -our contribution to the understanding of sensory integration- may benefit motion simulation. We have shown

[1] A similar set of concepts is used in the literature on multisensory perception, where Point of Mean Coherence (PMC) is referred to as Point of Subjective Equality, or PSE (e.g., Butler et al., 2010), and Coherence Zone Width (CZW) should be proportional to the Just-Noticeable Difference, or JND.

that semicircular canal and otolith information on self-tilt was integrated, which seems to be an effective way to make optimal use of their different sensitivities to different motion frequencies; and that integration of visual and inertial information strongly depends on specific characteristics of the visual stimulus.

Finally, our study on the effects of reduced gravity on spatial orientation showed that gravity must exceed a recognition threshold before being perceived as reference for upright. The results suggested that this threshold increases with age. This leads to the exciting conclusion that in future missions to Mars, young astronauts are more likely to know “up” from “down” when moving around on the surface of Mars.

References

Alais, D., & Burr, D. (2004). The Ventriloquist Effect Results from Near-Optimal Bimodal Integration. Current Biology, 14(3), 257-262.

Allison, R. S., Howard, I. P., Zacher, J.E. (1999). Effect of field size, head motion and rotational velocity on roll vection and illusory self-tilt in a tumbling room. *Perception, 28,* 299-306.

Andersen, T.S., Tiipana, K., & Sams, M. (2005). Maximum Likelihood Integration of rapid flashes and beeps. *Neuroscience Letters, 380,* 155-160.

Anderson, J.H., Blanks, R.H.I., Precht, W. (1978). Response Characteristics of Semicircular Canal and Otolith Systems in Cat. I. Dynamic Responses of Primary Vestibular Fibers. *Experimental Brain Research, 32,* 491-507

Angelaki, D.E., & Cullen, K.E. (2008). Vestibular System: The Many Facets of a Multimodal Sense. *Annual Review of Neuroscience, 31,* 125-150

Angelaki, D.E., Gu, Y., & DeAngelis, G.C. (2009). Multisensory integration: psychophysics, neurophysiology, and computation. *Current Opinion in Neurobiology, 19,* 452-458

Angelaki, D.E., Gu, Y., & DeAngelis, G.C. (2011). Visual and vestibular cue integration for heading perception in extrastriate visual cortex. *The Journal of Physiology, 859(4),* 825-833

Angelaki, D.E., McHenry, M.Q., Dickman, J.D., Newlands, & Hess, B.J.M. (1999). Computation of Interial Motion: Neural Strategies to Resolve Ambiguous Otolith Information. *Journal of Neuroscience, 19(1),* 316-327

Arrott, A.P., Young, L.R., & Merfeld, D.M. (1990). Perception of linear acceleration in weightlessness. *Aviation, Space, and Environmental Medicine, 61,* 319-326.

Bakker, N.H., Werkhoven, P.J., & Passenier, P.O. (1999). The effects of proprioceptive and visual feedback on geographical orientation in virtual environments. *Presence, 8(1),* 36-53.

Barnett-Cowan, M., & Harris, L.R. (2011). Temporal processing of active and passive head-movement. *Experimental Brain Research, 214,* 27-35.

Barnett-Cowan, M., Raeder, S.M., & Bülthoff, H.H. (2012). Persistent perceptual delay for head movement onset relative to auditory stimuli of different durations and rise times. *Experimental Brain Research, 220*, 41-50.

Beckers, N.W.M., Pool, D.M., Valente Pais, A.R., Van Paassen, M.M., & Mulder, M. (2010) Perception and Behavioral Phase Coherence Zones in Passive and Active Control Tasks in Yaw. *AIAA Modeling and Simulation Technologies Conference 13-16 August 2012, Minneapolis, Minnesota.*

Benson, A.J., Hutt, E.C.B., & Brown, S.F. (1989). Thresholds for the perception of whole body angular movement about a vertical axis. *Aviation, Space, and Environmental Medicine, 60(3),* 205-213.

Benson, A.J., Guedry, F.E., Parker, D.E., & Reschke. M.F. (1997). Microgravity vestibular investigations: perception of self-orientation and self-motion. *Journal of Vestibular Research, 7,* 453-457.

Benson, A.J., Hutt, E.C.B., & Brown, S.F. (1989). Thresholds for the Perception of Whole Body Angular Movement About a Vertical Axis. *Aviation, Space, and Environmental Medicine, 60(3),* 205-213.

Bischof, N. (1974). Optic-vestibular orientation to the vertical. In: Kornhuber H. (Ed.) Handbook of Sensory Physiology. Vestibular System. Psychophysics, Applied Aspects and general interpretations. Berlin-New York, Springer-Verlag, pp. 155-190.

Blake, A., Bülthoff, H.H., & Sheinberg, D. (1993). Shape from Textrue: Ideal Observers and Human Psychophysics. *Vision Research, 33(12),* 1723-1737.

Bles, W., & Groen, E.L. (2009). The DESDEMONA Motion Facility: Applications for Space Research, *Microgravity Science and Technology, 21(4),* 281-286.

Bos, J.E., & Bles, W. (2002). Theoretical considerations on canal–otolith interaction and an observer model. *Biological Cybernetics, 86,* 191-207.

Bos, J.E., Bles, W., & Groen, E.L. (2008). A theory on visually induced motion sickness. *Displays, 29,* 47-57.

Bresciani, J.-P., Dammeier, F., & Ernst, M.O. (2006). Vision and touch are automatically integrated for the perception of sequences of events. *Journal of Vision, 6,* 554-664.

Butler, J.S., Campos, J.L., Bülthoff, H.H., & Smith, S.T. (2011). The Role of Stereo Vision in Visual-Vestibular Integration. *Seeing and Perceiving, 24,* 453-470.

Butler, J.S., Smith, S.T., Campos, J.L., & Bülthoff H.H. (2010). Bayesian integration of visual and vestibular signals for heading. *Journal of Vision, 10(11): 23,* 1-13.

Clark. B., & Graybiel, A. (1949). Linear acceleration and deceleration as factors influencing nonvisual orientation during flight. *Journal of Aviation Medicine, 20,* 92–101.

Clemens, I. A. H., de Vrijer, M., Selen, L. P. J., van Gisbergen, J. A. M., & Medendorp, W. P. (2011). Multisensory Processing in Spatial Orientation: An Inverse Probabilistic Approach. *Journal of Neuroscience, 31(14),* 5365-5377.

Clément, G., & Reschke, M.F. (2008). Neuroscience in Space. Springer US, New York.

Clément, G., Moore, S.T. Raphan, T., & Cohen, B. (2001). Perception of tilt (somatogravic illusion) in response to sustained linear acceleration during space flight. *Experimental Brain Research, 138(4),* 410-418.

Correia Grácio, B.J., De Winkel, K.N., Groen, E.L., Wentink, M., & Bos, J.E. (2012).The Time Constant of the Somatogravic Illusion. *Experimental Brain Research.*

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

Cullen, K.E., & Roy, J.E. (2004). Signal processing in the vestibular system during active versus passive head movements. *Journal of Neurophysiology, 91,* 1919-1933.

Curthoys, I.S., & Oman, C.M. (1987). Dimensions of the Horizontal Semicircular Duct, Ampulla and Utricle in the Human. *Acta Oto-laryngologica, 103(3-4),* 254-261.

De Graaf, B., Bos, J.E., Groen, E., Tielemans, W., Rameckers, F., Clark, J.B., Mead, A.M., & Guedry, F.E. (1998). Otolith responses during centrifugation along three axes of orientation. Report 1402, Naval Aerospace Medical Research Laboratory, Pensacola, Fa.

De Graaf, B., Bos, J.E., Tielemans, W., Rameckers, F., Rupert, A.H., & Guedry, F.E. (1996). Otolith contribution to ocular torsion and spatial orientation during acceleration. Technical Memorandum 96-3, Naval Aerospace Medical Research Laboratory, Pensacola, Fa.

De Vrijer, M. (2010). Multisensory integration in spatial orientation. Dissertation, Radboud Universiteit Nijmegen, The Netherlands.

De Winkel, K.N., Soyka, F., Barnett-Cowan, M., Bülthoff, H.H., Groen, E.L., & Werkhoven, P.J. Integration of visual and inertial cues in the perception of angular self-motion (submitted)

De Winkel, K.N., Weesie, J., Werkhoven, P.J., & Groen, E.L. (2010). Integration of visual and inertial cues in perceived heading of self-motion. *Journal of Vision, 10(12): 1,* 1-10.

Di Luca, M. (2010). New method to measure end-to-end delay of virtual reality. *Presence, 19(6),* 569–584.

Dichgans, J., & Brandt, T. (1978). Visual-vestibular interaction: Effects on self-motion perception and postural control. In: Held, R., Leibowitz, R.H.W., & Tueber, H.L. (Eds.) Handbook of sensory psychology, Vol VIII: Perception. Springer-Verlag, Berlin, pp. 756-795.

Dickman, J.D., Angelaki, D.E., & Correia, M.J. (1991) Response Properties of Gerbil Otolith Afferents to Small Angle Pitch and Roll Tilts. *Brain Research, 556(2),* 303-310.

Doidge, N. (2007). The Brain That Changes Itself. Viking Press, US, pp. 427.

Draper, N.R., & Smith, H. (1966). Applied regression analysis. John Wiley, New York.

Dyde, R.T., Jenkin, M.R., & Harris, L.R. (2006).The subjective visual vertical and the perceptual upright. *Experimental Brain Research, 173,* 612-622.

Dyde, R.T., Jenkin, M.R., Jenkin, H.L., Zacher, J.E., & Harris, L.R. (2009). The effect of altered gravity states on the perception of orientation. *Experimental Brain Research, 194,* 647-660.

Einstein, A. (1907) Über das Relativitätsprinzip und die aus demselben gezogenen Folgerungen. *Jahrbuch der Radio-aktivität und Elektronik, 4*, 411-462.

Ernst, M.O. & Bülthoff, H.H. (2004). Merging the senses into a robust percept. *Trends in Cognitive Science, 8,* 162-169.

Ernst, M. O. (2005). A bayesian view on multimodal cue integration. In: Knoblich, G., Thornton, I.M., Grosjean, M., & Shiffrar, M. (Eds.) Perception of the human body from the inside out. Oxford University Press, New York, pp. 105–131.

Ernst, M.O., & Banks, M.S. (2002) Humans Integrate visual and haptic information in a statistically optimal fashion. *Nature, 415,* 429-433.

Fernandez, C., & Goldberg, J. (1971). Physiology of peripheral neurons innervating semicircular canals of squirrel monkey. II. Response to sinusoidal stimulation and dynamics of peripheral vestibular system. *Journal of Neurophysiology, 34(4),* 661-675.

Fernandez, C., & Goldberg, J. M. (1976). Physiology of Peripheral Neurons Innervating Otolith Organs of the Squirrel Monkey. III. Response Dynamics. *Journal of Neurophysiology, 39(5),* 996-1008.

Fetsch, C.R., Turner, A.H., DeAngelis, G.C., & Angelaki, D.E. (2009). Dynamic reweighting of visual and vestibular cues during self-motion perception. *The Journal of Neuroscience, 29(49),* 15601-15612.

Fischer, A.J.E.M. (1993). Het menselijk evenwicht. Anatomie, fysiologie en pathologie van een intrigerend zintuig. Academische Uitgeverij, Amersfoort.

Fisher, R.A. (1922). On the Mathematical Foundations of Theoretical Statistics. *Philosophical Transactions of the Royal Society A, 222,*309-368.

Fitzpatrick, R., & McCloskey, D.I. (1994). Proprioceptive, visual and vestibular thresholds for the perception of sway during standing in humans. *Journal of Physiology, 478,* 173-186.

Georgopoulos, A.P., Schwartz, A.B., & Kettner, R.E. (1986). Neural population coding of movement direction. *Science, 233,* 1416-1419.

Gibson, J.J. (1950). Perception of the visual world. Houghton Mifflin, Boston.

Gibson, J.J. (1961). Ecological Optics. *Vision Research, 1,* 253-262.

Glasauer, S., & Mittelstaedt, H. (1992).Determinants of orientation in microgravity. *Acta Astronautica, 27,* 1-9 .

Godwin, R. Apollo 14. (2000). The NASA Mission Reports, Apogee Books, Bulington, Ontario.

Godwin, R. Apollo 16. (2002). The NASA Mission Reports Volume One, Apogee Books, Bulington, Ontario.

Graybiel, A., & Kellogg, S. (1967). The inversion illusion in parabolic flight: its probable dependence on otolith function. *Aerospace Medicine, 38*, 1099-1103.

Graybiel, A., Clark, B., & MacCorquodale, K. (1947). The illusory perception of movement caused by angular acceleration and by centrifugal force during flight. I. Methodology and preliminary results. *Journal of Experimental Psychology, 37(2),* 170-177.

Green, A.M., & Angelaki, D.E. (2010). Multisensory integration: resolving ambiguities to build novel representations. *Current Opinion in Neurobiology, 20,* 353-360.

Green, D.M., & Swets, J.A. (1988). Signal Detection Theory and Psychophysics, Peninsula Publishing, Las Altos, California.

Greene, W.H. (2008). Econometric Analysis. 6th ed. Upper Saddle River, NJ: Prentice–Hall.

Groen, E.L., & Bles W. (2004). How to use body tilt for the simulation of self-motion. *Journal of Vestibular Research, 14*, 375-385.

Groen, E.L., Jenkin, H., & Howard, I.P. (2002). Perception of self-tilt in a true and illusory vertical plane. *Perception, 31,* 1477-1490.

Gu, Y., Angelaki, D.E., & DeAngelis, G.C. (2008). Neural correlates of multisensory cue integration in macaque MSTd. *Nature Neuroscience, 11(10),* 1201-1210.

Gu, Y., DeAngelis, G.C., & Angelaki, D.E. (2007). A functional link between area MSTd and heading perception based on vestibular signals. *Nature neuroscience, 10,* 1038-1047.

Guedry, F.E. (1974). Psychophysics of Vestibular Sensation. In: Kornhuber H. (Ed.), Handbook of Sensory Physiology. Vestibular System. Psychophysics, Applied Aspects and general interpretations. Springer-Verlag, Berlin-New York, pp. 31- 54.

Gundry, A.J. (1978).Thresholds of perception for periodic linear motion. *Aviation, Space, and Environmental Medicine, 49,* 679-686.

Harm, D.L., Parker, D.E., Reschke, M.F., & Skinner, N.C. (1998). Relationship between selected orientation rest frame, circular vection and space motion sickness. *Brain Research Bulletin, 47(5),* 497-501.

Harris, L.R., Jenkin, M.R.M., & Dyde, R.T. (2012).The perception of upright under lunar gravity. *Journal of Gravitational Physiology, 19*.

Helbig, H.B., & Ernst, M.O. (2007). Optimal integration of shape information from vision and touch. *Experimental Brain Research, 179,* 595-606.

Henn, V., Cohen, B., & Young, L.R. (1980). Visual-vestibular interaction in motion perception and the generation of nystagmus. *Neuroscience Research Bulletin, 18,* 457-651.

Harris, L.R., Jenkin, M., & Zikovitz, D.C. (2000). Visual and non-visual cues in the perception of self-motion. *Experimental Brain Research, 135,* 12-21.

Holly, J.E., Vrublevskis, A., & Carlson, L.E. (2008). Whole-motion model of perception during forward- and backward-facing centrifuge runs*. Journal of Vestibular Research, 18(4),* 171-186.

Howard, I. P. (1982). Human visual orientation. Wiley, New York.

Howard, I.P. (1997). Interactions within and between the spatial senses*. Journal of Vestibular Research, 7(4),* 311-345.

Howard, I.P., & Childerson, L. (1994). The contribution of motion, the visual frame and visual polarity to sensations of body tilt. *Perception, 23,* 753-762.

Howard, I.P., & Heckmann, T. (1989). Circular vection as a function of the relative sizes, distances, and positions of two competing visual displays. *Perception, 18(5),* 657-665.

Howard, I.P., Bergström, S.S., & Ohmi, M. (1990).Shape from shading in different frames of reference. *Perception, 19,* 523-530.

Janssen, M., Lauvenberg, M., van der Ven, W., Bloebaum, T., Kingma, H. (2011).Perception threshold for tilt. *Otology & Neurotology, 32,* 818-825.

Jonik, P.M., Valente Pais, A.R., Van Paassen, M.M., & Mulder, M. (2011) Phase Coherence Zones in Flight Simulation. *AIAA Modeling and Simulation Technologies Conference 08-11 August 2011, Portland, Oregon.*

Jürgens, R., & Becker, W. (2006). Perception of angular displacement without landmarks: evidence for Bayesian fusion of vestibular, optokinetic, podokinesthetic and cognitive information. *Experimental Brain Research, 174,* 528-543.

Jürgens, R., & Becker, W. (2011). Human spatial orientation in non-stationary environments: relation between self-turning perception and detection of surround motion. *Experimental Brain Research, 215,* 327-344.

Kang, K., Shapley, R.M., & Sompolinsky, H. (2004). Information Tuning of populations of Neurons in Primary Visual Cortex. *The Journal of Neuroscience, 24(15),* 3726-3735.

Kaptein, R.G., & Van Gisbergen, J.A.M. (2005). Canal and Otolith Contributions to Visual Orientation Constancy During Sinusoidal Roll Rotation. *Journal of Neurophysiology, 95,* 1936-1948.

Karnath, H.O., Ferber, S., & Dichgans, J. (2000). The origin of contraversive pushing: Evidence for a second graviceptive system in humans. *Neurology, 55(9),* 1298-1304.

Knill, D.C., & Pouget, A. (2004) The Bayesian brain: the role of uncertainty in neural coding and computation. *TRENDS in Neurosciences, 27(12),* 712-719.

Knill, D.C., & Saunders, J.A. (2003). Do humans optimally integrate stereo and texture information for judgments of surface slant? *Vision Research, 43,* 2539-2558.

Koenderink, J.J. (1986). Optic Flow. *Vision Research, 26(1),* 161-180.

Kontsevich LL, Tyler CW (1999). Bayesian adaptive estimation of psychometric slope and threshold. *Vision Research, 39,* 2729 – 2737.

Körding, K.P. (2007). Decision theory: What "should" the nervous system do? *Science, 318,* 606-610.

Körding, K.P., Beierholm, U., Ma, W.J., Quartz, S., Tenenbaum, J.B. & Shams, L. (2007). Causal Inference in Multisensory Perception. *PLoS ONE, 2(9)*.

Lackner, J.R, & Graybiel, A. (1978b). Some influences of touch and pressure cues on human spatial orientation. *Aviation, Space, and Environmental Medicine, 49,* 498-804.

Lackner, J.R. (1976).Influence of abnormal postural and sensory conditions on human sensorimotor localization. *Environmental Biology and Medicine, 2,* 139-177.

Lackner, J.R. (1992). Spatial orientation in weightless environments*. Perception, 21,* 803-812.

Lackner, J.R., & Graybiel, A. (1978a). Postural illusions experienced during z-axis recumbent rotation and their dependence on somatosensory stimulation of the body surface. *Aviation, Space, and Environmental Medicine, 49,* 484-488.

Lambrey, S., & Berthoz, A. (2003). Combination of conflicting visual and non-visual information for estimating actively performed body turns in virtual reality. *International Journal of Psychophysiology, 50,* 101-115.

Laurens, J., & Droulez, J. (2007). Bayesian processing of vestibular information. *Biological Cybernetics, 96(4),* 389-404.

Loe, P.R., Tomko, D.L., & Werner, G. (1973). The Neural Signal of Angular Head Position in Primary Afferent Vestibular Nerve Axons*. Journal of Physiology, 230,* 29-50.

Long, J.S. (1997). Regression Models for Categorical and Limited Dependent Variables. Sage, Thousand Oaks.

Macmillan, N.A., & Creelman, C.D. (2005). Detection theory: a user's guide. Lawrence Erlbaum Associates, Mahwah, New Jersey.

MacNeilage, P.R., Banks, M.S., Berger, D.R., & Bülthoff, H.H. (2007). A Bayesian model of the disambiguation of gravitoinertial force by visual cues. *Experimental Brain Research, 179(2),* 263-290.

Mayne, R. (1974). A Systems Concept of the Vestibular Organs. In Kornhuber H. (Ed.), Handbook of Sensory Physiology. Vestibular System. Psychophysics, Applied

Aspects and general interpretations. Springer-Verlag, Berlin-New York, pp. 493-580.

Melcher, G.A. & Henn, V. (1981). The latency of circular vection during different accelerations of the optokinetic stimulus. *Perception and Psychophysics, 30 (6),* 552-556.

Merfeld, D.M. (2011). Signal detection theory and vestibular thresholds: I. Basic theory and practical considerations. *Experimental Brain Research, 210,* 389-405.

Merfeld, D.M. (1995). Modeling human vestibular responses during eccentric rotation and off vertical axis rotation. *Acta Oto-Laryngologica, 115,* 354-359.

Merfeld, D.M., Park, S., Gianna-Poulin, C., Black, F.O., & Wood, S. (2005). Vestibular Perception and Action Employ Qualitatively Different Mechanisms. I., Frequency Response of VOR and Perceptual Responses During Translation and Tilt. *Journal of Neurophysiology, 94,* 186-198.

Merfeld, D.M., Zupan, L.H., & Gifford, C.A. (2001). Neural Processing of Gravito-Inertial Cues in Humans. II. Influence of the Semicircular Canals During Eccentric Rotation. *Journal of Neurophysiology, 85,* 1648-1660.

Mergner, T., & Rosemeier, T. (1998). Interaction of vestibular, somatosensory and visual signals for postural control and motion perception under terrestrial and microgravity conditions-a conceptual model. *Brain Research Reviews, 28(1-2),* 118-135.

Miller, I., & Miller, M. (2004). John E. Freund's Mathematical Statistics with Applications (7th Edition). Pearson Education.

Minor, L.B., & Goldberg, J.M. (1990). Influence of static head position on the horizontal nystagmus evoked by caloric, rotational and optokinetic stimulation in the squirrel monkey. *Experimental Brain Research, 82,* 1-13.

Mittelstaedt, H. (1986). The subjective vertival as a function of visual and extraretinal cues. *Acta Psychologica, 63(1),* 63-85.

Mittelstaedt, H. (1988). Determinants of space perception in space flight. *Advances in Oto-Rhino-Laryngology, 42,* 24-30.

Mittelstaedt, H. (1996). Somatic Graviception. *Biological psychology, 42,* 53-74.

Mittelstaedt, H. (1983). A new solution to the problem of the subjective vertical. *Naturwissenschaften, 70,* 272-281.

Müller, G.E. (1916.) über das Aubertsche Phaenomenon. *Zeitschrift für Psychologie, 49,* 109-244.

Ohmi, M. (1996). Egocentric perception through interaction among many sensory systems. *Cognitive Brain Research, 5,* 87-96.

O'Keefe, J., & Dostrovsky, J. (1971). The hippocampus as a spatial map: Preliminary evidence from unit activity in the freely moving rat. *Brain Research, 34,* 171-175.

Oman, C. (2007). Spatial Orientation and Navigation in Microgravity. In: Mast, F., & Jäncke, L. (Eds.) Spatial Processing in Navigation, Imagery and Perception. Springer US, New York, pp. 209-247.

Paige, G.D., & Tomko, D.L. (1991). Eye Movement Responses to Linear Head Motion in the Squirrel Monkey. I. Basic Characteristics. *Journal of Neurophysiology, 65,* 1170-1182.

Paloski, W.H., Oman, C.M., Bloomberg, J.J., Reschke, M.F., Wood, S.J., Harm, D.L., Peters, B.T., Mulavara, A.P., Locke, J.P., & Stone, L.S. (2008). Risk of sensory-motor performance failures affecting vehicle control during space missions: a review of the evidence. *Journal of Gravitational Physiology, 15,* 1-29.

Pavard, B., & Berthoz, A. (1977). Linear acceleration modifies perceived velocity of a moving visual scene. *Perception, 6(5),* 529-540.

Polit, D. (1996). Data Analysis and Statistics for Nursing Research. Appleton & Lange, Stamford, Connecticut.

Probst, T., Brandt, T., & Degner, D. (1986). Object-motion detection affected by concurrent self-motion perception. *Behavioral Brain Research, 22,* 1-11.

Raphan, T., & Cohen, B. (2002). The vestibule-ocular reflex in three dimensions. *Experimental Brain Research, 145,* 1–27.

Richter, E. (1980).Quantative study of human scarpa's ganglion and vestibular sensory epithelia. *Acta Otolaryngologica, 90,* 199-208.

Ringach, D.L., Shapley, R.M., & Hawken, M.J. (2002). Orientation Selectivity in Macaque V1: Diversity and Laminar Dependence. *The Journal of Neuroscience, 22(13),* 5639-5651.

Rock, I., & Victor, J. (1964). Vision and touch: An experimentally created conflict between the two senses. *Science, 143,* 594-596.

Rock, I., & Heimer, W. (1957).The effect of retinal and phenomenal orientation on the perception of form. *American Journal of Psychology, 70,* 493-511.

Roy, J.E., & Cullen, K.E. (2004). Dissociating Self-Generated from Passively Applied Head Motion: Neural Mechanisms in the Vestibular Nuclei. *Journal of Neuroscience, 24(9),* 2102-2111.

Sato, Y., Toyoizumi, T., & Aihara, K. (2007). Bayesian inference explains perception of unity and ventriloquism aftereffect: Identification of common source of audiovisual stimuli. *Neural Computation, 19,* 3335-3355.

Seidman, S.H., Telford, L., & Paige, G.D. (1998). Tilt perception during dynamic linear acceleration. *Experimental Brain Research, 119,* 307-314.

Serletis, D., Zalay, O.C., Valiante, T.A., Bardakjian, B.L., & Carlen, P.L. (2011). Complexity in Neuronal Noise Depends on Network Interconnectivity. *Annals of Biomedical Engineering, 39(6),* 1768-1778.

Shams, L., Ma, W.J., & Beierholm, U. (2005). Sound-induced flash illusion as an optimal percept. *Neuroreport, 16 (17),* 1923-1927.

Shirai, N., & Ichihara, S. (2010). Reduction in sensitivity to radial optic-flow congruent with ego-motion. *Vision Research, 62,* 201-208.

Si, X., Angelaki, D.E., & Dickman, J.D. (1997). Response Properties of Pigeon Otolith Afferents to Linear Acceleration. *Experimental Brain Research, 117,* 242-250.

Spence, C., & Squire, S. (2003). Multisensory Integration: Maintaining the Perception of Synchrony. *Current Biology, 13,* 519-521.

Tanner, T. (2008). Generalized adaptive procedure for psychometric measurement. *Perception 37: 93*.

Telford, L., Howard, I.P., & Ohmi, M. (1995). Heading judgments during active and passive self-motion. *Experimental Brain Research, 104,* 502-510.

Telford, L., Seidman, S.H., Paige, G.D. (1997). Dynamics of squirrel monkey linear vestibuloocular reflex and interactions with fixation distance. *Journal of Neurophysiology, 78,* 1775-1790.

Titterington, D.M., Smith, A.F.M., Makov, U.E. (1985). Statistical Analysis of Finite Mixture Distributions. John Wiley & Sons, New York.

Todd, J.T., Christensen, J.C., & Guckes, K.M. (2010). Are discrimination thresholds a valid measure of variance for judgments of slant from texture? *Journal of Vision, 10(2):20,* 1-18.

Trutoiu, L.C., Mohler, B.J., Schulte-Pelkum, J., & Bülthoff, H.H. (2009). Circular, linear, and curvilinear vection in a large-screen virtual environment with floor projection. *Computers & Graphics, 33,* 47-58.

Tschermak, A. (1931). Optischer Raumsinn[Optical sense of space]. In: Bethe, A., Bergmann, G., Emden, G., Ellinger, A. (Eds.) Handbuch de normalen und pathologischen 7(2). Springer, Berlin, pp. 824-1000.

Van Beers, R.J., Sittig, A.C., & Denier van der Gon, J.J. (1999). Integration of proprioceptive and visual position-information: An experimentally supported model. *Journal of Neurophysiology, 81,* 1355-1364.

Van Beuzekom, A.D., Medendorp, W.P., & Van Gisbergen, J.A.M. (2001). The subjective vertical and the sense of self orientation during active body tilt. *Vision Research, 41,* 3229-3242.

Van der Steen, F. A. M. (1998). Self-Motion Perception. Dissertation, Delft University of Technology, The Netherlands.

Valente Pais, A.R., Van Paassen, M.M., Mulder, M., & Wentink, M. (2010). Perception of Combined Visual and Inertial Low-Frequency Yaw Motion. *AIAA Modeling and Simulation Technologies Conference 2-5 August 2010, Toronto, Ontario Canada.*

Valente Pais, A.R., Van Paassen, M.M., Mulder, M., & Wentink, M. (2010). Perception Coherence Zones in Flight Simulation. *Journal of Aircraft, 47(6),* 2039-2048.

Vingerhoets, R.A.A., De Vrijer, M., Van Gisbergen, J.A.M., & Medendorp, W.P. (2009). Fusion of visual and vestibular tilt cues in the perception of visual vertical. *Journal of Neurophysiology, 101,* 1321-1333.

Vingerhoets, R.A.A., Van Gisbergen, J.A.M., & Medendorp, W.P. (2007). Verticality Perception During Off-Vertical Axis Rotation. *Journal of Neurophysiology, 97,* 3256-3268.

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

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

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,* 646-660.

Wertheim, A.H., & Reymond, G. (2007). Neural noise distorts perceived motion: the special case of the freezing illusion and the Pavard and Berthoz effect. *Experimental Brain Research, 180,* 569-576.

Witkin, H.A., & Asch, S.E. (1948).Studies in space orientation. III. Perception of the upright in the absence of a visual field. *Journal of Experimental Psychology, 38,* 603-614.

Wright, W.G., DiZio, P., & Lackner, J.R. (2005). Vertical linear self-motion perception during visual and inertial motion: More than weighted summation of sensory inputs. *Journal of Vestibular Research, 15,* 185-195.

Young, L.R., Dichgans, J., Murphy, R., & Brandt, Th. (1973). Interaction of optokinetic and vestibular stimuli in motion perception. *Acta Oto-laryngologica, 76,* 24-31.

Young, L.R., Oman, C.M., Watt, D.G.D., Money, K.E., & Lichtenberg, B.K. (1984). Spatial Orientation in Weightlessness and Readaptation to Earth's Gravity. *Science, 225,* 205-208.

Young, L.R., Oman, C.M., Watt, D.G.D., Money, K.E., Lichtenberg, B.K., Kenyon, R.V., & Arrott, A.P., M.I.T./Canadian vestibular experiments on the Spacelab-1 mission: 1. Sensory adaptation to weightlessness and readaptation to one-g: an overview. *Experimental Brain Research, 64,* 291-298.

Yuille, A.L., Bülthoff, H.H. (1996). Bayesian decision theory and psychophysics. In: Knill, D.C., & Richards, W. (1996). Perception as Bayesian Inference. Cambridge Univ. Press, Cambridge, pp. 123–61.

Yuval-Greenberg, S., & Deouell, L.Y. (2009). The dog's meow: assymetrical interaction in cross-model object recognition. *Experimental Brain Research, 193,* 603-614.

Zacharias G.L. & Young L.R. (1981). Influence of combined visual and vestibular cues on human perception and control of horizontal rotation. *Experimental Brain Research, 41 (2),* 159-171.

Zaichik, L.E., Rodchenko, V.V., Rufov, I.V., Yashin, Y.P., & White, A.D. (1999). Acceleration Perception. *American Institute of Aeronautics and Astronautics*, 512-520.

Zupan, L.H., Merfeld, D.M., & Darlot, C. (2002) Using Sensory Weighting to model the influence of canal, otolith and visual cues on spatial orientation and eye movements. *Biological Cybernetics, 86(3),* 209-230.

Samenvatting

Aan het begin van deze dissertatie hebben we een Bayesiaans kader beschreven om de waarneming van zelfbeweging en ruimtelijke oriëntatie te modelleren. Dit biedt een alternatief voor bestaande deterministische modellen die weliswaar de dynamica van de diverse zintuigsystemen en hun interacties beschrijven, maar geen rekening houden met onzekerheden in de zintuigsignalen. In tegenstelling tot deterministische modellen biedt een Bayesiaans kader een stochastische benadering waarmee kan worden beschreven hoe de verschillende zintuigsignalen door het centraal zenuwstelsel worden gewogen. In deze stochastische modellen wordt aangenomen dat de weging van zintuigsignalen proportioneel is aan de onzekerheid van die signalen. Het wordt verondersteld dat 'kennis' over die onzekerheid impliciet beschikbaar is in het centraal zenuwstelsel, in de vorm van variantie in de interne representatie van een extern signaal in neurale populaties.

De hypothese die in deze dissertatie centraal staat is dat een Bayesiaans kader geschikt is voor een modelmatige beschrijving van de interacties tussen visuele en inertiële signalen in de waarneming van ruimtelijke orientatie en zelfbeweging. Gegeven dat aan een aantal basale aannames wordt voldaan, komen de voorspellingen van het Bayesiaans kader overeen met die van een Maximum-Likelihood kader (MLI; zie hoofdstuk 2). MLI voorspelt dat de geïntegreerde waarneming een gewogen gemiddelde is van de afzonderlijke sensorische signalen en dat de onzekerheid van de geïntegreerde waarneming altijd kleiner is dan de onzekerheid van de afzonderlijke signalen. De reductie van die onzekerheid kan met MLI exact worden berekend.

Om bovenstaande hypothese te kunnen toetsen hebben we in onze experimenten de onzekerheid gemeten van de waarneming op basis van stimulatie van de afzonderlijke sensorische systemen ('unisensorisch'), en die vergeleken met de onzekerheid in de waarneming gebaseerd op gelijktijdige stimulatie van de betrokken sensorische systemen ('multisensorisch'). De hypothese voorspelt dat de onzekerheid het kleinst is in multisensorische condities.

In hoofdstuk 3 is de integratie onderzocht van informatie vanuit het **visuele systeem** en inertiële informatie vanuit de **otolieten** in het vestibulair systeem over de richting van **horizontale lineaire bewegingen**. De visuele stimuli waren gesimuleerde bewegingen door een zogenaamd 'starfield', dat bestond uit witte bollen tegen een donkere achtergrond. De visuele bewegingen werden getoond met een relatief groot blikveld. De inertiële beweging bestond uit daadwerkelijke fysieke translatie van de simulatorcabine over een lineaire baan met een lengte van 8 meter. In dit onderzoek kwam de data van slechts één proefpersoon overeen met de hypothese. De andere acht proefpersonen baseerden hun

responsies óf uitsluitend op vestibulaire informatie, óf afwisselend op visuele en vestibulaire informatie, maar integreerden de informatie niet. We veronderstellen dat de visuele stimulus niet overtuigend genoeg was om geïnterpreteerd te worden als een bron van informatie over zelfbeweging.

In hoofdstuk 4 is de integratie van **visuele informatie** en informatie vanuit de **halfcirkelvormige kanalen** in het vestibulaire systeem onderzocht over **draaibewegingen in een horizontaal vlak**. Omdat het voorgaande onderzoek suggereerde dat de visuele stimuli niet overtuigend genoeg waren, gebruikten we in dit onderzoek video-opnames van beweging van de omgeving. De resultaten lieten zien dat het gedrag van drie van de acht proefpersonen overeenkwam met de hypothese. De proefpersonen die geen integratie lieten zien vertrouwden óf uitsluitend op de visuele stimulus, óf afwisselend op visuele en vestibulaire informatie, in tegenstelling tot het voorgaande onderzoek waar in die gevallen de vestibulaire stimulus altijd dominant was. Dit is mogelijk te verklaren door het hogere realisme van de visuele stimulus.

In hoofdstuk 5 is onderzocht of inertiële informatie van de **halfcirkelvormige kanalen** en de **otolieten** over **kanteling in een verticaal vlak** op een statistisch optimale manier wordt geïntegreerd. Hierbij ging het dus om integratie van de twee deelsystemen binnen het vestibulaire systeem, in afwezigheid van visuele informatie. De halfcirkelvormige kanalen reageren op hoekversnellingen van het hoofd. Door hun dynamische eigenschappen geven zij dit aan het centraal zenuwstelsel door als informatie over de snelheid waarmee het hoofd kantelt. Tegelijkertijd geven de otolieten informatie over de oriëntatie van het hoofd ten opzichte van de zwaartekracht. Uit de resultaten van de unisensorische condities bleek dat de otolieten het meest gevoelig waren voor kanteling met een lage bewegings-frequentie (0.025Hz) en minder voor kanteling met een hoge bewegingsfrequentie (0.25Hz). Voor de kanalen was het omgekeerde het geval. De resultaten van de multisensorische conditie kwamen overeen met de voorspellingen van MLI: de onzekerheid was het kleinst wanneer beide systemen simultaan werden gestimuleerd, onafhankelijk van de bewegingsfrequentie. Dit toont aan dat het centraal zenuwstelsel in dit opzicht optimaal gebruik maakt van twee informatiebronnen.

In hoofdstuk 6 is de integratie van informatie van het **visuele systeem** en de **otolieten** over **kanteling in een verticaal vlak** onderzocht. Voor dit doel maakten we gebruik van een unieke gelegenheid om deel te nemen aan een campagne van de Europese Ruimtevaartorganisatie ESA. Hierbij werden met een verkeersvliegtuig paraboolvluchten uitgevoerd om verschillende zwaartekrachtniveaus te simuleren, namelijk 0g (gewichtloosheid), 0.16g (zwaartekrachtniveau op de Maan) en 0.38g (zwaartekrachtniveau op Mars). Onze verwachting was dat de weging van de otolietinformatie in de waarneming van kanteling omgekeerd evenredig zou zijn met

grootte van de zwaartekracht. De visuele stimuli waren door een computer gegenereerde stadsscenes, gepresenteerd in een Head-Mounted Display (HMD). De resultaten lieten geen integratie zien van visuele en otoliet-informatie, wat mogelijk te wijten is aan het relatief kleine blikveld van de HMD. Wel leverden de resultaten een opmerkelijke bevinding op over de invloed van de grootte van de zwaartekracht op de waargenomen kanteling: boven een bepaalde drempelwaarde gold de zwaartekracht als referentie voor de subjectieve verticaal, en domineerde als zodanig de waargenomen kanteling. Onder die drempelwaarde werd de zwaartekracht niet langer waargenomen als referentie voor verticaal, hoewel de kracht nog wel voelbaar was. Onder deze drempel werd de waarneming van kanteling plotseling gedomineerd door de lange lichaams-as (de ruggengraat). Hoewel er slechts bij een klein aantal proefpersonen werd gemeten, leek deze drempelwaarde bovendien toe te nemen met leeftijd. Dit suggereert dat bij toekomstige missies naar Mars jongere astronauten beter in staat zullen zijn zich te oriënteren dan oudere astronauten.

De bevindingen van de verschillende experimentele hoofdstukken van deze dissertatie bieden nogal uiteenlopende ondersteuning voor de hypothese dat de integratie van visuele en vestibulaire informatie over zelfbeweging en lichaamsoriëntatie beschreven kan worden met een Bayesiaans kader. In de gevallen dat integratie werd gevonden was deze statistisch optimaal, conform MLI. In de gevallen dat er geen integratie werd gevonden, vertrouwden de proefpersonen afwisselend op de verschillende bronnen van informatie. Er zijn geen duidelijke aanwijzingen voor suboptimale reductie van onzekerheid, en dus suboptimale integratie. De visuele stimuli die werden gebruikt in hoofdstukken 3, 4 en 6 varieerden in realisme, wat van invloed leek op integratie. Het wel of niet optreden van integratie lijkt daarmee af te hangen van specifieke karakteristieken van de visuele stimulus, zoals de grootte van het blikveld en de inhoud van de getoonde beelden.

Het onderzoek in hoofdstuk 5 naar integratie van twee inertiële bronnen van informatie liet vrijwel voor alle proefpersonen statistisch optimale integratie zien. Gezien het feit dat een lineaire versnelling met een lage frequentie door de otolieten wordt waargenomen als (illusoire) kanteling, lijkt integratie met informatie uit de halfcirkelvormige kanalen hier een juiste interpretatie van een – op zichzelf ambigue – stimulus te vergemakkelijken.

Dankwoord

“Het eerste wat me opviel”, zei mijn promotor tijdens onze laatste meeting van mijn AiO-schap over mijn dissertatie, “was hoe dun dit is.”. Ik antwoordde dat je voor een goed verhaal niet veel woorden nodig hebt. Hoewel ik een compacte dissertatie zelf een elegant eindresultaat vind, doet het slanke voorkomen geen eer aan de hoeveelheid werk die er in werkelijkheid achter zit. Deze dissertatie is het resultaat van de inspanningen van een heleboel mensen. Ik heb een geweldige tijd gehad, en hoewel iedere betrokkene daarvoor eigenlijk een persoonlijk woord van dank verdient, zal ik hier moeten volstaan met een beknopt dankwoord.

Allereerst wil ik mijn ouders bedanken. Jullie hebben me decennialang onderhouden, onderwezen en opgevoed. Dat zal niet altijd een makkelijke, prettige en/of dankbare taak geweest zijn, maar ik ben best tevreden met het resultaat. Ik hoop dat jullie het daarmee eens zijn.

Vervolgens wil ik mijn promotor Peter Werkhoven en copromotor Eric Groen bedanken.

Peter, ik had soms het idee dat je mijn werk niet begreep. Dat bleek echter aan mij te liggen: ik begreep het zelf niet. Je dwong me met gerichte vragen altijd dieper na te denken. Dat leidde bij mij zo nu en dan tot de nodige verwarring, maar uiteindelijk ben ik er telkens toch in geslaagd mijn denkraam open te breken, om het een beetje groter weer te sluiten. Ik ben je dankbaar voor alle verworven inzichten. Malle pietje en gekke henkie zal ik hierna nooit meer met jou in één alinea noemen.

Eric, doordat je mijn copromotor en dagelijks begeleider was hebben we veel met elkaar te maken gehad. Dat heeft zo nu en dan geleid tot felle discussies en gespannen trekjes in mijn gezichtsuitdrukking. Maar uiteindelijk ik ben trots op al onze publicaties, en dat is te danken aan jouw rode strepen, sturing, bemoedigende woorden en nimmer aflatende inzet. En dankzij jouw gave als lobbyist ben ik een van de eerste mensen die én gewichtloos én op de maan en Mars zijn geweest!

Een van mijn favoriete professoren tijdens mijn studie psychologie, Lex Wertheim, zei altijd dat je moet zwemmen in je data als je niet ziet hoe de vork in de steel zit. En dat heb ik gedaan. Ik ben telkens zonder zwembandjes in het diepe gesprongen, om erachter te komen dat ik niet kon zwemmen. Gelukkig was er dan Jeroen Weesie, badmeester in de statistiek, om me op het droge te slepen. Jeroen, bedankt voor alle lessen. Ik heb er zoveel van geleerd dat ik er vertrouwen in heb in de toekomst zelf mijn hoofd boven water te kunnen houden.

I would like to thank the scientific committee, consisting of Heinrich Bülthoff, Jeroen Smeets, Max Mulder, and Gilles Clément, who were burdened with judging my dissertation, for reading the work and providing feedback.

To the team of our parabolic-flight study, Eric Groen, Gilles Clément, Angie Bukley, Erik Laan, Wietse Ledegang, and Bernd de Graaf: I'm thankful for the (scientific) discussions we had, your feedback, the efforts you took to get the work published, and the fun we had during our week at Novespace. Erik and Wietse, you deserve an extra pat on the back for the courage you showed by getting on a plane that deliberately falls out of the sky, for a second time, after a technical failure. I'm glad we all survived!
Mijn collega's van TNO: Jelte Bos, Wouter Vos, Mark Houben, Wietse Ledegang, Sjoerd de Vries, Bart Kappe, Anne-Marie Brouwer, Ries Simons, en de technici Rob van de Pijpekamp, Ingmar Stel, Jan van de Kooij, Sjaak Kriekaard, Natasja van der Leden, Frank Geurtsen en Antoon Wennemers, jullie bedank ik voor alle hulp, het programmeerwerk, de inzet, en het delen van jullie kennis. Het was van onschatbare waarde met jullie te kunnen sparren over de problemen die ik tegenkwam tijdens mijn experimenten, en zonder jullie was de langdurige marteling van mijn proefpersonen niet mogelijk geweest.
De mensen van DESDEMONA: Bernd de Graaf, Mark Wentink, Suzanne Nooij, Ron van der Pols, en de operators Paul Bakker, Frank Geurtsen, en Fred en Amanda van het CML, dankzij jullie heb ik voor mijn werk de meest geavanceerde bewegingssimulator ter wereld kunnen gebruiken. Iedere keer dat ik vanuit de controlekamer naar die tonnen aan centrifugerend staal stond te kijken, was een onvergetelijke ervaring.
Stella Donker, Ignace Hooge en Maarten van der Smagt van Universiteit Utrecht, bedankt voor de ondersteuning en de inspiratie die jullie me hebben gegeven tijdens mijn studie.
I also want to thank the people I worked with at the Max Planck Institute: Florian Soyka, John Butler, Karl Beykirch, Michael Barnett-Cowan, and Paolo Robuffo Giordano. You're all brilliant scientists, and I'm grateful I was welcome at your institute and could experience the MPI firsthand.
De dilettanten bij TNO verdienen ook wat speciale aandacht. Tom Philippi, bedankt voor het vertrouwen dat je in mij had tijdens mijn afstuderen en de lovende woorden die er vast aan hebben bijgedragen dat ik deze promotieplaats kreeg. Bruno Correia Grácio, I'd like to thank you for the countless hours of help with my experiments, your explanations of engineering stuff, and your friendship. Our trip to the AIAA was more like a vacation than work. Sander, Marieke, Janne, Roald, Vanessa, Puck, Dylan, Martin, Benny van de UU, Rita van de TuD, en mijn afstudeerstudente Renske, bedankt voor al jullie hulp en gezelligheid.
Bruno en Robbert, op het moment dat jullie dit boekje onder ogen krijgen, moet mijn verdediging nog plaatsvinden. Mocht het onverhoopt tot een handgemeen komen, dan verwacht ik van jullie als paranimfen een adequate verdediging. Ik hoop dan ook dat jullie dit lezen temidden van een intensieve training.

Iris, ik ben je niet vergeten: ik wilde je graag tot het laatst bewaren. Je hebt prachtig artwork geleverd voor de omslag van mijn dissertatie. Je was een fijne collega, altijd in voor een spelletje matlablingo, whiteboardgalgje of een junky-Friday lunch. En buiten werktijd was het nog gezelliger. Tijd voor een nieuwe roadtrip, nu op de motor?

List of Publications

De Winkel, K.N., Correia Grácio, B.J., Groen, E.L., & Werkhoven, P.J. (2012) Integration of Otolith and Semicircular Canal Cues in the Perception of Lateral Body Tilt. *Submitted*.

De Winkel, K.N., Soyka, F., Barnett-Cowan, M., Bülthoff, H.H., Groen, E.L., & Werkhoven, P.J. (2012) Integration of Visual and Inertial Cues in the Perception of Angular Self-Motion. *Under revision for publication.*

De Winkel, K.N., Clément, G., Groen, E.L., & Werkhoven, P.J. (2012) The perception of verticality in Lunar and Martian gravity conditions. *Neuroscience Letters 529(1),* 7-11.

Correia Grácio, B.J., De Winkel, K.N., Groen, E.L., Wentink, M., & Bos, J.E. (2012) The Time Constant of the Somatogravic Illusion. *Experimental Brain Research*.

Soyka, F., De Winkel, K.N., Barnett-Cowan, M., Groen, E.L., & Bülthoff, H.H. (2011) Integration of visual and vestibular information used to discriminate rotational self-motion. *i-perception*.

De Winkel, K.N., Weesie, J., Werkhoven, P.J., & Groen, E.L. (2010) Integration of visual and inertial cues in perceived heading of self-motion. *Journal of Vision, 10(12): 1,* 1-10.

De Winkel, K.N., Correia Grácio, B.J., Groen, E.L., & Werkhoven, P.J. (2010) Visual-inertial coherence zone in the perception of heading. K.N. de Winkel, B.J. Correia Grácio, E.L. Groen, P.J. Werkhoven. *Conference Proceedings American Institute of Aeronautics and Astronautics.*

Curriculum Vitae

Ksander de Winkel was born on October 9, 1984, in Wageningen, The Netherlands. In 2003, he completed his secondary education at the 'Het Nieuwe Eemland College' in Amersfoort. After his secondary education, he studied Psychology at Utrecht University, with a minor in Methods and Statistics. For his master's thesis, he investigated fusion of visual and tactile information in the primary visual cortex, by means of electroencephalography (EEG), at the TNO Institute for Applied Scientific Research, in Soesterberg, The Netherlands.

Thereafter, he obtained a PhD-position at the department of Information and Computing Sciences at Utrecht University. For this position he worked on a daily basis at the TNO Institute for Applied Scientific Research. The topic of his research was multisensory perception of self-motion and spatial orientation. The project was supervised by Prof.dr. Peter Werkhoven and Dr. Eric Groen.

During his PhD-programme he carried out experiments in which the visual system and vestibular (sub-)system(s) were stimulated both in isolation from each other as well as simultaneously, making use of the unique capabilities of the DESDEMONA motion simulator at TNO. He worked at, and together with, the Max Planck Institute for Biological Cybernetics in Tuebingen, Germany, and participated in the first Joint-European Partial-G Parabolic Flight campaign, with a study on the effects of partial gravity on spatial orientation.