

Perception of self-motion:
Vection experiments in
multi-sensory Virtual Environments

Inaugural - Dissertation

zur
Erlangung des Grades eines Doktors
der Naturwissenschaften
in der Fakultät für Psychologie
der

RUHR-UNIVERSITÄT-BOCHUM

vorgelegt von:

Jörg Schulte-Pelkum

aus Tokyo

2007

Gedruckt mit Genehmigung der Fakultät für Psychologie der

RUHR-UNIVERSITÄT BOCHUM

Referent: Prof. Dr. Rainer Guski, Fakultät für Psychologie, Ruhr-Universität Bochum

Korreferent: Prof. Dr. Heinrich Bülthoff, Max-Planck-Institut für biologische Kybernetik,
Tübingen

Tag der mündlichen Prüfung: 18.07.2008

Contents

Zusammenfassung in deutscher Sprache	6
Summary	7
1 General introduction	8
1.1 Why study self-motion perception? Motivation and aim	9
1.2 Overview of the human senses contributing to self-motion perception	10
1.2.1 The visual system	10
1.2.2 The auditory system	11
1.2.3 The vestibular system	12
1.2.4 The somatosensory system	14
1.3 Human vestibular thresholds for the perception of passive, whole-body self-motion .	15
1.3.1 Thresholds for vestibular detection of rotational motion	17
1.3.2 Thresholds for vestibular detection of linear motion	18
2 General characteristics of vection	19
2.1 Visual parameters affecting vection	21
2.1.1 Size and retinal location of the stimulated visual FOV	21
2.1.2 Foreground-background separation between moving and stationary surfaces	22
2.1.3 Spatial frequency content of the moving visual stimulus	23
2.1.4 Optical velocity and direction of the visual stimulus	23
2.1.5 Effect of eye-movements on vection	24
2.1.6 Stereoscopic cues	25
2.2 Non-visual cues for vection	25
2.3 Multi-modal aspects of vection	26
2.4 Higher-level influences on vection	26
2.4.1 Might presence have an influence on vection?	28
2.5 Methodological considerations	30
2.5.1 “Subjective” vs. “objective” measures of vection	30
2.5.2 Vection experiments in VR	32
3 Circular vection experiments in VR	33
3.1 Experiment 1: Do natural, photorealistic visual stimuli enhance vection?	33
3.1.1 Introduction	33
3.1.2 Hypotheses	34
3.1.3 Methods	34
3.1.4 Results	38
3.1.5 Discussion and Conclusion	38
3.2 Experiment 2: The relation between the feeling of presence in VR and vection . . .	41
3.2.1 Introduction	41
3.2.2 Hypotheses	42
3.2.3 Methods	43
3.2.4 Results	44

CONTENTS

3.2.5	Discussion	48
3.2.6	Conclusions	51
4	Multi-sensory interactions in vection	53
4.1	Experiment 3: Visual-vibrational interactions in circular vection	53
4.1.1	Introduction	54
4.1.2	Hypotheses	54
4.1.3	Methods	54
4.1.4	Results	56
4.1.5	Discussion and Conclusions	56
4.2	Experiment 4: Visual-auditory interactions in circular vection	60
4.2.1	Introduction	60
4.2.2	Hypotheses	61
4.2.3	Methods	62
4.2.4	Results	62
4.2.5	Discussion and Conclusion	63
4.3	Experiment 5: Visual-vestibular interactions in forward linear vection	66
4.3.1	Introduction	66
4.3.2	Hypotheses	67
4.3.3	Methods	67
4.3.4	Results	69
4.3.5	Discussion and Conclusion	70
4.4	Experiment 6: Combined visual, vestibular, and vibratory effects in forward linear vection	73
4.4.1	Introduction	73
4.4.2	Hypotheses	73
4.4.3	Methods	74
4.4.4	Results	78
4.4.5	Discussion and Conclusion	78
5	Comparative discussion	81
5.1	Factors affecting vection onset latencies and compellingness of vection	83
5.1.1	Perceptual and cognitive influences on vection	86
6	Theoretical frameworks of perception	88
6.1	Indirect and direct theories of perception	88
6.2	Vection viewed through indirect and direct perspectives on perception	89
6.3	Visual-vestibular circuit models of self-motion perception	90
6.4	Maximum likelihood estimation (MLE) and <i>Bayesian</i> sensory integration	92
6.4.1	Bayesian integration and self-motion perception	95
6.5	An integrative view on self-motion perception	97
7	Conclusions	101
7.1	VR-specific recommendations for enhanced vection and future research	101

CONTENTS

References	103
8 Appendix	113
Acknowledgements	113
Ehrenwörtliche Erklärung	114

Zusammenfassung in Deutscher Sprache

In dieser Arbeit wurde untersucht, wie visuelle, auditorische, vestibuläre, und somatosensorische Sinnesinformationen zur menschlichen Wahrnehmung von Eigenbewegung beitragen, und welche multi-sensorischen Interaktionsprozesse dabei eine Rolle spielen. Moderne Virtual Reality (VR) Technologie wurde eingesetzt, um bisher unerforschte multi-sensorische Integrationsprozesse bei der Eigenbewegungswahrnehmung zu untersuchen. Eine leitende Frage war, wie perzeptuelle, multi-sensorische Prozesse auf der einen Seite, und kognitive Faktoren, wie z.B. das Präsenzerleben in VR, die Wahrnehmung von Eigenbewegung beeinflussen, und wie diese Prozesse genutzt werden könnten, um Eigenbewegungssimulationen zu optimieren. Ausgehend von einer detaillierten Literaturübersicht über die relevanten Bereiche wurden in sechs psychophysischen Experimenten multi-sensorische und kognitive Aspekte der Vektions-Illusion untersucht. Vektion ist eine Illusion von Eigenbewegung, die unter bestimmten Umständen bei stationären Beobachtern auftritt, wenn z.B. ein Zug langsam vom Nachbargleis losrollt. Zunächst wurde validiert, dass mit modernen VR Setups vergleichbare Resultate in Vektionsexperimenten zu erzielen sind wie bei klassischen psychophysischen Versuchsaufbauten, wie z.B. optokinetische Trommeln. Im nächsten Schritt wurde gefunden, dass natürliche, realistische visuelle Szenen stärkere Vektion hervorrufen als artifizielle Stimuli, und dass natürliche Szenen ebenfalls zu höheren Präsenzratings führten. Desweiteren wurden signifikante Korrelationen zwischen Vektionsantworten und dem Präsenzerleben in VR gefunden. Die multi-sensorischen Experimente zeigten übereinstimmend, dass Vektion durch konsistente multi-modale Stimulation verstärkt wurde, verglichen zu unimodaler visueller Stimulation. Kohärente multi-sensorische Stimulation verstärkte ebenfalls das Präsenzerleben in VR. Im Einzelnen wurden visuell-auditive, visuell-somatosensorische, visuell-vestibuläre, und visuell-vestibulär-somatosensorische Interaktionen untersucht. Die Befunde aus den Experimenten zeigen auf, dass sowohl perzeptuelle bottom-up Faktoren, wie auch kognitive top-down Faktoren bei der Wahrnehmung von Eigenbewegung in VR eine Rolle spielen. Aufgrund dieser Ergebnisse wird diskutiert, wie Bewegungssimulationen in VR dahingehend optimiert werden können, um den perzeptuellen Realismus von Bewegungssimulationen zu erhöhen. Schließlich werden die Ergebnisse im Kontext verschiedener psychologischer Theorien zur Wahrnehmung diskutiert, und ein integratives konzeptionelles Modell zur Eigenbewegungswahrnehmung wird vorgeschlagen.

Summary

This dissertation investigated how visual, auditory, vestibular, and somatosensory information contribute to the perception of self-motion. Utilising the versatility of modern Virtual Reality (VR) technology, multi-sensory interactions in self-motion perception that have not been investigated previously were studied. One guiding question was how these multi-sensory processes and also cognitive factors, such as spatial presence in VR, contribute to self-motion perception, and how these can be applied in order to improve self-motion simulation. Six psychophysical experiments are reported that investigated multi-sensory and cognitive aspects of thevection illusion, which is an illusion of self-motion in stationary observers. First, it was shown that the results ofvection experiments using VR technology were comparable to studies that used traditional psychophysical apparatus, such as optokinetic drums. Then, it was found that natural, realistic visual stimuli induce strongervection than artificial stimuli, and also increase presence in VR. Following this, the question was asked how consistent multi-sensory information can intensify the perceived magnitude of this illusion. The interactions between visual, vestibular, auditory, and somatosensory stimuli (vibrations) during self-motion perception were systematically studied. The main finding is that the self-motion illusion can be significantly increased by adequate multi-sensory stimulation, compared to a purely visual baseline, and that also cognitive, higher-level factors, such as spatial presence in the simulated Virtual Environment can have a significant influence onvection. Based on these results, it is discussed how perceptually optimised VR environments might improve self-motion simulation. The results of this thesis and several other recent studies in the field of self-motion perception are discussed in the context of different theoretical approaches to multi-sensory integration and self-motion perception. Finally, an integrative conceptual model about self-motion perception is proposed that incorporates the findings from the current thesis and other recent studies on this topic.

1 General introduction

The ability to freely move around in an environment and to navigate and avoid obstacles is of essential importance for all mobile animals. The animal needs to be able to perceive self-motion and object-motion in the environment accurately in order to successfully interact with the environment. Under natural circumstances, it is effortless for humans to perceive and control self-motion and to distinguish this from object-motion (i.e., to perceive whether “I am moving in the world” vs. “Something in the world or the environment is moving”) and to further distinguish this from cases when changes in the visual scene are due to eye-or head-movements, or a combination of any of those possibilities. The brain integrates sensory information from visual, vestibular, auditory, kinesthetic, and somatosensory sensors to fulfill this task, and the perceptual system is very accurate and robust. Only in very rare and special situations, our senses are “fooled”. One such case is the perception of illusory self-motion, a phenomenon also known as “vection”. The vection illusion frequently occurs to train passengers awaiting the departure of their train. If you are seated in a stationary train and see a train from a neighbouring track pull out from the station, you often get the compelling illusion that the train you are sitting in has started to move in the opposite direction of the other train. Another example is when you are standing on a bridge over a river and stare onto the flowing water, you can get the impression after a while that you are moving with the bridge in the opposite direction of the stream. It is particularly interesting that the vection illusion also occurs with same strength and intensity if the observer has objective knowledge that he or she will not move. Especially in cases when the moving visual stimulus covers a large amount of the visual field of view (FOV), the illusion can be very compelling. The sensation of vection can have exhilarating effects, and theme park fun-rides and wide-screen 3D movies (e.g., IMAX theatres) attract many people who want to experience this illusion.

The vection phenomenon provides a unique opportunity to scientifically investigate the processes involved in human self-motion perception, and how the different sensory modalities interact in this perceptual integration process. Most of the studies in this field have so far concentrated on the visual and vestibular modalities, while other senses, such as the auditory or somatosensory systems, have only scarcely been investigated in the context of self-motion perception (Dichgans & Brandt, 1978; Warren & Wertheim, 1990). The current thesis studies how the combined stimulation of multiple sensory modalities influences the perception of self-motion, making use of psychophysical methods and the versatility of modern Virtual Reality (VR) technology. While vection induced by purely visual stimuli can already be very compelling under certain circumstances, only little is known so far about the extent to which other sensory modalities contribute to vection. In most natural situations, multi-sensory information is available to humans during self-motion. It is thus conceivable that multi-sensory interactions may also have an effect on vection. If such interactions can be revealed, this would have important implications both for our theoretical understanding of self-motion perception, and also for applied aspects, such as motion simulator design. One goal of the research presented here is to clarify whether self-motion simulation in VR can be improved in the sense of increasing realism and believability of simulated motion by adequate multi-sensory stimulation. In order to minimise simulation effort and costs, the emphasis will be to move the observer as little as possible physically (i.e., reduce vestibular stimulation to a minimum) but still

enable realistic self-motion perception by means of consistent multi-modal stimulation of the non-vestibular human senses for self-motion.

1.1 Why study self-motion perception? Motivation and aim

While the perception of self-motion seems effortless and trivial to us in everyday life, we only have limited scientific understanding of how the brain accomplishes this task. From all sensory modalities, the visual and vestibular systems are considered the most important senses for self-motion perception, and a large body of research has investigated how those sensory modalities contribute to this perceptual process (Howard, 1986a, 1986b; Warren & Wertheim, 1990). Still, there remain many open questions about the exact nature of visual-vestibular interaction. This becomes clear when one compares, for example, the effort and money spent on high-end motion simulators, such as flight simulators that pilots use for training purposes, with the achieved outcome. Despite large and expensive machines and highly trained engineers who design and control the machines, the motion simulation is often times not very convincing, and in the worst cases, users can get motion sick. Part of the problem may be attributed to the technological limitation of any kind of motion simulator being only able to approximate, but never produce realistic acceleration forces. This is related to the fact that the motion range of the simulator is always limited, so that one has to 'cheat'. This results in perceptual conflicts between the visual and vestibular modalities of the user, which is believed to be the primary cause of the high incidence rate of motion sickness (Reason & Brand, 1975), even though there are also a number of other theories that emphasise different aspects of motion sickness (Stoffregen, Hettinger, Haas, Roe, & Smart, 2000; Treisman, 1977). Also, our understanding of the perceptual and neuronal processes for self-motion perception is still insufficient, and attempts to define the optimal simulation parameters in professional training settings such as pilot training have to rely on trial-and-error procedures. For example, in the field of flight simulation, experienced pilots' judgements are used to calibrate the simulation parameters until the motion to be simulated "feels right" for the experienced user.

The current thesis aims to extend the scope of investigation further to other sensory modalities and see whether the self-motion illusion can be increased by adding somatosensory and auditory cues to the classically investigated visual cues, and at the same time reducing vestibular cues to a minimum in simulations. The idea behind this approach is that the more information about self-motion is presented consistently to multiple sensory modalities, the more realistic the self-motion illusion should become. It is surprising that the multi-sensory nature of self-motion perception has received only little research attention so far in thevection literature. Another important topic of the current thesis is how cognitive, higher-level factors can influencevection. It is proposed thatvection is not only bottom-up driven, but also top-down influences exist, especially in VR environments. Using Virtual Reality technology, a novel self-motion simulation approach is explored that concentrates on both low-level and high-level factors. Ultimately, such an approach, which concentrates more on perceptual realism, rather than physical realism of the stimulus, might lead to a novel, cost-efficient way of simulating self-motion. Another aim of this thesis is to extend our theoretical understanding of self-motion perception. The results of sixvection experiments will be discussed within current theoretical frameworks of multi-sensory interactions, highlighting strengths and weaknesses of different theories and approaches in the domain of self-motion perception.

1 GENERAL INTRODUCTION

1.2 Overview of the human senses contributing to self-motion perception

When we move around in the environment under natural conditions, either by active locomotion, head movements, or by passive transportation in a vehicle, the motion is almost always perceived by multiple sensory systems that provide rich and oftentimes redundant information about how we are moving in space. The visual, auditory, vestibular, kinaesthetic, and somatosensory systems contribute to the perception of self-motion. This section provides a very brief overview over the basic physiological principles involved in human self-motion perception for the most important sensory systems, and their respective relation to thevection illusion. This is by no means a comprehensive review, but is meant to provide a very basic introduction of the sensory physiology underlying self-motion perception. For a comprehensive review, see Kandel, Schwartz, & Jessell (2000).

1.2.1 The visual system

The visual system is stimulated by structured light falling onto the retina at the back of each eye. Light-sensitive sensory cells on the retina called rods and cones react to light by sending neural signals to the primary visual cortex (V1), and from there to higher visual areas. When an observer moves in space, the visual scene that is projected on the retina is characterised by, among other things, systematic and lawful transformations called optic flow. This “optic flow field” (Gibson, 1979) contains information about the environment, such as 3D-layout, relative distances etc., and uniquely specifies the direction and speed of self-motion in a static environment. For example, the focus of expansion (FOE) of an optic flow field specifies the heading of the moving observer, and the gradient of vectors of single points in the flow field contain information about speed and distance relative to the surrounding surfaces (Gibson, 1979). Figure 1 shows optic flow fields resulting from linear forward and curvilinear motions on a panoramic screen.

Several studies have shown that optic flow containing only randomly distributed dots can be sufficient for human observers to perceive self-motion, to determine speed and heading of self-motion accurately, and to discriminate distances travelled (Bremmer & Lappe, 1999; Warren, Kay, Zosh, Duchon, & Sahuc, 2001; Warren & Hannon, 1990). Several neuro-physiological studies have found that motion-sensitive brain areas such as middle temporal (MT) and medial superior temporal area (MST) in the parietal cortex of macaque monkeys are especially sensitive to optic flow stimuli (Bremmer, Duhamel, Hamed, & Graf, 2000), and that these areas show different characteristic patterns of activation for different optic flow stimuli. The same has been found in the homologue brain areas in humans using functional magnetic resonance imaging (fMRI) (Smith, Wall, Williams, & Singh, 2006). Other optical variables that are informative for self-motion are dynamic occlusion and uncovering of surfaces placed at different distances in the environment, and motion parallax, which also contains relative distance information. It is generally believed that optic flow is the most important source of visual information to control locomotion (Warren et al., 2001), even though some authors have recently found evidence that for heading perception, other information, such as the location or direction of a target is used (Rushton, Harris, Lloyd, & Wann, 1998).

Vision has been shown to be the dominant sensory modality for the perception of self-motion. The predominance of the visual system is illustrated by the “classical”vection case which was first reported by Ernst Mach in 1875. In this case, a stationary observer surrounded by a drum



Figure 1: Optic flow for simulated straight (left) and curvilinear (right) motion trajectories projected on a panoramic screen.

that is painted inside with vertical black and white stripes perceives self-rotation after viewing the rotating drum for some 3 to 20 seconds. Even though vestibular signals are missing that would naturally accompany self-rotation, visually inducedvection occurs in earth-stationary observers. In everyday life, self-motion illusions often happen to train passengers when observing a train from the neighbouring track pull out of the station, or to car drivers when a large vehicle moves slowly right next to their own car. One interesting aspect is that even if the observer has objective knowledge that he/she will not move physically, the illusion still occurs with the same intensity as with naïve observers. The occurrence of thevection phenomenon has been interpreted as a result of neural convergence between the visual and vestibular modalities (Howard, 1982). Recently, it has been found that neurons in area VIP in the macaque brain show similar activation when stimulated visually or vestibularly (Bremmer, Klam, Duhamel, Ben Hamed, & Graf, 2002).

Mostvection studies in the literature used abstract visual stimuli to induce illusory self-motion, typically consisting of black and white stripes or random dots distributed in the visual field (Dichgans & Brandt, 1978; Howard, 1986a). In the current thesis, more natural, realistic visual stimuli were used as visualvection stimuli, in order to investigate whether such stimuli can increase thevection illusion, compared to artificial stimuli.

1.2.2 The auditory system

The auditory system transduces sound energy into neural signals. The hair cells in the cochlea in the inner ears transduce mechanical forces into electrical impulses, according to the frequency and amplitude of the stimulus (see Figures 2, 3). From the cochlea, the acoustic nerve transmits the neural signals to the primary auditory cortical area in the superior temporal gyrus.

The auditory system can pick up changes in overall sound intensity, changes in inter-aural time differences, inter-aural intensity differences, and changes in the frequency spectrum of the auditory scene (Blauert, 1997). In this thesis, virtual auditory displays are used to simulate auditory self-motion information. For this purpose, HRTFs (Head-Related Transfer Functions) will be used. Technical details about HRTFs will be provided in section 4.2.

1 GENERAL INTRODUCTION

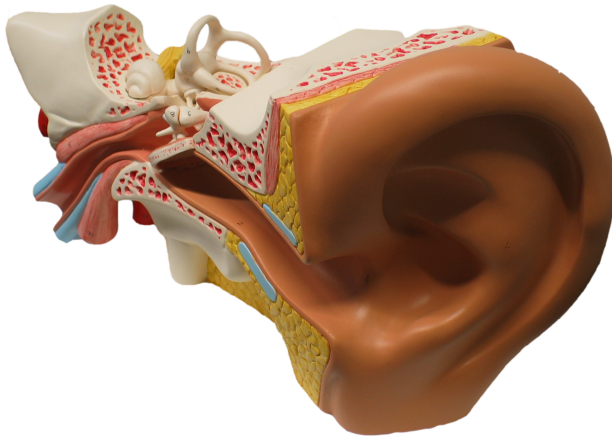


Figure 2: Inner ear housing the labyrinth with auditory and vestibular functional parts.

In the vection literature, there are only very few reports on auditorily induced vection. Dodge (1923) reported that some attempts to induce vection by a moving sound field were successful, even though clear data were not obtained. Later, Lackner (1977) was successful in inducing circular vection in blindfolded participants. Lackner rotated, using different rotational velocities, an array of six hidden loudspeakers around seated participants, playing either white noise at 100Hz or 1000Hz for six minutes. Depending on stimulus combinations, about 25% - 75% of the participants reported the sensation of self-rotation. Notably, nobody reported vection when visual information about the stationary environment was accessible. This finding seems to reflect that auditory information is weighted less than visual information during self-motion perception. Similar results were obtained recently in a study by Larsson, Västfjäll, & Kleiner (2004), who used the HRTF technology which was identical to that used in the current thesis (see chapter 4.2).

1.2.3 The vestibular system

Until the early 19th century, it was commonly believed that changes in blood distribution or felt pressure on the skin give rise to the sensation of body movement. Subsequently, neuroanatomical studies in the 1840s revealed that a labyrinth-like organ in the inner ear gives rise to perception of body movement. Nowadays, it is well established that the human vestibular system registers linear and angular acceleration. The two vestibular organs are located close to the cochlea in each ear, and arranged in a mirror-symmetric fashion to each other in the horizontal plane. Each of the two organs contains five vestibular sensors: the three semicircular canals, which measure angular acceleration, and the two otolithic organs, named saccule and utricle, which measure linear acceleration, stemming from inertial or gravitational acceleration (see Figure 3). All vestibular organs are filled with a fluid called endolymph. The three semicircular canals are approximately orthogonal to each other and referred to as the anterior, posterior, and horizontal canals. The endolymph inside the canals can move relative to each of the surrounding tubes. Each canal houses an “ampulla”, inside which a gelatinous structure called cupula is situated which contains the sensory hair cells. Angular acceleration along the axis of canal orientation causes a deflection of the cupula with the hair cells relative to the endolymph due to inertia of the fluid. This mechanical force is transduced to electric neural signals in the vestibular nerve. The hair cells get activated when deflected in one direction, and

inhibited when deflected in the opposite direction. One important feature of the semicircular canals is that in the frequency range of normal human (and animal) head and body movements, the neural response dynamics are actually proportional to angular velocity, rather than angular acceleration (Fernandez & Goldberg, 1971). This is due to the hydrodynamics of the duct-cupula-endolymph systems. The deflection of the cupula and hence the firing rate of the sensory cells is proportional to angular velocity of the head and not to angular acceleration. This means that even though angular acceleration is the effective physical stimulus acting on the sensory organ, angular velocity is what is sensed during normal human head movements. However, it would be incorrect to state that the vestibular canals register angular velocity directly. Rather, the result of angular velocity is filtered by a dynamical system with high-pass characteristics. The following example illustrates this mechanism: After about five seconds of constant velocity rotation (e.g., on a rotating chair), the endolymph flows back to the resting position, and the sensory cells decrease the firing rate back to resting level, even though the observer is still rotating. The vestibular nuclei circuitry in the brainstem extends this time constant to about 15 seconds. If other sensory information, such as visual information, is missing in this case, the sensation of motion will gradually diminish. It is commonly believed that these high-pass characteristics of the vestibular canals are crucial for the occurrence ofvection as well as aftereffects due to self-motion, such as the rotation aftereffect.

The otolithic organs contain an otolithic membrane filled with gelatinous material consisting of calcium carbonate stones (otoconia). Analogous to the cupula in the canals, the sensory hair cells in the macula of the utricle and saccule are deflected by linear acceleration due to the shear forces and the resulting change in the sensory neuron's membrane potential emits a neural signal. The utricle is orientated in a horizontal plane in the normal upright head position, and the saccule is oriented in a vertical (sagittal) plane. Each otolithic hair cell has a collection of stereocilia that extend into the endolymph and one kinocilium. Deflection of the stereocilia towards the kinocilium increases the firing rate, whereas deflection away from the kinocilium inhibits neural firing. The distribution of the orientation of the cilia in the utricle and saccule show a morphological axis of polarity which defines the respective motion axis to which the sensory cell is sensitive. In contrast to the semicircular canals, the otolith system is not as heavily damped by viscous forces, and the neural response of the otolithic sensory cells are approximately proportional to the linear acceleration of the head (Benson, 1990). Accordingly, vestibular threshold values are commonly reported as accelerations (m/s^2) for linear motion and velocities (deg/s) for rotational motion.

The vestibular system's main functions are the maintenance of balance for upright stance, and the control of a number of ocular reflexes, such as the vestibulo-ocular reflex (VOR), which stabilises a fixated target on the retina during head movements. A dense network of vestibular nuclei in the brainstem and projections to the cerebellum, as well as a few small cortical areas in the insula are involved in fulfilling these functions. Most of the projections from the vestibular nerve lead to the brain stem and to the cerebellum, and only very few to the cortex. This property is also reflected in our everyday life in that normally, we are not consciously aware of vestibular sensations. Only in unusual situations, like stopping abruptly after prolonged spinning, or pathologies in the vestibular organs, do we specifically become aware of vestibular sensations.

1 GENERAL INTRODUCTION

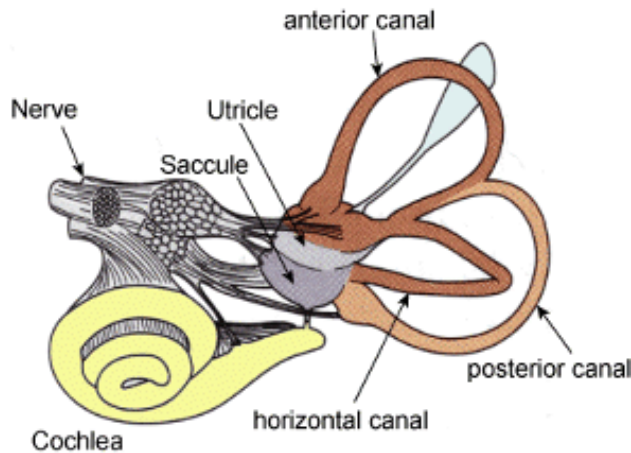


Figure 3: Human labyrinth showing the three semi-circular canals and the otoliths, consisting of utricle and saccule (adapted from Kandel et al., 2000).

1.2.4 The somatosensory system

The somatosensory system includes various sensory receptors, such as sensations of touch or pressure, vibrations, temperature, pain, the sensations of muscle movement and joint position including posture and body movement are subsumed in this system. The first four senses are referred to as cutaneous (skin), the latter as kinesthesia (movement) senses. The mechanical sensors of the cutaneous system register temperature and pressure changes on the skin, whereas spindle receptors signal changes of tension and length of the muscles and tendons, as well as joint positions. Furthermore, the sensation of inner organs are related to somatosensation, for example, the 'gut feeling' when one is accelerated rapidly, like on a roller coaster.

When we are accelerated forward in a vehicle, the somatosensory system registers change of pressure in the back as we are pressed into the seat by inertia, and vibrations from the vehicle. Oftentimes, it is difficult or even impossible to disambiguate vestibular and somatosensory contributions of self-motion (Guedry, 1974). In Sections 4.1 and 4.4, vection experiments will be reported that attempt to exploit this ambiguity of sensory signals by using vibrations in order to minimise simulation effort and cost. The fact that vibrations applied to certain points can induce compelling self-motion illusions has been demonstrated by Lackner (1992). Specifically, when the achilles tendons of a standing observer who is strapped to a wall are vibrated, he or she will perceive paradoxical sensations of constant forward tilting. Similarly, Dichgans & Brandt (1978) reported that prolonged tactile stimulation of the hands or other parts of the body can elicit vection (e.g., when a blindfolded participant touches the inside of a rotating optokinetic drum). Recently, the relative importance of somatosensory information compared to vestibular information has been highlighted by Bringoux, Nougier, Barraud, Marin, & Raphel (2003), who showed that thresholds for the detection of slow orientation changes with respect to gravity are significantly increased if all body pressure cues are eliminated by a pressure suit. Furthermore, it was also found that these thresholds were not increased in bilateral labyrinthine-defective patients if body pressure cues were available, compared to a group of healthy participants with functioning vestibular organs (Bringoux, Schmerber, Nougier, Dumas, Barraud, & Raphel, 2002). Another interesting finding was reported by Lackner & DiZio (2000a)

who observed that in micro-gravity environments (e.g., during weightless period during parabolic flight), a slight touch by some contact surface on the top of the head makes blindfolded people feel upside down, while a touch under the feet will make them feel right side up. As will be further elaborated in the next section, these findings have important implications for psychophysical studies that aim to measure thresholds for the detection of self-motion, most of which explicitly try to eliminate all non-vestibular cues.

Oftentimes, the term *proprioception* is used instead of *somatosensation*, but since proprioception can also include vestibular sensation, I will use the term somatosensation from now on to differentiate between vestibular and somatosensory contributions. The distinction between vestibular and somatosensory sensation will become important in several multi-sensoryvection experiments to be reported later in this thesis (cf. Sections 4.3 and 4.4).

1.3 Human vestibular thresholds for the perception of passive, whole-body self-motion

The section above provided a brief overview about the human senses involved in the perception of self-motion. Among all human senses, the vestibular and visual systems are commonly viewed as the most important senses for self-motion perception. This section will briefly review human vestibular thresholds for the detection of self-motion that have been measured in psychophysical studies. Most of the studies used vestibular stimulation in blindfolded participants. Participants were moved passively in darkness and requested to report onset and/or direction of perceived self-motion. Thresholds have been determined for linear motion along the x-, y-, and z-axes, and thresholds for rotational motion were obtained for rotations around the yaw, pitch, and roll axes.¹

It should be noted that threshold values measured by passive whole-body movements of observers bear some methodological limitations: First, phylogenetically, the sensory systems have evolved to detect self-motion caused mostly by active movement, such as head movements or locomotion, rather than passive transportation, which is a form of motion that phylogenetically speaking became common only very recently. Thus, the measurement method imposes a somewhat artificial situation, in which part of the normally correlated non-vestibular self-motion information (e.g., somatosensation, efference copy, reafference, etc.) is missing. Even in the case of passive whole-body movement of a blindfolded observer, self-motion perception is a multi-modal experience, since somatosensory and auditory cues usually accompany vestibular cues. It is likely that the absence of normally correlated inputs has a considerable influence on the response dynamics of the sensory system in a self-motion detection task. One could assume that under natural conditions, detection thresholds might be lower than those obtained in restricted environments in which studies on vestibular psychophysics have traditionally been conducted. Even though it seems nearly impossible to completely dissociate the vestibular contribution to self-motion perception from non-vestibular cues, most of the researchers in the field try to minimise non-vestibular motion cues as much as possible to measure purely vestibular thresholds. For example, Benson, Spencer, & Stott (1986) had participants wear a large visor and gloves to prevent wind being felt on the skin. However, it is nearly

¹

• In this thesis, I will use a head-centric coordinate system, where the x-axis points straight ahead, the y-axis points sideways, forming an imagined line between the ears, and the z-axis points up vertically.

1 GENERAL INTRODUCTION

impossible to completely eliminate somatosensory information during passive whole-body motion: The mechanoreceptors on the skin and in the muscles react to pressure changes during acceleration (e.g., the feeling of being pressed into the seat during forward acceleration). Even if one uses cushions or pads, it is not clear if this completely eliminates pressure cues. Thus, the threshold values for self-motion always stem from a combined stimulation of vestibular, somatosensory, and sometimes other receptors under natural, ecological conditions. One exception is the case of upright on-axis yaw rotation around the earth-vertical, when a seated observer is rotated on-axis on a turntable at moderate velocities. Here, the observer will notice little change in the force environment during the rotation.

The importance of non-vestibular cues on human perception of body orientation relative to gravity has recently been highlighted by Bringoux et al. (2003). In this study, participants were rotated very slowly at $0.005^\circ/\text{s}$ (i.e., below vestibular threshold) in the pitch dimension in total darkness. Their task was to indicate when they were exactly upright, or to detect a deviation from verticality, or from the horizontal supine position. In a within-subjects design, they were either strapped to the rotating platform, or held completely immobilised in a depressurised body cast. In the latter condition, no noticeable changes of tactile pressure cues due to changes in orientation were available. Results showed that thresholds were significantly higher in all conditions when the body cast was used. In another study with a similar experimental protocol, Bringoux et al. (2002) found no difference in the detection thresholds for pitch and roll rotations between labyrinthine-defective patients who had absolutely no vestibular inputs and healthy participants when strapped to a rotatable chair such that cutaneous pressure cues were available. This indicates that for the patients, the somatosensory cues were sufficient under these conditions to completely compensate for the lack of vestibular information from the otoliths. These findings highlight the importance of non-vestibular contributions to the perception of body orientation in the absence of vision. It is very likely that similar mechanisms are at work during self-motion perception. Consequently, the efforts spent to measure pure vestibular motion detection thresholds are likely to result in thresholds that are higher than human sensitivity to detect passive self-motion under ecological conditions, when somatosensory and vestibular cues are both available.

The threshold values for the detection of passive self-motion reported in the literature vary up to one order of magnitude in extreme cases, from 0.035 to $4.0^\circ/\text{s}^2$ for rotational acceleration around an earth-vertical axis (Clark, 1967). The enormous variability might partly be related to the issues mentioned above. Furthermore, there seem to be considerable individual differences in vestibular motion detection thresholds. One must also keep in mind that self-motion detection thresholds do not only depend on individual vestibular sensitivity, but also on attentional factors, the experimental setup, experimental instructions, and stimulation method (periodic vs. sustained stimulation, frequency and duration of stimulation, etc.). With periodic stimulation, lower thresholds are found compared to sustained stimulation (Berthoz & Droulez, 1982). One reason for this seems to be that periodic stimuli are predictable. For instance, if an observer is moved periodically forward and backward, the motion amplitude and turning points become predictable due to the regularity of stimulation. Another typical finding is that thresholds depend on the duration and frequency of stimulation: At frequencies above 1 Hz, vestibular sensitivity increases, whereas below 0.1 Hz, sensitivity is very low (Benson, Hutt, & Brown, 1989). This pattern of findings seems to reflect the hydrodynamics and sensory processing of the vestibular system (cf. section 1.2.3 on page 12). It is also known

that the rate of change of acceleration (“jerk”) is a crucial parameter for the detection of self-motion (Guedry, 1974). During strong changes in acceleration, it is likely that somatosensory contributions are increased relative to vestibular cues in self-motion detection. In sections 4.3 and 4.4, two vection experiments will be reported that try to use jerks in order to increase vection in VR.

Given these issues, psychophysically obtained self-motion detection thresholds always need to be interpreted in the context of the various limiting factors listed above. One desirable alternative could be to investigate vestibular thresholds using signal detection theory (SDT). According to SDT, the concept of an absolute sensory threshold per se is inadequate: Considering that neural processing always contains background noise, such as spontaneous neural activity, sensory signals need to be filtered out to be “detected” (Green & Swets, 1966). Thus, the ratio between signal and noise is one determining factor, and another equally important factor is the subjective criterion of the observer whether a certain level of neural activity is “interpreted” as a signal. The subjective criterion varies depending on the situation: For example, if it is very important not to miss even only one signal, then the observer would rather report having noticed a movement, even if in doubt, and thus use a very liberal criterion, taking into account possible false alarms. In contrast, for example, if the penalty for a false alarm is very high, the observer would use a very conservative criterion and thus put up with many missed signals. The problem with the SDT method is that threshold measurements require a large number of trials, presented with different background noise conditions, probabilities of motion, and instructions, which may be why so far no attempts have been made to measure vestibular sensitivity using SDT (Guedry, 1974). In summary, it should be clear by now that the concept of an “absolute” threshold for self-motion is problematic.

The limitations of the vestibular system described above usually do not pose any problems for humans, because under natural conditions, we hardly experience sustained accelerations when we locomote. Things are different, though, in special circumstances, such as aviation. There are a number of well documented somato-gravic illusions that pilots experience, such as the gravito-inertial illusion. Here, pilots who are accelerated at very high rates in fighter jets feel like they are pitched upwards, even though the aircraft is nearly level. This is due to the fact the resultant change of the gravito-inertial force vector acting on the vestibular apparatus is misinterpreted as being caused by pitch rotation and not by forward acceleration. This becomes especially dangerous under bad visibility conditions. If pilots pull down the aircraft’s nose to compensate for the illusory pitch rotation, the aircraft crashes into the ground or into the sea. It is estimated that about 10% of the accidents in military aircraft aviation and about 35% in general aviation are caused by spatial disorientation of the pilots (Benson, 1990; Leibowitz, 1988).

In the following, experimental reports from the literature on detection thresholds for passive whole-body self-motion will be briefly reviewed.

1.3.1 Thresholds for vestibular detection of rotational motion

A good review on thresholds for upright yaw rotations is provided by Guedry (1974). In a more recent study by Benson et al. (1989), participants were blindfolded and seated on a turntable. Care was taken to eliminate other potential motion cues from other sensory modalities by playing white noise via headphones, covering all skin parts to eliminate perception of wind, seating the participants on cushions, and restrain relative body- and head movement using a five-point-harness and pillows. During yaw rotations, participants sat upright and cross-legged on the turntable. Stimulation time

1 GENERAL INTRODUCTION

was held constant at 3.3 seconds with a cosine bell velocity trajectory. For pitch rotations, participants lay on their right side, and for roll motion, they lay supine. This was done in order to limit vestibular stimulation to the canals only, since during pitch and roll movements in upright orientation, the otoliths are also stimulated. As mentioned in the subsection above, this method produces an artificial stimulus condition that does not correspond to natural sensory stimulation under ecological conditions. For technical reasons, the axis of rotation was always 20 cm from the centre of the head. Using a staircase method, threshold velocities of $1.58^{\circ}/s$ were found for yaw rotations, and $2.07^{\circ}/s$ for pitch, and $2.04^{\circ}/s$ for roll rotations. Interestingly, in a repeated threshold measurement of yaw rotations with exactly the same method and the same participants, a lower threshold of $1.20^{\circ}/s$ was found, even though no feedback was provided. This finding seems to reflect that the human sensitivity to self-motion is able to adapt to unusual conditions of sparse stimulation. Furthermore, thresholds were found to depend on stimulus duration such that they increased significantly with increased stimulus duration, up to $5.25^{\circ}/s$ at 20 seconds duration.

1.3.2 Thresholds for vestibular detection of linear motion

Thresholds for linear motion were reviewed by Berthoz & Droulez (1982), and Benson et al. (1986) reported a more recent study. Using the identical experimental setup as in the above mentioned experiment, Benson et al. (1986) investigated thresholds for x-, y-, and z-axis linear motion. Using a cosine bell velocity trajectory with 3 seconds duration, mean threshold acceleration values of $\pm 0.0625 \text{ m/s}^2$ for x-axis, $\pm 0.0569 \text{ m/s}^2$ for y-axis, and $\pm 0.154 \text{ m/s}^2$ for z-axis linear motion were found. For x- and y-axis motion, participants were tested upright, aligned with gravity, whereas for z-axis motion, they lay supine, which means that the otoliths were in an unusual orientation to gravity. It is possible that the body orientation towards gravity influences thresholds. However, Lepecq, Giannopulu, Mertz, & Baudonniere (1999) found comparable threshold values for z-axis motion using an “elevator-seat”, which kept head and world coordinates in alignment. They used a trapezoidal velocity profile, with 20 ms acceleration, 5 seconds constant velocity, and 20 ms deceleration. Individual values between 0.128 and 0.363 m/s^2 for upward accelerations were found, with a mean value of 0.216 m/s^2 in 23 adult participants. Thus, in the natural upright orientation and with a different acceleration profile, slightly higher thresholds were found compared to Benson et al. (1986).

The characteristics of vestibular self-motion sensation discussed above are thought to be closely related to the vection phenomenon, which will be reviewed in more detail in the following section.

2 General characteristics of vection

As already noted earlier, the visual and vestibular modalities are generally believed to be the most important senses for self-motion perception. The fact that visual information plays an important role in self-motion perception is nicely illustrated by the fact that visual stimuli alone can induce illusions of self-motion in stationary observers, a phenomenon commonly referred to as vection. This section provides an overview over the most relevant findings about vection. For more comprehensive reviews, see Dichgans & Brandt (1978) and Howard (1986a). Also, Hettinger (2002) gives an overview about recent VR-related work on vection.

The self-motion illusion has been investigated for more than a century. Ernst Mach (1875) was the first to report that stationary observers perceive illusory self-motion after viewing a uniformly moving visual stimulus that covers a large part of the visual field for a prolonged time. The direction of perceived self-motion is always in the opposite direction of the visual stimulus. The term “vection” is attributed to Fischer & Kornmüller (1930) and Tschermak (1931).

Vection has been shown to occur for all motion directions and along all motion axes. *Linear vection* can occur for forward-backward, up-down, or linear sideways motion (i.e., along linear paths of the x-, y-, and z-axis). *Circular vection* can be induced for upright rotations around the vertical (yaw) axis, and similarly for the roll axis (frontal axis along the line of sight, like in a “tumbling room”), as well as around the pitch axis (an imagined line passing through the body from left to right). The latter two forms of circular vection are especially nauseating, since they include a strong conflict between visual orientation and gravitational cues. During pitch and roll vection, observers have the paradoxical sensation of constant rotation and at the same time a sustained deviation from the gravitational vertical (Howard, 1982).

The best documented type of vection is circular vection around the yaw-axis (Dichgans & Brandt, 1978). In this special situation where the observer perceives self-rotation around the vertical axis, there is no effect of gravitation that might interfere, since the body orientation always remains aligned with gravity during motion. In classical circular vection experiments, participants are seated inside a rotating drum painted with black and white vertical stripes inside, a device called an optokinetic drum. After the drum starts to rotate, the latency from stimulus onset until the participant reports perceiving self-rotation is measured. Initially, observers “correctly” perceive object motion of the drum and themselves as stationary. After some seconds (1-14, depending on different stimulus parameters), the observer starts to feel self-rotation, with a gradual build-up of self-rotation velocity. At the same time, an inverse gradual decrease of drum velocity is perceived (Wong & Frost, 1978). Vection is said to be *saturated* when the observer perceives only self-motion in an earth-stationary drum, and most observers reach this state after prolonged visual stimulation. The time it takes from onset of visual motion until the observer experiences self-motion is called vection onset latency, and this measure is one of the main dependent measure in vection research.

Apart from the onset latency of the illusion, the strength of the illusion is measured either by the duration of the illusion, or by some indication of perceived speed or intensity of rotation (e.g., by magnitude estimation; Stevens, 1957), or by letting the participant physically counter-rotate the seat until the sensation of self-rotation vanishes. When drum rotation stops abruptly, observers feel a reversal of turning direction. This negative aftereffect can last up to a minute after prolonged

2 GENERAL CHARACTERISTICS OF VECTION

exposure to a circular vection stimulus (Dichgans & Brandt, 1978). A few studies measured just the mere occurrence of vection (i.e., yes-no responses, % vection responses).

Similar characteristics of the timecourse of vection build up have been reported for the case of linear vection by Berthoz & Droulez (1982). They report that after a prolonged initial onset latency of up to 20 seconds on the very first trial, in consecutive trials, on average, most participants report vection after 1 to 2 seconds at linear image velocities between 0.2 and 1 m/s. The most common method to induce linear vection in a laboratory is to use two monitors or screens facing each other, with the participant's head between the two screens. The sagittal plane of the head is aligned parallel to the screens so that they cover a large part of the peripheral visual field of the observer (Berthoz, Pavard, & Young, 1975; Johansson, 1977; Lepecq, Jouen, & Dubon, 1993). Optic flow stimuli specifying linear motion induce linear vection in the opposite direction of visual motion. Johansson (1977) reported the "elevator illusion" where observers feel upward or downward motion while fixating a stationary target in front while viewing optic flow stimuli in the periphery. For some observers, this illusion occurred even if they looked outside the building through a window instead of fixating the target. For those observers, the whole building and environment seemed to move upward or downward.

It has also been reported that the direction of linear motion seems to affect onset latencies: Berthoz et al. (1975) noted that the thresholds for backward vection were lower than for forward vection. They assumed that this result reflects normal human behaviour (i.e., we perceive forward motion very often and are thus very used to this), but since we hardly ever see linear backward motion, our sensitivity might be lower. Giannopulu & Lepecq (1998) compared onset latencies for forward-backward and upward-downward vection and found that the latter produced shorter onset times. They argued that this might be due to the fact that visual-vestibular conflict is less pronounced for vertical motion than for forward- and backward motion. In another study, Kano (1991) found that onset latencies for vertical linear vection were significantly shorter than for forward and backward vection when observers were seated upright, but this difference disappeared when they observed the identical stimuli in a supine position. It is possible that this might be related to different utricular and macular sensitivities, but it remains unclear how retinal and gravitational reference frames interact during vection.

These general vection characteristics are commonly interpreted as a result of the convergence of visual and vestibular neural processing (Dichgans & Brandt, 1978; Howard, 1982). While the optokinetic response has low-pass characteristics and is sensitive to constant velocity stimuli, the vestibular response has high-pass characteristics and is sensitive to acceleration. The visual and vestibular systems thus have a complementary frequency response and, under natural conditions, they give accurate estimates about self-motion. As was briefly outlined in section 1.2.3, the response dynamics of the vestibular system play an important role for the vection phenomenon. First, let us consider how motion is perceived in the case of real physical displacement of an observer in darkness who is seated on a rotating chair. At rotation onset, the hair cells in the cupula/macula get deflected, and the resulting vestibular neural signals let the observer perceive self-rotation. After some 10 seconds of constant velocity motion, the hair cells return to their resting position, and the vestibular signal decays. Without any concurrent visual or other cues about self-motion, motion will no longer be perceived, even though the observer is still rotating in the dark. Now, let us consider the case of visually induced vection. When, for example, the optokinetic drum starts to rotate, the observer

2 GENERAL CHARACTERISTICS OF VECTION

initially “correctly” perceives drum rotation because of the missing vestibular signal. After some time of constant velocity drum rotation, a vestibular signal would gradually decay during real body rotation, and the moving visual stimuli becomes more dominant and induces vection. It makes sense that in the case of vection, the visual stimulus gets increasingly dominant according to the decay of vestibular signals, since it is the only remaining reliable source of information about self-motion. Since we evolved in a world where the surrounding visual environment usually does not move, it is reasonable to perceive this uniform environmental motion to be caused by self-motion. This is the common notion used to explain the onset delay and timecourse of gradual build up of vection.

One typical finding in vection research is that there are large inter-individual differences for onset delays and perceived vection intensity. One possible explanation for this was put forward by Lepecq et al. (1999) who found that participants with lower vestibular thresholds to detect motion (i.e., vestibularly highly sensitive persons), report the onset of vection later than participants who are vestibularly less sensitive. In two experiments, they measured vection onset latencies and also vestibular thresholds for self-motion and found a negative correlation, such that the higher the vestibular thresholds, the lower the vection onset latencies. Their explanation is that vestibularly sensitive persons register a stronger and longer lasting visual-vestibular conflict than vestibularly less sensitive persons.

2.1 Visual parameters affecting vection

In this section, the most important findings from the literature about visual factors that influence the vection illusion will be reviewed.

2.1.1 Size and retinal location of the stimulated visual FOV

Using an optokinetic drum, Brandt, Dichgans, & Koenig (1973) found that visual stimuli covering large FOVs induce stronger circular vection and shorter onset latencies, and that stimulation of the entire FOV results in strongest vection. Limiting the FOV systematically increased onset latencies and reduced vection intensities. It was also found that a stimulus with a FOV of 30° viewed in the periphery of the visual field induces vection at levels comparable to full field stimulation, whereas the identical 30° stimulus viewed in the central FOV did not induce vection. This observation led to the conclusion that the visual periphery dominates self-motion perception, and that the central FOV is more important for the perception of object motion. However, this view was later challenged by Andersen & Braunstein (1985) and Howard & Heckmann (1989). Andersen and Braunstein showed that a centrally presented visual stimulus showing an expanding radial optic flow pattern that covered only 7.5° was sufficient to induce forward linear vection when viewed through an aperture. In pilot experiments for this study, the authors had found that in order to perceive self-motion under these conditions, participants had to believe that they were in an environment where they could actually be moved in the direction of perceived vection. Accordingly, participants were standing on a movable booth and looked out of a window to view the optic flow pattern. This observation is very interesting, since it indicates a cognitive influence on vection. Issues related to cognitive influences on vection will be dealt with in greater detail in section 3.2.

Howard & Heckmann (1989) proposed that the reason Brandt et al. (1973) found a peripheral dominance was due to a confound of misperceived foreground-background relations: Howard and

2 GENERAL CHARACTERISTICS OF VECTION

Heckmann found that when the moving stimulus is perceived to be in the foreground relative to a static background, it will not induce vection. They suspected that this might have happened to the participants in the Brandt et al. study, and they provided support for their claim in an experiment by placing the moving visual stimulus stereoscopically either in front or in the back of the plane of the rotating drum. Their data showed that if a central display is perceived to be in the background, it will induce vection. Thus, the original idea of peripheral dominance for self-motion perception should be reassessed (for more on this issue, see Section 2.1.2). However, the general notion that larger FOVs are more effective at inducing vection does hold true. For virtual reality applications, this means that larger displays are better suited for inducing a compelling illusion of self-motion.

Different types of displays have been used to induce linear vection: In the classical studies by Berthoz et al. (1975) and Johansson (1977), two peripheral displays were used (see section 2). With these displays, rather short vection onset latencies of about 2-3 seconds are reported. Other studies that used a frontal projection screen found onset latencies that were considerably higher (Palmisano, 1996). This might be attributed to the fact that the structure of optic flow for linear forward motion is such that there is much more motion in the far periphery than in the central part, where motion around the focus of expansion (FOE) in the optic flow field is minimal.

2.1.2 Foreground-background separation between moving and stationary surfaces

As mentioned in the subsection above, a moving stimulus must be perceived to be in the background in order to induce vection. A number of studies have investigated this effect (Howard & Heckmann, 1989; Howard & Howard, 1994; Ohmi, Howard, & Landolt, 1987), and all have found a consistent effect of the depth structure of the moving stimulus on vection. Only moving stimuli that are perceived to be in the background will reliably induce vection. If a stationary object is seen behind a moving stimulus, no vection will occur (Howard & Howard, 1994). Dichgans & Brandt (1978) have proposed that the very occurrence of vection might be due to our inherent assumption of a stable environment: When we see a large part of the visual scene move in a uniform manner, especially if it is at some distance away from us, it is reasonable to assume that this is caused by ourselves moving in the environment, rather than the environment moving relative to us. The latter case occurs only in very rare cases in every day occasions, such as in the train illusion, where our brain is fooled into perceiving self-motion. It has been shown that stationary objects in the foreground that partly occlude a moving background will increase vection (Howard & Howard, 1994), and that a foreground that moves slowly in the opposite direction to the background will also facilitate vection (Nakamura & Shimojo, 1999). It has also been shown that if two depth-separated random dots move in orthogonal directions, the slowly moving foreground can induce vection in it's motion direction (Nakamura & Shimojo, 2000). This illusion was termed *inverted vection* because of the reversed direction of perceived motion direction, which is normally always opposite to the direction of the stimulus. The authors hypothesised that a mis-registered eye-movement signal might cause this illusion (see also Subsection 2.1.5).

In Section 3.2, I will present new data from experiments conducted for the current thesis that extend this finding and discuss implications for self-motion simulation from an applied perspective.

2 GENERAL CHARACTERISTICS OF VECTION

2.1.3 Spatial frequency content of the moving visual stimulus

Diener, Wist, Dichgans, & Brandt (1976) observed that moving visual patterns with high spatial frequencies are perceived to move faster than similar visual patterns with lower spatial frequencies, even though both move at identical angular velocities. This means that a vertical grating pattern with, for example, 20 contrasts (such as black and white stripes) per given visual angle will be perceived to move faster than a different pattern with only 10 contrasts within the same visual angle. Palmisano & Gillam (1998) revealed that there is an interaction between the spatial frequency of the presented optic flow and the retinal eccentricity in vection. While high spatial frequencies produce most compelling vection in the central FOV, peripheral stimulation results in stronger vection if lower spatial frequencies are presented. This finding contradicts earlier notions of peripheral dominance (see Section 2.1.1) and shows that both high- and low frequency information is involved in the perception of vection, and that mechanisms of self-motion perception differ depending on the retinal eccentricity of the stimulus. In the context of VR, this implies that fine detail included in the graphical scene may be beneficial in the central FOV, and that stimuli in the periphery of the user's gaze might be rendered at a lower resolution and fidelity.

2.1.4 Optical velocity and direction of the visual stimulus

Howard (1986a) and Brandt et al. (1973) reported that the intensity and perceived speed of self-rotation in circular yaw vection is linearly proportional to the optical velocity of the optokinetic stimulus up to values of approximately $90^\circ/\text{s}$. As detailed in section 2.1.3, the perceived velocity interacts with the spatial frequency of the stimulus. While Brandt et al. (1973) report that the vection onset latency for circular vection is more or less constant for optical velocities up to $90^\circ/\text{s}$, others report that very slow movement below vestibular threshold results in faster vection onset (Wertheim, 1994). This might be due to differences in their methods. For instance, Brandt et al. (1973) accelerated the optokinetic drum in darkness to a constant velocity and measured the vection onset latency from the moment the light was switched on, whereas the studies where faster vection onset was found for slow optical velocities typically used sinusoidal motion with the drum always visible during visual acceleration and deceleration.

When observers are exposed to visual stimuli rotating around the roll or pitch axis, a paradoxical perception of both continuous rotation and at the same time a sustained, constant deviation from the vertical orientation occurs (Howard, 1982). The conflicting gravitational cues prevent observers to perceive full 360° rotations, whereas in zero gravity, observers do feel complete pitch and roll rotations (Lackner & DiZio, 2000b). For roll rotations, vection was found to saturate after about 18 seconds, and a constant tilt of about 15° on average was reported (Howard, 1982). These cases in which gravitational cues are in conflict with visual cues are known to be very nauseating.

For the case of linear vection, it has also been found that the magnitude of vection depends on the velocity of the moving visual stimulus. Berthoz et al. (1975) found a more or less linear relationship between perceived self-velocity and stimulus velocity, up to a certain level where an upper limit of the sensation of vection was reached. Interestingly, it has been found that thresholds for backward and downward vection are lower than for forward and upward vection (Berthoz & Droulez, 1982). The authors assumed that this result reflects normal human behaviour in that we perceive forward motion very often and are thus very used to this, but since we hardly ever see linear backward motion,

2 GENERAL CHARACTERISTICS OF VECTION

our sensitivity might be higher. In another study, Kano (1991) found that onset latencies for vertical linear vection is significantly shorter than for forwards and backwards vection when observers are seated upright. This difference disappeared, however, when they observed the identical stimuli in a supine position. In the supine position, vection towards the own feet (retinally “down”) and towards the back (retinally “back”; gravitationally “down”) was found to develop faster than vection towards one’s own head (retinally “up”) and towards the ceiling (gravitationally “up”). It is possible that this might be related to different utricular and macular sensitivities of the vestibular system, but it remains unclear how retinal and gravitational reference frames interact during vection.

2.1.5 Effect of eye-movements on vection

There are a number of ocular reflexes that compensate for retinal motion (also known as “retinal slip”) during self-motion. They all have in common that they move opposite to the direction of self-motion in order to keep the retinal image more or less static. Different types of eye reflexes use visual, vestibular, and proprioceptive information. If the head undergoes rotational or linear movement, the vestibulo-ocular reflex (VOR) stabilises a fixated target in the environment using visual and vestibular input signals from the semicircular canals and/or otoliths to cancel the retinal slip. Under pure visual motion stimulation, the optokinetic reflex (OKR) elicits eye movement patterns that track individual elements in the scene passing by with a slow-phase nystagmus and jump forward with a saccade to fixate a new element. This reflexive eye movement occurs in natural situations (e.g., when one looks out of the window from a moving bus).

It has long been recognised that eye movements influence the vection illusion. Mach (1875) was the first to report that if observers fixate a stationary target, vection will develop faster than when the eyes follow the stimulus. This finding has been replicated many times (Brandt et al., 1973; Becker, Raab, & Jürgens, 2002b; de Graaf, Wertheim, & Bles, 1991). It is also known that the intensity of vection is increased when the eyes fixate a stationary target. Becker et al. (2002b) investigated this effect in an optokinetic drum by systematically varying the instructions of how to watch the stimulus. In one condition, participants had to follow the stimulus with their eyes, thus not suppressing the optokinetic nystagmus (OKN). In other conditions, they either had to voluntarily suppress the OKN by fixating a stationary target that was presented on top of the moving stimulus, or stare through the moving stimulus. Results showed that vection developed fastest when the eyes fixated a stationary fixation point, and slightly less rapidly when participants stared through the stimulus. Vection took longest to develop when the eyes moved naturally, following the stimulus motion. The authors concluded that this effect is due to the fact that the retinal signal (retinal slip) is processed with higher gain than the efferent eye movement signals (i.e., efference copy and proprioceptive information about eye movements). This explanation is related to the Aubert-Fleischl phenomenon, where the subjective velocity of a moving target is slower when pursued with the eyes than when the eyes are stationary (Aubert, 1886). This is commonly interpreted such that eye-movement signals about eye velocity are underestimated by the brain. An interesting observation from (Becker et al., 2002b) is that the mere intention to visually track one moving object seems to reduce vection, compared to a condition where observers were instructed to stare through the moving pattern.

Another interesting finding about possible influences of eye-movements on vection was reported by Nakamura (2004) who found that a slowly moving foreground stimulus moving in front of a fast background stimulus moving in orthogonal direction can induce *inverted* vection. In this case, vec-

2 GENERAL CHARACTERISTICS OF VECTION

tion is perceived in the same direction as the foreground stimulus (see subsection 2.1.2). Nakamura concluded that this effect seems to be caused by a mis-registered eye-movement, which results from OKN-suppression when observers fixate a stationary target.

2.1.6 Stereoscopic cues

It has been found that stereoscopic cues specified by binocular disparity can influence both linear and circular vection. Perceived vection speed for forward linear vection is increased and vection onset latency is reduced if optic flow is presented stereoscopically (Palmisano, 1996, 2002). When forward translation through a cloud of dots is simulated stereoscopically, dynamic stereoscopic information is available, since the disparities of optic flow elements increase exponentially as the simulated dots approach the observer. Also, inter-ocular velocity differences of the dots are available as additional stereoscopic information about self-motion. These factors were shown to be effective when they were added to non-stereoscopic optic flow stimuli that contained changing size cues. In a different study, Wist, Diener, Dichgans, & Brandt (1975) found that circular vection around the yaw axis is enhanced if the perceived distance of the rotating optokinetic drum is increased using Pulfrich filters, even though in this case the angular velocity of the rotating stimulus is independent of stimulus distance. The authors concluded that in this special situation of yaw-axis rotation, the brain erroneously applies the velocity constancy mechanism, which holds true only for linear motion between observer and elements in the environment, but not for circular motion.

2.2 Non-visual cues for vection

The vast majority of the vection literature has been concerned with visually induced vection. The question whether other sensory modalities can also trigger vection has received only little attention. For example, Lackner (1977) demonstrated that a rotating sound field generated by an array of loudspeakers can induce circular vection and nystagmus in blindfolded participants. Recent experiments demonstrated that auditory vection can also be induced by headphone-based auralization using generic head-related transfer functions (HRTFs), both for rotations and translations (Larsson et al., 2004; Vålstam, Larsson, Västfjäll, & Kleiner, 2004, 2005). Several factors were identified as being effective for inducing auditory vection: The realism of the acoustic simulation and increasing numbers of sound sources were found to enhance vection. Larsson et al. (2004) observed also a cognitive or top-down influence: Acoustic landmarks, which are sound sources that are typically associated with stationary objects (e.g., church bells) demonstrated a higher capacity for inducing auditory vection than artificial sounds (e.g., pink noise) or sounds that are typically related to moving objects (e.g., foot steps). It is important to keep in mind, however, that auditory vection occurs only in about 25-70% of blindfolded participants and is far less compelling than visually induced vection, which can be indistinguishable from actual motion (Brandt et al., 1973). Thus, auditory cues alone seem to be a relatively weak cue for self-motion perception.

Furthermore, it has been reported that tactile stimuli can induce vection: Dichgans & Brandt (1978) found that if a blindfolded person touches the inside of a rotating drum with the hands or with a large part of the body, vection can be elicited. Also, if one arm is passively rotated in darkness around an observer, vection and nystagmus can be induced (Brandt, Büchele, & Arnold, 1977). These findings indicate that vection can be induced by tactile and somatosensory stimulation. Re-

2 GENERAL CHARACTERISTICS OF VECTION

ports on vection induced by modalities other than vision, audition, or touch are sparse. One exception is the “podokinesthetic” self-motion illusion where blindfolded participants walk on the edge of a rotating turntable. They hold a handle and walk in order to stay in place (Becker, Nasios, Raab, & Jurgens, 2002a). After stepping in place for some while, participants get the illusion of walking in circles.

2.3 Multi-modal aspects of vection

The main focus of the current thesis is the investigation how multiple sensory modalities interact during self-motion perception. As already pointed out in section 1.2, virtually every sense is stimulated when we actively move in a natural environment. It is thus surprising that little work has been devoted to the multi-modal aspects of self-motion perception. Only some aspects of the interaction between visual and vestibular modality have been investigated, which reflects the fact that these two senses are considered the most important senses for self-motion (Howard, 1986a; Wong & Frost, 1981). The aim of the current thesis is to broaden the view by extending the analysis to more than two sensory modalities. In this thesis, vection experiments will be reported that investigated the contributions of visual, auditory, somatosensory, and vestibular information to the vection illusion, and the multi-sensory interactions between those modalities during self-motion perception. These experiments were conducted using VR technology.

The question of how sensory information from different modalities is fused or combined into one “percept” has been a matter of debate over centuries. Recently, the view that this sensory integration process can be best understood within a Bayesian approach which uses maximum-likelihood estimation (MLE) has become very popular in the literature (Knill & Richards, 1996; Ernst & Banks, 2002; Ernst & Bühlhoff, 2004; Alais & Burr, 2004). The main proposition of this approach is that signals from all sensory modalities are weighted according to each modality’s reliability: The higher the variance of a sensory signal, the lower the reliability (reliability = inverse variance), and accordingly, a lower weight will be assigned in the multi-modal combination process. The linear sum of all weighted sensory inputs result in the final percept. While this approach has been applied to visual-haptic and visual-auditory integration, recent attempts to model self-motion perception were only partly successful. Furthermore, it has been found that statistically optimal fusion only occurs in a narrow parameter space. Rosas, Wagemans, Ernst, & Wichmann (2005) found that texture and haptic cues were not combined in MLE manner when a large parameter space was tested.

Chapter 6 will give a more detailed discussion of the literature, including experimental results from this thesis.

2.4 Higher-level influences on vection

The previous subsections summarised research that investigated the influence of low-level, bottom-up parameters on illusory self-motion perception. In the following, several studies will be reviewed which indicate that not only low-level factors, but also cognitive, higher-level factors might play an important role in the perception of illusory self-motion, especially in a VR context. It was proposed in 1978 that the very occurrence of vection might be linked to our inherent assumption of a stable environment (Dichgans & Brandt, 1978). Perhaps this is why the background of a vection-inducing stimulus is typically the dominant determinant for vection and modulator of the strength of vection

2 GENERAL CHARACTERISTICS OF VECTION

(Howard & Heckmann, 1989; Ohmi et al., 1987). In daily life, the more distant elements comprising the background of visual scenes are generally stationary and therefore any retinal movement of those distant elements is more likely to be interpreted as a result of self-motion (Nakamura & Shimojo, 1999). In the study by Andersen & Braunstein (1985) referred to in subsection 2.1.1, the authors remark that pilot experiments had shown that in order to perceive linear forwardvection, participants had to believe that they could actually be moved in the direction of perceivedvection. Accordingly, participants were asked to stand in a movable booth and to look out of a window to view the optic flow pattern. Similarly, a study by Lackner (1977) demonstrated that circularvection can be induced by a rotating soundfield in blindfolded participants when participants were seated on a chair that could be rotated. Note that by making participants believe that they could be moved physically, Andersen and Braunstein were able to elicitvection with a visual FOV as small as 7.5° , and Lackner was able to inducevection by presenting a moving sound field. Under these conditions of limited or weak sensory stimulation, cognitive factors seem to become more and more relevant. This idea is consistent with the Bayesian framework of perception, which states that sensory information is combined with cognitive, higher-level information called *priors*. The higher the sensory reliability, the less influence priors will have, and the lower sensory reliability, the higher the influence of priors will have. A more detailed discussion of the Bayesian approach to perception will be provided in Section 6.

It is possible that cognitive factors generally might have a much stronger influence onvection than previously believed. For example, the cognitive manipulations in the experiments might not have been powerful enough, or sensory stimulation might have been so strong that ceiling level was already reached, which is likely to be the case in an optokinetic drum that covers the full visible FOV. One study that explicitly investigated cognitive influences on linearvection was done by Lep-ecq, Giannopulu, & Baudonniere (1995). They found that seven year old children perceivevection earlier when they are previously shown that the chair they were seated on can physically move in the direction of simulated motion, even though this never happened during the actual experiment. Interestingly, thisvection-facilitating influence of pre-knowledge was not present in eleven year old children. A study by Kitazaki & Sato (2003) showed thatvection can also be modulated by attentional factors. When observers were asked to specifically pay attention to one of two simultaneously presented upward and downward optic flow fields of different colours, the non-attended flow field was found to determinevection direction. The participants were asked to fixate on a fixation point, so it is unlikely that this effect can be explained by different eye movements patterns, and also not on foreground-background relationships of the optic flow stimulus.

Another interesting effect was revealed by Wertheim, Mesland, & Bles (2001) who demonstrated a clear cognitive influence on the vestibular perception of linear forward-backward motion. They had blindfolded participants judge what kind of movement they perceived after they were moved forward and backward sinusoidally at different frequencies and amplitudes on a linear sled. One group of participants who were completely naïve to what kind of motion device they were seated on reported both sensations of tilt as well as forward-backward motion. In another group, all participants had seen the linear sled before climbing onto it, and 100% of their responses were forward-backward. This finding indicates that prior knowledge about what kind of motion is possible on the motion device helps to disambiguate otolith signals about linear acceleration. Since the otoliths are inertial sensors, they can not physically distinguish whether a change of the gravito-inertial force vector

2 GENERAL CHARACTERISTICS OF VECTION

is caused by forward acceleration or tilt of the body, since they can only detect the change of the resultant gravito-inertial force vector (see Section 1.2.3). This is known as tilt-translation ambiguity (Merfeld & Zupan, 2002), and it is not too surprising that in such cases of sensory ambiguity, other supplementary information, such as prior knowledge about possible physical motion direction, improves perceptual performance. However, one has to consider the possibility that this effect was due to a reporting bias and not a perceptual effect.

More than ten years ago, Wann & Rushton (1994) stressed the importance of an ecological context and a naturalistic optic array for studying self-motion perception. Vection research has, however, traditionally used abstract stimuli like black and white striped patterns or random dot displays, and only recently have more naturalistic stimuli become more common in self-motion research (van der Steen & Brockhoff, 2000; Riecke, Schulte-Pelkum, Avraamides, von der Heyde, & Bülthoff, 2006; Mohler, Riecke, Thompson, & Bülthoff, 2005). One might expect that more natural looking stimuli have the potential of not only inducing stronger vection, but also more intensive feelings of “presence” - the feeling of being and acting in the virtual environment (see also the following subsection). Consequently, it seems appropriate to consider possible interactions between presence and vection. Even though presence is typically not assessed or discussed in vection studies, it is conceivable that presence might nevertheless have influenced some results. For example, Palmisano (1996) found that forward linear vection induced by a simple random dot optic flow pattern was increased if stereoscopic information was provided, compared to nonstereoscopic displays. Even though presence was not measured in this experiment, it is generally known that stereoscopic displays increase presence (Freeman, Avons, Meddis, Pearson, & IJsselstein, 2000; IJsselstein, de Ridder, Freeman, Avons, & Bouwhuis, 2001). In another study, van der Steen & Brockhoff (2000) found unusually short vection onset latencies, both for forward linear and circular yaw vection. They used an immersive VR setup consisting of a realistic cockpit replica of an aircraft on a motion simulator with a wide panoramic projection screen. Visual displays showed highly realistic scenes of landscapes as would be seen from an airplane. Even though presence was not assessed here, it is possible that the presumably high level of presence might have contributed to the strong vection responses of the observers.

2.4.1 Might presence have an influence on vection?

Presence is a conceptualisation of the phenomenon that humans can have the experience to be in a different place than where they are physically located. On such occasions, one has the feeling of being present and acting within a mediated environment, such as a story told in a book, in the cinema or theatre plays, or even just day dreaming. For most VR applications, creating a good sense of presence for the user is an explicit goal. VR users who are experiencing a computer-generated world using some immersive technology can get a very compelling illusion of being and acting in the simulated environment, instead of being present in the real environment.

Several different definitions for presence have been suggested in the literature. IJsselstein (2004) and Sadowski & Stanney (2002) give detailed reviews on different conceptualisations, definitions, and measurement methods. The current thesis will use the definition by Witmer & Singer (1998) which states that “presence is defined as the subjective experience of being in one place or environment, even when one is physically situated in another”. Presence thus refers to the psychological response of a user who is exposed to a mediated environment, such as a movie in a theater, a virtual environment, or even a book. While some authors use the terms *presence* and *immersion* inter-

2 GENERAL CHARACTERISTICS OF VECTION

changeably, we will follow the distinction suggested by Slater & Wilbur (1997) that immersion refers to the technological and physical aspects of the VR setup that enables users to experience a VR environment. For example, a panoramic, wide-screen VR projection setup or a head-tracked HMD would be the technological requisites to create *immersive VR* that allows users to *feel present* in the computer-simulated environment.

Now, even with the most sophisticated current immersive VR technology, a simulated environment will never be seriously mistaken as reality by any user. One important aspect to consider here is that in any VR application, the user is always confronted with two, possibly competing, egocentric representations or reference frames. There is the real environment (i.e., the physical room where the VR setup is situated), and there is the computer-generated virtual environment (VE), which provides an intended reference frame or representation that might interfere with the real world reference frame, if, for example, users notice delays in the tracking and updating of the simulated scene. This emphasises the importance of reducing users' awareness of the physical surroundings, which has already been recognised by many researchers and VR designers. If not successful, a perceived conflict between competing egocentric reference frames arises, which can critically disrupt presence (Slater & Steed, 2000; IJsselsteijn, 2004).

Presence has been conceptualised as a multi-dimensional construct, and is usually measured with questionnaires where users are asked to provide subjective ratings about the degree to which they felt present in the VR environment after exposure (IJsselsteijn, 2004). This method has been criticised by Slater (2004) who argued that questionnaires alone are insufficient to assess whether presence actually had any influence on the mental activity or behaviour of the participant, and that other, preferably physiological methods should be established. However, this has not been successful yet, and the vast majority of researchers still rely on post-test rating scales. In the current thesis, we used the IPQ Presence Questionnaire by Schubert, Friedmann, & Regenbrecht (2001). Using factor analyses, the authors extracted three factors that constitute presence based on a sample of 246 participants. These three factors were interpreted as *spatial presence* (i.e., the relation between the own body and the VE as a space), *involvement* (i.e., the amount of attention devoted to the VE), and *realness* (i.e., the extent to which the VE is accepted as reality).

In the following, we will review a selection of papers that investigated presence in the context of self-motion perception. So far, few studies have addressed this question explicitly. For example, Slater, Steed, McCarthy, & Maringelli (1998) found a significant positive correlation between extent and amount of body movement and subjective presence in virtual environments presented on HMDs. Depending on the task condition, one group of participants had to move their head and body a lot, while the other could do the task without much body movement. The group that had to move more showed much higher presence ratings in the post-experimental presence questionnaires. It is plausible that the more an observer wearing an HMD experiences perceptual consequences of his or her own body movements in the simulated environment, the more he or she will feel present in the VE and not in the real world.

There are several studies that investigated the influence of stereoscopic presentation on presence and vection: Freeman et al. (2000) and IJsselsteijn et al. (2001) found that presence and postural responses were increased when observers watched a stereoscopic movie that was shot from the windshield of a rally car, as compared to a monoscopic version of the film. Vection, however, was not improved by the stereoscopic presentation. In the studies by Freeman et al. and IJsselstein et al.,

2 GENERAL CHARACTERISTICS OF VECTION

presence was assessed with only one post-test question. Participants were simply asked to rate how much they felt present in the displayed scene by placing a mark on a scale depicting a continuum between the extremes “not at all there” and “completely there”. Since presence is conceptualised as a multi-dimensional construct, it is possible that assessing presence by only one item was too coarse to reveal a correlation with vection. This motivated us to perform a more fine-grained analysis on the possible relations between presence and vection using the IPQ presence questionnaire.

In the preceding two subsections, we reviewed the relevant literature on vection and presence, and extracted a number of observations which indicate that there might be a relationship between cognitive, higher-level aspects and self-motion perception. Since VR is increasingly being used as a standard tool in vection research, it seems worthwhile to investigate possible connections between presence and vection. Previous studies that failed to show such a connection have the limitation that presence was assessed only coarsely (Freeman et al., 2000; IJsselsteijn et al., 2001). Furthermore, there are studies where only vection and not presence was measured, even though factors that are known to influence presence, such as stereoscopic viewing, were manipulated (Palmisano, 1996). Given these circumstances, a more detailed investigation was warranted. The basic assumption is that different dimensions which in sum constitute presence might have differential influences on different aspects of the self-motion illusion. This question will be addressed in section 3.2.

2.5 Methodological considerations

2.5.1 “Subjective” vs. “objective” measures of vection

In vection experiments, observers are typically asked to give subjective judgements about perceived self-motion. The variables of interest are the onset latency, (i.e., the time from onset of the visual motion until self-motion is reported), the intensity or perceived velocity of vection, and the latency until vection saturates, (i.e., until observers perceive the moving visual stimulus as earth stationary and exclusively perceive self-motion). The most commonly used measurement method for vection intensity is magnitude estimation (Stevens, 1957). It has been shown that observers are able to reliably rate the intensity of perceived self-motion illusions: Kennedy, Hettinger, Harm, Ord, & Dunlap (1996) showed that magnitude estimates for the vection illusion are stable and reliable over repeated sessions across different days, and that there is a slight tendency towards longer vection latencies with repeated exposure, which seems to reflect a sensory adaptation process. It has also been reported that vection latency and also motion sickness symptoms are greater for males than for females (Kennedy et al., 1996; Darlington & Smith, 1998). Other authors have used vection duration as a measure of vection intensity (Andersen & Braunstein, 1985). Here, the duration of perceived self-motion during a trial is recorded. Perceived vection speed has been measured also by magnitude estimation, or by having participants press a button every time they thought that they had turned 90° (Dichgans & Brandt, 1978; Becker, Jürgens, & Boß, 2000).

While magnitude estimation is accepted and used by the majority of researchers in the field, some have criticised this method as being not objective and suggestible to experimental demands. Alternative methods to measure vection objectively that have been suggested include determining the extent to which the perception of physical motion on a linear motion cart is influenced by visually induced vection (Carpenter-Smith, Futamura, & Parker, 1995). In this study, visual and inertial stimulation were controlled independently, and observers’ task was to report whether they perceived

2 GENERAL CHARACTERISTICS OF VECTION

forward or backward movement on the cart after each trial. The observed shifts of the PSE (point of subjective equality) caused by different visual motion conditions (forwards/backwards/stationary) was taken as a new, objective measure of vection. While this method has the advantage of being less subjective than magnitude estimation, it has the disadvantage that many trials are necessary to obtain the PSE, and vection onset latencies are not measured. An alternative approach was taken by Lepecq et al. (1993), who had participants point towards memorised targets in the laboratory after exposition to linear vection stimuli. Participants pointed to the invisible targets as if they had really travelled in space, which was taken as evidence that illusory self-motion can trigger the same spatial orientation processes as real motion does. Yet another different approach was taken by van der Steen (1998). Here, vection per se was not measured, but participants were asked to judge whether a briefly presented visual motion (0.2sec) was matched to a physical motion presented by a motion simulator, or whether it was moving faster or slower than the physical motion. Sinusoidal forward and backward motion was presented, and observers were instructed to estimate if the visual stimulus was stationary in the laboratory, or moving. Using this method, so called “coherence zones” between the visual and vestibular modality were established: It was shown that a certain amount of mismatch between visual and vestibular stimulation goes unnoticed during self-motion. This has important implications for motion simulator design, which was the main motivation in that study.

Other studies investigated physiological correlates of self-motion perception. For example, it was found that the slow phase gain of OKN (opto-kinetic nystagmus) is slightly reduced and the so-called beating field of the eyes is shifted into the direction of perceived self-motion when observers perceive vection, compared to a situation where they perceive object motion under exactly the identical optokinetic stimulation (Thilo, Guerraz, Bronstein, & Gresty, 2000, 2002). The authors interpreted this to be caused by an anticipatory attentional mechanism that moves spatial attention into the direction of self-motion.

Other studies used fMRI, positron emission tomography (PET) and visually evoked potentials (VEP) to measure brain activation patterns associated with self-motion perception (Kleinschmidt, Thilo, Buchel, Gresty, Bronstein, & Frackowiak, 2002; Brandt, Bartenstein, Janek, & Dieterich, 1998; Thilo, Kleinschmidt, & Gresty, 2003). All of those studies found a deactivation of early visual processing areas as well as vestibular cortical areas (PVC - parieto-insular vestibular cortex) during vection, compared to phases where the identical visual stimulation led to object motion perception. The only brain area that showed increased activity exclusively during vection was the area around the cerebellar nodulus (Kleinschmidt et al., 2002). These findings indicate that neural correlates are related to perceptual changes between object motion and self-motion perception exist and they extend the scope of observation from subjective reports to correlated changes of neural activity. Furthermore, they support the notion that methods using subjective reports are a valid, reliable method, given that care is taken that participants are not biased cognitively by instructions or other influences.

In the current thesis, magnitude estimation will be used to measure vection, not only for reasons of practicality, but also because it has been shown to be a reliable and valid method (Kennedy et al., 1996). Of course, one has to keep in mind possible specific influences, such as task demands, experimental instructions, etc. For example, Palmisano & Chan (2004) showed that vection responses were faster if participants were instructed to report *onset of self-motion perception* by pressing a button in a vection trial, compared to a second group that was instructed to report the *offset of object motion perception* by releasing a button which they had pressed at onset of object motion (i.e., when

2 GENERAL CHARACTERISTICS OF VECTION

object motion turned into self-motion perception). In the vection experiments of the current thesis, care was taken to avoid systematic biases due to instructions as much as possible. Presence was measured using the standardised IPQ Presence Questionnaire by Schubert et al. (2001).

2.5.2 Vection experiments in VR

The experiments reported in the thesis were conducted using Virtual Reality (VR). Traditionally, most psychophysical experiments on vection used optokinetic drums with abstract visual stimuli, such as random dots or black and white stripes to induce vection. While this approach has the advantage of providing high stimulus control, it has the disadvantage of low ecological validity. With recent technological advancement, VR has become a suitable tool for investigating the vection phenomenon. Using advanced VR technology, it is possible to provide a high level of sensory realism to the participants in the experiments, while at the same time keeping good experimental control over the psychophysical stimulus parameters. It is well documented that VR can effectively and reliably induce vection (Hettinger, 2002). The main question addressed in this thesis is whether increasing the realism across multiple sensory modalities and believability of the stimulus can enhance the vection illusion. Most of the studies so far have concentrated on how single modalities - mainly vision - influence vection (Dichgans & Brandt, 1978; Howard, 1982; Lackner, 1977). In this thesis, emphasis is placed on how visual, auditory, and somato-sensory modalities influence vection, while reducing vestibular stimulation to a minimum. The reasoning behind this is that vestibular stimulation, which is very cost intensive since it requires large, expensive motion simulators, might become redundant or at least less important if all other senses are stimulated in a coherent, realistic way that is consistent with real self-motion. If successful, such an approach might in future allow for a lean and elegant way of simulating self-motion.

3 Circular vection experiments in VR

In this section, two experiments on circular vection will be reported. These two experiments serve two purposes: One is to establish VR as a method for vection experiments, comparing the results with traditional studies that used optokinetic drums. The second and main purpose is to investigate whether highly realistic natural visual stimuli induce stronger vection than abstract stimuli.

3.1 Experiment 1: Do natural, photorealistic visual stimuli enhance vection?

This experiment investigated whether visually induced circular vection is increased if a natural, photorealistic scene is used as the inducing stimulus, compared to an artificial image. Two visual scenes were compared: One was a photorealistic scene of a market place, and the other was a scrambled version of the same image. In the scrambled image, no information about spatial layout and pictorial depth structure were available. The two visual scenes were tested at rotational velocities of 20, 40, and 60°/s. Three dependent measures of vection were obtained: Vection onset latency, vection intensity, and the convincingness of the feeling of self-motion. Results showed that all dependent measures of vection were increased for the natural scene in that onset latencies were shorter and intensity and convincingness ratings were higher. Also, higher rotational velocities induced stronger vection. The results support the hypothesis that natural, realistic scenes induce stronger vection than artificial stimuli. However, further experiments are needed in order to control for confounding factors to allow for clear conclusions about the origin of this effect.

3.1.1 Introduction

In this experiment, we investigated whether a photorealistic image of a scene that contains coherent spatial information about pictorial depth and scene layout (e.g., linear perspective, relative size, texture gradients etc.) can induce vection more easily than a comparable stimulus with the same image statistics but without relative depth and scene layout. The underlying idea is that the photorealistic scene might facilitate vection by providing the observers with a convincing reference frame for the simulated environment so that they can feel "spatially present" in that scene. As was already indicated in section 2.3 on page 26, the concept of frames of reference becomes important for self-motion experiments in VR. In any VR application, users are confronted with multiple reference frames: There is the physical surrounding in the lab, including the simulation setup (visualisation, motion platform, etc.), and also, there is the computer-generated simulated environment. VR designers aim to mask out the physical surrounding as much as possible, in order to make the user feel present in the virtual environment. However, oftentimes, these two reference frames are simultaneously present and thus a conflict between frames of reference arises, which disrupts presence (IJsselstein, 2004; Slater & Steed, 2000). For the case of vection in virtual environments, this means that the better observers accept the virtual scene instead of the physical surrounding (i.e., the simulation setup - as the "real world"), the less conflict between the two competing reference frames should arise and hence vection in the virtual scene should be enhanced. The underlying line of thought was originally put forward by Dichgans & Brandt (1978), who stated that the occurrence of vection might be related to our underlying assumption of a stable environment. Since we evolved

3 CIRCULAR VECTION EXPERIMENTS IN VR

in a world where the environment does not move around by itself, it is reasonable to perceive self-motion if we see a large part of the visual scene move in a uniform manner. Now, for the case of VR, it is reasonable to assume that a photorealistic, natural virtual environment might increase the impression of a stable environment, compared to an abstract stimulus, such as random dots. Under this assumption, vection in VR should be increased if a photorealistic image of a natural scene is used as the vection inducing stimulus, compared to an abstract stimulus that does not contain visual information about spatial layout, which may more likely be interpreted as object motion.

3.1.2 Hypotheses

If photorealistic scenes have higher vection-inducing potency than similar images without pictorial depth information and spatial layout, we would expect shorter vection onset times and higher intensities of perceived vection for stimuli of the former kind. To test this hypothesis, we generated two kinds of visual stimuli. One was a photorealistic 360° roundshot of the Tübingen Market Place (Fig. 4, top). The second stimulus was a mosaic-like scrambled version of the same picture, where all parts of the picture were shuffled and reassembled at random positions (4, bottom). This had the effect that all information about pictorial depth and scene layout was eliminated, even though individual small parts of the scene, (e.g., a window), remained recognisable. A natural, photorealistic scene is a lot richer than an artificial scene, such as random dots or Gabor patches, in that it contains a lot more visual information about pictorial depth, scene layout, relative size, etc., and it is difficult to control for all visual factors independently. For the current experiment, we aimed to keep the visual features constant for the two visual conditions on a pixel-level, while manipulating the availability of scene layout and pictorial depth information. This was done using the scrambling technique described above. As can be seen in 4, the scrambling changed one important aspect of the image statistics. Compared to the natural scene, many high contrast edges were added to the scrambled picture, which is known to facilitate vection. Psychophysical studies have shown that increasing the contrast and spatial frequencies increases perceived speed (Distler, 1996) and facilitates vection (Dichgans & Brandt, 1978; Palmisano & Gillam, 1998). This fact works against the hypothesis that eliminating scene layout and pictorial cues should impair vection.

3.1.3 Methods

Eighteen participants (ages between 19 and 46 years, nine females) who were naïve to the purpose of the experiment were recruited through the subject database at the Max-Planck Institute (MPI). All had normal or corrected-to-normal vision and no history of vestibular disorders. All gave their informed consent and were paid at standard rates.

STIMULI AND APPARATUS. The experiment was carried out in the Motion Lab at the MPI for Biological Cybernetics in Tübingen. The Motion Lab houses a 6 degree of freedom (6-dof) Stewart motion platform with a mounted projection screen (see Figure 5). The visual stimuli were created using roundshot photographs and standard 3D modelling software. The photo roundshots that had been taken on the market place were digitalised and then mapped onto a virtual cylinder. The virtual camera was placed in the centre of the cylinder, and self-rotation was simulated by rotating the camera around the yaw axis. The visual stimulus was projected onto the projection screen. Participants were seated in a darkened cabin at a distance of 106 cm to the screen, which

3 CIRCULAR VECTION EXPERIMENTS IN VR

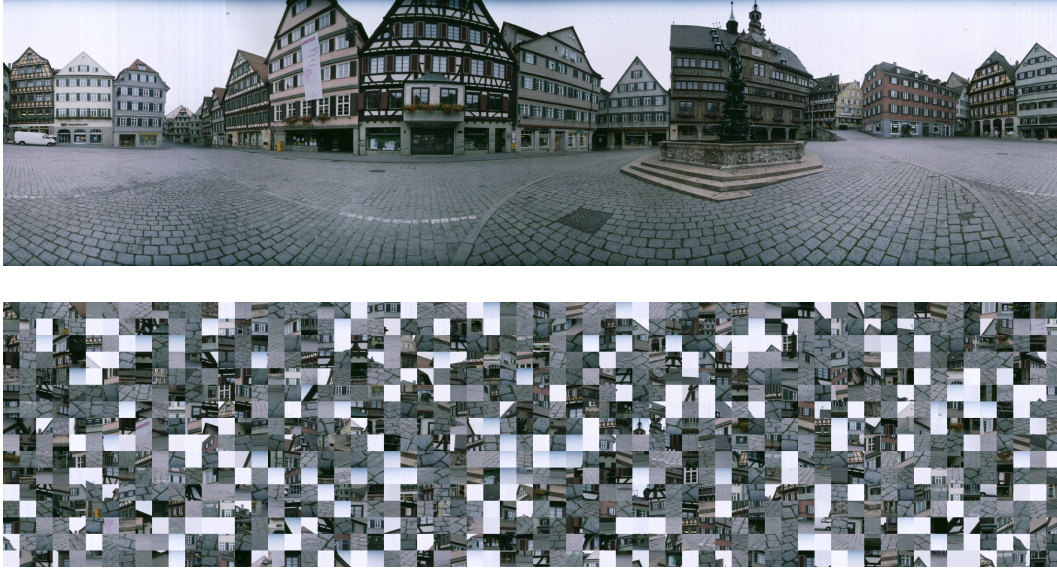


Figure 4: Top: 360° roundshot of the Tübingen Market Place. Bottom: Scrambled version of the same picture.

resulted in a $84^\circ \times 63^\circ$ FOV. The screen was curved cylindrically and had a curvature radius of 2 m. Physical and simulated fields of view (FOV) were matched², and care was taken to fade out the physical reference frame of the simulation setup as much as possible by masking all light apart from the projector with black light-absorbing cloth. A SONY LCD projector was used with a pixel resolution of 1024×768 (XGA) and an image update rate of 60 Hz. The projected image was non-stereoscopic. Figure 5 shows a schematic of the experimental setup.

PROCEDURE. Participants were told that they would see moving images, and that their task was to continuously report whether they perceived a moving image or self-motion. They were told that at the beginning of the trial, they would probably see a moving image and perceive no self-motion, and that this might change after a while and they might perceive self-motion. They were also instructed that these two perceptual states might well flip from one to another throughout the experiment. Participants were asked to continuously report what they perceived, and they were instructed to watch the stimuli "as relaxed and naturally" as possible (for example as they would do when looking out of the window of a moving train or bus). They were also told not to suppress any reflexive eye movements - thus, not to suppress the optokinetic reflex (OKR), and neither to stare through the screen nor to fixate on a static point on the screen, but to concentrate on the image in the central part of the projection screen. We did not use a fixation point, even though it is known that a fixation point reduces vection onset times (Becker et al., 2002b). The main reason being that from an applied perspective for self-motion simulation, it is important to investigate vection-inducing parameters under "natural" viewing conditions, even though this method reduces experimental control. Furthermore, this method reduced the perceived flicker and ghost images

²The physical FOV relates to the size of the viewing angle of the projected image. It is defined by the absolute size of the projection setup and the viewing distance of the observer. The simulated FOV (sometimes referred to as geometric FOV), in contrast, is a property of the computer-generated simulation. It is defined by the geometry of the viewing frustum, (i.e., the angle of the opening with which the virtual eye sees the simulated scene). If, for example, the simulated FOV is larger than the physical FOV, this corresponds to a wide angle effect.

3 CIRCULAR VECTION EXPERIMENTS IN VR

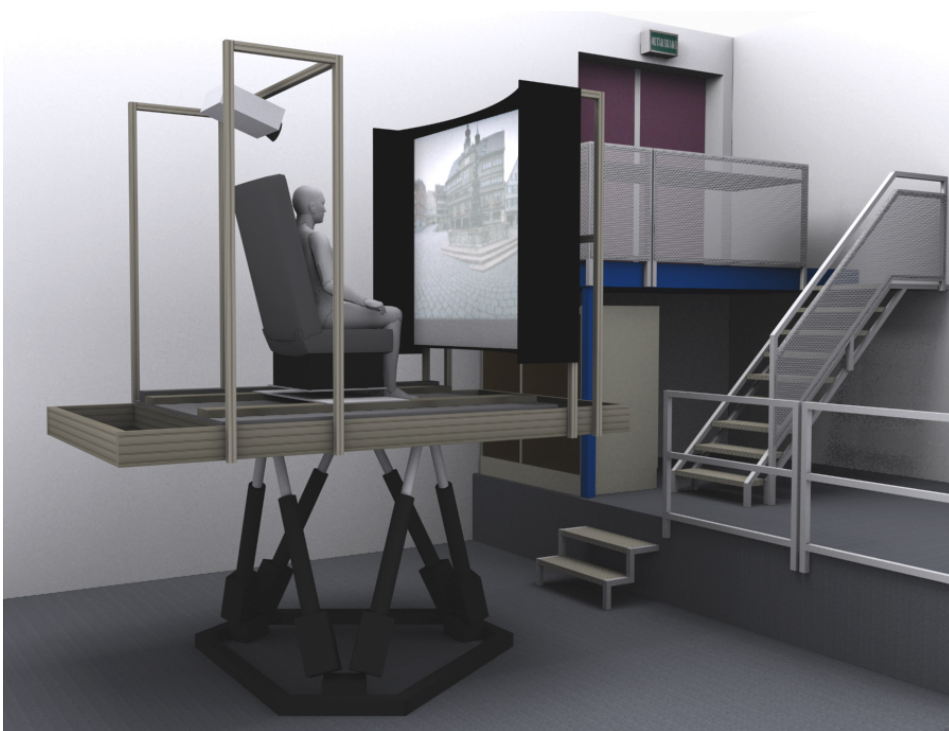


Figure 5: Schematic of Motion Lab. Participants were seated in the cabin which was mounted on a 6-dof motion base. The cabin was covered with black cloth on all sides. The schematic shows a flat projection screen - in Experiments 1 through 4, a curved screen was used, Experiment 5 used a flat screen as depicted. Participants were aware that the motion platform would not physically move during Experiments 1, 2, and 4.

3 CIRCULAR VECTION EXPERIMENTS IN VR

which arise due to the 60 Hz projection if gaze is held fixed at one point on a moving projected image.

Participants started trials by a button press, upon which the static image started to rotate around the observer. They used a Microsoft force-feedback joystick to give their vection responses. As soon as participants felt self-motion, they deflected the joystick into the direction of perceived self-motion. Three measures of vection were obtained: The time from onset of visual motion to onset of perceived self-motion was taken as the *vection onset latency*. The amount of joystick deflection was taken as an indication of *vection intensity*: The stronger the perceived vection, the more they deflected the joystick. If vection was “saturated”, (i.e., if participants perceived the moving scene as earth stationary and felt that they were rotating at the speed of the visual stimulus), they deflected the joystick to the maximum. Care was taken that participants understood that they were not to rate the perceived vection velocity, but the intensity. That means that the joystick would be fully deflected if saturated vection was perceived, no matter whether a fast, slow, or medium rotational velocity was presented. After each rotation, subjects additionally rated the *convincingness* of perceived self-motion using a 0 - 100 scale in steps of 10 (0 = “no perceived self-motion at all”, 100 = “very convincing sense of self-rotation”: the image is perceived as earth-stationary, all motion is perceived as self-motion).

Before the practice session started, all participants were first shown a rotating vection stimulus for two minutes for instructional purposes, until they reported a strong sense of self-rotation. This was done because none of the participants had previously experienced vection in a laboratory, and it has been reported that the vection illusion can take a very long time to develop when exposed to a vection stimulus in the laboratory for the very first time, (up to 20 seconds) (Berthoz & Droulez, 1982, page 182). In the practice session, participants were trained to indicate the onset and strength of perceived vection and the convincingness of rotation: If no self-motion was perceived, the joystick was not deflected at all. As soon as self-motion was perceived, they started to move the joystick to the direction of perceived motion, and if saturated vection was perceived, the joystick was deflected to the maximum. Before all trials, their hand rested right next to the joystick lever in order to prevent unintended joystick deflection. As long as no self-motion was perceived, the joystick was not touched. Any intermediate vection intensity was indicated accordingly by joystick deflection angle. After each trial, the convincingness of perceived self-rotation was rated using the 0-100 scale. All participants were trained until they were confident in their understanding of all three response measures.

DESIGN. In the experiment, two factors were varied in a 2×3 within-subject design: Realistic vs. scrambled visual stimuli were compared across three rotational velocities of 20, 40, and 60°/s. Constant acceleration times of 3 seconds and deceleration times of 6 seconds were used in all trials. The duration of constant velocity rotation was 60 sec. Leftward and rightward rotations always were alternated in order to prevent motion sickness. Each rotation stopped automatically after a maximum of 60 seconds. If the joystick was fully deflected for 10 consecutive seconds, the rotation stopped earlier. This was done to prevent motion sickness caused by prolonged exposure. Participants completed 36 experimental trials after the practice session. The market scene and scrambled image stimuli were run in two blocks separated by a ten minutes break. This was done to avoid potential carryover-effects between the two kinds of visual stimuli across trials, since it is known that after prolonged sensation of vection, aftereffects can last for several minutes (Hettinger, 2002;

3 CIRCULAR VECTION EXPERIMENTS IN VR

Welch, 2002). Within the blocks, the presentation order of the different rotational velocities was randomised. Presentation order of the market scene and scrambled image was balanced across participants and also across gender, which means that half of the participants experienced the real scene first, and half of these participants were female.

3.1.4 Results

Three separate 2×3 repeated-measures ANOVAs for vection onset time, vection intensity, and convincingness showed the following effects:

For VECTION ONSET TIME as the dependent variable, the main effects of the visual scene and rotational velocity were significant: $F(1,14) = 9.12$, $p < 0.01$ and $F(1.38,19.25) = 14.72$, $p < 0.001$, respectively (F-values for the factor velocity were Greenhouse-Geisser-corrected to adjust for violated sphericity-assumptions; hence non-integer F-values). As can be seen in Figure 6 (left), mean vection onset times were always longer for the scrambled image than for the natural market scene. This effect is especially pronounced for slower rotations. At all velocities, paired-samples t-tests showed significant differences between the two visual scenes. It can also be seen that vection onset time was shortened with higher velocities, and this effect was stronger for the scrambled image. The interaction between image degradation and velocity is significant at $F(1.42,19.92) = 4.30$, $p < 0.05$. The presentation order of the two visual conditions had no significant effect on any of the dependent variables: For vection onset time, F-values were $F(1,17) = 0.30$, $p = 0.59$ and $F(1,17) = 1.53$, $p = 0.23$ for the intact market and the scrambled image, respectively.

VECTION INTENSITY. The plot in the centre of Figure 6 shows the results for vection intensity. Vection intensity was measured by joystick deflection angle: 0% corresponds to no perceived vection at all, 100% corresponds to saturated vection. The ANOVA revealed the same pattern of results as for vection onset times: Both main effects of visual scene and rotational velocity were significant ($F(1,14) = 5.92$, $p = 0.003$, and $F(1.31,18.39) = 18.87$, $p < 0.001$, respectively). The natural market scene induced higher vection intensities than the scrambled scenes across all rotational velocities (again, F-values for the factor velocity were Greenhouse-Geisser-corrected to adjust for violated sphericity-assumptions). There was no significant effects of session, nor were there any interactions.

CONVINCINGNESS RATINGS. The same pattern of results was found for the convincingness ratings of perceived self-motion. The right graph in Figure 6 shows the convincingness data on a 0-100 scale. As can be seen, convincingness ratings were always higher for the market scene, and higher velocities were rated as more convincing ($F(1,14) = 12.02$, $p < 0.01$ and $F(1.21,16.99) = 19.05$, $p < .001$, respectively).

Previous studies have reported gender differences in circular vection (see Darlington & Smith (1998)). However, in our study, we found no gender differences at all for any of the dependent measures. It is unclear whether this is due to the rather small sample size, or perhaps due to the different methodology employed in this experiment.

3.1.5 Discussion and Conclusion

In the classical vection literature, vection onset latencies between 3 and 20 seconds are reported for circular vection (Dichgans & Brandt, 1978; Howard, 1982). In these studies, participants sat in the dark inside an optokinetic drum, and were suddenly exposed to the rotating stimulus by lights being

3 CIRCULAR VECTION EXPERIMENTS IN VR

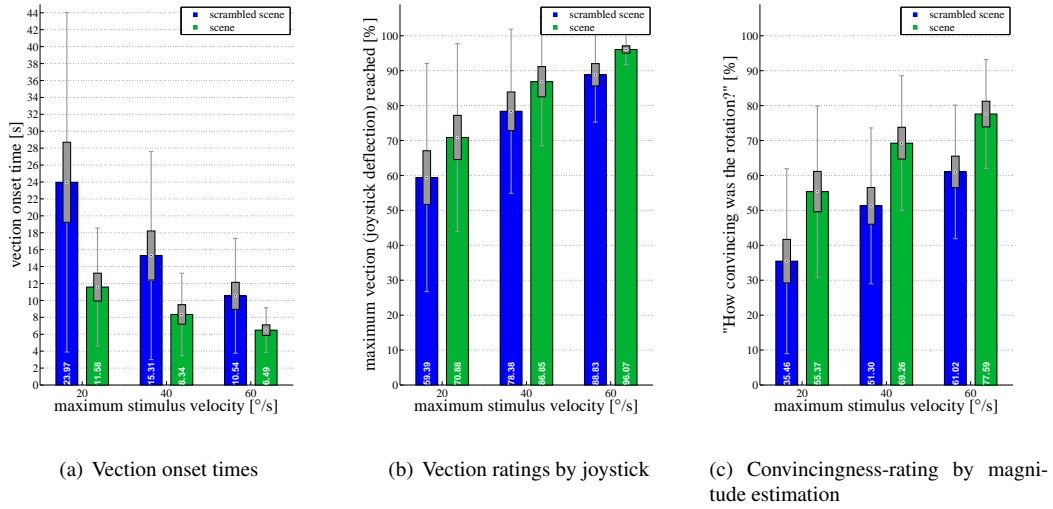


Figure 6: Left: Mean vection onset times. Centre: Maximum perceived vection intensity, measured by % joystick deflection. Right: Mean convincingness ratings for perceived self-motion. Boxes show standard error of the mean, whiskers depict one standard deviation.

switched on after the drum had reached constant rotational velocity. The current study differs from the classical studies with respect to the apparatus and stimulus presentation: Here, VR technology was used, and observers were exposed to an initially static stimulus, which started to accelerate until constant rotational velocity was reached. Furthermore, the stimulated FOV in this study was only $84^\circ \times 63^\circ$, whereas optokinetic drums usually provide full field visual stimulation. Even though these differences should reduce vection and thus increase vection onset latencies for the current experiment, it is notable that the natural scene produced vection onset times that are still relatively short with values between 6.5 and 11.6 seconds and are thus comparable to the values reported in classical studies. Interestingly, onset latencies for the scrambled image ranged between 10.5 and 24 seconds, which is considerably longer than the onset latencies produced by the natural scene. This indicates that on the one hand we can assume that the VR setup used in this experiment provides an adequate technology to conduct vection experiments and on the other hand we might preliminarily conclude that natural, highly realistic images have the potential to significantly increase vection in VR. The finding that the onset latencies for the natural scene are considerably shorter than the scrambled stimulus relate well to a study by van der Steen & Brockhoff (2000), where mean vection onset latencies as short as 3 seconds were reported. In that study, a realistic cockpit-replica on a motion simulator with a large FOV ($142^\circ \times 110^\circ$) spherical projection screen that showed a highly realistic natural landscape was used. The enhanced foreground-background separation between the cockpit window and the projection screen might have facilitated vection (see Nakamura & Shimojo (1999)). Most importantly, the cockpit-replica provided a very convincing reference frame for the simulated visual environment. The simulated Virtual Environment was a realistic landscape viewed from a cockpit, which means that both the physical surroundings and the simulated environment matched perfectly. These combined effects might have contributed to the extremely short vection onset latencies found in that study.

3 CIRCULAR VECTION EXPERIMENTS IN VR

In conjunction with the above mentioned finding from van der Steen & Brockhoff (2000), the finding from Experiment 1 that highly realistic natural visual scenes can enhance vection compared to abstract stimuli raises the question how low-level and high-level factors influence self-motion perception. The underlying logic in this experiment was that a natural, realistic scene should increase the impression of being surrounded by a stable environment. If the idea of Dichgans & Brandt (1978) is correct, the inherent assumption of a stable environment should increase vection, compared to an abstract, scrambled image which is devoid of information about spatial layout and depth. Even though our data support the hypothesis that vection is increased by a natural image, other confounding factors need to be controlled in order to allow for clear conclusions about the origin of this effect. By scrambling the image, we intended to remove all visual information that makes the scene recognizable as a place extending in three dimensions which surrounds the observer. Indeed, the scrambled image looked like a mosaic-like 2D surface, rather than a three-dimensional space. On a pixel-level, the scrambled scene contained the identical information as the natural scene, but it also contained many additional high-contrast vertical edges that moved with the visual scene as a result of the scrambling method. This altered low-level feature might have influenced the vection responses. Note, however, that from what is known so far, increasing the number of high-contrast edges should have the effect of *increasing* vection, rather than *reducing* it (see 2.1.3 on page 23 and 3.1.2 on page 34). One possibility to consider why increasing spatial frequency did not strengthen vection in our case is that image content and spatial frequency might interact in some way. The effect of spatial frequency was so far only investigated using black and white stripes and not using natural stimuli (Diener et al., 1976; Distler, 2003).

Another visual factor that might possibly have affected vection is pictorial depth. In the natural scene, all visual cues that specify spatial depth, such as perspective, texture gradients, and relative size were globally consistent and informative about the spatial layout of the depicted scene. This was not the case for the scrambled image, where only local information about these cues were available in each patch or mosaic, but no globally consistent information for the whole scene. It is possible that not only low-level, but also high-level factors contributed to the results. The natural, photorealistic scene might make observers feel to be “present” in the simulated scene or to be “at a place” instead of viewing an image. If the scene starts to move, the self-motion illusion might be increased because of our inherent assumption of a stable environment. An artificial stimulus, on the other hand, might be perceived rather as object motion, and vection might take longer to develop. Now, it has long been shown that artificial stimuli can induce compelling vection - the whole classical vection literature is based on experiments that used black-and-white stripes or random patterns to induce vection. Still, there remains the possibility that also high-level factors might influence vection such that vection is increased by realistic, natural stimuli of an environment.

Because of the scrambling method used in this experiment, it is difficult to clearly separate low-level and high-level factors in the two visual stimuli. Experiment 2 was designed in order to better control for this. Furthermore, the issue of possible cognitive influences on vection will be investigated systematically by measuring presence and relating this measure to vection data.

3.2 Experiment 2: The relation between the feeling of presence in VR and vection

This study investigated whether the self-motion illusion induced by moving visual stimuli (circular vection) in VR is influenced by the type of visual stimulus used. This was done by presenting either a photorealistic image of a natural scene or scrambled versions of the same stimulus. The latter were created by either scrambling image parts in a mosaic-like manner or by slicing the original image horizontally and randomly reassembling it. The natural image led to faster, stronger, and more convincing vection than any of the scrambled stimuli. Furthermore, the natural scene led to higher presence ratings as indicated by scores on spatial presence questionnaires. This suggests that there might be a direct link between spatial presence and self-motion perception. We posit that stimuli depicting naturalistic scenes provide observers with a convincing reference frame for the simulated environment which enables them to feel "spatially present". This, in turn, might be an important factor that can facilitate the self-motion illusion. This result has important implications for both our understanding of self-motion perception and for the field of motion simulation applications.³

3.2.1 Introduction

In Experiment 1, we found evidence that circular vection in VR is enhanced if highly realistic, natural images are used, instead of a scrambled stimulus that does not contain information about spatial depth and scene layout. Experiment 2 was designed to further investigate this finding by controlling for potential low-level artefacts, and by exploring whether the feeling of presence in a Virtual Environment influences the vection illusion.

As already mentioned previously, traditional vection experiments used abstract visual stimuli that can only be interpreted as a 2D surface, such as black and white stripes or random dots. While the influence of low-level stimulus features on vection has been investigated extensively, there are hardly any studies that explicitly studied cognitive, higher-level influences on vection (see 2.1 on page 21), apart from a few exceptions. For example, Lepecq et al. (1995) showed that vection onset latencies can be influenced by a cognitive manipulation in children. In an experiment, children who were aware being seated on a movable chair reported onset of forward linear vection faster than children who knew they were sitting on a chair that was solidly mounted to the ground. However, the probability that observers report vection was not influenced by this manipulation. In another study, Kitazaki & Sato (2003) found an attentional modulation for vection: They presented superimposed upward and downward moving optic flow stimuli of differing colours, and participants were asked to pay attention to one of the colours. They found that the non-attended flow pattern determines vection direction. That is, if participants pay attention to the "red" upward flowfield, upward linear vection is perceived, since the downward moving flowfield induces upward vection. This effect was found only when the two flowfields were presented in the same depth plane. When presented in different depth planes, the stimulus in the background determined vection direction, irrespective of attentional instruction. However, it remains unclear how differing eye-movement patterns might have influenced this result: It is known that in order to attend to an object, one has to fixate and do pursuit eye-movements to follow the object. Unfortunately, eye-movements were not monitored in that study.

³Parts of this study have been published earlier in Riecke et al. (2006).

3 CIRCULAR VECTION EXPERIMENTS IN VR

Apart from these studies that explicitly addressed cognitive aspects during self-motion perception, there are also some anecdotal reports about possible cognitive effects: Andersen & Braunstein (1985) mention that forward linear vection could be induced with a stimulated FOV as small as 7.5° , provided that participants stood in a movable booth that could be moved in the direction of simulated optic flow. Thus, participants had to be made believe that motion in the direction of simulated movement was physically possible in order to elicit vection with that stimulus set. A similar finding was obtained by a recent VR study about auditory circular vection (Larsson et al., 2004). In that study, circular vection was induced by an auditory stimulus played over headphones using generic HRTFs (Head-Related Transfer Functions). Notably, participants had to be blindfolded and seated on a rotating chair (which never rotated in the experiment) in order to elicit vection, and four “fake” loudspeakers were placed in each corner of the experimental room. Only under these circumstances, auditory vection could be elicited.

In summary, these findings support the idea that not only low-level, bottom-up features of the vection stimulus, but also cognitive factors may influence the vection illusion. In Experiment 2, the idea is tested that the reason why the natural image induces stronger vection might be that this stimulus creates a stronger feeling of presence in the simulated environment. As already elaborated on in Section 2.4.1 on page 28, presence is defined as “a state of consciousness, the (psychological) sense of being in the virtual environment” (Slater & Wilbur, 1997). *Presence* thus denotes the psychological phenomenon that observers who are exposed to a mediated world, such as TV, cinema, or a VE, feel like they are being (and sometimes acting) in this mediated world and not in the real world. In distinction, the term *immersion* is defined as “a description of a technology, the extent to which the computer displays are capable of delivering an inclusive, extensive, surrounding, and vivid illusion of reality to the senses of a human participant” (Slater & Wilbur, 1997). While most researchers share this distinction, sometimes these terms are used interchangeably. For the sake of clarity, I will comply with the definition by Slater & Wilbur (1997) to differentiate between technical equipment that enables a psychological response of feeling present in a VE.

3.2.2 Hypotheses

In order to control for possible low-level artefacts in Experiment 1, a second kind of visual stimulus was created: Instead of scrambling the original image, the market place image was sliced horizontally, and the slices were reassembled such that no consistent scene could be recognised anymore (see Figure 7, middle row). This had the same effect of destroying the scene consistency as the scrambled stimulus did, without adding more vertical high-contrast edges in the image. If the finding from Experiment 1 of reduced vection with the scrambled stimulus was due to the additional vertical edges, this effect should diminish with the sliced stimulus. As noted earlier, increasing the number of vertical edges should have increased vection, rather than reducing it in Experiment 1. The sliced stimuli were created in order to test how removing information about spatial layout of the scene without influencing the number of vertical edges affects vection.

Furthermore, the scrambling/slicing severity of the visual stimulus was varied in this experiment in four levels (see Fig. 7): this was done by scrambling or slicing the original scene such that 2, 8, or 32 slices or 2×2 , 8×8 , or 32×32 mosaics were visible in the $54^\circ \times 40^\circ$ image, respectively. The vection inducing power of the intact scene was compared to the degraded stimuli. The reasoning behind this manipulation is based on the idea that there might be a relation between vection and

3 CIRCULAR VECTION EXPERIMENTS IN VR

the feeling of presence in VR (see section 3.1.1): The more a person feels present in a Virtual Environment, the more one should accept the simulated scene as the earth-stationary surroundings one is acting and moving in (i.e., as the primary reference frame), and one should be more inclined to perceive vection when the visual scene is moved. On the other hand, if presence is low, the moving scene should rather be interpreted as object motion, since the primary reference, i.e., the physical surroundings, is stationary. Thus, vection and presence should be strongest in the intact scene, and should decrease with image degradation. Specifically, it was assumed that the more the scene is degraded, the more presence should be disrupted. We thus predict a *gradual* decrease of both vection and presence from the intact scene to the three stimuli in the order of the respective image degradation.

Finally, if there is a relation between presence and vection, one would expect a correlation between the two measures. Specifically, we expect that stimuli that induce higher presence should reduce vection onset times and increase vection intensity and convincingness ratings.

3.2.3 Methods

Twelve naïve participants (5 of which were female) took part in this experiment. Participants' ages ranged from 17 to 34 years (mean 23.8, SD 5.0). All had normal or corrected-to-normal vision, and they were paid at standard rates for participation.

STIMULI AND APPARATUS. Stimuli and apparatus were similar to Experiment 1, with the following exceptions: For technical reasons, a JVC D-ILA projector (DLA-SX21S) with a pixel resolution of 1400×1050 and an update rate of 60 Hz was used in this study, and participants were seated at a distance of 1.8 metres from the centre of the screen. This was done to seat participants as close as possible to the radius of the screen curvature. As a result, the stimulated FOV in this study was $54^\circ \times 40^\circ$. Physical and simulated FOV were matched. The stimuli were various versions of the Tübingen market place: The natural, intact scene was that used in Experiment 1, and three scrambled and sliced versions were also included (see Figure 7). The scrambled stimulus with 8×8 mosaics was identical to the one used in Experiment 1. Participants used the identical joystick that was used in Experiment 1 to give their vection responses. In order to keep the duration of the experiment manageable, only rotations of 20° and $40^\circ/\text{s}$ were presented in this experiment.

PROCEDURE. The procedure was similar to Experiment 1: Participants received the same instructions and practice session before the experiment, and the same vection measures (vection onset time, vection intensity and convincingness) were obtained with the identical response method using the joystick. Between trials, there was a pause for 20 seconds before the next trial started in order to reduce the interference of potential motion aftereffects. Participants completed two sessions of 32 trials each. After every 16 trials, participants were given a short break. In one session, the scrambled stimuli and natural scene were shown, in the other, the sliced stimuli and the natural scene were presented. The presentation order of the two sessions was balanced across participants. After each session, participants filled out the standardised 14-items Igroup Presence Questionnaire (IPQ, (Schubert et al., 2001)) on a 7-point-scale for all four visual scenes they had been exposed to, i.e., the intact scene, and the three scrambled and sliced versions. In total, eight sets of presence ratings were obtained from each participant. Finally, in order to monitor simulator sickness, participants completed the SSQ Simulator Sickness Questionnaire (Kennedy, Lane, Lilienthal, Berbaum,

3 CIRCULAR VECTION EXPERIMENTS IN VR

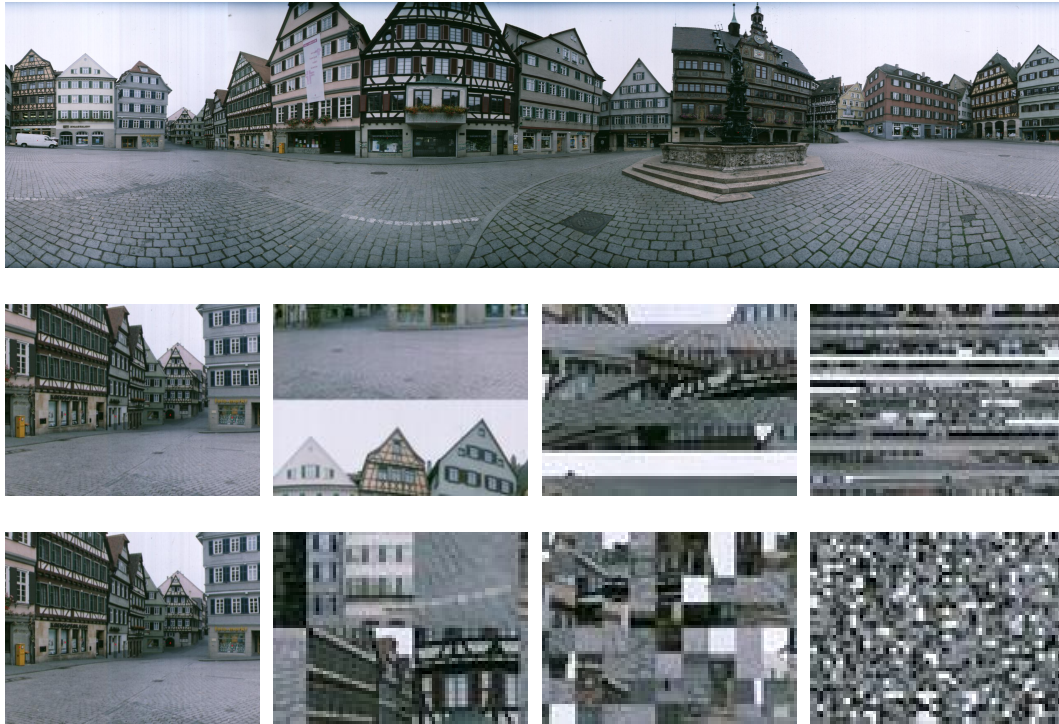


Figure 7: **Top:** 360° roundshot of the Tübingen Market Place. **Middle:** 54x40.5° view of the 4 stimuli used in session A: original image and 2, 8, and 32 slices per 45° FOV. **Bottom:** 54x40.5° view of the 4 stimuli used in session B: original image and 2x2, 8x8, and 32x32 mosaics per 45°x45° FOV.

& Lawrence, 1992) before and after the experiment. In total, the experiment lasted about one and a half hours.

DESIGN. The experiment was run in a 2 (session: scrambled, sliced) \times 4 (image degradation severity) \times 2 (rotation velocity) \times 2 (turning direction) within-subject design with two repetitions per condition.

3.2.4 Results

Separate repeated-measures ANOVAS were calculated for three dependent variables - vection onset times, vection intensity, and convincingness ratings. The presence questionnaire data were subjected to factor analysis in order to explore the dimensionality of presence. Finally, correlation analyses were performed between vection and presence data.

VECTION ONSET TIMES. The ANOVA revealed two significant main effects: First, rotation velocity had a significant influence on reported vection onset latency. At 40°/sec, vection was reported faster than at 20°/sec ($F(1,10) = 23.9$, $p < .001$). Second, image degradation had a significant effect ($F(3,30) = 6.23$, $p < .01$). Here, only the intact scene produced significantly reduced onset times, compared to all scrambled or sliced images. There were absolutely no statistically significant differences between the sliced and scrambled stimuli for vection onset latency: Both scrambling and slicing led to the same increase of vection onset time at all image degradation levels. There was no effect of session and gender, and no significant interactions were observed between image degradation and any other factors.

3 CIRCULAR VECTION EXPERIMENTS IN VR

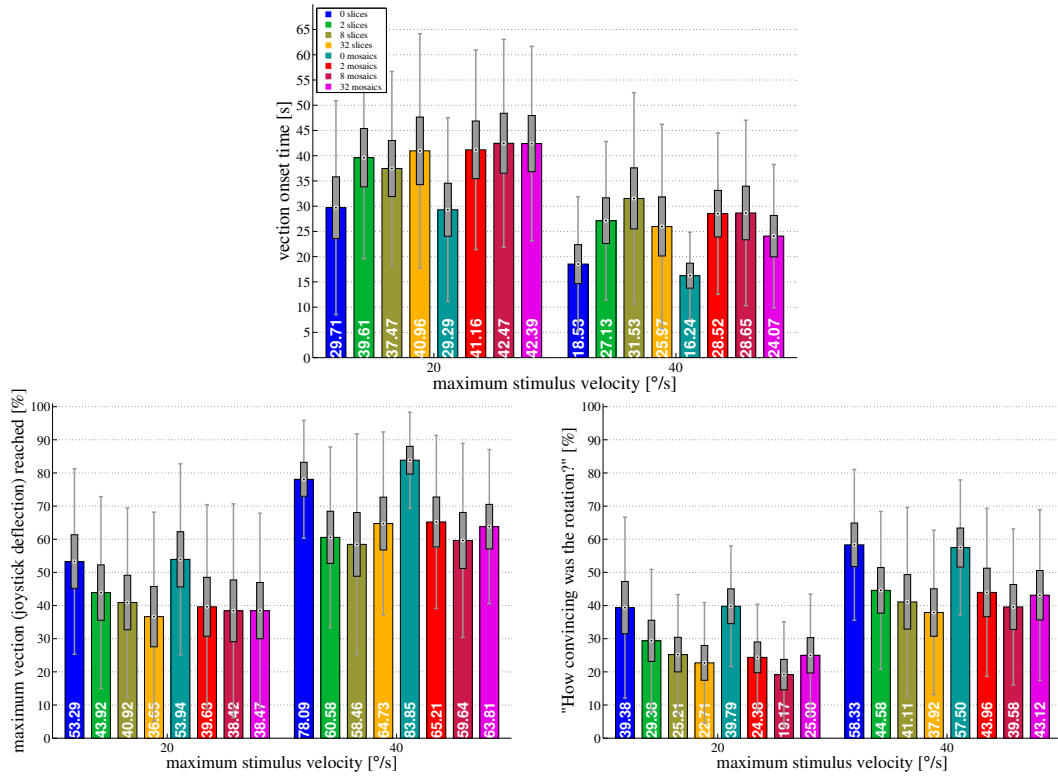


Figure 8: **Top:** Plotted are mean vection onset times for each of the 16 experimental conditions of Experiment 1. The left and right group of eight bars represent the low and high velocity conditions (20°/s and 40°/s, respectively). Boxes and whiskers depict one standard error of the mean and one standard deviation, respectively. The eight visual stimulus conditions are explained in the figure legend at the top. For the sake of clarity in all three data plots, overall between-subject differences in vection responses were removed using the following normalization procedure: Each data point per participant was divided by the ratio between the mean performance of that participant across all conditions and the mean of all participants across all conditions. Note that this procedure was only applied for the data plotting and not for any of the statistical tests. **Bottom left:** Mean perceived vection intensity, quantified as the maximum joystick deflection reached. **Bottom right:** Mean convincingness ratings.

VECTION INTENSITY. The results for vection intensity showed the same pattern as that for onset latency: The only significant main effects were rotation velocity and image degradation ($F(1,10) = 42.0$, $p < .001$ and $F(3,30) = 8.29$, $p < .001$, respectively): Higher vection intensity was reported for 40°/sec rotations and for the intact market scene (all pairwise t-tests significant at $p < .05$). Again, all degraded stimuli resulted in significantly decreased vection intensity, and no difference was observed between scrambled and sliced stimuli across all image degradation levels (see Figure 8). Also, neither the effect of session or any of the interactions were significant.

CONVINCINGNESS RATINGS. The analysis of convincingness ratings showed results that matched the pattern of results of the other two measures: Participants rated vection induced by rotations at 40°/sec as more convincing, compared to the 20°/sec rotation ($F(1,10) = 23.7$, $p < .001$). Furthermore, vection induced by the natural, intact scene was rated as more convincing than any of the degraded images ($F(3,30) = 41.4$, $p < .001$; all pairwise p 's $< .05$). Again, none of the interactions nor the session had any significant effect (see Fig. 8).

QUESTIONNAIRES. Figure 9 shows the results of the IPQ Presence Questionnaire. The questionnaire data had a high reliability in our sample ($\alpha=.91$). Mean presence scores were computed

3 CIRCULAR VECTION EXPERIMENTS IN VR

for all stimuli shown in the experiment, (i.e., for the natural, intact scene and all sliced and scrambled stimuli, see Figure 9, top left). A repeated-measures ANOVA with session (horizontal slices vs. mosaic-like scrambles) and image degradation (intact, 2, 8, 32 slices/mosaics) as independent factors showed a significant effect only for image degradation ($F(3,18) = 21.5$, $p < .005$). Post-hoc tests revealed that the intact scenes were rated significantly higher in terms of presence, compared to all degraded stimuli at all degradation levels (Bonferroni-corrected $p < .005$), while no significant differences were found between the different degradation levels, neither for the sliced, nor for the scrambled images. This indicates that the weakest level of image degradation with two slices or mosaics was sufficient to impair presence significantly, and no further decrease in presence was observed for the stronger image degradation levels (see Figure 9). The IPQ Presence Questionnaire (Schubert et al., 2001) consists of four subscales: realism, presence, space, and involvement. The mean scores of each subscale are presented in Figure 9.

The questionnaire data was subjected to factor analysis in order to examine the dimensionality and structure of presence, and a two-dimensional structure of presence was revealed: Factor 1 contained items about spatial presence and perceived realism of the simulated scene, such as the sense of acting in the virtual environment, while Factor 2 contained items about attentional aspects and involvement in the scene, e.g., how much attention was captured by the simulated virtual scene vs. by the physical surrounding of the observer. The content of Factor 1 was interpreted as spatial presence, and Factor 2 as the attentional component of presence. In Figure 9, mean scores of the factor values of Factors 1 and 2 are shown in the bottom right and middle graphs for all eight visual stimuli. Factor 2 coincides with the original subscale “involvement”, while Factor 1 represents a compound scale containing all remaining subscales.

CORRELATION ANALYSES. In the next step, the relation between presence and vection in VR was investigated by calculating correlations between the three vection measures (onset latency, intensity, and convincingness) and the overall mean presence scores. Table 1 shows the correlation coefficients of paired-samples correlations (r) and the corresponding p -values. For the overall presence scores, a significant correlation between presence ratings and convincingness ratings was found: $r = .772$, $p < .005$ (see Table 1). Also, there was a significant negative correlation between vection onset time and presence: With increasing presence ratings, vection onset times were shortened ($r = -.616$, $p < .05$). No significant correlations were observed between presence and vection intensity. In a more detailed analysis, the vection measures were correlated with the factor values of Factors 1 and 2. Here, an interesting dissociation was revealed: While convincingness of vection was positively correlated with Factor 1 (spatial presence) with $r = .579$, $p < .05$, vection onset latency correlated negatively with Factor 2 (involvement) with $r = -.620$, $p < .05$. This indicates that these two measures of vection are related to two different aspects of presence: The more attention or involvement is captured by the virtual environment (VE), the faster vection seems to occur, while the more a person feels to be present in the VE, the convincingness of vection is enhanced. No significant correlations were observed between the vection intensity measures and any of the presence scales (see Table 1). Overall, it should be pointed out that despite the rather small sample size of only 12 participants, the correlations are quite substantial.

SIMULATOR SICKNESS QUESTIONNAIRE. The SSQ data showed slightly increased mean scores for reported symptoms of simulator sickness after the experiment ($0.336 \pm 0.049SE$ vs. $0.173 \pm 0.048SE$ on a 0-3 point rating scale). However, all participants were comfortable completing the experiment.

3 CIRCULAR VECTION EXPERIMENTS IN VR

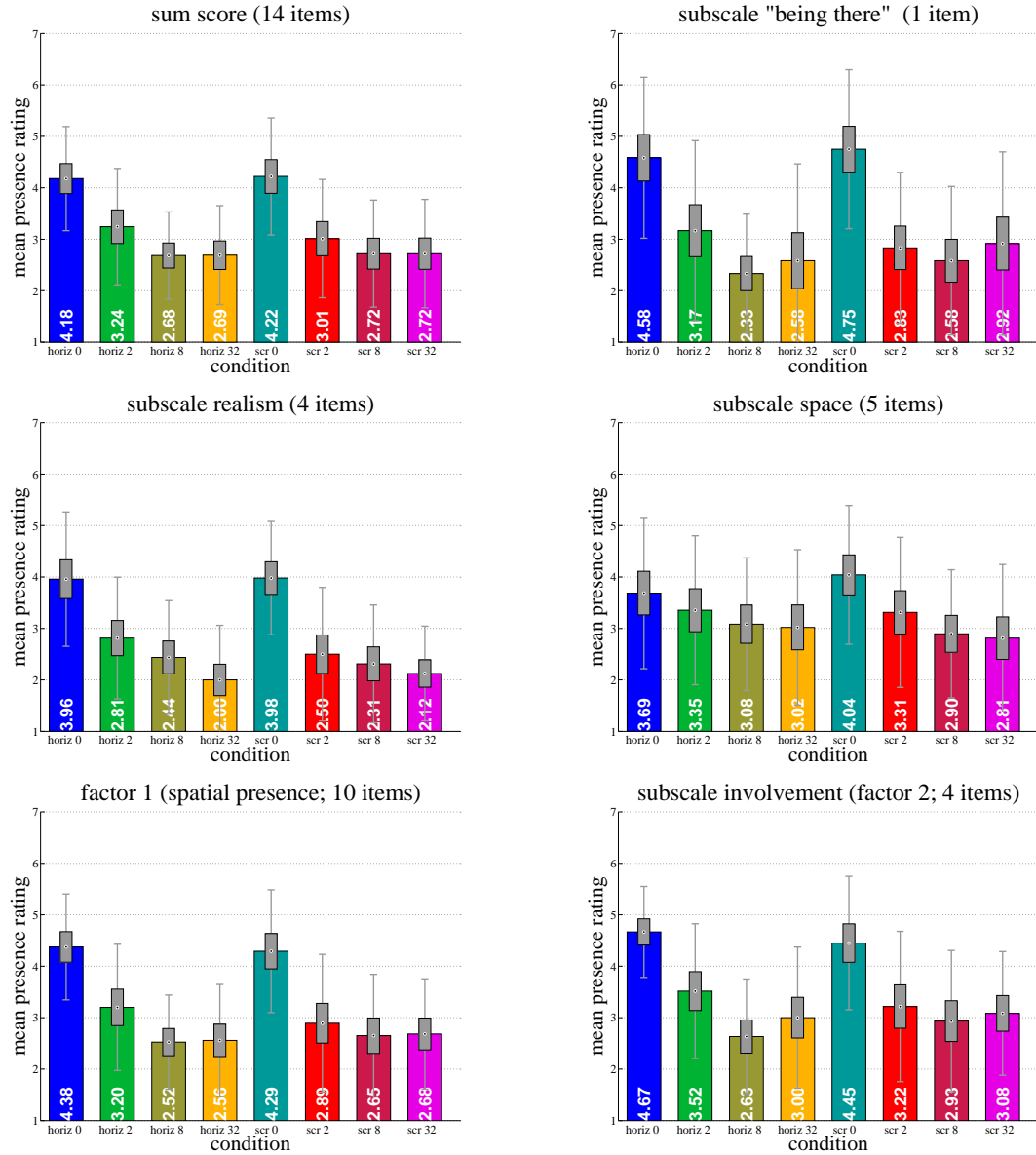


Figure 9: Presence ratings for the eight different visual stimuli. The top left plot shows the mean sum score over all 14 items. These were split up into four subscales: “Being there”, realism, space, and involvement. The involvement subscale coincides with Factor 2 of the factor analysis, and the remaining 3 subscales constitute Factor 1. Note the qualitatively similar pattern of results for all scales: Only the intact scenes (horiz 0 and scr 0) yielded high presence ratings, while all scene scrambling reduced presence consistently.

3 CIRCULAR VECTION EXPERIMENTS IN VR

With mean values below 0.4, mean sickness ratings after the experiment were very low in general, which indicates that the setup and stimulus parameters chosen were adequate.

3.2.5 Discussion

Two important findings were obtained from Experiment 2: First, only the natural market scene induced vection significantly faster and stronger, and increased compellingness of vection, compared to all degraded scenes. Importantly, no differences were observed between scrambled and sliced stimuli, and also, the level of image degradation had no effect on vection.

The finding from Experiments 1 and 2 that natural, realistic visual stimuli enhance vection do not seem to be due to low-level features of the stimulus caused by the increased high-contrast vertical edges of the scrambled stimuli, since the sliced images produced exactly the same results. On the other hand, the predicted gradual decrease of vection with increasing image degradation was not observed: There was absolutely no difference between the different image degradation levels on vection for the scrambled, nor for the sliced stimuli. Only the intact natural scene showed a significantly different pattern of enhanced vection relative to any other stimuli. The same pattern of results was also found for the presence ratings: Only the intact, natural scene obtained high presence ratings, while all sliced and scrambled images resulted in significantly reduced presence, and no difference was observed for presence between the different levels of stimulus degradation. This indicates that destroying the scene consistency only with two slices or 2×2 mosaics was sufficient to reduce presence and vection significantly. The finding that highly realistic images increase visually induced vection compared to abstract stimuli indicate that there might be a stronger influence on vection of parameters that are not exclusively related to low-level features of the visual stimulus. A natural scene (in our study, a market place scene) with many objects, buildings, streets etc. contains much richer visual information than an artificial stimulus, such as random dots. Perspective, texture gradients, occlusion, spatial layout of surfaces, known size of objects and apparent distance are but some cues that are informative about motion and space. One important motion cue that was not available in the natural scene is motion parallax - this was because a pure rotation of the camera in the virtual scene was simulated. Motion parallax (sometimes referred to as motion perspective in generalised terms) is a very powerful cue for motion under natural viewing conditions (Cutting & Vishton, 1995). One might speculate that motion parallax should increase vection in a natural stimulus. It is thus surprising that a natural stimulus without parallax enhanced vection in Exp. 1 and 2. From the results so far, it is difficult to pinpoint which visual features in the natural scene increased vection. What we can say is that a rich, natural visual scene increased vection, compared to an artificial scrambled image of the same scene.

The second important finding from Experiment 2 is that significant correlations were observed between presence ratings and vection measures: With increased presence, vection onset latencies were reduced and convincingness was increased. To our knowledge, this is the first time that such a correlation that suggests a connection between vection and presence in VR has been reported. This finding provides further indication that cognitive factors, such as the feeling of presence, might influence vection. This finding is especially interesting given that the use of VR for research, entertainment, and training applications has been increasing steadily in recent years. The factor analysis revealed a two-dimensional structure of presence, and an interesting dissociation was observed: while the factor “spatial presence” correlated positively with vection convincingness, the factor “at-

3 CIRCULAR VECTION EXPERIMENTS IN VR

		overall presence sumscore			factor 1 (spatial presence)			factor 2 (involvement)		
		onset time	convincingness	vection intensity	onset time	convincingness	vection intensity	onset time	convincingness	vection intensity
horizontal										
presence rating	r	-.289	.665*	.209	.072	.702*	.088	-.403	.173	.166
0 slices	p	.36	.018	.515	.824	.011	.786	.194	.591	.607
presence rating	r	-.675*	.892**	.302	-.135	.657*	-.043	-.743**	.424	.602*
2 slices	p	.016	.001	.208	.676	.020	.894	.006	.169	.038
presence rating	r	-.790**	.625*	.313	-.197	.393	-.260	-.779**	.340	.721**
8 slices	p	.002	.03	.321	.540	.206	.415	.003	.279	.008
presence rating	r	-.319	.302	-.004	.172	.176	-.470	-.530	.149	.512
32 slices	p	.312	.34	.991	.594	.583	.123	.076	.644	.089
scrambled										
presence rating	r	.086	.776**	-.1	.387	.756**	-.257	-.407	.222	.202
0 slices	p	.791	.003	.757	.214	.004	.420	.189	.487	.528
presence rating	r	-.448	.596*	.364	-.006	.355	.056	-.606*	.347	.431
2 slices	p	.144	.041	.245	.984	.258	.862	.037	.270	.161
presence rating	r	-.728**	.753**	.385	-.283	.584*	-.023	-.579*	.193	.556
8 slices	p	.007	.005	.216	.374	.046	.943	.048	.549	.061
presence rating	r	-.086	.646*	-.057	.184	.379	-.340	-.346	.500	.344
32 slices	p	.79	.02	.869	.568	.225	.280	.271	.098	.274

Table 1: Table of paired-samples correlations between vection measures (vection onset time, convincingness, and vection intensity) and presence ratings for the different visual stimuli (sum scores (left) and factor 1 (middle) and factor 2 (right) of the factor analysis). N = 12. Note that convincingness ratings correlated highly with the factor 1 (spatial presence) but not with factor 2 (involvement). Conversely, vection onset time was negatively correlated with factor 2 (involvement), but not with factor 1 (spatial presence). Vection intensity was only moderately correlated with factor 2 (involvement), but not at all with factor 1 (spatial presence).

3 CIRCULAR VECTION EXPERIMENTS IN VR

tention/involvement” correlated negatively with vection onset latency. This suggests that the two dimensions of presence affect the vection measures differentially. A very similar result was recently obtained in a study by Wright, DiZio, & Lackner (2006). Here, vertical sinusoidal oscillation was simulated on an HMD which induced vertical vection. The visual stimulus shown on the HMD was a highly realistic lab environment that was recorded with a camcorder that was mounted on a elevator seat which was capable of moving up and down for 1.6m. In a within-subject design, participants were tested in two different environments where the visual stimulus either matched or did not match the testing environment. They were either seated on the elevator seat, viewing the same lab environment from the correct perspective on the HMD, or in a different room, seated on an office chair that was obviously not capable of moving up and down and the surrounding environment was different from the one viewed on the HMD. The elevator seat never moved in the experiment, which means that no physical motion was applied at any time. Results showed that the compellingness of vection was rated much higher in the motion-compatible condition on the elevator seat, whereas no difference was found for the vection onset latencies between the two testing environments. The authors proposed that vection onset latency and compellingness of self-motion are driven by two dissociable mechanisms. Even though presence was not directly assessed in this study, the authors speculated that the highly immersive visual image that exactly matched the lab environment of the elevator-seat might have increased the compellingness of self-motion via a cognitive mechanism. On the other hand, the fact that vection onset latencies were unaffected whether the visual stimulus matched or did not match the testing environment led the authors to the conclusion that onset latency may be rather affected by a mechanism that is mainly driven by visual input. The observation that vection onset latency and compellingness of perceived self-motion are affected by dissociable mechanisms is consistent with the results of Experiment 2 in the current study, where we found that the attentional aspect of presence correlated with vection onset latency, while compellingness of vection correlated with the spatial aspect of presence. Future motion simulation applications might benefit if the relation between presence and vection can be further clarified. From an applied perspective, for example, vection onset times might be further reduced if the user’s attention could be captured more effectively by the VE.

In the current study, the observers’ eye-movements were not controlled, even though it is known that eye-movements can influence vection (Becker et al., 2002b). As mentioned in section 3.1.3 on page 34, participants were required not to suppress their reflexive eye movements for two reasons: First, for applied purposes, unrestricted, natural eye movements reflect the more natural viewing conditions. Second, fixating one’s gaze on a spot results in image flicker and ghost images due to the technological limitation of 60Hz projection. Because no suitable eye tracking device was available, it remains unclear whether the degraded visual stimuli lead to different eye-movements. However, it seems unlikely that the two degraded stimuli with two slices or four mosaics could have changed the eye-movement patterns so drastically to elicit such a strong difference in the vection responses. Given that the pattern of results from the presence ratings was exactly parallel to the vection data, and that significant correlations were found between vection and presence data, it appears likely that a common mechanism underlying the two responses exists. A plausible candidate for this would be the mechanism that enables spatial orientation or spatial updating process. A natural scene with rich and structured visual information about spatial layout can make the observer feel to be in the simulated space, which means being surrounded by the simulated environment, especially in an

3 CIRCULAR VECTION EXPERIMENTS IN VR

immersive VR environment. Our natural ability to orient ourselves and keep track of where we are and how we move in the environment is at work without conscious effort whenever we are moving around. This mechanism which is commonly referred to as *spatial updating* (Klatzky, Loomis, Beall, Chance, & Golledge, 1998). It has been shown in several studies that this mechanism can also be activated in a virtual environment. For example, one study showed that spatial updating in VR is possible from purely visual information under certain conditions (Riecke, von der Heyde, & Bühlhoff, 2005). In this study, the identical setup and even exactly the same visual stimulus (roundshot photograph of Tübingen market place) was used as in the current experiment. Participants in that study were aware that their task was to update locations of previously learned targets during simulated self-motion. It is very likely that in the current experiment, the spatial updating mechanism was automatically activated during the vection task by the natural scene which contained many identifiable objects, buildings, etc. and made observers feel to be at a place. This assertion is based on the observation that whenever we are in a place or room and move around, this mechanism is automatically activated and cannot be consciously “switched off”. This is nicely illustrated in a little self-experiment: If one remembers several targets in the immediate surroundings, such as the location and orientation of the window, the door, and the table, etc. relative to yourself, and then you turn and move around for a bit, and are then asked to point to these targets *as if you had not* moved, this is very cognitively demanding and performance is usually bad. Thus, whenever we perceive self-motion through an environment, this mechanism seem to operate. Presence ratings indicated that observers did feel being surrounded by the natural scene, but not by the artificial scenes. The latter were consistently perceived as two-dimensional images and not as a scene or place surrounding oneself, as was confirmed in post-experimental interviews. It is thus possible that the feeling of *not* being in a place vs. the feeling of being at a place extending in 3-D space made a difference. This assertion is admittedly speculative at this stage, but appears worthwhile investigating in future research.

3.2.6 Conclusions

First, we can conclude from Experiments 1 and 2 that the VR setup used is suitable to investigate vection, since results were comparable to classical studies that used optokinetic drums. Remarkably, vection onset latencies were considerably low given that only a relatively small FOV of $84^{\circ} \times 63^{\circ}$ or $54^{\circ} \times 45^{\circ}$ was stimulated, respectively, in contrast to full-field stimulation that is provided in optokinetic drums. Also, simulator sickness was no problem in these two experiments: All participants were able to complete the experiment, and sickness ratings in the SSQ were very low in Experiment 2.

However, further experiments are necessary to clarify what exactly are the relevant stimulus parameters that increase vection with a naturalistic visual stimulus and dissociate between low-level and high-level influences. Furthermore, a significant correlation between the feeling of presence in a Virtual Environment and vection was found. Earlier studies that had failed to find a correlation between presence and vection have used coarse measurement methods both for vection and presence (Freeman et al., 2000; IJsselstein et al., 2001). The current study used more fine grained measurements and found that the attentional aspect of presence correlated negatively with vection onset latency, and vection intensity and convincingness correlated positively with spatial presence. This suggests that there might be a link between the perception of self-motion and presence in VR. It

3 CIRCULAR VECTION EXPERIMENTS IN VR

is possible that both vection and presence are affected by a common mechanism. One possible candidate is the spatial updating mechanism which is activated by the impression of moving in a simulated virtual space which surrounds the observer. This mechanism is unlikely to be activated by watching a 2D image such as the scrambled or sliced stimuli used in Experiment 2. The finding that vection onset latency and compellingness of vection are differentially affected by different mechanisms is in agreement with results of Wright et al. (2006), who found similar results. From the congruent findings of both studies, it is feasible to assume that vection is affected by two mechanisms, one perceptual, the other cognitive. An interesting observation is that these two aspects of vection can be differentially influenced. Wright et al. (2006) found that vection onset latencies were not affected by a cognitive manipulation, but the cognitive compellingness ratings about vection did depend on whether the visual vection stimulus did or did not match with the testing environment.

This finding yields important implications for motion simulation applications: If the relation between self-motion perception and the feeling of presence in VR is further elucidated, future motion simulators might benefit greatly from this finding and make motion simulations more effective and cost-efficient. Especially, possible interactions between low-level perceptual factors, and high-level cognitive influences should be investigated in more detail.

This might eventually lead to a novel motion simulation approach, which brings together both aspects in order to maximise vection and that will perhaps be able to reduce simulation cost.

In the first two experiments, we found that vection is increased by using a natural scene, compared to an artificial image devoid of consistent, global spatial information. In the following experiments, we will investigate possible multi-sensory influences on vection. The idea is that if visual, auditory, and somatosensory information consistently provide ecologically valid information about self-motion, the lacking vestibular information might be overwhelmed by the sensory system. The question we ask in the following experiments is how much visually induced vection is increased by adding other sensory modalities, and what this tells us about perceptual organisation of self-motion perception.

4 Multi-sensory interactions in vection

When we move or locomote in the natural environment, virtually all senses are activated in characteristic patterns of stimulation specifying self-motion. We see optical flow, hear objects pass by, feel the wind on the skin or pressure changes in the back when accelerated strongly in a vehicle, and the vestibular organs react, even though we only rarely become conscious of vestibular sensation (see section 1.2.3 on page 12). When we actively locomote, we have proprioceptive, somatosensory and vestibular information. In the self-motion literature, emphasis has been placed on the visual and vestibular modalities, since those are considered the most important senses for self-motion perception. However, Lackner (1977) for example has shown that a rotating soundfield can induce circular vection and nystagmus in about 66% of blindfolded participants. In this chapter, four new experiments will be reported that investigate multi-sensory interactions during circular and linear vection: In particular, the interaction between visual-vibrational, visual-auditory, visual-vestibular, and also combined tri-modal effects of visual, vestibular, and vibratory stimulation on vection were investigated. The contributions of these modalities to self-motion perception have been investigated only in a few studies so far. The main question asked here is whether vection can be enhanced if consistent information about self-motion is presented to multiple modalities, compared to the visual-only baseline.

4.1 Experiment 3: Visual-vibrational interactions in circular vection

This study investigated whether the visually induced perception of illusory self-motion (vection) can be influenced by vibrational cues applied to the somatosensory senses. Additionally, the notion was tested that vection can be cognitively influenced by making participants believe that they might actually be moved physically. In the experiment, circular vection was induced in 24 participants who were seated on a 6 degree-of-freedom (dof) motion platform and viewed a naturalistic scene of a market place displayed on a curved projection screen. Participants task was to report vection onset by moving a joystick. The convincingness of the illusion was rated by magnitude estimation. Three factors were varied: The presence or absence of vibrations, the velocity profile of the visual stimulus (3 or 12 seconds to reach a constant velocity of $30^\circ/\text{s}$), and the cognitive framework (whether physical motion was possible or not). Vibrations significantly reduced vection onset times and increased convincingness ratings. Trials with a short acceleration period were found to induce vection significantly faster than trials with a long acceleration period. There was no effect of the cognitive manipulation on the vection responses, however, even though 67% of the participants did believe that they had physically moved on some trials. We conclude that the main hypothesis was corroborated: Vibrational cues can enhance vection in absence of vestibular stimulation. These results show that multi-sensory interactions between visual and vibrational cues occur in the perception of self-motion, which has important implications for both the theory of self-motion perception and for motion simulator design.

4 MULTI-SENSORY INTERACTIONS IN VECTION

4.1.1 Introduction

The current experiment investigated whether visually induced vection can be enhanced if concurrent vibrational cues are applied.⁴ There are several possible ways how vibrations might increase vection: One possibility is that the somatosensory information arising from vibrations applied to the skin might bias the brain to perceive self-motion, since in many occasions in which passive movement is experienced (e.g., when we are driving in a vehicle), we often experience vibrations. Another possibility is that somatosensory information could reduce the amount of visual-vestibular conflict and thus increase vection (cf. sections 1.2.3 on page 12 and 1.2.4 on page 14). As mentioned earlier, in many cases, it is nearly impossible to separate out somatosensory and vestibular components during acceleration. In this experiment, we examine the hypothesis that vibrations that are below vestibular threshold applied to the observer's seat can enhance vection. The underlying assumption is that the presence of vibrations might "mute" the veto of the vestibular system which would normally detect the visual-vestibular conflict and thus prevent a quicker onset of vection.

Additionally, we also tested whether prior knowledge that one might really be moved affects the onset of vection. Lepecq et al. (1995) have demonstrated that children developed forward linear vection faster when they were shown that the chair they were sitting on might actually move, even though they never moved in the experiment. In the present study, we tested whether adults' susceptibility to vection can also be cognitively influenced by instructing the participants and demonstrating to them that in one condition, they might really be moved on a 6dof motion platform, whereas in another condition, the platform was switched off and they were explicitly told that physical motion was impossible.

4.1.2 Hypotheses

Hypothesis 1 states that vibrations applied concurrently with a congruent visual stimulus will enhance vection: Vection onset latencies should be reduced, and perceived intensity and convincingness of vection should be increased. Hypothesis 2 states that prior knowledge that one might really be moved physically should enhance vection in the same way as mentioned above.

4.1.3 Methods

Twenty-four participants took part in this experiment (mean age 23.9, 12 male). All had normal or corrected-to-normal vision and no history of vestibular dysfunction. They were paid at standard rates.

STIMULI AND APPARATUS. The stimuli and apparatus used in this experiment were almost identical to those used in Experiment 2. Participants were seated in front of the $54^\circ \times 40^\circ$ projection screen at a distance of 1.8 metres from the centre of the screen, and the setup of the projection system as well as the response measures using the joystick were identical as before. The onset latency of vection and vection intensity were measured via joystick response, and convincingness ratings were obtained using magnitude estimation. The visual stimulus consisted of the natural, photo-realistic image of the Tübingen market place. Circular vection was induced by rotating the visual scene around the yaw axis at a velocity of $30^\circ/\text{sec}$. In 50% of all trials, vibrations applied by

⁴Parts of this study were presented at IMRF (International Multisensory Research Forum) in Barcelona, 2004

4 MULTI-SENSORY INTERACTIONS IN VECTION

shakers were presented simultaneously to visual motion. Vibrations were generated by three shakers (frequency range 15 - 90 Hz) mounted below the seat and floor panel of the motion platform where the participants' feet were resting. The vibrational stimulus was created by mixing and overlaying several recorded sounds of flowing water, transforming them to the frequency range of the shakers, and playing them back on the shakers. This resulted in a continuous, broad frequency vibration. The volume was set to a level that the participants clearly felt tactile vibrations, but the vibrations were sub-threshold for the vestibular system.

PROCEDURE. The procedure was similar to Experiment 2. After participants received exactly the same instructions (see Experiment 2, section 3.2.3 on page 43) and a practice session with the motion platform switched off, the experimental trials were started. Participants initiated trials by pressing a button on the joystick, upon which the scene started to rotate. Participants gave vection responses using the joystick. There was a 15 second pause between trials to avoid motion after-effects and motion sickness. For the same reason, the rotation of the stimulus was stopped automatically if maximum joystick deflection was sustained for 10 seconds. Otherwise, the rotation continued for 60 seconds. Leftward and rightward rotations always alternated. The experiment was run in two blocks with a break in between. In one block, the platform was turned off, and participants were informed that physical motion was impossible. In the other block, the platform was turned on, and participants were told that physical motion was possible. Before participants started this block, they were shown from the outside that the platform can really move. A standard motion trajectory of the platform that included translations and rotations along all six degrees of freedom was demonstrated while participants were standing outside the safety zone of the motion platform. Then, the platform was switched off again so that participants could enter the cabin on the motion platform (see Figure 5 on page 36). During this block, participants were required to put on safety belts. Then, the platform was switched on, and the motion platform moved upwards for about 30cm to the default position, and participants could notice the upward movement of the cabin. During the whole experiment, this was the only real motion of the platform. In order to ensure that participants used the same response criterion in both platform off and on conditions for vection, they were instructed during the practice trials always to pay attention whether they felt self-motion according to the visual stimulus. Since none of the participants had experienced vection in a laboratory before, they were shown a rotating stimulus for about two minutes before the practice trials started, until all of them reported a strong sense of self-motion. They were then instructed to take this feeling of self-motion as a reference for all experimental trials.

DESIGN. Three independent factors were varied: The presence or absence of vibrations, the velocity profile of the visual stimulus (3 or 12 seconds to reach a constant rotational velocity of $30^\circ/\text{s}$, which correspond to $10^\circ/\text{s}^2$ or $2.5^\circ/\text{s}^2$, respectively), and the cognitive framework (whether physical motion was possible or not). The two velocity profiles were used in order to vary the magnitude of the visual-vestibular conflict. The higher visual acceleration induces a stronger visual-vestibular conflict than the low visual acceleration. If vibrations reduce the amount of perceived visual-vestibular conflict during vection, we should observe a larger enhancement of vection in the strong-conflict situation (3 seconds acceleration) than in the weak-conflict situation (12 seconds acceleration). All factors were balanced and presented in a within-subject design in two blocks on the same day, separated by a short break. In one block, the platform was on, in the other, it was turned off. The presentation order of platform on/off blocks was counterbalanced across participants.

4 MULTI-SENSORY INTERACTIONS IN VECTION

Within each block, the presentation order of acceleration profiles and vibrations on/off trials was randomised. Each condition was repeated three times, resulting in a total of 48 trials. The whole experiment took about one and a half hours.

4.1.4 Results

DATA PROCESSING. Pre-analyses showed that vection onset times varied largely between participants: Some participants perceived vection after 2 seconds on average, others took up to 20 seconds on average. Furthermore, for all participants, the very first trials per condition were found to yield much longer onset times and higher variabilities than all following trials. For this reason, the vection onset data were pre-processed in two steps: First, all first-time presentations of each condition were considered as additional practice trials and were excluded from data analysis. In the second step, the between-participant variability was removed by a normalisation process: For each condition per participant, the relative response time was calculated by dividing each value per condition of a participant by the ratio between that participant's mean response time in that condition and the mean response time across all 24 participants in the corresponding condition. Using this procedure, the mean onset time per condition averaged across all participants, remains identical, but the inter-individual differences are removed.

From the processed data, separate repeated-measures ANOVAs were calculated for vection onset time, vection intensity, and for convincingness ratings. The data are plotted in Figure 10. The ANOVA results show the following main effects: Vibrations reduced the vection onset times significantly ($F(1,21)=8.34$, $p < .01$), and increased the convincingness of the illusion ($F(1,21)=7.82$, $p < .05$). The short acceleration time (3 seconds) was found to reduce vection onset times significantly ($F(1,21)=26.74$, $p < .001$), compared to the long acceleration time (12 seconds). The interaction between visual acceleration and vibrations was not significant ($F(1,21)=1.75$, $p=.20$). There was no significant effect of the cognitive manipulation: Vection responses and convincingness ratings did not differ at all between the platform on and off conditions ($F(1,21)=.001$, $p=.974$ for onset time and ($F(1,21)=.032$, $p=.575$) for convincingness, respectively). There was also no effect of the presentation order of platform on and off conditions. The vibrations did not have any influence on vection intensity measured by joystick deflection.

A post-experimental interview revealed that 67% of the participants did believe that they had actually been moved on some trials in the platform-on condition (see Figure 11), but not in the platform-off condition. Interestingly, though, there was absolutely no effect of the cognitive manipulation on the vection responses, on vection onset time nor on the convincingness ratings. In most of the cases where real motion was perceived, vibrations were present. Participants reported that they imagined being on the market place and attributed the vibrations to some sort of locomotion device. This was the case for both the platform on and off conditions.

4.1.5 Discussion and Conclusions

One of the main questions of the current experiment was whether tactile vibrations that are presented concurrently to a visual stimulus influence the vection illusion. It was found that vibrations reduced vection onset latencies and increased convincingness of vection. Hypothesis 1 was thus supported by the data. The magnitude of the vection enhancing effect of vibrations was equal in all condi-

4 MULTI-SENSORY INTERACTIONS IN VECTION

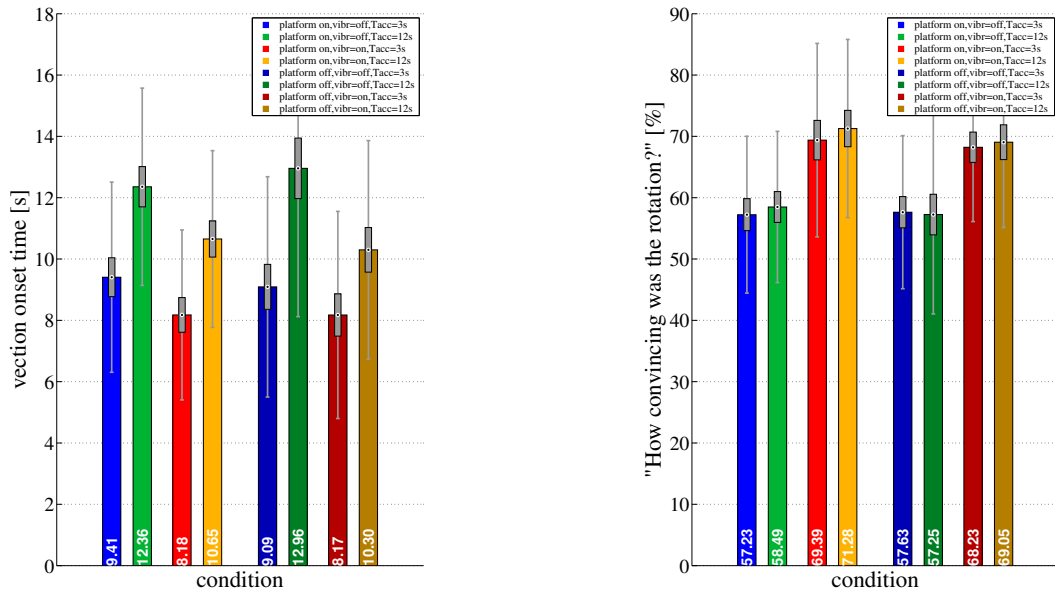


Figure 10: Left: Normalised mean vection onset times: Shorter acceleration times yielded quicker onset times and trials with vibrations show shorter onset latencies than trials without vibrations. There is no difference between platform on (4 left bars) and platform off conditions (4 right bars). Boxes show one standard error, whiskers denote one standard deviation. Right: Normalized mean convincingness ratings: Trials with vibrations show much higher ratings. No difference between short and long accelerations were found. There is also no difference between platform on (4 left bars) and platform off conditions (4 right bars).

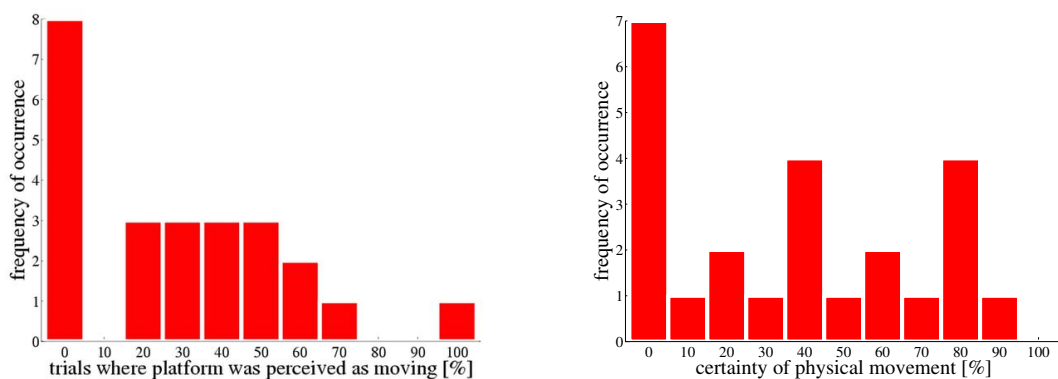


Figure 11: Histogram of number of responses how often participants perceived that the motion platform had moved in the platform on condition (left), and rated certainty about physical displacement on a 0-100% scale (right). In the left plot, eight participants said the platform never moved (0%), and one participant said that it always moved (100%). In the right plot, seven participants were absolutely sure that the platform never moved (0%), and four participants rated the certainty that the platform moved to be 80%.

4 MULTI-SENSORY INTERACTIONS IN VECTION

tions (platform on/off; high/low visual-vestibular conflict), and no interaction was observed between vibrations and visual acceleration times, as well as the platform on/off condition. The presence of vibrations generally increased vection across all conditions for the same amount (about 10-15%). The reason why we did not observe any interaction between visual acceleration time (i.e., large or small visual-vestibular conflict) and vection might be that the mere presence of vibrations was sufficient to silence the veto of the vestibular system which would otherwise prevent quick vection onset due to the visual-vestibular mismatch. However, no effect of the vibrations was observed for vection intensity ratings. It is possible that this measurement method using the degree of joystick-deflection was too noise-sensitive in this case, since it might have been rather demanding for the observers to indicate the perceived intensity of vection using this method. Note, however, that we did find effects in Experiments 1 and 2 for vection intensity using this method. Therefore, it might be that perceived vection intensity is not affected by the presence of vibrations, contrary to vection onset latency and convincingness.

Another finding of the current study is related to the acceleration time of the visual stimulus: Two different visual accelerations were used in order to manipulate the magnitude of visual-vestibular conflict, since we expected an interaction of this factor with the vibrations. While no such interaction could be observed, a main effect of acceleration time was found: The strong acceleration of the visual stimulus ($10^\circ/\text{s}^2$) that lasted for 3 seconds was found to yield quicker vection onset times than the weak acceleration ($2.5^\circ/\text{s}^2$) which lasted for 12 seconds. This finding is inconsistent with Melcher & Henn (1981) who had found the shortest vection onset latencies at visual accelerations of $5^\circ/\text{s}^2$ and longer delays of vection onset for accelerations of $10^\circ/\text{s}^2$ and above. However, comparisons between these two studies need to be done with caution, since in their study, an optokinetic drum was used, while an immersive VR setup with a much smaller FOV of the stimulated visual area was used in the present studies. According to Dichgans & Brandt (1978), the ideal stimulus to induce vection is a large, uniformly moving visual stimulus moving at constant velocity. In our 3 seconds acceleration condition, the rotating stimulus reached constant velocity earlier than in the 12 seconds acceleration condition. However, the average image velocity during a trial was also slower in the 12 seconds acceleration condition. Given that in Experiments 1 and 2, we had found shorter vection latencies for faster rotations, it is questionable whether the strength of the visual-vestibular conflict is really the primary determining parameter for the onset latency of vection. The results from Experiments 1, 2 and 3 indicate that vection develops fastest if the initially static stimulus accelerates at relatively high rate to reach constant velocity quickly. For example, in Experiment 1, we found the shortest average vection onset latencies of 6.5s for the condition where the stimulus accelerated in 3 seconds to $60^\circ/\text{s}$. This issue will be addressed in more detail later in the general discussion.

The second question asked here was whether participants can be cognitively biased in their vection reports. The cognitive manipulation used in this experiment of whether physical motion was possible or not did not have an effect on the vection responses. There was absolutely no difference at all in any of the vection measures between the platform on and off conditions. This is especially interesting since 67% of the participants did believe that the motion platform had really moved in some trials in the platform on condition, whereas no participant reported that for the platform off condition. This means that the vection responses were absolutely not affected by whether the observers thought they were moving physically or not. It seems that for the majority of the participants the self-motion illusion induced by the VR setup was so compelling that it was impossible to dis-

4 MULTI-SENSORY INTERACTIONS IN VECTION

tinguish the illusion from actual motion. On the other hand, eight participants correctly stated in the interview after the experiment that they were sure the platform never moved. Interestingly, there was no difference in the response behaviour for vection between these two groups. The fact that the current study failed to find a cognitive influence on vection is contrary to the studies that had found such an effect (Lepecq et al., 1995; Palmisano & Chan, 2004). Palmisano et al. found an effect of experimental demand in vection: One group was instructed to report the onset of vection by pressing a button as soon as the impression that they were watching a moving image turned into a sensation of self-motion, while another group was instructed to report the onset of object-motion by pressing the button and to release the button when the sensation of object motion changed to a sensation of self-motion. The latter group reported the onset of vection later than the former group. There is a methodological difference between these two studies and the current one, which is that here, a within-subject design was used, while the other two studies both used a between-subject design. While between-subject designs have the advantage that possible order- and carry-over effects are eliminated, they have the disadvantage that the large inter-individual differences concerning vection responses can obscure the results. Since it is well known that vection responses vary largely between individuals (Kennedy et al., 1996), the current study used a within-subject design. It is possible that the cognitive manipulation of the current study was not strong enough to be effective in a within-design. Indeed, at least eight participants noticed that the motion platform never moved. However, as already stated above, no difference in any vection responses could be observed between the observers who thought they had really moved and those who thought they had not moved. This finding is consistent with a recent experiment by Wright et al. (2006) who also found that vection onset latencies were unaffected whether the visual stimulus matched or did not match the spatial, environmental context. Lepecq et al. (1995) also did not find any effect of prior knowledge for adults, but only for children. So, despite the methodological limitation in the current study, it might be tentatively concluded that adults' vection responses are not affected by prior knowledge that they might actually move physically. One possible reason for this could be that the Virtual Reality setup with the large visual FOV used in this experiment induced a very strong, compelling sense of self-motion, and this might have caused a ceiling effect in the "platform off" condition.

The question remains whether the main finding from the current experiment that vibrations enhanced vection is a perceptual result of multi-sensory interaction, or whether observers were cognitively biased by the vibrations and reported that vection developed faster and was more convincing. From the current data, it is difficult to draw definite conclusions. Notably, the originally intended cognitive manipulation to bias observers cognitively by telling them the platform would really move did not work. Furthermore, observers stated consistently in the post-experimental interview that the presence of vibrations made the impression that one is moving more compelling. Actually, most cases where participants thought that they were really moving during the "platform on" block were trials with vibrations. Some observers stated that the vibrations supported the imagination that they were moving on some kind of rotating disc in the simulated market place scene. In many everyday situations (e.g., when we drive), we experience vibrations during self-motion. It might well be that due to this fact, the co-occurrence of visual motion and vibrations lead the observers to perceive stronger vection. While no conclusive answer can be derived from the data, the finding that vibrations increase vection are highly relevant for motion simulation applications and motion simulator

4 MULTI-SENSORY INTERACTIONS IN VECTION

design. This fact will be discussed later in section 5 together with the results of all multi-sensory experiments.

4.2 Experiment 4: Visual-auditory interactions in circular vection

The current study investigated visuo-auditory interactions for circular vection by testing whether adding an acoustic landmark that moves together with the visual stimulus enhances vection. Twenty observers viewed a photorealistic scene of a market place that was projected onto a curved projection screen (FOV $54^{\circ} \times 40^{\circ}$). In each trial, the visual scene rotated at $30^{\circ}/s$ around the earth-vertical axis. Three conditions were randomized in a within-subject-design: No-sound, mono-sound, and spatialised-sound (moving together with the visual scene) played through headphones using a generic head-related transfer function (HRTF). We used sounds of flowing water, which matched the visual depiction of a fountain that was visible in the market scene. Participants indicated vection onset by deflecting the joystick in the direction of perceived self-motion. The convincingness of the illusion was rated on an 11-point scale (0-100%). Only the spatialized-sound that moved according to the visual stimulus increased vection significantly: Convincingness ratings increased from 60.2 % (mono-sound) to 69.6% (spatialised-sound), and the latency from vection onset until saturated vection decreased from 12.5 sec. (mono) to 11.1 sec. (spatialised-sound). In addition, presence ratings assessed by the IPQ Presence Questionnaire were slightly but significantly increased. Average vection onset times, however, were not affected by the auditory stimuli. We conclude that spatialised-sound that moves concordantly with a matching visual stimulus can enhance vection. The magnitude of the effect was, however, rather small (15%). These results indicate visual-auditory interaction effects and have important implications for our understanding of multi-modal cue-integration during self-motion.

4.2.1 Introduction

In Experiment 3, it was found that vibrations that are presented simultaneously with visual motion can increase vection. It is likely that similar multi-sensory interactions during self-motion perception also occur between other modalities. The current experiment investigated whether visually induced vection can be enhanced by adding matching auditory stimuli.⁵ It has been shown that a rotating soundfield can induce circular vection and nystagmus in blindfolded observers (Lackner, 1977). In that study, an array of speakers playing either white noise or a sinus tone was rotated physically around the seated participants. However, vection induced by auditory stimuli was much weaker than visually induced vection, and only in 50 to 70% of the trials did the illusion occur. A similar result was recently obtained by Larsson et al. (2004) and also by Völjamäe et al. (2004) who used headphone-based auralization with generic head-related transfer functions (HRTFs). Both rotational and translational vection could successfully be induced using this technology in blindfolded participants. Furthermore, it was found that both the realism of the acoustic simulation and increasing numbers of sound sources enhanced auditory vection. The finding that auditory induced vection was less frequent and less compelling than visually induced vection might reflect the fact that auditory information receives less weight in the sensory integration process in the brain during self-motion.

⁵Parts of this study were presented at the 8th International Workshop on Presence in London, 2005

4 MULTI-SENSORY INTERACTIONS IN VECTION

A similar conclusion was drawn by Marme-Karelse & Bles (1977) who found that only a few participants perceived auditory induced circular vection, and that also the audiokinetic after-effect was much weaker and ended after only 3 seconds, while the optokinetic aftereffect lasted for about 20 seconds. One reason for this might be that in the spatial domain, visual information is generally more accurate and reliable, whereas auditory information is generally more reliable in the temporal domain (Goldstein, 1999).

There is one study which investigated the influence of moving auditory cues on self-motion perception under near-threshold vestibular stimulation: In this study, Schinauer, Hellmann, & Höger (1993) had blindfolded participants sit on a swivel chair and judge the direction and the strength of perceived self-rotation. Auditory stimulation was provided via headphones that played pre-recorded sounds of four sound sources (typewriter, splashing water, stereophonic music) distributed in the room. The binaural recordings were done using two microphones placed in the ears of a dummy head that was rotated at the speed of vestibular stimulation ($16^\circ/\text{sec}$). The authors found that in the condition where auditory motion direction was compatible with physical motion, responses were significantly increased compared to conditions where the auditory scene either did not move or moved in the “wrong” direction (incompatible conditions). These results show that dynamic auditory information can influence perceived self-motion in cases where self-motion information from other senses (in this case, vestibular) is ambiguous or at near-threshold levels.

In the current experiment, we investigated if visually induced vection can be further increased using spatialised dynamic sound. The rationale for the experiment is that despite the relatively weak auditory effects found previously, vection might be enhanced as a result of synergistic sensory interaction effects if consistent visual and auditory information is presented in a VR motion simulation setup.

4.2.2 Hypotheses

Two hypotheses were tested. First, we investigated the influence of adding non-spatialised auditory cues. The prediction of Hypothesis 1 is that adding auditory information that semantically fits to the visual scene should increase vection, even though the spatial and dynamical aspects of the sound are in conflict with the moving visual scene. To address this issue, a no-sound condition (purely visual baseline) was compared with a condition where a simple mono rendering of flowing water could be heard. This sound was intended to mimic the sound of the fountain on the market place scene that was used as the visual stimulus. A stationary sound of flowing water was played through the headphones, which did not move concordantly with its’ visual counterpart. If we found an enhancement of vection despite the conflict between the moving visual stimulus and stationary auditory stimulus, this would suggest a general bias which would be likely to have a cognitive origin. Second, we tested the influence of adding spatialised acoustic information. Hypothesis 2 predicts that only correctly spatialised moving sound sources added on top of visual stimuli will enhance visually induced vection. This hypothesis was tested by comparing the above mentioned mono-sound condition with a proper spatialised acoustic rendering of the correct location of the landmark using a generic HRTF. Due to the HRTF, the acoustic landmark appeared properly externalised and spatialised, and the distance of the visual and auditory representation of the fountain to the observer was always matched correctly. In this condition, the fountain was always audible (as we have omni-directional hearing), even when the visual counterpart was outside of the current field of view, and it always

4 MULTI-SENSORY INTERACTIONS IN VECTION

moved concordantly with the visible fountain. If vection is enhanced for this condition, but not in the mono-sound situation, this would imply that matching spatio-temporal information across the visual and auditory channel is necessary to increase vection.

4.2.3 Methods

Twenty participants (eight male, ages 18-36, mean 27.8) who were naïve to the purpose of the experiment took part in the study. All had normal or corrected-to-normal vision and no history of vestibular disorders. It was also confirmed that all participants were able to correctly localise the spatialised sound source.

STIMULI AND APPARATUS. The physical setup and visual stimulus used in this experiment were identical to Experiment 3 (see Section 4.1.3 on page 54). The natural, photorealistic scene of the market place was used, and viewing distance, projection devices, visible FOV ($54^{\circ} \times 40^{\circ}$) and vection measures were exactly the same as used before. The visual scene accelerated within 3 seconds to $30^{\circ}/s$ and stopped after 46 seconds. The same force-feedback joystick was used to record vection responses. Force-feedback was used in order to always recenter the joystick after each trial. Participants wore active noise-cancelling headphones (Sennheiser HMEC 300) that either played no sound, a mono-rendering static sound of flowing water, or spatialised acoustic rendering of flowing water that rotated around the observer using a generic HRTF. This stimulus was created using a hardware-based rendering technique with a Lake DSP system (Huron engine).

PROCEDURE. The same experimental procedure was followed as in Experiment 3. Prior to the practice session, all participants were exposed to a rotating visual scene for maximally two minutes, until all reported a strong sensation of self-rotation. After that, participants were instructed exactly the same as in Experiment 3, and the practice session started. After eight practice trials, all participants reported that they were confident in their vection responses. Again, vection onset latency, vection intensity, and convincingness of rotation were recorded. Participants always started trials by pressing a button, at which time the visual scene started to rotate. Between trials, there was a 15 seconds break so that potential motion after-effects would subside and the occurrence of motion sickness would be prevented. The static visual scene remained visible during that pause, and also the sound kept playing if present in the trial. Leftward and rightward rotations always alternated in order to prevent motion sickness. In total, the experiment lasted one and a half hours.

DESIGN. Three conditions were presented in a within-subject-design. The no-sound, mono-sound, and spatialised sound conditions were presented in randomised order, and the visual scene was always the same. Participants completed 46 trials each in a 3 auditory conditions \times 4 repetitions \times 2 turning directions \times 2 sessions design. Between the two sessions, a short break was taken to prevent fatigue. Within both sessions, the order of conditions was randomised. After the experiment, spatial presence was assessed using the IPQ Presence Questionnaire (Schubert et al., 2001) for all three auditory conditions.

4.2.4 Results

Separate within-subject repeated-measures ANOVAS with the factor audio (no audio, mono, spatialised) were computed for vection onset time, vection intensity, vection saturation time, and convincingness ratings. Vection saturation time was computed from raw data and describes the time it

4 MULTI-SENSORY INTERACTIONS IN VECTION

takes for vection to saturate from the moment vection is first reported. The only significant main effects of the auditory conditions on vection responses was found for convincingness ($F(2,34)=8.638$, $p < .005$) and for vection buildup time ($F(2,34)=4.069$, $p < .05$). Vection intensity and vection onset times were unaffected by the auditory manipulation ($F(2,34)=2.286$, $p=.117$ and $F(2,34)=.217$, $p=.081$, respectively). Results of the four vection measures across the three sound conditions are summarised in Figure 12. The results of paired t-tests (Bonferroni-corrected) are indicated at the top inset of each plot. There was a slight trend towards increased convincingness ratings for the mono sound condition, but this was not statistically significant. All other vection measures showed no difference between the no sound and mono sound condition. Comparing the mono condition with the spatialised sound condition demonstrates, however, a small but consistent significant vection-facilitating effect of the sound spatialisation. The strongest effect was observed for the convincingness ratings (16% increase) and the vection saturation time (12% decrease). The other vection measures show only small and non-significant effects, albeit in the predicted direction according to Hypothesis 2. A similarly small, but consistent advantage for the spatialised sound can be observed for the presence ratings, which are summarised in Figure 13. This effect reached significance for the presence sum score and the spatial presence sub-scale. In addition, the realism sub-scale showed a marginally significant effect. The other presence sub-scales did not show any significant effects.

4.2.5 Discussion and Conclusion

In this experiment, a small but significant enhancement of vection was observed when spatialised auditory information was added to visual information: Vection saturation time was reduced, and convincingness of self-motion was increased. However, vection onset latency and vection intensity were unaffected. Thus, the data support Hypothesis 2: Only properly spatialised auditory information added to visual information increased vection. No support was found for Hypothesis 1 which predicted an enhancement of vection by adding non-spatialised mono-sound that meaningfully matched to the visual scene. Even though adding mono-sound showed a slight but non-significant trend for increased convincingness of the self-motion sensation, neither the presence ratings nor any of the other vection measures were affected. This argues against a non-specific benefit of just adding audio cues which could bias observers to report stronger vection. Only when the sound source was actually perceived to originate from the same location as its visual counterpart did we observe a significant increase of both vection and presence, which argues for a specific facilitation due to the spatialisation of the sound source. This indicates that vection can be increased if consistent multi-sensory motion information is presented to the perceptual system. The fact that we found only a small enhancement of vection in two of four vection measures might reflect that visual information is weighted much stronger in the sensory integration process during self-motion than auditory information. Furthermore, the natural, photorealistic visual scene used in this experiment had been found to have a strong vection-inducing potency in Experiments 1, 2 and 3, so that a ceiling effect might have obscured a stronger auditory effect. Also of interest, there was only one auditory sound source combined with a full visual scene, and the sound of the water flowing in the fountain was rendered “dry”, without room acoustics of the market place (i.e., without reflected sound). Using the identical audio rendering technique with HRTFs, auditory vection has been found to be increased when several sound sources were used instead of only one, and realistic room acoustics simulation also increased auditory vection, compared to unechoic renderings (Larsson et al., 2004). It is thus

4 MULTI-SENSORY INTERACTIONS IN VECTION

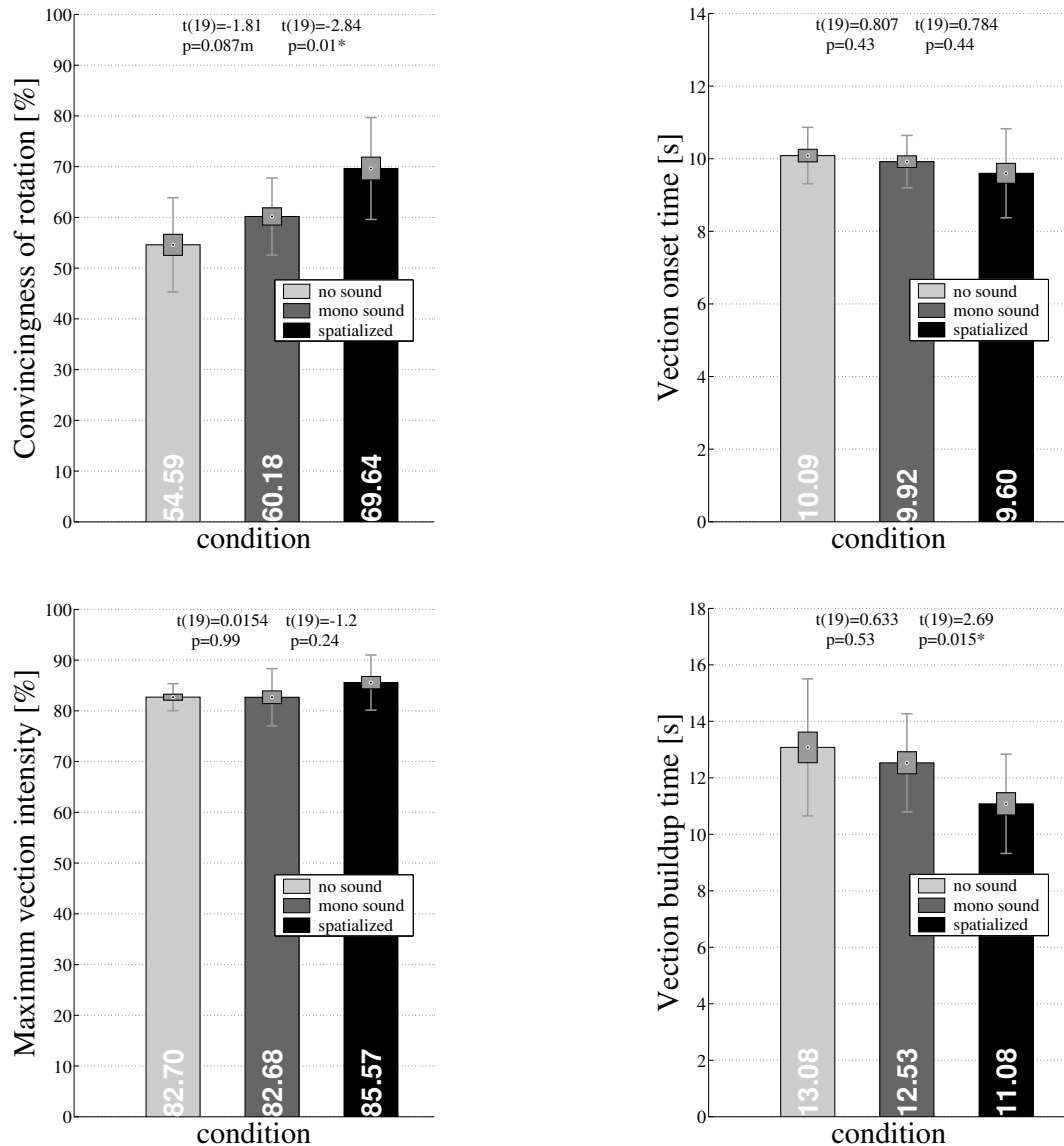


Figure 12: Mean of the four vection measures, averaged over 20 participants. Boxes indicate one standard error of the mean, whiskers depict one standard deviation. The results of pairwise comparisons between the three sound conditions using paired t-tests are indicated in the top inset of each plot. Note the small but consistent vection-facilitating effect of the proper spatialised auditory rendering of the fountain sound (black bars on the right) as compared to simple mono display (middle dark-grey bars). There were no significant differences between using mono sound and no sound at all.

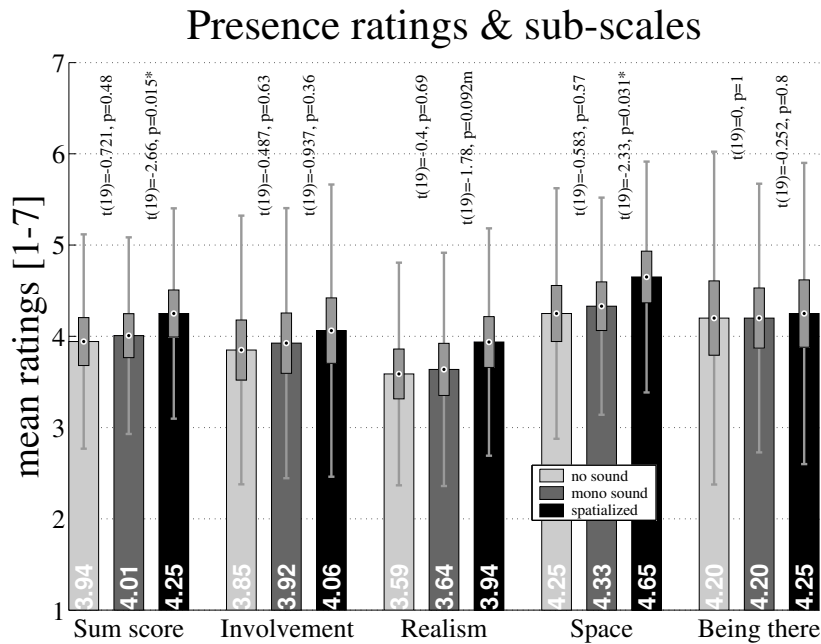


Figure 13: Presence ratings for the three sound conditions. The sum score over all 14 items of the Igroup Presence Questionnaire (left three bars) were split up according to the four original sub-scales described by Schubert et al.: Involvement, realism, space, and being there. Even though the effect was quite small (<6%), the presence ratings were consistently higher for the spatialised sound condition.

possible that using more sound sources and rendering acoustic reflections and later reverberations in the simulated scene might show a stronger vection enhancement by acoustic information.

Similar to Experiment 2, we found an increase of both presence and vection in the most natural, realistic experimental condition where the sound moved together with the visual counterpart. From the current data, it is not possible to clarify whether there is a causal relationship or mediation between presence and vection. However, the significant positive correlation between presence and vection strength in Experiment 2 and the similar result from the current study indicate that there might be a closer link than previously thought between vection and higher-level influences such as presence. It is, of course, also conceivable that the increased multi-sensory realism using versatile VR technology increased both vection and presence. Further systematic research is needed to investigate the relationship between sensory realism, presence in VR, and vection.

4 MULTI-SENSORY INTERACTIONS IN VECTION

4.3 Experiment 5: Visual-vestibular interactions in forward linear vection

This experiment investigated how vestibular information can enhance visually induced forward linear vection. Forward vection was induced by showing an optic flow stimulus which simulated motion through a long street in Virtual Tübingen. Vestibular stimulation was provided using a 6 dof motion platform. Two factors were manipulated: Visual acceleration and vestibular acceleration. Visual acceleration was varied between 1.2 m/s^2 and 12 m/s^2 . After the acceleration, the visual speed was 6 m/s in all trials. In 50% of trials, the motion platform delivered an initial “kick” by moving forward for a small amount, either for 1 cm (moderate kick) or 3 cm (strong kick). Vection onset time, vection intensity, and convincingness of vection was measured. It was hypothesised that vection should be most enhanced if visual and vestibular accelerations match, since visual-vestibular conflict is eliminated. Results showed that vection was significantly increased by vestibular stimulation. However, no interaction between strength of the vestibular kick and visual acceleration was observed. It is concluded that for the range of accelerations tested here, the mismatch between visual and vestibular accelerations did not interfere with the general benefit of adding vestibular stimulation to increase vection.

4.3.1 Introduction

The experiments reported so far have investigated circular vection around the vertical (yaw) axis. Circular vection was found to be increased when the used stimuli induced a higher feeling of presence in the Virtual Environment (Exp. 2 & 4). Furthermore, vection was increased if the rotating visual scene was supplemented by tactile vibrations (Exp. 3) or by a spatialised sound source that moved concordantly with the visual scene (Exp. 4). The following two experiments investigated whether similar multi-sensory benefits would be observed for forward linear vection. Experiment 5 explores how minimal motion cueing using a Stewart motion platform can be used to enhance self-motion illusions in Virtual Environments.

It is generally believed that the visual and vestibular modalities are the predominant senses for the perception of self-motion (Warren & Wertheim, 1990). Most high-fidelity motion simulators used for pilot training or driving research combine a six-degree-of-freedom (6dof) motion platform for vestibular stimulation with a visual display to simulate self-motion. The most commonly used type of motion platform is the so-called Stewart motion platform which features six independently movable legs that support and move the platform. The current experiment was conducted using such a Stewart platform in the Motion Lab at the MPI for Biological Cybernetics. It has previously been shown for the case of circular vection that a short initial physical acceleration which coincides with the visual motion can speed up the occurrence of vection (Wong & Frost, 1981). In that study, participants were seated on a rotatable chair inside an optokinetic drum which could both be rotated independently, and their task was to report self-rotation as soon as it was perceived. It was found that corroborating chair rotation at the onset of drum rotation (i.e., turning the chair in opposite direction as the optokinetic drum) reduced vection onset and saturation latencies. The authors took these results as support for the notion that the initial visual-vestibular conflict is responsible for the delay of vection onset, and that reducing this conflict by an initial acceleration in the opposite direction to the visual stimulus speeds up the development of vection. The authors found no difference when the

4 MULTI-SENSORY INTERACTIONS IN VECTION

chair was rotated into the “wrong” direction, (i.e., when the chair turned into the same direction as the visual stimulus) compared to the purely visual condition where no chair rotation was presented. This means that a “correct” vestibular stimulus enhanced vection, whereas a “wrong” vestibular stimulus did not reduce vection.

The present experiment aimed to investigate whether a similar enhancing effect can be found for the case of forward linear vection.

4.3.2 Hypotheses

The hypothesis was tested that adding an initial forward acceleration to the visual stimulus should speed up the onset of vection. More specifically, it was assumed that this facilitating effect should be strongest if the magnitude of visual and physical acceleration match, since in this condition, the visual-vestibular conflict should be almost eliminated.

4.3.3 Methods

Fourteen participants (7 male, 7 female, mean age 23.3 years) with normal or corrected-to-normal vision and no history of vestibular dysfunction participated in this study. They were recruited from the subject database at the MPI and paid at standard rates. Participation was completely voluntary.

APPARATUS AND STIMULI: The setup used in this experiment was identical to that used in the previous experiments, with a few exceptions (see schematic in Figure 5 on page 36). Other than the four preceding experiments, the visual stimulus consisted of a long street in Virtual Tübingen, whereas a roundshot photograph of the real Tübingen market place was used for the first four experiments. Participants were seated on the motion platform and viewed a forward optic flow stimulus simulating motion through a long, straight street in the virtual model of Tübingen (see Figure 14), which is a highly detailed copy of the historical old town realised at the MPI for Biological Cybernetics (see also <http://virtual.tuebingen.mpg.de>). The model is based on high quality pictures of real houses that are mapped onto a true-to-scale 3 dimensional model. In this study, a flat projection screen and a projector with 1400x1050 pixels resolution was used (JVC D-ILA SX 21), which was more suitable to project undistorted forward linear optic flow than a curved screen. The participants sat at a distance of 108 cm from the screen, which resulted in a field of view of 75° horizontally and 55° vertically. Again, a standard Microsoft Force-feedback joystick was used to record participants' vection responses, and they wore headphones with active noise cancellation (Sennheiser Noisegard HMEC 300) through which we played a looping sound of flowing water to mask any external sounds. Additionally, we were able to give participants verbal instructions through the headphones via microphone. Participants were seated in the cabin mounted on top of the motion platform. The cabin was completely closed with black curtains and the lights of room were switched off in order to prevent the participants from seeing the surroundings of the lab. Physical motion was delivered using the 6 degrees of freedom motion platform (Motionbase Maxcue). One third of the trials were presented without the platform moving, and two thirds of the trials were presented with platform motion. In those trials that contained physical motion, a “jerk” that coincided with the start of the visual acceleration was applied, moving the participants forward by a very short distance (1 or 3 cm) during a short time period of about one second. Figure 15 illustrates the acceleration profiles of the two stimuli as recorded with accelerometers attached to a participant's head. The accelerations

4 MULTI-SENSORY INTERACTIONS IN VECTION



Figure 14: Participant seated on motion platform viewing the optic flow stimulus that was created using the Virtual Tübingen model. Participants wore active noise-cancellation headphones and used a Microsoft Force-Feedback Joystick to give the vection responses. During the experiment, the curtains were closed and lights were switched off.



Figure 15: Physical acceleration profiles, recorded using an accelerometer attached to a participant's head. The left profile shows the 1 cm platform travel condition and the right one the 3 cm platform travel condition. The vertical red line shows the point at which the motion started.

gave participants the sensation of a gentle kick from the back, similar to what can be sometimes felt when a vehicle starts moving. The platform was moved back to its original position during the convincingness rating and the 6 seconds pause that followed each trial. A simple spring model with a high damping value which ensured a very smooth, sub-threshold motion was used for the resetting of the platform.

PROCEDURE. The experimental procedure and instructions were almost identical as those in the previous experiments. First, participants were exposed to a simulated forward movement through Virtual Tübingen, and all participants reported a feeling of self-motion after a while. Then, they were instructed to report the onset, the intensity and convincingness of vection using exactly the same method as in all previous experiments. The only difference here was that participants moved the joystick forwards to report vection, since they were asked to indicate the direction of perceived self-motion. Care was taken to ensure that participants reported the onset of vection and not the noticed real motion of the motion platform. This was done by instructing them to judge whether they felt they were moving according to the visual stimulus. All participants stated they were confident giving the vection responses after six practice trials. It took roughly one hour to complete the experiment.

4 MULTI-SENSORY INTERACTIONS IN VECTION

DESIGN. Two independent factors were varied in a within-subject factorial design: The length of the visual acceleration phase, and the strength of the kick generated by the motion platform. The visual speed during the constant velocity phase was the same across all the trials (6 m/s), so varying the length of the acceleration phase also varied the acceleration value. Two visual acceleration times were used: 0.5 and 5 seconds, resulting in visual accelerations of 12 m/s^2 and 1.2 m/s^2 , respectively. The second factor was the strength of the platform kick. To produce the kicks, the platform was programmed to move by a very short distance, but as fast as possible, in order to produce a very short but relatively strong physical acceleration. The strength was varied by changing the length travelled by the motion platform. Three conditions were used: no movement at all, 1cm and 3cm motion. These motion trajectories corresponded to acceleration values of about 1 or 2 m/s^2 , respectively. The platforms movements were smoothed out by software filters provided by the manufacturer. All factors were balanced and each combination was presented 4 times in a random order. We expected the strongest enhancement ofvection for the cases when visual and physical acceleration profiles matched, (i.e., when the strong jerk is coupled with the higher visual acceleration and when the soft jerk is coupled with the slow visual acceleration).

4.3.4 Results

Since large inter-individual differences invection onset latencies were found, the variability between participants was removed by a normalisation process: For each condition per participant, the relative response time was calculated by dividing each value per condition of a participant by the ratio between that participant's mean response time in that condition and the mean response time across all 14 participants in the corresponding condition. Using this procedure, the mean onset time per condition, averaged across all participants, remains identical, but the inter-individual differences are removed.

Separate repeated-measures ANOVAs with the factors platform motion (no motion, small kick, large kick) and visual acceleration were calculated for the convincingness rating,vection onset time, maximumvection intensity and the time to reach the maximum intensity value. Results are plotted in Figure 16.

VECTION ONSET TIME: The platform motion had a significant main effect on thevection onset time ($F(2,26)=8.726$, $p < .005$). However, there was no significant influence of the strength of the kick. The overall meanvection onset time decreased from 14.3 s for the no kick condition to 6.7 s and 7.8 s for the smaller and larger kicks, respectively. No significant effect was observed for the visual acceleration time ($F(1,13)=1.05$, $p=.324$). There was no significant interaction between visual acceleration time and platform motion ($F(2,26)=2.089$, $p=.144$). The results are represented in the top left graph in Figure 16.

CONVINCINGNESS RATINGS: The platform kicks also showed a significant influence on the convincingness rating ($F(2,26)=6.991$, $p=.004$). The presence of the kicks increased the overall mean convincingness of the perceived self-motion from 31.6% for the purely visual condition to 48.1% and 49.5% for the 1 cm and the 3 cm kicks, respectively. Again, no significant effect of visual acceleration time was observed ($F(1,13)=1.281$, $p=.278$). The data shows no influence of the strength of the kicks on convincingness ratings (for the 0.5s acceleration phase, $t(13)=-1.26$, $p=.23$, for the 5 sec. acceleration phase, $t(13)=0.00001$, $p=.99$). The plots are shown on the top right graph in Figure 16.

4 MULTI-SENSORY INTERACTIONS IN VECTION

VECTION INTENSITY: The presence of kicks had a significant effect on the maximum deflection of the joystick ($F(2,26)=5.223$, $p=.012$), which was used as an indication of the perceived strength of vection. The overall mean value increased from 28.3% of the joystick range to 46% and 47.6% for the smaller and larger kick, respectively. Again, no significant effect of visual acceleration time was observed ($F(1,13)=0.946$, $p=.348$). The data shows no influence of the strength of the kick on vection intensity ratings (for the 0.5s acceleration phase, $t(13)=-2.16$, $p=.05$, for the 5s acceleration phase, $t(13)=1.45$, $p=.17$). Results are plotted in the bottom left graph in Figure 16.

VECTION BUILDUP TIME: In this experiment, some observers did not develop saturated vection. It is thus more adequate to use the term “vection buildup time”, rather than saturation time. The acceleration value was found to significantly affect the time it took the participants to reach the maximum vection strength during the trial ($F(1,13)=16.632$, $p=.001$). The mean values were 14.36 s for the 0.5 s acceleration phase, and 19.58 s for the 5 s acceleration phase. However, the acceleration had no effect on the vection onset time. The data is represented in the bottom right graph in Figure 16.

4.3.5 Discussion and Conclusion

In the current experiment, we found that forward linear vection can be enhanced if corroborating vestibular stimulation is provided. Small initial accelerations (kicks) that were applied by moving the motion platform 1 or 3 cm were sufficient to speed up vection onset by more than 50%. These results are consistent with earlier findings that found a similar effect on circular vection (Wong & Frost, 1981). While this result was to be expected, one surprising observation was that we did not find any effect of the magnitude of the kicks: Both the 1 and 3 cm platform motion reduced vection onset latencies, increased intensity and convincingness ratings for the same amount. Notably, there was no two-way interaction between kick magnitude and visual acceleration time, which we had predicted. This seems to indicate that exact quantitative matching of visual and vestibular acceleration magnitude was not so critical. In the debriefing after the experiment, it turned out that some participants had not noticed that two different visual accelerations were used. Two participants were actually unaware that two different platform accelerations had been presented in the experiment. This is surprising since the difference of the two platform accelerations are clearly above threshold, and in a pre-experiment, lab-members of the institute were able to easily notice the difference of the two platform accelerations in the absence of visual stimulation. It is possible that participants were absorbed so much in the task of reporting whether they felt to be moving according to the visual stimulus that these differences went unnoticed for some of them. One also has to consider the fact that the 3 cm trajectory did not precisely match the strong visual acceleration in physical terms, which was due to technical limitations of the motion platform: The platform was programmed to move as fast as possible for 3 cm from one frame to the next at 100 Hz update rate, but due to the mass of the cabin and limited power of the electric step motors in the platform legs, the actual acceleration that the platform instantiated was less strong than intended. The actual acceleration as measured using the accelerometer attached to a participant’s head is also only an approximation, due to damping caused by the seat and possible minimal neck movements of the observer. All participants had been instructed to lean their head to the back of the seat, and for safety reasons, no rigid head fixation could be used in this experiment. Despite these limitations, it appears feasible to conclude that it was more important that there was a noticeable platform kick that *qualitatively*

4 MULTI-SENSORY INTERACTIONS IN VECTION

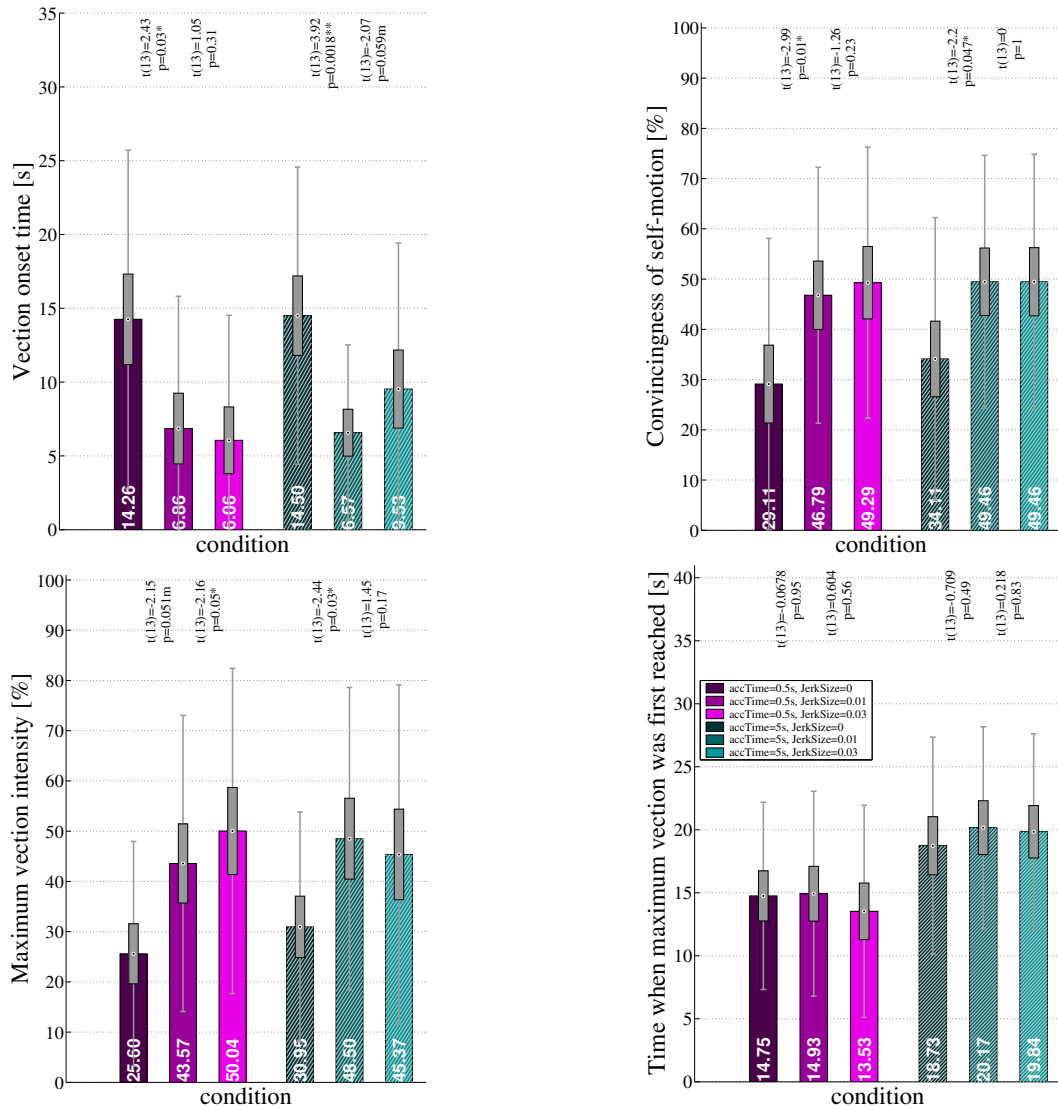


Figure 16: Mean vection data for the 14 participants: Top left: Vection onset latencies. Top right: Convincingness ratings. Bottom left: Maximum vection intensity. Bottom right: Vection buildup time. Boxes represent one standard error of the mean, whiskers indicate one standard deviation.

4 MULTI-SENSORY INTERACTIONS IN VECTION

matched the visual stimulus in order to enhance vection, but that the *quantitative* matching of visual and physical acceleration did not play any important role, at least within the range of values tested here. It is known that certain amounts of mismatch between visual and physical acceleration can not be detected by the human perceptual system (van der Steen, 1998). This is mainly due to biophysical mechanisms of the vestibular organs, as explicated in Section 1.2.3 on page 12. This is a very important “feature” of our perceptual system for motion simulation applications, and the exploration of the exact dimensions of the so-called “coherence zones” within which visual-vestibular conflicts go unnoticed is a pressing question for motion simulation design. It is to be assumed that the coherence zones vary depending on a number of factors, such as the direction and type of motion (i.e., linear or rotatory), frequency and duration of stimulation, and acceleration. The surprising finding from the current study that vection onset is largely reduced by a corroborative kick, even if there is a fairly substantial mismatch between the visual and physical acceleration. This finding indicates that quite large amounts of visual-vestibular mismatch seem to be tolerable for the case of forward linear vection. These results support the notion that in cases of passive self-motion, the vestibular system acts similar to a gating mechanism which helps to disambiguate whether visual motion is due to self-motion or environmental motion. A qualitatively matching vestibular stimulus seems to be sufficient to substantially shorten the delay until vection develops. A more detailed discussion will be provided in Section 6.

The results of the current experiment also show that vection onset latencies for forward linear vection are considerably higher than for circular vection around the yaw axis, and also perceived vection intensity and convincingness were higher for circular vection. This will be discussed in more detail in section 5, together with the results of the following experiment.

4.4 Experiment 6: Combined visual, vestibular, and vibratory effects in forward linear vection

This study investigated whether forward linear vection can be enhanced by minimal vestibular and vibrational cues. In the experiment, linear vection was induced in 8 participants who were seated on a wheelchair and watched a simulated forward motion in a photo-realistic virtual environment (Virtual Tübingen) displayed on a large panoramic projection screen. Participants' task was to report the occurrence and strength of the self-motion illusion by moving a joystick. The convincingness of the self-motion illusion was rated by magnitude estimation. Three levels of vibrations/kicks were applied: No vibration, brief kick at visual motion onset, and a kick followed by continuous vibration throughout the duration of visual motion. The acceleration profile of the visual stimulus was also varied (0.5 or 10 sec to reach a constant velocity of 8m/s). The trials where vibrations were paired with kicks lead to significantly reduced vection onset times and increased convincingness ratings. Trials with a short acceleration period were found to induce vection significantly faster than trials with a long acceleration period. On average, vection was most intensive and convincing in trials with kicks combined with continuous vibration throughout the trial. Results indicate that vibrational and small vestibular cues do enhance vection. Especially, if events like acceleration or deceleration were accompanied with increasing or decreasing vibration intensities, this effect was most pronounced. While it remains unclear whether the vection-facilitating effect of the kicks and vibrations are due to low-level, perceptual effects, or due to a moderating factor of cognitive interpretation, or maybe a combination of both, this result indicates the potential of our lean and elegant self-motion simulation paradigm that relies on presenting consistent multi-modal cues. These results indicate specific visual-vestibular interactions in the perception of self-motion, which has important implications for both the theory of self-motion perception and for the design of lean and elegant motion simulators.

4.4.1 Introduction

In this experiment, we examined the combined effect of vestibular and vibratory cues on visually induced vection. In Experiment 3, we had found that circular vection was significantly enhanced if vibrations were added, and in Experiment 5, forward linear vection was significantly increased when vestibular (and somatosensory) stimulation was provided using minimal motion cueing. In the current study, we investigated whether simultaneous presentation of vibrational and vestibular cues on top of visual stimulation add up to further decrease vection onset latencies and increase intensity and convincingness ratings. Similar to Experiment 5, we varied the level of visual-vestibular conflict by using two different visual accelerations. In contrast to all previous experiments, we used a large, cylindrical, panoramic projection screen to investigate visual-vibrational-vestibular interactions under conditions where visual stimulation is very intense and probably close to ceiling level. If we still find an enhancement of vection by vibrations and vestibular stimulation, this would indicate the high potential of these additional cues to further enhance vection in VR.

4.4.2 Hypotheses

Hypothesis 1 predicts that vection should be enhanced if concurrent vestibular and vibrational cues are provided on top of visual stimulation. Hypothesis 2 states that vection will develop faster in the high visual acceleration condition, as was found in Experiment 5. Hypothesis 3 predicts an

4 MULTI-SENSORY INTERACTIONS IN VECTION

interaction between visual acceleration and vestibular/vibrational stimulation such that there should be a stronger effect in the strong conflict situation. Note that we failed to observe such an interaction in Experiment 5. The current experiment explores whether this interaction occurs under modified experimental conditions.

4.4.3 Methods

Eight observers (3 female, mean age 24.5 years) who were naïve to the purpose of the experiment volunteered to participate in the study in exchange for monetary compensation (approved by the local ethics committee).

STIMULI AND APPARATUS: This study was conducted in the PanoLab at the Max Planck Institute for Biological Cybernetics in Tübingen (see Figure 17). Participants viewed the identical visual stimulus as in Experiment 5, which was a simulated forward motion through a long street in the Virtual Tübingen model. Contrary to Experiment 5, a large panoramic projection screen was used. Participants sat in the exact centre of the cylindrical projection screen at a distance of 3.5 m, and the field of view (FOV) was 220° horizontal \times 50° vertical (see Figure 18). The visual stimulus was created by mapping digital photographs of all individual houses onto the polygons which had been modelled in accurate scale. Three digital projectors (JVC D-ILA SX 21) projected from the ceiling via mirrors onto the entire screen, and both software and physical edge blending techniques were used to ensure that the overlapping zones between two projectors created no artefacts such as increased brightness or unsmooth image motion. Participants were seated in a wheelchair which was put on a podium such that observers' eye height was at about 1.65 m above ground. Vestibular, vibrational and somatosensory stimulation was provided by the experimenter who manually moved and shook the wheelchair. Before the study started, the experimenter was trained extensively to manually generate consistent physical motion stimuli. To ensure that all participants received similar vestibular and vibrational stimuli, motion data from all trials were recorded at a sampling rate of 100 Hz using a small accelerometer that was attached to the participants' head (see Figure 18, right), and trials in which the applied kicks and vibrations exceeded a preset tolerance margin could be identified offline to be later excluded from analysis. Vection responses were recorded using exactly the same method and device as all previous experiments.

PROCEDURE: At the beginning of a trial, participants viewed a static scene. Upon a button press by the participant, the virtual camera started to move forward along the depicted road, showing an expanding optic flow field that simulated linear forward motion at 8m/s. The participants task was to report vection onset, i.e., the moment when they start to feel self-motion, by moving a force-feedback joystick to the direction of their perceived self-movement - thus, in this case, forward. The more self-motion they perceived, the more they were to deflect the joystick. The convincingness of the illusion was rated by magnitude estimation after each trial on a 0-100% scale (0% = not convincing at all, i.e., no perceived self-motion at all, only motion of the visual stimulus; 100% = very convincing self-motion and visual stimulus perceived as earth-stationary). Since none of the participants had ever experienced vection in the laboratory before, they were first presented with the stimulus and asked what they perceived. All participants reported perceiving a sensation of self-motion after some time. Then, participants received 6 practice trials before the actual experiment began. The maximum duration of one trial was 40 seconds. At the end of each trial, the visual stimulus decelerated until stop. There was a pause of 8 seconds between trials to avoid the occurrence of motion aftereffects

4 MULTI-SENSORY INTERACTIONS IN VECTION

and motion sickness. Participants were requested to watch the stimulus naturally, as when, e.g., driving a car, and neither to stare through the screen nor to fixate on any static point on the screen. Furthermore, they were asked to concentrate on the image in the central part of the projection screen and not to move their head during the trials. To stabilise head position, a head-rest attached to the wheelchair was used (see Figure 18, right). Vestibular and vibratory cues were created by the experimenter gently shaking the wheelchair. The visual-only baseline condition was compared to two different kinds of vibrations: A) a small kick at the beginning and the end of visual motion, where the wheelchair was pushed gently forward for about 1 cm to mimic acceleration and braking, and B) the identical kicks followed by continuous vibration throughout the duration of the whole trial. The experimenter was trained carefully to consistently apply diffuse, small kicks within a range of accelerations between 0.05 and 0.03 m/s^2 at about 2 Hz in the kick-trials, and constant vibrations of about 0.02g at about 6 Hz in all vibration-present trials. The kicks were applied by pushing the wheelchair gently forward, both for the acceleration and deceleration phases. The vibrations were applied by shaking the handles attached at the back of the wheelchair forward and backward with both hands at a very fast rate, with small amplitudes of about 1 mm. The kicks and vibrations were matched to the visual acceleration profile: For the strong visual acceleration, a stronger kick was applied, and vibration intensity was constant throughout the trial. For the slow visual acceleration, a gentle kick was combined with gradually increasing vibration intensity. Acceleration data from the accelerometer was recorded and analysed off-line to exclude trials from data analysis in cases where acceleration limits had been exceeded. The left figures in 19 and 20 show representative plots of the physical accelerations of a strong-kick- and a strong-kick-and-constant-vibration-trials, respectively. Time is plotted on the x-axis, and on the y-axis, the acceleration along forward-backward axis is shown in g. The original data points are shown in red, and a low-pass filtered signal at a sampling rate of 10 Hz in blue on top of that. Low-pass filtering was applied to exclude high-frequency noise from the signal. To identify the fundamental frequencies of the kicks and vibrations, we performed complex Digital Fourier Transformations (DFT) on the raw data. The graphs on the right in Figures 19 and 20 show the results.

DESIGN: Two independent factors were varied in a within-subject factorial design: First, the presence or absence of vibrations was manipulated in three levels: One third of the trials were presented without vibrations, one third had a brief kick at the beginning and end of visual motion, and in the remaining cases, the initial kick was followed by continuous vibrations during the whole visual motion and a final braking kick. The second factor varied was the velocity profile of the visual stimulus. Participants initiated trials with a button press, upon which the static visual scene accelerated for 0.5 or 10 seconds to reach a constant velocity of 8m/s. This corresponds to accelerations of 16 m/s^2 or 0.8 m/s^2 , respectively. According to the visual-vestibular conflict theory, vection should develop faster in low-conflict situations, i.e., with the lower acceleration. On the other hand, Dichgans & Brandt (1978) have shown that the ideal visual pattern to induce vection should move at constant velocity. In our study, the constant velocity phase of the visual stimulus is reached faster with the higher acceleration, which should speed up the development of vection if constant velocity is the crucial parameter. All factors were balanced and presented in a within-subject design. The presentation order of acceleration profiles and type of vibrations was randomised. Each condition was repeated four times, yielding four blocks of six conditions each.

4 MULTI-SENSORY INTERACTIONS IN VECTION



Figure 17: Schematic of PanoLab. The cylindrical screen with 3.5m radius shows a $220^\circ \times 50^\circ$ panoramic view using three projectors. Observers have an undistorted view of the projected image from the centre of the cylinder.



Figure 18: Left: Participant sitting on the wheelchair and viewing the street in Virtual Tübingen on the panoramic projection screen. The wheelchair is put on a podest so that the participants eyes were at the height of the screens centre. The physical and simulated FOV were matched to be identical at $220^\circ \times 50^\circ$. Participants sat at a distance of 3 metres away from the screen. The wheelchairs wheels were tightened by expanders to allow only small physical motions of maximally 5cm in each direction. Right: Participant wearing the accelometer on the head, with the headposition stabilised on a headrest. Participants used a joystick to give vection ratings. Walkman-headphones playing sound of flowing water were used to mask external noise.

4 MULTI-SENSORY INTERACTIONS IN VECTION

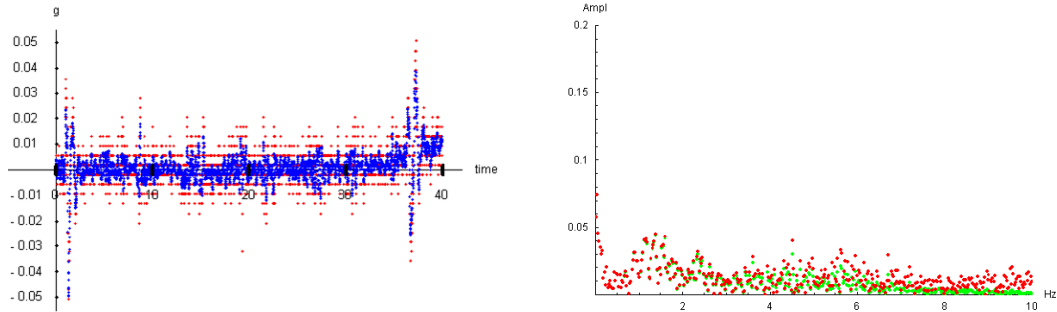


Figure 19: Left: Acceleration data from a representative trial with a strong kick at the beginning and end of visual motion. Red dots indicate original data from the accelerometer attached to the participant's head, sampled at 100 Hz, blue dots show low-pass filtered data at a sampling rate of 10 Hz (which eliminates high-frequency noise). As can be seen, the initial kick happens at 2 seconds, with an amplitude of 0.05g, and the braking kick with an amplitude of 0.04g at about 38 seconds. These kicks correspond to visual motion onset and offset of the trial. Right: Fourier-transformed acceleration data. Red dots show original data submitted to DFT, green dots show low-pass filtered data passed to DFT. As can be seen, the low-pass filtered signal captures well the fundamental frequencies, and the main frequency of the kick is at about 1.5 Hz.

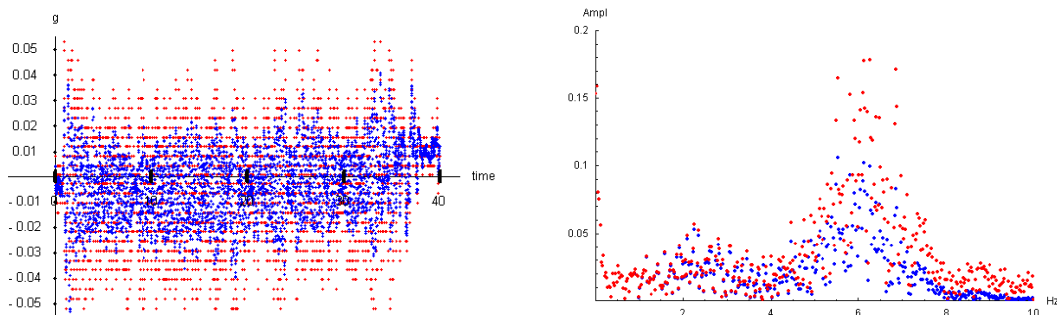


Figure 20: Left: Acceleration data from a representative trial with a kick followed by continuous vibrations. Red dots indicate original data from the accelerometer attached to the participant's head, sampled at 100 Hz, blue dots show low-pass filtered data at a sampling rate of 10 Hz (which eliminates high-frequency noise). As can be seen, the initial kick happens at 2 seconds, with an amplitude of 0.05g (forward), and the braking kick with an amplitude of 0.04g at about 38 seconds. The vibrations are at about 0.02g to 0.03g. These kicks correspond to visual motion onset and offset of the trial. Right: Fourier-transformed acceleration data. Red dots show original data submitted to DFT, blue dots show low-pass filtered data passed to DFT. As can be seen, the low-pass filtered signal captures well the fundamental frequencies, and the main frequency of the vibrations are at about 6 Hz, the kick is at about 2 Hz.

4 MULTI-SENSORY INTERACTIONS IN VECTION

4.4.4 Results

DATA PROCESSING: Pre-analyses showed that vection onset times varied largely between participants: Some participants perceived vection after only 2.5 seconds on average; others took up to 10 seconds on average. For this reason, the data was pre-processed in two steps: First, all first-time presentations of each condition (i.e., the first block of eight trials) were considered as additional practice trials and were excluded from data analysis. In the second step, the between-participants variability was removed by a normalisation process: For each condition per participant, the normalised response time was calculated by dividing each response value per condition of a participant by the ratio between that participant's mean response time and the mean response time across all 8 participants. Using this procedure, the mean onset time per condition, averaged across all participants, remains identical, but the systematic inter-individual differences are removed.

The results of all vection data are plotted in Figure 21. To quantify the effects, separate repeated-measures ANOVAs were calculated for vection onset time, convincingness ratings, and maximum vection intensity. Results show the following main effects:

VECTION ONSET TIME: Vibrations and kicks reduced the vection onset times significantly ($F(2,14)=4.118$, $p < .05$). The stronger visual acceleration of 16m/s^2 (corresponding to the shorter acceleration phase of 0.5s) was found to reduce vection onset times significantly ($F(1,7)=35.255$, $p < .005$), compared to the weaker acceleration of 0.8m/s^2 (longer acceleration phase of 10s). The interaction between visual acceleration and kicks and vibrations was marginally significant ($F(2,14)=3.065$, $p=.079$) for vection onset times.

VECTION INTENSITY: For maximum vection intensity, the main effect of kicks and vibrations was marginally significant ($F(2,14)=3.37$, $p=.064$). Visual accelerations had no significant effect on maximum vection intensity ($F(1,7)=.869$, $p=.382$). There was a significant interaction between the kicks and vibrations and visual acceleration time for maximum vection intensity ($F(2,14)=4.294$, $p < .05$): The slow visual acceleration paired with kicks and vibrations resulted in the highest vection intensity, compared to all other conditions.

CONVINCINGNESS RATINGS: Vibrations and kicks increased the convincingness of the illusion ($F(2,14)=9.306$, $p < .001$). No significant main effects were observed for visual accelerations on convincingness ($F(1,7)=.182$, $p=.682$). The interaction between kicks/vibrations and visual acceleration time was marginally significant for convincingness ratings ($F(2,14)=3.387$, $p=.063$). As can be seen in Figure 21, convincingness was highest for trials with slow visual acceleration with kicks and vibrations.

VECTION BUILDUP TIME: The stronger visual acceleration of 16m/s^2 (corresponding to the shorter acceleration phase of 0.5s) reduced vection buildup time, i.e., the time it took until the maximum perceived vection intensity was reached ($F(1,7)=5.905$, $p < .05$). Here, the presence of vibrations/kicks had no significant effect ($F(2,14)=0.991$, $p=.396$).

4.4.5 Discussion and Conclusion

Consistent with earlier findings in Experiments 3 and 5, we found that vibrations and kicks enhance the vection illusion. Vibrations paired with kicks significantly reduced the vection onset latencies and increased the convincingness of illusory self-motion in VR. This effect was more pronounced in the slow visual acceleration condition. With the large FOV of 220° , it is probable that ceiling level

4 MULTI-SENSORY INTERACTIONS IN VECTION

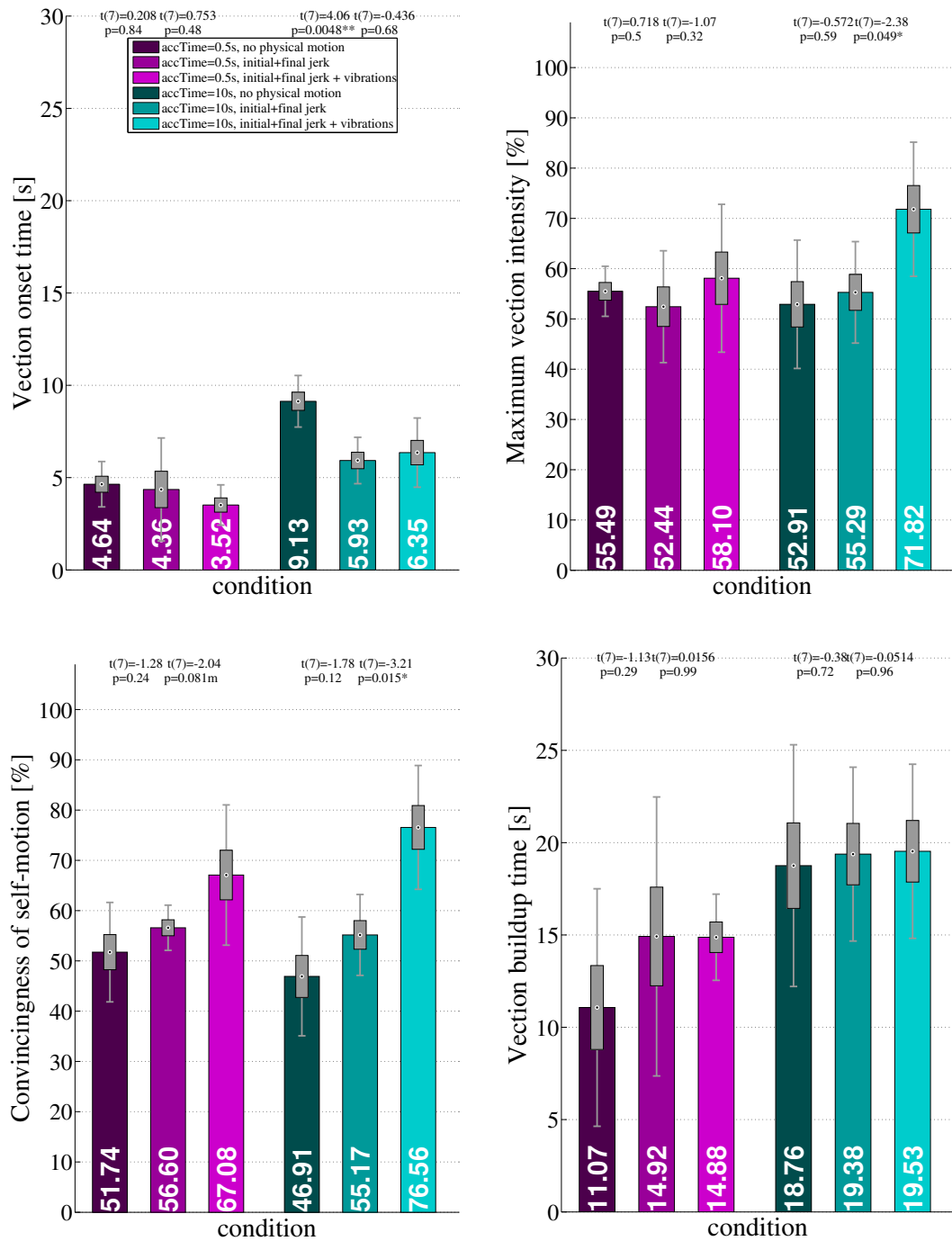


Figure 21: Top left: Mean vection onset times: Shorter acceleration times yielded quicker onset times (left 3 bars), and trials with kicks and/or vibrations show shorter onset latencies than trials without vibrations/kicks. Top right: Maximum vection intensities: The highest values are achieved by a slow visual acceleration with kicks and vibrations. There was no main effect. Bottom left: Mean convincingness ratings: Trials with kicks and vibrations show much higher ratings. No difference between short and long accelerations were found. The highest convincingness ratings were achieved by the slow visual acceleration paired with kicks and vibration. Bottom right: Mean vection buildup time. Boxes show one standard error, whiskers denote one standard deviation. All plots show normalised data. Inlets on top of the figures indicate results from pair-wise comparisons between factor levels.

4 MULTI-SENSORY INTERACTIONS IN VECTION

was reached in the fast visual acceleration condition. Furthermore, it was again found that short acceleration periods with higher visual accelerations induce vection faster than low visual accelerations. One interesting finding was that the highest convincingness and vection intensity ratings were obtained by the slow visual acceleration paired with kicks and vibrations. Results of post-experimental interviews indicate that in this case, participants judged the gradually increasing intensity of the vibrations that reflect the slowly increasing speed of self-motion to be very convincing. Some participants reported that they associated the high-frequency vibrations with the wheels running over the cobblestone of the simulated street in Virtual Tübingen. This indicates that consistent multi-modal stimuli supported the interpretation of the situation “I am moving”, and might have led to stronger vection responses. Even though it remains unclear whether this result was mainly due to low-level, perceptual effects, (i.e., multi-modal cues being integrated to a coherent and natural, realistic percept of self-motion), or whether the high-level, cognitive interpretation biased the responses of the participants to report stronger vection, this finding provides important insights for possibilities to achieve lean and elegant self-motion simulation. We were able to show that by matching the stimuli for two modalities over time with correlated changes in intensity for events like acceleration or deceleration, vection and convincingness can greatly be enhanced. It is also interesting to note that while vection onset times were fastest with high visual accelerations, the most convincing vection was obtained in trials where a slow visual acceleration was coupled with kicks and vibrations. As mentioned above, in these trials, there was more corresponding change over time between the visual and vibrational stimuli, since the vibration intensity increased throughout the 10 seconds visual acceleration time. It seems plausible that the more coherent we are able to stimulate several modalities, the more intensive the self-motion illusion should be.

Another interesting finding is that with the large horizontal FOV of 220° , linear vection was already very strong and compelling when only a visual stimulus was presented. Compared to Experiment 5, where we used a FOV of $75^\circ \times 58^\circ$ that resulted in mean vection onset latencies of 15 seconds and convincingness ratings of 30% in the visual-only condition, the present study found mean onset latencies of 5 to 9 seconds and convincingness ratings of 50% in the comparable condition with a 220° FOV. Interestingly, the benefit of the kicks/vibrations was larger for the small FOV of $75^\circ \times 58^\circ$ in Experiment 5. When kicks were added, both setups resulted in roughly similar mean vection onset times: 3.5 seconds for the 220° screen, and 6 seconds for the 75° screen. This finding has important implications for the design of possible low-cost motion simulators. This means that if one can not afford a large panoramic projection system that requires large space for the screen, several projectors, and complex geometry correction measures, a compact system with a 75° projection that is capable of applying vibrations and small kicks can significantly improve the self-motion illusion.

These findings also highlight the importance of an ecologically plausible motion metaphor in self-motion simulation.

5 Comparative discussion

In this chapter, the most important results of all experiments conducted for the current thesis will be recapitulated and discussed in a comparative manner. Furthermore, the implications and also the limitations of the studies are discussed.

This thesis investigated how visual, auditory, somatosensory, and vestibular information interact during self-motion perception. Multi-sensory VR environments were used to induce the vection illusion and to examine how the different sensory modalities contribute to the perception of self-motion. One guiding question in the current thesis was how perceptual, low-level factors and cognitive, high-level factors interact during vection.

The results from six experiments about circular vection around the yaw axis and forward linear vection showed that vection in VR can be enhanced by consistent multi-modal stimulation if natural, realistic stimuli are used. The addition of auditory, somatosensory, and/or vestibular information that was consistent with visual motion increased vection as well as presence in the VE, compared to the purely visual baseline.

In Experiments 1 and 2, it was found that visually induced circular vection around the yaw axis is enhanced if a natural image of a place was used, compared to various scrambled or sliced versions of the same image which were degraded such that no consistent information about spatial layout was available. For example, in Experiment 1, for the 20°/s rotations, mean vection onset latencies were reduced (11.6 s vs. 24.0 s), vection intensity was increased (70.9% vs. 59.4%), and also the compellingness of vection (55.4% vs. 35.5%) was enhanced with the natural stimulus. In Experiment 2, presence was found to be increased in the natural scene compared to the scrambled and sliced scenes, and significant correlations were observed between the vection data and presence ratings. Importantly, there was a differential effect of spatial and attentional aspects of presence on vection: While the spatial aspect of presence (i.e., the impression that one is situated in the simulated environment) correlated positively with vection intensity and compellingness ratings of vection, the attentional aspect (i.e., how much attention was captured by the VE) correlated negatively with vection onset latency such that the more attention was devoted to the VE, the more vection onset latencies were reduced. To the author's best knowledge, this is the first time that correlations between presence and vection were reported. Earlier studies that failed to observe such correlations used very coarse measurement methods for both vection and presence (Freeman et al., 2000; IJsselstein et al., 2001).

Experiments 3 to 6 investigated multi-sensory interactions in vection. In Experiment 3, it was found that circular vection was enhanced by adding tactile vibrations that were applied to the observer's seat. Vibrations reduced vection onset latencies for 10-15% on average, irrespective of the strength of visual-vestibular conflict, which was varied using either a smooth or an abrupt initial acceleration of the visual stimulus. It is possible that somatosensory cues (here: tactile vibrations) are capable of partly substituting the missing vestibular signals to reduce vection onset latency, even if the vibrations are below vestibular detection thresholds. Importantly, the cognitive manipulation of telling the participants that in one block, the platform would move, but not in the other, did not show any effect on the vection responses, even though more than half of all participants believed that the platform moved in some trials. Thus, the perceptual manipulation showed robust effects, whereas the cognitive manipulation did not have any biasing effects on the vection responses. There are two

5 COMPARATIVE DISCUSSION

other studies that used similar cognitive manipulations. One study by Lepecq et al. (1995) found that only children, but not adults were biased to report vection faster when seated on a movable office chair, compared to when they were seated on a solidly mounted seat. On the other hand, Wright et al. (2006) found that compellingness ratings of vection were increased when the simulated VR environment exactly matched the testing environment from the observers' perspective, compared to a situation where they were tested in a different lab. However, vection onset latencies were unaffected by this manipulation. Wright et al. (2006) proposed that dissociable mechanisms, one perceptual, one cognitive, seem to affect different aspects of vection. The perceptual factors (i.e., conditions of sensory stimulation) seem to affect vection onset latency, while cognitive factors (such as presence) seem to mainly affect judgements about compellingness of vection.

In Experiment 4, evidence for visual-auditory interactions in vection were found. The compellingness of circular vection was increased by moving sounds that moved concordantly with the visual stimulus - in our case, the sound of flowing water was associated to the image of a fountain. However, this effect was rather small, and vection onset times and vection intensities were unaffected by the auditory manipulation. This finding is consistent with earlier studies that showed that auditory vection is less compelling than visually induced vection (Lackner, 1977; Larsson et al., 2004). It has been shown, though, that under conditions of near-threshold vestibular stimulation, moving sounds that are compatible with physical motion can increase the number of self-motion reports (Schinauer et al., 1993). On the other hand, spatialised sound did increase presence ratings by a small, but significant amount. It appears that for self-motion perception, auditory information has less weight than visual information, even though small effects for cognitive ratings about vection compellingness and presence were found. This finding is somewhat reminiscent of the findings by Wright et al. (2006). It appears that additional dynamic auditory information can serve to enhance presence and convincingness of vection in VR, but is not very instrumental to reduce vection onset latencies if reliable visual information is available.

Experiment 5 showed that forward linear vection is substantially increased by brief initial vestibular stimulation during visual motion onset. The platform kicks reduced mean vection onset latencies from 14.3 s to 6.1 s and increased mean vection compellingness from 29% to 49%. This finding is consistent with results from Wong & Frost (1981) who found a similar effect in circular vection. Furthermore, the results indicated that exact quantitative matching of visual and vestibular acceleration magnitude was not necessary and that a qualitatively matching vestibular kick applied simultaneously to the visual motion onset was sufficient to enhance vection. Even if physical and visual acceleration magnitudes were in conflict, vection onset latencies were nevertheless reduced for more than 50%.

In Experiment 6, it was found that combined visual, vibrational and vestibular stimulation produced the fastest vection onset times (mean = 3.4 s) and highest compellingness ratings (mean = 76%) for forward linear vection. In this experiment, a large, panoramic screen with 220° FOV was used, and visually induced vection can be assumed to be close to ceiling level under such conditions of almost full-field stimulation. Still, a significant enhancement of vection was observed when vibrations and kicks were added to visual motion, which was more pronounced for slow visual accelerations that took 10 s to reach constant velocity. In this condition, the intensity of vibrations was gradually increased in order to reflect increasing speed. It is likely that this matching of multi-modal

stimuli with correlated changes in intensity over time increased vection and compellingness, since the stimulus properties reflected natural patterns that occur during acceleration or deceleration.

5.1 Factors affecting vection onset latencies and compellingness of vection

In the current thesis, circular vection around the yaw axis and forward linear vection were investigated. Forward linear translations and yaw-rotations are the most frequent types of motion for terrestrial locomotion and navigation and are thus most relevant for motion simulations.

In this thesis, three measures of vection were employed: Vection onset latency, vection intensity, and compellingness ratings. Additionally, presence that observers felt during exposure to the VEs was assessed in Experiments 2 and 3 using a standardised presence questionnaire by Schubert et al. (2001). One interesting general finding was that the measures of vection were differentially affected by different factors in some cases. For example, in Experiment 4, spatialised moving sounds were found to increase the compellingness of vection and to increase presence, but onset latencies were unaffected. On the other hand, visual, vibrational, and vestibular stimulation strongly influenced vection onset times and also vection compellingness. As can be seen in Table 2, the strongest effect for the reduction of vection onset latency and increase of compellingness was found in Experiment 5 for linear vection when vestibular stimulation in form of an initial kick was added. Under the tested conditions, the kick reduced onset latencies by 55% and increased compellingness by 45%. A similar result was reported by Wong & Frost (1981) for circular vection. In Experiment 6, where the FOV was about five times larger than Experiment 5, the kicks reduced vection onset by 30% and increased compellingness by 36%. Thus, with a smaller FOV, the benefit of initial physical accelerations was more pronounced.

Earlier informal reports and also our own observations so far indicated that forward linear vection is less compelling and takes longer to develop than circular vection. Linear vection for up- and down motion was found to develop faster and to be more compelling than forward-backward linear vection (Kano, 1991; Giannopulu & Lepecq, 1998). It has been argued that this advantage for upward- downward vection is due to the fact that visual-vestibular conflict is less pronounced in this situation. To the best knowledge of the author, there has so far not been a comparative study that systematically investigated differences between linear and circular vection. In the current study, both linear and circular vection were investigated, but due to confounding factors such as FOV, speed of visual motion and acceleration times between the experiments, comparisons between the experiments are difficult. Despite these limitations, the summarised results shown in Table 2 allow for some tentative comparisons: For all circular vection experiments, compellingness ratings are clearly higher than for linear vection. For example, in Experiments 2 to 4 that used a smaller FOV of $54^\circ \times 40^\circ$, compellingness ratings for circular vection were about 20% higher than in Experiment 5, which used a larger FOV of $75^\circ \times 55^\circ$ for forward linear vection. Given that it is generally accepted that the larger the FOV, the stronger and more compelling vection should be (see, e.g., Dichgans & Brandt, 1987), this result indicates that circular vection might indeed be more compelling than linear vection. The fact that in Exp. 6, which had the largest FOV ($220^\circ \times 50^\circ$), linear vection compellingness reached only 46%, which is clearly below all values reached for circular vection, is a clear indication for this. In the other condition where visual acceleration time was only 0.5

5 COMPARATIVE DISCUSSION

seconds, compellingness ratings were higher at 55% (denoted as Exp. 6b in Table 2). It is likely that this is due to the short visual acceleration time, as will be discussed further below.

Comparisons for onset latencies are more problematic here, since not only the FOV, but also the acceleration times differ between the linear and circular vection experiments. All experiments in which visual acceleration times were manipulated (i.e., Exp. 3, 5 & 6) consistently showed a very strong effect such that the shorter the acceleration time (i.e., the earlier constant velocity is reached), the faster vection onset occurred. Thus, the somewhat longer onset latencies for linear vection in Exp. 5 as compared to Exp. 3 and Exp. 4 should be interpreted with caution. The relatively long onset latency of 9.1 seconds with an acceleration time of 10 seconds in Exp. 6 is quite intriguing, since with a large FOV of $220^\circ \times 50^\circ$, one would expect that vection should be as strong as it could possibly be. In this context, two findings from Exp. 4 and Exp. 6 provide strong evidence that the acceleration time and profile of the visual motion seem to have a stronger influence on vection onset latency than FOV, both for linear and circular vection: First, with a comparably small FOV of $54^\circ \times 40^\circ$ and an acceleration time of 3 seconds, mean vection onset in Exp. 4 was 10 seconds for circular vection, which did not differ much from Exp. 6 where a mean onset latency of 9.1 seconds was found with acceleration times of 10 seconds and a FOV of $220^\circ \times 50^\circ$. Furthermore, in the comparable condition in Exp. 6 where acceleration time was only 0.5 seconds, a short onset latency of only 4.6 seconds was found (see Figure 21). This finding is also important from an applied perspective for motion simulation, since fast vection onset is a desirable feature in most applied cases.

In order to answer the question how linear and circular vection differ perceptually and where the differences stem from, further research is needed. To devise a “fair” psychophysical test between linear and circular vection, one must control for all the parameters mentioned above. One problematic issue is how to equate the visual speeds of linear and rotational motion. The optic flow fields of linear forward translations and circular yaw rotations differ substantially from one another in that the rotations yield uniform, *lamellar* flow where all motion vectors have the same length and direction, whereas the forward translations yield a *radially expanding* pattern where visual motion increases from zero at the FOE towards the periphery. It is thus conceivable that rotational visual motion has more vection inducing potency due to the increased amount of motion energy, compared to linear forward motion. Even if one carefully adjusted the overall motion energy across the whole FOV for lamellar and radial flow, it is unclear how the uneven distribution of low velocities in the centre and the higher velocities in the periphery in the radially expanding pattern would be weighted by the visual system. For example, it has been shown that human speed judgements about forward translation depend on angular declination of the fixation point relative to the optical flowfield (Pretto, Vidal, & Chatziastros, 2007). From this result, one might predict that vection onset in linear vection should depend on where the fixation point is placed in the flow field.

The question of whether differential vestibular sensitivities of semicircular canals and otoliths play a role should also be addressed in this context.

In summary, the results of the experiments conducted in this thesis show a clear and consistent effect that vection can be increased by ecologically valid multi-sensory stimulation and by using natural, realistic visual stimuli. The question how much of these effects can be ascribed to perceptual and cognitive influences will be discussed in the next section.

Table 2: Selective vection results from Experiments 1-6. Vel. denotes velocity during constant velocity motion of the visual stimulus, T_{acc} denotes acceleration times of the visual stimulus until constant velocity was reached. Mean vection onset latencies (vis.) and compellingness ratings (vis.) from the purely visual baseline are shown, and the columns Reduction OL (onset latency) and Gain CO (compellingness) show the mean benefit in absolut values and in % relative to the visual baseline for each experimental manipulation.

	FOV	Type	Vel.	T_{acc}	Onset latency vis.	compellingness vis.	Reduction OL	Gain CO	manipulation
Exp. 1	$84^\circ \times 63^\circ$	circular	$40^\circ/s$	3s	8.3s	69%	15.3s (-46%)	+ 35%	natural vs. artificial image
Exp. 2	$54^\circ \times 40^\circ$	circular	$40^\circ/s$	3s	16.2s	58%	28.5s (-43%)	+ 23%	natural vs. artificial image
Exp. 3	$54^\circ \times 40^\circ$	circular	$30^\circ/s$	3s	9.4s	57%	8.2s (-14%)	69 (+21%)	vibrations
Exp. 4	$54^\circ \times 40^\circ$	circular	$30^\circ/s$	3s	10.1s	55%	9.6s (-6%)	70 (+15%)	moving sounds
Exp. 5	$75^\circ \times 55^\circ$	linear	6m/s	5s	14.5s	34%	6.6s (-55%)	50 (+45%)	platform kick
Exp. 6	$220^\circ \times 50^\circ$	linear	8m/s	10s	9.1s	46%	6.4s (-30%)	77 (+36%)	kick + vibrations
Exp. 6b	$220^\circ \times 50^\circ$	linear	8m/s	0.5s	4.6s	52%	3.5s (-24%)	67 (+29%)	kick + vibrations

5 COMPARATIVE DISCUSSION

5.1.1 Perceptual and cognitive influences on vection

In the current thesis, two classes of manipulations of the vection stimuli were applied. One class was concerned only with the visual parameters (e.g., naturalness, acceleration time, velocity), and the other with combined multisensory stimulation. A third type of manipulation was a cognitive manipulation applied in Exp. 3 in order to investigate possible cognitive bias effects on vection.

The general effect that vection is enhanced by consistent multisensory stimulation is likely to be mainly of perceptual origin. The manipulation that revealed the strongest effect on reducing vection onset latencies and increasing compellingness was adding vestibular kicks to the visual stimulation (Exp. 5). This finding was to be expected from what is known from visual-vestibular interactions, even though it was somewhat surprising that exact quantitative matching of visual and vestibular acceleration magnitude was not necessary in order to achieve the effect. Rather, a qualitatively matching vestibular kick that was delivered together with visual motion onset was sufficient for the current purpose of reducing vection onset latency and increasing compellingness. Another interesting observation is that this effect was larger for the medium size FOV in Exp. 5 than in the large FOV in Exp. 6. This finding is especially relevant for motion simulation applications, since it indicates that setups with small- to medium sized FOVs benefit most from minimalistic motion cueing, which can be delivered by low-cost setups, such as commercially available motion seats that are used in the gaming and entertainment industry. It is noteworthy that even mere vibrations applied to the seat in Exp. 3 were effective, even though the effect was considerably smaller than vestibular stimulation. The combined visual, vestibular and vibrational condition in Exp. 6 indicates that even under conditions of large FOV visual stimulation where visual vection can be assumed to be at ceiling level, there was still a significant enhancement of vection by adding kicks and vibrations (see Figure 21 on page 79). The finding that adding spatialised moving sounds only had a small but significant effect on vection compellingness, but not on vection onset latency indicate that for self-motion perception, especially for vection, vision is far more dominant than the acoustic channel, such that in the combined visual-auditory case, only a very small gain is added by the dynamic acoustic stimulation.

The finding from Exp. 1 and 2 that using natural images instead of artificial, scrambled images increased vection substantially is more difficult to interpret. One way of explaining the results would be to say that they are purely perceptual, since the scrambling and slicing applied to the scene changed low-level image statistics. This is only partly true, since on a pixel scale, the images contained the same pixels, only the spatial distribution differed. The artifact of increased number of high-contrast edges of the scrambled images should have increased vection, not reduce it. One could also argue that the rich visual information in the natural scene about spatial layout which was specified by many optical variables, such as texture gradients, perspective, occlusion, eye-height information etc. were not available and thus visual information was less reliable. An artificial visual stimulus is ecologically speaking informationally underspecified, which may be one reason why vection is reduced. On the other hand, assuming that presence in VR is a cognitive phenomenon, one could also argue that the effect is of cognitive origin, since the natural scene yielded also higher presence ratings, and vection and presence data were significantly correlated. Of course, correlations should not be interpreted in a causal way. Due to the experimental design, it cannot be concluded whether presence influenced vection responses or vice versa, and if so, whether in a direct or indirect manner. However, one likely possibility is that one common underlying mechanism influenced

5 COMPARATIVE DISCUSSION

both presence andvection, which is related to the feeling of being at a place, which enables spatial updating. The natural scene depicted a place extending in three dimensional space, and the fairly large FOV enabled observers to feel present and being surrounded by the simulated scene. As it was already described in 3.2.5 on page 48, it has been shown in an earlier study which used exactly the same setup and the identical natural, realistic visual stimulus that spatial updating is possible under such conditions (Riecke et al., 2005). It is possible thatvection responses depend on whether observers perceive themselves being at a place or in the simulated scene, or whether they perceive an artificial, underspecified stimulus which does not occupy 3-D space. This assertion is admittedly speculative at this stage and should be investigated further in future research.

It thus appears reasonable to assume that thevection enhancing effect of natural stimuli might have both perceptual and cognitive roots. Another rather clear indication for cognitive influences onvection is related to the finding that the attentional aspect of presence as measured using the IPQ Presence Questionnaire correlated significantly withvection onset latency in Experiment 2. The more attention was captured by the simulated VR environment, the fastervection onset became. In an earlier study, Kitazaki & Sato (2003) had also found thatvection was modulated by attention. In that study, however, presence was not assessed. Future research should test in more detail how different attentional aspects, such as top-down, focused attention or bottom-up driven shifts of attention (e.g., caused by an unexpected event in the scene) relate to both presence andvection.

The finding that the cognitive manipulation in Exp. 3 did not show any biasing effects on thevection responses was unexpected. It is possible that the manipulation was not strong enough to show an effect. Even if they were told and also shown that the motion platform was turned off, observers still might have thought that the platform could move in principle. A clearer design with two different setups, one of which would apparently never be capable of moving, would be necessary to clarify this issue. In thevection literature that investigated cognitive biasing effects, there is one null-finding by Lepecq et al. (1995) who found that adults were unaffected by a comparable manipulation (only 7-year-old children showed a bias effect of reportingvection onset faster), whereas Wright et al. (2006) found a positive effect such thatvection compellingness was increased, butvection onset was unaffected. Further investigations are needed to obtain a more conclusive answer to this question.

6 Theoretical frameworks of perception

In this section, I will sketch two major theories of perception, and discuss how these two different classes of theories relate to the vection phenomenon and self-motion perception. Furthermore, I will argue that despite vastly different theoretical points of views, the theories are not mutually exclusive and both afford important contributions for our understanding of self-motion perception. An integrative view of self-motion perception that encompasses aspects of both theories will be outlined.

6.1 Indirect and direct theories of perception

There are different classes of theories about the nature of perception. Two of the major theories are referred to as the *indirect* and the *direct* theory of perception. The classical view is that perception is indirect in the sense that we cannot access distal objects out in the world directly, but only through our senses. From the proximal sense data, a representation of the distal object needs to be generated in the brain in order to have any experience of it. The central idea is that sensory information is highly ambiguous, and some a priori knowledge about the world is necessary to resolve the ambiguity. Thus, some kind of (unconscious) inference (or “taking into account”) is necessary in order to make sense out of sensory data (von Helmholtz, 1910). One of the main problems of perception dealt with in this class of theory is how the three dimensional outside world is reconstructed from the projected 2-D image on the retina. Different models have been proposed to explain how this computationally demanding task may be accomplished by the brain. Among the most prominent of those is the multi-stage model of visual perception by Marr (1982). Marr proposed three stages of visual perception, each of which employs certain algorithms that extract image features from the successively reconstructed scene, starting with the retinotopic projection in primary visual cortex (V1). A similar view is taken by Gregory (1985), who states that perception can be thought of as some kind of multi-stage problem solving or hypothesis testing.

One biological mechanism that is often quoted in this class of theories is the principle of *reafference*, which states that the comparison of two neural signals is a basic principle for perception, especially for motion perception (von Holst & Mittelstaedt, 1950). For example, the visual system can dissociate between object motion, self-motion, and a combination of both by subtracting the extraretinal signals about the motor commands about eye movements (efference copy) from the retinal image motion. If the eyes move (i.e., rotate in the head) in a stationary environment, for example, retinal motion and extraretinal signals will indicate a zero difference and a stable environment will be perceived.

Contrary to this view, the direct perception theory (also known as the ecological theory of perception), which was put forward by James Gibson, states that perceivers have direct access to environmental information and that no mental calculations or a reconstructed representation of the world is necessary (Gibson, 1979). According to this view, the visual world is not chaotic and ambiguous, but highly structured and rich in information. All information necessary is contained in the visual world in form of certain invariant features, and observers can pick up these invariants to guide behaviour. Gibson stressed the importance of recognising that perception is an active process, that motion of the observer creates useful information instead of complicating the matter, and that perception and action are not separable from one another and should be studied together. Two central

concepts in the direct perception theory are *invariants* and *affordances*. Invariants are higher-order patterns of structured light (or energy). For example, when we move towards a surface, a typical radially expanding pattern of global motion will result. The lawful transformation patterns in the optic array were termed *optic flow field*, and they contain important action-relevant information: The optic flow field specifies the direction and relative velocity of motion, and the central focus of expansion (FOE) in the flow field specifies the exact heading direction (Gibson, 1950). The second central concept in direct perception is the affordance. Affordances stress the importance of the relation between the perceiver and environment. Affordances contain action-relevant environmental information for the perceiver and are specified by invariants. We can thus directly perceive what actions are possible (afforded) in an environment. For example, whether an object affords grasping, throwing, whether a surface affords walking, standing, or sitting on, or whether we can pass through an opening in a wall, etc. Animals with different action capabilities and dimensions will perceive different affordances in the same environment.

6.2 Vection viewed through indirect and direct perspectives on perception

It is interesting to note that despite of the vastly differing theoretical positions, the direct and indirect theories of perception take quite similar approaches to explain the vection phenomenon. The indirect view is based on the inherent assumption of a stable world that humans appear to have (see, e.g., Dichgans & Brandt, 1987). There is a number of indirect models of self-motion perception which model visual-vestibular integration in terms of cue-conflict estimation. These models explain the gradual development of vection using the visual-vestibular conflict theory. For example, Wertheim (1994) proposed a functional model in which he modified the reafference principle by introducing a so-called *reference signal* which combines retinal, extraretinal, and vestibular signals. By comparing this reference signal with the retinal signal, object-motion, self-motion, and stationarity in the environment can be determined. When an observer moves in space, visual and vestibular components of the reference signal will equal the retinal motion, and self-motion is perceived. For vection, the case is different, because the vestibular component of the reference signal is missing, since the vestibular senses are not stimulated. As a consequence, object-motion will “correctly” be perceived first. The gradual buildup of vection is ascribed to the gradual change of vestibular and visual components in the reference signal: Soon after constant velocity (of the visual stimulus) is reached, the vestibular “veto” in the reference signal decreases which initially prevented perception of self-motion and the visual component increases. At this time, vection will be perceived. More models of this kind will be discussed in more detail in the following section.

The direct view, on the other hand, states that vection is a case of informational underspecification. Now, as an important side note, according to the Gibsonian view of perception, illusions are not adequate for perceptual research since they result from unnatural, impoverished situations, which only seldom occur in nature, but frequently under informationally impoverished lab conditions. For the case of vection, vestibular and somatosensory information that usually co-vary with the optic flow during real self-motion are absent at the onset of visual motion, and hence the situation is informationally underspecified. The question how internal, physiological processes (such as visual-vestibular interaction) are related to perception are usually not in the focus of investigation in

6 THEORETICAL FRAMEWORKS OF PERCEPTION

the Gibsonian community. Even though Gibson stressed that information is multiple and picked up concurrently by different senses (Gibson, 1966, 1979), Gibsonian-inspired research on this topic is rare. In general, multi-sensory research has started to receive research attention only from the early 1990s, and most of these studies were inspired by the Bayesian models of the brain.

One model that contains elements of both indirect and direct perception was proposed by Cutting (1986). Cutting stated that information is often multiply specified within one or between different sensory modalities. For example, linear movement above a surface can be specified by different optical variables, such as edge rate (which depends on ground speed and texture density) or global flow rate (which depends on altitude above ground) (Larish & Flach, 1990). Of course, other sensory modalities also can pick up self-motion information, such as vestibular, somatosensory (vibrations, wind on the skin, etc.) or auditory information (acoustic flow, Doppler-effect, pitch and loudness of engines or wheels, etc.). The model proposed by Cutting suggests that multiple information about one object or event may be either *integrated* in a weighted (additive) manner, *selected* (i.e., using one source and ignoring another) or *integrated but also selected* (two integrated and a third ignored) (Cutting, 1991). According to this view called *directed perception*, there is a many-to-one mapping between information and environmental entities, not a one-to-one mapping, which means that object properties are multiply specified. While this approach extends the direct perception view to multiple sensory modalities and also makes a step towards indirect perception theory, it remains rather unspecific about how integration or selection is done and what mechanisms or laws govern the integration and selection processes, which is, as already stated above, not in the focus of ecologically motivated perception research.

It should be noted here that in order to understand the phenomenon of vection and self-motion perception, both direct and indirect theories of perception have explanatory value. The main difference between the two classes of theories is in the focus of analysis. The indirect view is more concerned with understanding the neural and computational processes inside the perceiver, whereas direct theory considers the informational coupling between a perceiver and the environment, and how action-relevant information is specified by invariants and affordances. This issue will be discussed in more detail in Section 6.5.

Given that self-motion is usually perceived through multiple sensory modalities, an adequate model of self-motion perception should also consider more than just the visual modality. While traditional psychophysics has been investigating the functionality of different sensory modalities in separation from one another, more and more studies are now devoted to the question how the brain integrates multi-modal information for perception and action, for example, in object recognition, face recognition, object localisation, and also self-motion perception. Several models have been proposed to explain how the brain integrates sensory information from multiple modalities, most of which are related to the indirect theory of perception. The following sections will review some models that are most relevant for the domain of self-motion perception.

6.3 Visual-vestibular circuit models of self-motion perception

In general, visual and vestibular signals are thought to be the most relevant for self-motion perception (Dichgans & Brandt, 1978; Howard, 1982). Neurophysiological studies have found evidence that visual and vestibular cues are jointly processed, and the vestibular nuclei were identified as neural

6 THEORETICAL FRAMEWORKS OF PERCEPTION

structures that respond in a similar manner to both visual and vestibular stimulation. Based on findings from both neurophysiological and psychophysical experiments, several models about how visual and vestibular input signals are combined in the brain have been proposed. Many of these models have in common that they understand visual-vestibular integration as a dynamic control system, with visual and vestibular sensory signals as input variables and perception of self-motion in space, and in some models also reflexive eye-movements (e.g., VOR, OKN) as output variables (Zacharias & Young, 1981; Mergner, Schweigart, Kolev, Hlavacka, & Becker, 1995; van der Steen, 1998; Telban & Cardullo, 2005; Zupan, Merfeld, & Darlot, 2002). Most of those models contain elements that model the dynamics of the visual and vestibular systems and some transformations that generate predictions based on one sensory modality for a signal in the other sensory modality (e.g., “if visual acceleration is V , the vestibular signal should be X ”). This is accomplished by so-called *internal models* that include “knowledge” about sensor dynamics, body dynamics, and physical relationships. In a third stage, the input signals of the two modalities are compared by addition/subtraction, and if a conflict is registered, perception is biased towards vestibular inputs, while otherwise, visual inputs tend to dominate, depending on stimulation frequency and amplitude.

In a self-motion model proposed by van der Steen (1998), the concept of neural filters was introduced, which transduce the sensory input signal into a perceptual variable. This conceptual model does not explicitly model how visual-vestibular cue conflict is estimated, but rather states that the visual information “attracts” the vestibular estimate of self-motion, and this attraction leads to a specific magnitude of the optokinetic signal. According to the authors, this neural filter can be represented as a first-order low-pass filter, and computer simulations of the model were able to reproduce the gradual build-up of vection as observed in psychophysical vection experiments. The model parameters such as time constants for optokinetic and vestibular filters were estimated based on psychophysical vection experiments.

Further models have been proposed by Mergner et al. (1995) and Telban & Cardullo (2001). They are also based both on psychophysical data and model how visual-vestibular conflict is estimated: Based on the visual motion signal, an *expected* vestibular signal is calculated and then compared to the *real* vestibular signal. If the conflict estimator “detects” a conflict, the percept is thought to be biased towards the vestibular signal. The dynamics of the vestibular system are also included, and thus visual motion information dominates after a time period of constant velocity (see also Section 2). Telban & Cardullo (2001) propose two different models, one for linear and one for circular vection. The only difference between the two is that for linear vection, the otolith dynamics are modelled, and for circular vection, the semicircular canals dynamics. By estimating the time constants and adaptation operators according to experimental data on vection onset latencies, the authors were able to replicate typical results for vection onset latencies found in the literature for linear and circular vection.

One of the most sophisticated and comprehensive models of self-motion perception, eye movements, and spatial orientation was published by Zupan et al. (2002). In this model, which the authors call “sensory weighting model”, the sensory integration process is modelled in three stages: First, the sensory signals from each modalities are augmented such that they cover the whole frequency range by using information from other sensors, a process called “frequency completion”. For example, the signals from the vestibular system with their high-pass characteristics would be augmented by the low-pass content which can be estimated based on visual motion information. In a second

6 THEORETICAL FRAMEWORKS OF PERCEPTION

step, before sensory information gets combined, sensory estimates with different physical meanings are transformed into an “intermediate estimate” which has a common, head-centred reference frame. For both the first and second steps, internal models about sensor dynamics as well as models for physical relationships are necessary. Finally, in a third step, a combined estimate is calculated according to maximum likelihood estimation (MLE, see next subsection). A number of phenomena, such as vection onset latencies, reflexive eye movements, and spatial orientation in space can be simulated quite successfully in this model with proper adjustment of the free parameters.

Another model was proposed by Reymond, Droulez, & Kemeny (2002) who considered perception of self-motion based on integrated sensory information to be a dynamic optimisation process. In the model, self-motion variables for angular and linear velocity, linear acceleration and gravity acceleration are estimated based on sensory inputs. Self-motion perception results from a compromise between sensory information and so-called coherence constraints, which represent internal models of physical laws of motion. For example, if visual and canal signals both indicate rotation, the mutually predicted and measured signals match and the coherence constraint is met. In this case, the signals are fused using a weighted average. In a dynamic optimisation process, a number of cost functions are minimised by using a weighted sum of all corresponding cost functions for all constraints.

Many of these models are qualitatively very similar and can explain some experimental findings from vection research, such as the gradual buildup and timecourse of vection. However, as Mergner et al. (1995) note, self-motion perception does not seem to solely depend on the stimulus presented in psychophysical experiments, but also on many other factors that can hardly be controlled, such as attention, expectations, methods, apparatus and task, and eye movements, to name just a few. Accordingly, such models can only explain a limited aspect of self-motion perception, and can not sufficiently account for the variability in the individual data that is observed between and even within studies. There is only one model by Jürgens & Becker (2006) which explicitly includes higher-level, cognitive influences to the kinds of low-level models described above. This model will be described in more detail section 6.4.1 on page 95, after MLE is briefly reviewed.

6.4 Maximum likelihood estimation (MLE) and *Bayesian* sensory integration

Recently, the view that the sensory integration process can be best understood within a Bayesian approach has become increasingly popular in the literature (Knill & Richards, 1996; Ernst & Bühlhoff, 2004). This probabilistic approach stems from the idea that the brain always needs to operate in a state of sensory uncertainty, since all of our sensory channels are non-ideal in the sense that they are all associated with sensory noise. Thus, the brain needs to somehow estimate the *real* value of one physical variable (e.g., the size and/or distance of an object). Since many physical properties of the surrounding world and of one’s own body can be sensed through several sensory modalities, it is likely that biological systems have developed some successful method of estimating the real physical values from combined sensory stimulation. Otherwise, we would not be able to successfully interact with an environment and act in a goal-directed manner. One obvious method to estimate the values would be to calculate the numerical average of all available signals concerning one variable. In this case, however, the accuracy of the respective sensory channel would be disregarded. The

6 THEORETICAL FRAMEWORKS OF PERCEPTION

statistically optimal way to obtain a reliable and robust estimate of the “real” value of the physical variable in question is to weight each input signal according to its sensory accuracy or reliability in the averaging process. This is achieved by the MLE method. In MLE, a weighted average of the available signals is calculated, where the weight for each modality is determined by their respective accuracy. In more precise terms, the sensory weight is inversely proportional to the signal variance σ^2 . The more noisy a signal is, the less reliable it is and less weight is given to the signal. Expressed in a formula, the optimal combined estimate $s_{1,2}$ of a physical variable measured by two sensory signals s_1 and s_2 can be calculated by:

$$s_{1,2} = s_1 * \omega_1 + s_2 * \omega_2$$

where ω_1 and ω_2 are the weights of the cues specified by their respective variances:

$$\omega_1 = \frac{\sigma_1^2}{\sigma_1^2 + \sigma_2^2}$$

$$\omega_2 = \frac{\sigma_2^2}{\sigma_1^2 + \sigma_2^2}$$

It is assumed that $\omega_1 + \omega_2 = 1$.

The variance of the combined estimated signal $s_{1,2}$ is given by:

$$\sigma_{1,2}^2 = \frac{\sigma_1^2 * \sigma_2^2}{\sigma_1^2 + \sigma_2^2}$$

One necessary condition is that the noise of the different signal channels are “known” to the system and that they are statistically independent from one another and have a normal distribution. The Bayesian theory assumes that this is valid, since the sensory modalities are separate from one another, receive their input signals from different sensory organs and are processed in separate neural structures, at least in early processing stages.

Due to its statistical optimality and rigorous mathematical formulation, the MLE approach has drawn substantial research attention recently (Ernst & Bühlhoff, 2004), and a number of studies found evidence that MLE principle seems to be at work in human multi-sensory integration. For example, Ernst & Banks (2002) studied the integration of visual and haptic information on the perception of object size. They reduced the reliability of visual information by blurring the visual stimulus, and found that the weight of the haptic signal was increased the more the visual signal was deteriorated, fulfilling quantitative predictions by the $1/\sigma^2$ law. Similarly, Knill & Saunders (2003) used a visual slant estimation task and found evidence that stereo and texture information are weighted according to the respective sensory reliabilities as predicted by MLE. In another study, Alais & Burr (2004) investigated the ventriloquist effect and concluded that the ventriloquist effect results from near-optimal audio-visual integration for localisation in the sense of MLE. As noted earlier, though, Rosas et al. (2005) showed that MLE can only be observed in a certain parameter space that is tested.

While the MLE sensor fusion algorithm is only concerned with bottom-up sensory information, the sensory integration process is called *Bayesian* if not only bottom-up sensory signals, but also prior assumptions or prior knowledge are taken into account. For example, we inherently *know* that the environment around us does not move by its own, or that light usually comes from above, that gravity pulls us to the ground, etc. This type of cognitive, top-down information might result from

6 THEORETICAL FRAMEWORKS OF PERCEPTION

innate, “built-in” values, or from daily experience, or from some cognitive expectation. Bayesian integration assumes that prior knowledge and bottom-up sensory information is integrated according to the Bayes’ theorem (Yuille & Bülthoff, 1996):

$$P(A|S) = \frac{P(S|A) * P(A)}{P(S)}$$

In this formula, S denotes the measured sensory signal and A is the underlying ground truth which is to be estimated, (e.g., “I am moving” vs. “An object is moving”). The *posterior probability* $P(A|S)$ (i.e., the probability that A is true given that S is observed) is calculated by multiplying the *likelihood function* $P(S|A)$ (the probability to obtain signal S in situation A) with the *prior distribution* $P(A)$, which specifies the relative probability that A occurs in the world. The denominator $P(S)$ is a normalisation factor which is necessary to make the posterior a true probability distribution.

Put into simple words, the a priori information (*Bayesian prior*) modifies the bottom-up MLE in order to take into account top-down information regarding the probability that a registered sensory signal about a stimulus in the world is “correct”. In this way, the number of possible interpretations for ambiguous sensory signals can be constrained. The Bayesian prior can be conceived of as an additional signal that is integrated with bottom-up sensory signals according to the MLE method. Yuille & Bülthoff (1996) who applied Bayesian decision theory to vision had proposed that without prior information or assumptions about the world, the visual system can not get reliable estimates of the viewed three-dimensional scene based on two-dimensional retinal images. Using the Bayesian framework, they modelled how different visual cues such as shape, texture, shading, slant of objects etc. can be optimally integrated in a Bayesian framework, and how this integration process can explain psychophysical data from experiments about stereo vision, shape perception from shading, and shape perception from texture.

The Bayesian framework has recently been successfully applied to describe human perception and performance in a number of multi-modal perceptual tasks, such as visual-haptic discrimination of object size (Ernst & Banks, 2002), or visual-auditory interactions (Alais & Burr, 2004). According to the Bayesian theory, perception is concerned with three stages of processing: The first stage is concerned with sensory estimation, then, the influence of prior knowledge needs to be accounted for, and in the third step, a final decision-making process to make a perceptual judgement is required. The decision process is modelled using some gain/loss function in relation to a defined goal. Depending on the task, the “penalty” for a mistaken perceptual judgement is different. For example, the question whether an approaching object is on collision course to oneself in real life will be judged with different criteria than, for example, in typical psychophysical experiments, where observers’ task is to make perceptual judgements about some perceptual entity that have no direct relevance for the observer. Usually, judgements such as “stimulus 1 is larger/faster/brighter than stimulus 2” are required in some kind of 2AFC (2 alternative forced choice) or 2IFC (2 interval forced choice) task. It appears appropriate to assume that the decision rule used highly depends on the anticipated “penalty” - this idea is already very familiar from Signal Detection Theory (cf. Section 1.3).

An important open question that remains is “How do the priors get into the brain?” Even though it is a reasonable assumption that the brain has some form of “knowledge” about the probabilistic nature of sensory stimulation patterns, the models make no explicit statement on how this is learned

6 THEORETICAL FRAMEWORKS OF PERCEPTION

or represented by the brain. Also, the questions how priors evolve, how flexible and general they are, and whether there is only one prior, or several competing priors for one perceptual entity, are unanswered yet. Furthermore, there are different ideas about the extent that sensory fusion is mandatory or not, and how strong or weak the fusion is (Clark & Yuille, 1990). For example, there is an unresolved debate that has started more than a century ago about whether sensory information is merged into a single amodal percept as a result of cue combination, or whether the brain has access to the single cues also after the fusion process.

One important point to consider is that the influence of priors becomes stronger the higher the sensory uncertainty is. Many psychophysical experiments, especially those that deliberately introduce cue conflict situations, create such situations. It may thus be that the influence of priors is somewhat over-represented due to the research methods applied in most cue-combination studies.

For the case ofvection, this has an interesting implication, since the influence of priors, which can be considered as cognitive, high-level information, should be capable of enhancing the self-motion illusion. This will be more thoroughly addressed in Section 6.5.

6.4.1 Bayesian integration and self-motion perception

Recent attempts to apply the Bayesian approach to self-motion perception, however, have been only partly successful: For example, Berger (2007) found only weak evidence for MLE for visual and vestibular cues in the perception of passive upright yaw rotations. In a series of experiments, the participants' task was to either actively turn themselves back to the original position after a passively presented rotation, or to identify which of two consecutively presented rotations was larger in a 2IFC paradigm. In the classical Bayesian manner, the rotations were first presented under single cue conditions (visual only, vestibular only), and then, based on the variances of the two modalities, the sensory weights in the combined visual-vestibular condition were predicted. The empirically obtained sensory weights for the combined visual-vestibular condition showed only weak correlation with the theoretically predicted weights. For some conditions, when observers noticed a visual-vestibular conflict, the predicted reduction of the variance for the combined case was not observed. The results for the 2IFC experiment were best accounted for by assigning a constant visual weight of .86, without any re-weighting of the visual and vestibular cues.

In another study, Wright, DiZio, & Lackner (2005) found that visual-vestibular integration for linear upwards motion does not always follow a MLE rule. Rather, vision was found to almost completely dominate contradicting vestibular signals when the visual stimulus showed a highly realistic representation of the testing environment, even if vestibular stimulation was far above threshold. In that study, participants viewed video-taped scenes of the lab from an ego-centric perspective on an HMD while being physically moved up and down on an "elevator seat". When visual motion amplitude was large and the visual scene was a highly accurate representation of the testing environment, vision was almost completely dominant over discordant vestibular signals. It was concluded that this finding can not be explained by a model that assumes linear summation of weighted sensory inputs. The findings were also incompatible with sensory conflict theories, which predict that conflicting vestibular signals above threshold should impair visualvection (Young, Dichgans, Murphy, & Brandt, 1973). Proponents of the MLE school might counter this finding by suggesting that the results might be explained by a changing influence of priors, or even a change in strategy depending

6 THEORETICAL FRAMEWORKS OF PERCEPTION

on the visual context. There is, however, no model yet than can convincingly explain how such changes in strategy might be triggered.

An interesting approach that takes a step forwards in addressing this issue was proposed recently by Jürgens & Becker (2006). They investigated the perception of self-rotation from visual, vestibular, and proprioceptive cues in an optokinetic drum with a turning platform inside, both of which could be controlled independently. Thus, they were able to present all those cues one by one or in all possible combinations. Participants' task was to estimate the angular displacements that they experienced during the different sensory stimulation conditions. It was found that estimates of self-rotation angles increased in accuracy and decreased in variability from mono-modal to bi-modal to tri-modal stimulation. Also, with tri-modal stimulation, responses became independent of rotation velocity, whereas in the mono- and bi-modal conditions, desired turning angles were overshoot at higher velocities and undershoot at lower velocities. The authors propose a Bayesian sensory fusion model that explicitly includes the influence of priors. It is basically a MLE model that treats the priors like an extra sensory input: After the sensory signals from visual, vestibular, somatosensory and other modalities are fused in an MLE manner, the weighted influence of the prior is combined with the fused sensory signals to form a "compound sensory signal". The weight of the priors is determined according to the uncertainty of the sensory signals: The higher the sensory uncertainty is, regardless of whether it is due to ambiguous or noisy sensory signals, the higher the weight of the prior becomes; conversely, the lower the sensory uncertainty, the lower the prior's weight. The authors were able to qualitatively predict the psychophysical data using this model. However, Bayesian models have generally failed to make quantitative predictions about how the sensory weights and weights of priors change. One problem is that it is not possible to directly estimate the value of the prior in psychophysical experiments, or to determine what the "default value" for the prior is. It is, thus still not clear how the *Bayesian observer* sets and adapts the values of the priors. It has been argued that it is plausible that the brain assumes sensory signals to be in the natural range of everyday experience, perhaps a result of perceptual learning. For example, for human speed perception, Stocker & Simoncelli (2006) proposed a Bayesian model which assumes a prior for low speeds. With this model, they were able to predict psychophysical data where visual speed perception seems to be biased towards lower speeds when contrast of the visual stimulus is reduced. In general, it is observed that participants in psychophysical experiments show a tendency to give responses that are close to the mean of the range of values tested. This phenomenon is known as a compression towards the mean and becomes more pronounced the more uncertain or ambiguous sensory information is (see, e.g., Jürgens & Becker, 2006).

One interesting observation is that the MLE-inspired experiments on visual-vestibular integration were done in open loop, due to methodological requirements. A recent study by Wilkie & Wann (2005) investigated visual and vestibular contributions in a closed-loop steering task. The participants' task was to steer actively towards a moving target while retinal flow, gaze angle, and vestibular stimulation presented passively via a rotating chair were varied independently. Results showed that vestibular stimulation had little or no influence on steering performance when "correct" physical rotations corresponding to the steering action were presented, and even more surprising, when phase and amplitude of vestibular stimulation was conflicting with visual information, performance decreased only slightly. It seems that vestibular information was completely disregarded by the observers for this task. This is remarkable since turning rates were above threshold and in the

range of natural steering manoeuvres. The authors concluded that visual and vestibular signals were not integrated by the observers in this task. It is possible that at higher steering rates, a vestibular influence might have occurred. This finding is in contrast to other studies that were done in open-loop which stress the importance of vestibular stimulation, such as Bertin & Berthoz (2004) who found that observers were better able to reproduce passively travelled curved and linear trajectories correctly when vestibular information was available. One difference between these two studies apart from the open-loop vs. closed-loop design is that for the online steering control task towards a target used by Wilkie & Wann (2005), position- and orientation-based information is less important, compared to the task of reproducing passively travelled trajectories employed by Bertin & Berthoz (2004). This may be true for the case of upright yaw rotations, since this does not involve changes in orientation relative to gravity, whereas the case is different for pitch- or roll-rotations. In a close-loop experiment by Berger, Terzibas, Beykirch, & Bühlhoff (2007) on helicopter hovering control, non-pilot participants were better able to stabilise a simulated helicopter in a hovering position when pitch- and roll-rotations were delivered by a 6-dof motion platform, compared to a purely visual baseline. It appears that in this case, vestibular (and other body cues, e.g., somatosensory) were useful for the task. The Bayesian framework does not have a convincing explanation why in the helicopter hovering task, visual and vestibular information are integrated, but not in the steering task which includes only yaw rotations. In this realm, an ecologically oriented analysis might prove more promising in order to understand the dynamics of perception and action (Warren, 2006).

6.5 An integrative view on self-motion perception

In the sections above, the theoretical positions of direct and indirect perception were described, as well as how they concentrate on different levels of analysis. Several models about visual-vestibular interaction were discussed which can be considered as instantiations of the indirect theory. Here, I will argue that despite the theoretical and philosophical issues that seem to make these two views absolutely incompatible, both can equally contribute to our understanding of self-motion perception, once one puts the philosophical debate aside. Depending on whether the research question is directed to physiological, neuronal processes that underlie perception or whether the question is about what information is used for certain perceptual situations or tasks, such as steering a vehicle is being addressed, there are different approaches and levels of analysis that are adequate.

The aim of the current thesis is to extend our understanding of self-motion perception by analysing how bottom-up and top-down influences affect the perception of self-motion. Thevection illusion was used in to investigate multi-sensory contributions to self-motion perception. One important distinction that one should make is between *active* and *passive* self-motion. Vection is in most instances a case of passive self-motion perception, since only afferent information (visual, auditory, somatosensory, etc.) is available.⁶ Under natural conditions, comparable situations arise only seldom, such as in the train illusion. The distinction between active and passive self-motion is also relevant for self-motion simulations. In more physiological terms, the distinction between *ideothetic*, i.e., self-generated, and *allothetic*, i.e., externally generated motion put forward by Mittelstaedt & Mittelstaedt (2001) is useful. In VR, ideothetic information (especially vestibular and propriocep-

⁶One exception is the special case of podokineticvection when blindfolded observers step in place on a rotating disk such that there is no physical displacement, and only efferent information from proprioception is available (Becker et al., 2002a). After stepping in place for several seconds, observers perceive rotationalvection.

6 THEORETICAL FRAMEWORKS OF PERCEPTION

tive information) is generally impoverished, unless one walks around in a VE wearing a head-tracked HMD (head-mounted display). The perceptual systems that are tuned to perceive self-motion have evolved in an environment where almost all self-motion is self-initiated. The fact that natural ideothetic information is missing in most cases of vection and also during passive transportation and also during motion simulation does make a big difference in the pattern of sensory stimulation. Especially the vestibular modality is differentially affected by active or passive stimulation. For example, it has been shown that active vs. passive head-rotations at the same frequencies and amplitudes evoke different VOR (vestibulo-ocular reflex) responses such that in the passive case, where the head of an observer is turned by a torque motor, VOR gains are lower than in active turns (Leigh & Zee, 2006). Furthermore, most physical motions that we passively experience in vehicles have different frequency characteristics than our natural head- and body-movements. This might be one reason why realistic self-motion simulation with adequate vestibular stimulation is so difficult. In general, it is possible that in the range of natural turning rates that occur during passive navigation, such as steering a car or bicycle, vestibular information may be less important than other cues for active control of self-motion, even if the rotational motion is clearly above threshold. This could be a reason why Wilkie & Wann (2005) found no vestibular influence on the active steering task.

Under such informationally underspecified conditions of sensory stimulation, it is reasonable to assume that sensory processing is altered. One consequence can be that other sources of information, including cognitive knowledge, are included under such conditions by the perceptual system in order to increase reliability. In this regard, it is interesting to observe that, in recent times, the influence of priors or other cognitive factors have been receiving more and more attention in perception research.

The current thesis showed that the feeling of presence in a Virtual Environment, which can be considered a cognitive factor, can also affect vection. Furthermore, consistent multi-modal sensory stimulation resulted in both higher levels of reported presence and increased self-motion perception.

Irrespective of whether the fusion of sensory and cognitive information follows the principles of Bayesian integration or not, the idea that also cognitive factors can influence perception, especially in VR settings, seems to be conceivable and more research attention should be devoted to this question.

In the following, a conceptual model of self-motion perception that integrates the findings of the current thesis and other related studies will be proposed.

Figure 22 shows the relations how self-motion perception depends on the pattern of sensory stimulation, motor commands, and cognitive factors. The five senses considered here all can contribute to self-motion perception, either in an active case when motion is self-generated (i.e., ideothetic), or in a passive case, when sensory motion information is allothetic, like in the case of vection and also in most self-motion simulation applications. It is proposed that sensory information is integrated, and the thickness of the arrows which connect the sensory modalities with the integration stage indicate their respective relative weights in the sensory integration process. Note that these weights are usually estimated in passive self-motion conditions. As can be seen, vision and vestibular information have the highest weights, followed by proprioception, somatosensation and audition. It should be noted that this model does not make any explicit statement about whether sensory integration follows the rules of MLE, or whether it results in an amodal percept or representation. The arrows are bi-directional in order to reflect the experimental findings that sensory weights are flexible and adaptable in the multi-sensory integration process. In the schematic, the five senses are not interconnected with one another before the integration stage, even though this is likely the case, like studies

6 THEORETICAL FRAMEWORKS OF PERCEPTION

on prism adaptation or sensory substitution have shown. Such low-level interaction processes are subsumed graphically in the integration stage.

The upper left branch with the blue arrows represents the active case of self-motion, in which the natural perception-action loop is uninterrupted. The motor commands which are executed lead to self-motion, which is perceived, and the new sensory information resulting from that perception generates new motor commands. In this model, the refference feeds back to the integration stage, but this does not imply that refference operates only on the integrated percept.

The upper right branch represents the case of passive self-motion. Here, the perception-action loop is interrupted, and information about self-motion is allothetic. The thick line connecting vision with the integration stage indicates the visual dominance that is typically found in vection research. Vestibular information is the second strongest modality for self-motion. The current thesis showed that cognitive factors, such as attention and presence, can have an influence on perceived self-motion, even in the case of purely visually induced vection (Exp. 1 & Exp. 2). Cognition can have an influence on the sensory integration process, as suggested, e.g., by Bayesian models, which use the concept of priors. One example of this can be found in the self-motion model proposed by Jürgens & Becker (2006). On the other hand, if the pattern of sensory stimulation is incongruent or unnatural, this might in turn have an effect on cognition, such as reducing presence in a VE. It is also possible that cognition might have a direct influence on self-motion perception, without necessarily influencing the sensory integration process. One such case would be reporting bias, where observers might be led to report the onset of vection faster due to experimental demand or instructions (see, e.g., Palmisano & Chan, 2004). It could also be that if strong vection is perceived in a VE, this might in turn lead to a higher sense of presence. This idea that vection and presence in VR might mutually influence each other was suggested earlier by Schulte-Pelkum & Riecke (2007).

The fact that perceived self-motion was increased by providing ecologically valid, multi-sensory patterns of stimulation indicated that higher-order, cross-modal invariants might be picked up by perceivers in VR if they are provided. One has to state that current VR is still quite far away from being able to provide a degree of sensory realism anywhere close to being indistinguishable from reality, but little steps that were taken into this direction in the current thesis showed measurable improvement for the case of vection.

The current model is useful to conceptualise the findings about how perceptual and cognitive factors influence the perception of self-motion. This model is also useful in the special context of VR. Some tentative recommendations about how to improve self-motion simulation in VR that can be derived from the results of the current thesis and from this model will be sketched in Section 7.1.

6 THEORETICAL FRAMEWORKS OF PERCEPTION

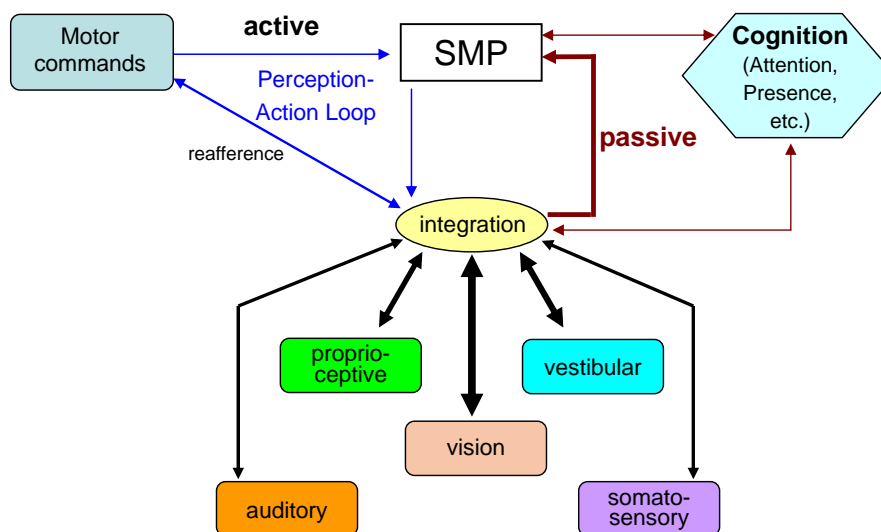


Figure 22: Conceptual model of human self-motion perception (SMP). SMP can either result from active or passive self-motion, as represented in the left and the right branches, respectively. The left branch with the blue arrows represents the natural perception-action loop and includes motor commands. The right branch with arrows in dark red represents the passive case of self-motion, e.g., the case of vection. Cognitive influences, such as attention or presence, can also have an effect. This cognitive influence may either directly affect SMP, or the sensory interaction process. The thickness of the arrows connecting each sensory modality with the integration stage represent the relative weights of the modalities in the integration process. See text for more details.

7 Conclusions

In the current thesis, multi-sensory interactions during self-motion perception in VR were investigated, and it was found that vection is increased by consistent multi-sensory stimulation, compared to the purely visual baseline. It was also found that cognitive factors, such as presence or attention, can have an influence on vection, even though the strongest effects were found in the multi-sensory visual-vestibular condition. The results indicated that natural stimuli which provide a high degree of sensory realism enhanced both vection and presence in VR. In the following, an outlook is provided that focuses on VR-specific issues how one might improve self-motion simulation based on the current results.

7.1 VR-specific recommendations for enhanced vection and future research

From the results obtained in this dissertation, some tentative recommendations can be derived how vection might be enhanced in VR setups. Knowledge of this kind is not only useful for gaming and entertainment applications, but also for researchers using VR technology for self-motion perception, since VR is becoming more and more a standard tool for psychophysical research.

Consistent with the vection literature, vision was found to be the most influential modality. A novel finding is that using a highly realistic visual scene substantially increased vection, compared to artificial stimuli. Furthermore, larger FOVs induced stronger vection, which was also to be expected. Generally, adding consistent stimulation from non-visual modalities increased vection. If one has to make a cost-benefit calculation, adding minor vestibular kicks and vibrations that qualitatively match to a natural, realistic visual stimulus presented on a medium-sized FOV setup would be most beneficial. For example, motion seats which are already used in gaming and entertainment industry might be suitable for this purpose. This kind of minimal motion cueing seems to be most effective if one aims at reducing vection onset latencies and increasing vection compellingness. Thus, a mid-sized FOV visual projection, combined with minimalistic motion cueing and possibly dynamic auditory information presented via headphones seems to be a reasonable alternative to full-blown 6-dof motion base simulator systems. Of course, the issue would be different if the purpose of the simulation is to train a specific skill, such as controlled emergency braking in a driving simulator or specific flight manoeuvres in aviation. In such situations, even full-blown motion simulators that aim to simulate inertial forces as realistically as possible still face unsolved problems.

Another factor that was found to have a strong influence on vection onset latency is the motion profile of the visual stimulus. The shorter the acceleration time (i.e., the time the static visual stimulus took to accelerate to reach constant velocity), the faster vection onset occurred. Also, the higher the speed of constant velocity motion, the stronger vection was, at least for the range of values tested here (i.e., 20 to 60°/sec for yaw rotations and 6 to 8m/s for forward translation). This finding is somewhat incompatible with the classical visual-vestibular cue-conflict models by Zacharias & Young (1981) and Mergner et al. (1995), since a higher visual acceleration implicates larger conflict. It is possible that the transient signal of the vestibular system decays quicker if visual acceleration time is short, which would then be also reflected in the case of vection where the vestibular *veto* due to the missing vestibular signal would decay faster. Further progress in modern neural imaging

7 CONCLUSIONS

technology is necessary in order to be able to test this hypothesis empirically on human subjects, since it is currently not possible to correlate activation of the vestibular nuclei to visual stimulation patterns.

Adding moving spatialised sound using generic HRTFs slightly increased vection compellingness and presence, but did not affect vection onset. Still, this resource should not be neglected, since it can be implemented at relatively low cost. It should also be noted that a stronger auditory influence might be found if the auditory rendering quality is increased such that individual HRTFs are used, many objects distributed in the whole scene are audible, and room acoustics (i.e., reflected sound) are simulated (Larsson, Västfjäll, & Kleiner, 2003). Technology of this kind is currently developing at a fast pace, and its' usefulness should be further investigated in the context of immersive VR.

There are two main shortcomings in the current experiments that limit the conclusions that can be drawn from the results. One is that all experiments were purely passive and observers did not have any control over the simulated motion, other than pressing a button in order to start a trial. This was required by the measurement method applied. Both from an applied and theoretical point of view, it would be desirable to investigate how the extent of possible active control of movements in the VE influences vection of VR users, when speed and direction of motion can be freely chosen. The other limitation is that no experiment investigated the combined effect of *all* sensory modalities together in one experiment. Experiment 6 tested the effect of combined visual, vestibular, and vibratory stimulation and found a significant effect such that tri-modal stimulation increased vection, compared to bi-modal stimulation. It is possible that stimulating all modalities tested here in a perceptually realistic and consistent manner might have a synergistic effect of further increasing vection. In order to overcome these two limitations, still some technical progress, both for software and hardware issues, is necessary. However, with moderate optimism, one might expect that truly multi-sensory VEs with a high degree of real-time interactivity might become available in the near future. A technology of that kind would open up new possibilities of conducting both open-loop and closed-loop experiments that provide nearly as rich sensory information as in the real world. Both the research community as well as motion simulation applications for training and entertainment would largely benefit from such developments.

References

- Alais, D., & Burr, D. (2004). No direction-specific bimodal facilitation for audiovisual motion detection. *Cognitive Brain Research*, **19**(2), 185–194.
- Andersen, G. J., & Braunstein, M. L. (1985). Induced Self-Motion in Central Vision. *Journal of Experimental Psychology-Human Perception and Performance*, **11**(2), 122–132.
- Aubert, H. (1886). Die Bewegungsempfindung. *Pflügers Archiv*, **39**, 347–370.
- Becker, W., Jürgens, R., & Boß, T. (2000). Vestibular perception of self-rotation in different postures: a comparison between sitting and standing subjects. *Experimental Brain Research*, **131**(4), 468–476.
- Becker, W., Nasios, G., Raab, S., & Jurgens, R. (2002a). Fusion of vestibular and podokinesthetic information during self-turning towards instructed targets. *Experimental Brain Research*, **144**(4), 458–474.
- Becker, W., Raab, S., & Jürgens, R. (2002b). Circular vection during voluntary suppression of optokinetic reflex. *Experimental Brain Research*, **144**(4), 554–557.
- 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., Spencer, M. B., & Stott, J. R. R. (1986). Thresholds for the detection of the direction of whole-body, linear movement in the horizontal plane. *Aviation, Space, and Environmental Medicine*, **57**(11), 1088–1096.
- Benson, A. (1990). Sensory functions and limitations of the vestibular system. In A. R. Warren (Ed.), *Perception & control of self-motion* (pp. 145–170). Hillsdale, New Jersey: Lawrence Erlbaum.
- Berger, D. (2007). *Sensor fusion in the perception of self-motion*. Ph.D. thesis, Universität Ulm, Fakultät für Informatik.
- Berger, D., Terzibas, C., Beykirch, K., & Bühlhoff, H. (2007). The role of visual cues and whole-body rotations in helicopter hovering control. In *Proceedings of the AIAA Modeling and Simulation Technologies Conference and Exhibit, Hilton Head, USA*, No. AIAA-2007-6798.
- Berthoz, A., & Droulez, J. (1982). Linear self motion perception. In A. H. Wertheim, W. A. Wagenaar, & H. W. Leibowitz (Eds.), *Tutorials on Motion Perception* (pp. 157–199). New York: Plenum.
- Berthoz, A., Pavard, B., & Young, L. R. (1975). Perception of linear horizontal self-motion induced by peripheral vision (linearvection) - basic characteristics and visual-vestibular interactions. *Experimental Brain Research*, **23**(5), 471–489.
- Bertin, R. J. V., & Berthoz, A. (2004). Visuo-vestibular interaction in the reconstruction of travelled trajectories. *Experimental Brain Research*, **154**(1), 11–21.

- Blauert, J. (1997). *Spatial hearing: The psychophysics of human sound localization*. Cambridge, Mass.: MIT Press.
- Brandt, T., Bartenstein, P., Janek, A., & Dieterich, M. (1998). Reciprocal inhibitory visual-vestibular interaction - Visual motion stimulation deactivates the parieto-insular vestibular cortex. *Brain*, **121**, 1749–1758.
- Brandt, T., Büchele, W., & Arnold, F. (1977). Arthrokinetic nystagmus and ego-motion sensation. *Experimental Brain Research*, **31**(2), 331–338.
- Brandt, T., Dichgans, J., & Koenig, E. (1973). Differential effects of central versus peripheral vision on egocentric and exocentric motion perception. *Experimental Brain Research*, **16**, 476–491.
- Bremmer, F., Duhamel, J.-R., Hamed, S., & Graf, W. (2000). Stages of self-motion processing in primate posterior parietal cortex. In M. Lappe (Ed.), *Neuronal processing of optic flow*, Vol. 44 of *International review of Neurobiology* (pp. 173–198). New York, Boston, Tokyo: Academic Press.
- Bremmer, F., Klam, F., Duhamel, J. R., Ben Hamed, S., & Graf, W. (2002). Visual-vestibular interactive responses in the macaque ventral intraparietal area (VIP). *European Journal of Neuroscience*, **16**(8), 1569–1586.
- Bremmer, F., & Lappe, M. (1999). The use of optical velocities for distance discrimination and reproduction during visually simulated self motion. *Experimental Brain Research*, **127**(1), 33–42.
- Bringoux, L., Nougier, V., Barraud, P., Marin, L., & Raphel, C. (2003). Contribution of somesthetic information to the perception of body orientation in the pitch dimension. *Quarterly Journal of Experimental Psychology Section A - Human Experimental Psychology*, **56**(5), 909–923.
- Bringoux, L., Schmerber, S., Nougier, V., Dumas, G., Barraud, P., & Raphel, C. (2002). Perception of slow pitch and roll body tilts in bilateral labyrinthine-defective subjects. *Neuropsychologia*, **40**, 367–372.
- Carpenter-Smith, T. R., Futamura, R. G., & Parker, D. E. (1995). Inertial acceleration as a measure of linear vection: An alternative to magnitude estimation. *Perception & Psychophysics*, **57**(1), 35–42.
- Clark, B. (1967). Thresholds for the perception of angular acceleration in man. *Aerospace Medicine*, **38**, 443–450.
- Clark, J. J., & Yuille, A. L. (1990). *Data Fusion for Sensory Information Processing Systems*. Kluwer Academic Publishers, Boston.
- Cutting, J., & Vishton, P. (1995). Perceiving Layout and Knowing Distances: The Integration, Relative Potency, and Contextual Use of Different Information about Depth. In W. Epstein & S. Rogers (Eds.), *Perception of Space and Motion*, Handbook of Perception and Cognition (pp. 69–117). San Diego: Academic Press.
- Cutting, J. (1986). *Perception with an eye for motion*. Cambridge, MA: MIT Press.

- Cutting, J. (1991). Four ways to reject directed perception. *Ecological Psychology*, **3**(1), 25–34.
- Darlington, C. L., & Smith, P. F. (1998). Further evidence for gender differences in circularvection. *Journal of Vestibular Research-Equilibrium & Orientation*, **8**(2), 151–153.
- de Graaf, B., Wertheim, A. H., & Bles, W. (1991). The Aubert-Fleischl Paradox Does Appear in Visually Induced Self-Motion. *Vision Research*, **31**(5), 845–849.
- Dichgans, J., & Brandt, T. (1978). Visual-Vestibular Interaction: Effects on Self-Motion Perception and Postural Control. In R. Held, H. W. Leibowitz, & H.-L. Teuber (Eds.), *Perception*, Vol. VIII of *Handbook of Sensory Physiology* (pp. 756–804). Berlin Heidelberg: Springer.
- Diener, H. C., Wist, E. R., Dichgans, J., & Brandt, T. (1976). Spatial Frequency Effect on Perceived Velocity. *Vision Research*, **16**(2), 169–176.
- Distler, H. (1996). Psychophysical Experiments in Virtual Environments. In *Virtual Reality World 96 Conference Documentation*. München 1996: Computerwoche Verlag AG.
- Distler, H. K. (2003). *Wahrnehmung in Virtuellen Welten*. Ph.D. thesis, Justus-Liebig-Universität Gießen.
- Dodge, R. (1923). Thresholds of rotation. *Journal of Experimental Psychology*, **6**, 107–137.
- Ernst, M. O., & Banks, M. S. (2002). Humans Integrate Visual and Haptic Information in a Statistically Optimal Fashion. *Nature*, **415**, 429–433.
- Ernst, M. O., & Bühlhoff, H. H. (2004). Merging the senses into a robust percept. *Trends in Cognitive Sciences*, **8**(4), 162–169.
- Fernandez, C., & Goldberg, J. (1971). Physiology of peripheral neurons innervating semicircular canals of the squirrel monkey. II. Response to sinusoidal stimulation and dynamics of peripheral vestibular system. *Journal of Neurophysiology*, **34**(4), 661–675.
- Fischer, M. H., & Kornmüller, A. E. (1930). Optokinetisch ausgelöste Bewegungswahrnehmung und optokinetischer Nystagmus [Optokinetically induced motion perception and optokinetic nystagmus]. *Journal für Psychologie und Neurologie*, 273–308.
- Freeman, J., Avons, S. E., Meddis, R., Pearson, D. E., & IJsselstein, W. I. (2000). Using behavioral realism to estimate presence: A study of the utility of postural responses to motion stimuli. *Presence - Teleoperators and Virtual Environments*, **9**(2), 149–164.
- Giannopulu, I., & Lepecq, J. (1998). Linear-vection chronometry along spinal and sagittal axes in erect man. *Perception*, **27**(3), 363–372.
- Gibson, J. J. (1950). *The perception of the visual world*. Boston: Houghton Mifflin.
- Gibson, J. J. (1966). *The senses considered as perceptual systems*. Boston: Houghton Mifflin.
- Gibson, J. J. (1979). *The ecological approach to visual perception*. Houghton Mifflin, Boston.
- Goldstein, E. B. (1999). *Sensation and perception* (5th edition). Brooks/Cole.

- Green, D., & Swets, J. (1966). *Signal Detection and Psychophysics*. New York: Wiley.
- Gregory, R. (1985). Perceptions as hypotheses. *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences*, **290**(1038), 181–197.
- Guedry, F. E. (1974). Psychophysics of vestibular sensation. In H. H. Kornhuber (Ed.), *Handbook of Sensory Physiology: The vestibular system*, Vol. VI (2) (pp. 2–154). Springer.
- Hettinger, L. J. (2002). Illusory Self-motion in Virtual Environments. In K. M. Stanney (Ed.), *Handbook of Virtual Environments* (pp. 471–492). Lawrence Erlbaum.
- Howard, I. P. (1982). *Human visual orientation*. Chichester, New York, Engl.: J. Wiley.
- Howard, I. P. (1986a). The perception of posture, self motion, and the visual vertical. In K. R. Boff, L. Kaufman, & J. P. Thomas (Eds.), *Sensory processes and perception*, Vol. 1 of *Handbook of human perception and performance* (pp. 18.1–18.62). New York: Wiley.
- Howard, I. P. (1986b). The vestibular system. In K. R. Boff, L. Kaufman, & J. P. Thomas (Eds.), *Sensory processes and perception*, Vol. 1 of *Handbook of human perception and performance* (pp. 11.1–11.30). New York: Wiley.
- Howard, I. P., & Heckmann, T. (1989). Circular vection as a function of the relative sizes, distances, and positions of 2 competing visual-displays. *Perception*, **18**(5), 657–665.
- Howard, I. P., & Howard, A. (1994). Vection - the contributions of absolute and relative visual motion. *Perception*, **23**(7), 745–751.
- IJsselstein, W., de Ridder, H., Freeman, J., Avons, S. E., & Bouwhuis, D. (2001). Effects of stereoscopic presentation, image motion, and screen size on subjective and objective corroborative measures of presence. *Presence - Teleoperators and Virtual Environments*, **10**(3), 298–311.
- IJsselstein, W. A. (2004). *Presence in Depth*. Ph.D. thesis, Technische Universiteit Eindhoven, Netherland.
- Johansson, G. (1977). Studies on visual perception of locomotion. *Perception*, **6**(4), 365–376.
- 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**(3), 528–543.
- Kandel, E., Schwartz, J., & Jessell, T. M. (2000). *Principles of Neural Science*. McGraw-Hill Companies, The Hague.
- Kano, C. (1991). The Perception of Self-Motion Induced by Peripheral Visual Information in Sitting and Supine Postures. *Ecological Psychology*, **3**(3), 241–252.
- Kennedy, R. S., Hettinger, L. J., Harm, D. L., Ord, J. M., & Dunlap, W. P. (1996). Psychophysical scaling of circular vection (CV) produced by optokinetic (OKN) motion: Individual differences and effects of practice. *Journal of Vestibular Research-Equilibrium & Orientation*, **6**(5), 331–341.

- Kennedy, R. S., Lane, N. E., Lilienthal, M. G., Berbaum, K. S., & Lawrence (1992). Profile Analysis of Simulator Sickness Symptoms: Application to Virtual Environment Systems. *Presence - Teleoperators and Virtual Environment*, **1**(3), 295–301.
- Kitazaki, M., & Sato, T. (2003). Attentional modulation of self-motion perception. *Perception*, **32**(4), 475–484.
- Klatzky, R. L., Loomis, J. M., Beall, A. C., Chance, S. S., & Golledge, R. G. (1998). Spatial updating of self-position and orientation during real, imagined, and virtual locomotion. *Psychological Science*, **9**(4), 293–298.
- Kleinschmidt, A., Thilo, K. V., Buchel, C., Gresty, M. A., Bronstein, A. M., & Frackowiak, R. S. J. (2002). Neural correlates of visual-motion perception as object- or self-motion. *Neuroimage*, **16**(4), 873–882.
- Knill, D., & Richards, W. (Eds.). (1996). *Perception as Bayesian Inference*. New York: Cambridge University Press.
- Knill, D. C., & Saunders, J. A. (2003). Do humans optimally integrate stereo and texture information for judgments of surface slant?. *Vision Research*, **43**(24), 2539–2558.
- Lackner, J. R. (1977). Induction of Illusory Self-Rotation and Nystagmus by a Rotating Sound-Field. *Aviation Space and Environmental Medicine*, **48**(2), 129–131.
- Lackner, J. R., & DiZio, P. (2000a). Artificial gravity as a countermeasure in long-duration space flight. *Journal of Neuroscience Research*, **62**(2), 169–176.
- Lackner, J. R., & DiZio, P. (2000b). Human orientation and movement control in weightless and artificial gravity environments. *Experimental Brain Research*, **130**(1), 2–26.
- Lackner, J. (1992). Multimodal and motor influences on orientation: Implications for adapting to weightless and virtual environments. *Journal of Vestibular Research*, **2**(4), 307–322.
- Larish, J. F., & Flach, J. M. (1990). Sources of optical information useful for perception of speed of rectilinear self-motion. *Journal of Experimental Psychology: Human Perception and Performance*, **16**(2), 295–302.
- Larsson, P., Västfjäll, D., & Kleiner, M. (2004). Perception of Self-motion and Presence in Auditory Virtual Environments. In *Proceedings of Seventh Annual Workshop Presence 2004*, pp. 252–258. Available: www.kyb.mpg.de/publication.html?publ=2953.
- Larsson, P., Västfjäll, D., & Kleiner, M. (2003). On the Quality of Experience: A Multi-Modal Approach to Perceptual Ego-Motion and Sensed Presence in Virtual Environments. In U. Jekosch & S. Möller (Eds.), *First ISCA Tutorial & Research Workshop on Auditory Quality of Systems*, pp. 97–100 Akademie Mont-Cenis, Germany.
- Leibowitz, H. (1988). The human senses in flight. In E. Wiener & D. Nagel (Eds.), *Human factors in aviation*, (pp. 83–109). San Diego: Academic Press.
- Leigh, R., & Zee, D. (2006). *The neurology of eye movements*. Oxford: Oxford University Press.

- Lepecq, J. C., Giannopulu, I., & Baudonniere, P. M. (1995). Cognitive effects on visually induced body motion in children. *Perception*, **24**(4), 435–449.
- Lepecq, J. C., Giannopulu, I., Mertz, S., & Baudonniere, P. M. (1999). Vestibular Sensitivity and Vection Chronometry Along the Spinal Axis in Erect man. *Perception*, **28**(1), 63–72.
- Lepecq, J. C., Jouen, F., & Dubon, D. (1993). The Effect of Linear Vection on Manual Aiming at Memorized Directions of Stationary Targets. *Perception*, **22**(1), 49–60.
- Mach, E. (1875). *Grundlinien der Lehre von den Bewegungsempfindungen*. Leipzig, Germany: Engelmann.
- Marme-Karelse, A. M., & Bles, W. (1977). Circular vection and human posture II: Does the auditory-system play a role?. *Agressologie*, **18**(6), 329–333.
- Marr, D. (1982). *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information*. San Francisco: W.H. Freeman.
- Melcher, G. A., & Henn, V. (1981). The latency of circular vection during different accelerations of the optokinetic stimulus. *Perception & Psychophysics*, **30**(6), 552–556.
- Merfeld, D. M., & Zupan, L. H. (2002). Neural processing of gravitoinertial cues in humans. III. Modeling tilt and translation responses. *Journal of Neurophysiology*, **87**(2), 819–833.
- Mergner, T., Schweigart, G., Kolev, O., Hlavacka, F., & Becker, W. (1995). Visual-vestibular interaction for human ego-motion perception. In T. Mergner & F. Hlavacka (Eds.), *Multisensory Control of Posture* (pp. 157–167). New York, London: Plenum.
- Mittelstaedt, M. L., & Mittelstaedt, H. (2001). Idiothetic navigation in humans: estimation of path length. *Experimental Brain Research*, **139**(3), 318–332.
- Mohler, B. J., Riecke, B. E., Thompson, W. B., & Bühlhoff, H. H. (2005). Measuring Vection in a Large Screen Virtual Environment. In *ACM SIGGRAPH Symposium on Applied Perception in Graphics and Visualization*, (APGV), pp. 103–109 La Coruña, Spain.
- Nakamura, S. (2004). Effects of spatial arrangement of visual stimulus on inverted self-motion perception induced by the foreground motion: examination of OKN-suppression hypothesis. *Vision Research*, **44**(16), 1951–1960.
- Nakamura, S., & Shimojo, S. (1999). Critical role of foreground stimuli in perceiving visually induced self-motion (vection). *Perception*, **28**(7), 893–902.
- Nakamura, S., & Shimojo, S. (2000). A slowly moving foreground can capture an observer's self-motion - a report of a new motion illusion: inverted vection. *Vision Research*, **40**(21), 2915–2923.
- Ohmi, M., Howard, I. P., & Landolt, J. P. (1987). Circular Vection as a Function of Foreground-Background Relationships. *Perception*, **16**(1), 17–22.
- Palmisano, S. (1996). Perceiving self-motion in depth: The role of stereoscopic motion and changing-size cues. *Perception & Psychophysics*, **58**(8), 1168–1176.

- Palmisano, S. (2002). Consistent stereoscopic information increases the perceived speed of vection in depth. *Perception*, **31**(4), 463–480.
- Palmisano, S., & Chan, A. Y. C. (2004). Jitter and size effects on vection are immune to experimental instructions and demands. *Perception*, **33**(8), 987–1000.
- Palmisano, S., & Gillam, B. (1998). Stimulus eccentricity and spatial frequency interact to determine circular vection. *Perception*, **27**(9), 1067–1077.
- Pretto, P., Vidal, M., & Chatziastros, A. (2007). The effect of gaze direction and field-of-view on speed constancy. In H. Bülthoff, A. Chatziastros, H. Mallot, & R. Ulrich (Eds.), *Proceedings of the 10th Tübinger Perception Conference* Kirchentellinsfurt, Germany. Knirsch Verlag.
- Reason, J., & Brand, J. (1975). *Motion Sickness*. London: Academic Press.
- Reymond, G., Droulez, J., & Kemeny, A. (2002). Visuovestibular perception of self-motion modeled as a dynamic optimization process. *Biological Cybernetics*, **87**(4), 301–314.
- Riecke, B. E., Schulte-Pelkum, J., Avraamides, M., von der Heyde, M., & Bülthoff, H. H. (2006). Cognitive Factors can Influence Self-Motion Perception (Vection) in Virtual Reality. *ACM Transactions on Applied Perception (TAP)*, **3**(3), 194–216.
- Riecke, B. E., von der Heyde, M., & Bülthoff, H. H. (2005). Visual cues can be sufficient for triggering automatic, reflex-like spatial updating. *ACM Transactions on Applied Perception (TAP)*, **2**(3), 183–215.
- Rosas, P., Wagemans, J., Ernst, M., & Wichmann, F. (2005). Texture and haptic cues in slant discrimination: reliability-based cue weighting without statistically optimal cue combination. *Journal of the Optical Society Of America A-Optics Image Science and Vision*, **22**(5), 801–809.
- Rushton, S., Harris, J., Lloyd, M., & Wann, J. (1998). Guidance of locomotion on foot uses perceived target location rather than optic flow. *Current Biology*, **8**(21), 1191–1194.
- Sadowski, W., & Stanney, K. (2002). Presence in Virtual Environments. In K. M. Stanney (Ed.), *Handbook of Virtual Environments* (pp. 791–806). Lawrence Erlbaum.
- Schinauer, T., Hellmann, A., & Höger, R. (1993). Dynamic acoustical stimulation affects self-motion perception. In A. Schick (Ed.), *Contributions to Psychological Acoustics. Results of the 6th Oldenburg Symposium on Psychological Acoustics*, Contributions to Psychological Acoustics (pp. 373–385). Oldenburg: Bibliotheks- und Informationssystem der Carl von Ossietzky Universität Oldenburg.
- Schubert, T., Friedmann, F., & Regenbrecht, H. (2001). The experience of presence: Factor analytic insights. *Presence - Teleoperators and Virtual Environments*, **10**(3), 266–281.
- Schulte-Pelkum, J., & Riecke, B. E. (2007). An integrative approach to presence and self-motion perception research. In W. IJsselstein, J. Freeman, & F. Biocca (Eds.), *Handbook of Presence* (). Lawrence Erlbaum. accepted.

- Slater, M. (2004). How colorful was your day? Why questionnaires cannot assess presence in virtual environments. *Presence - Teleoperators and Virtual Environments*, **13**(4), 484–493.
- Slater, M., & Steed, A. (2000). A virtual presence counter. *Presence - Teleoperators and Virtual Environments*, **9**, 413–434.
- Slater, M., Steed, A., McCarthy, J., & Maringelli, F. (1998). The influence of body movement on subjective presence in virtual environments. *Human Factors*, **40**(3), 469–477.
- Slater, M., & Wilbur, S. (1997). A framework for immersive virtual environments (FIVE): Speculations on the role of presence in virtual environments. *Presence - Teleoperators and Virtual Environments*, **6**(6), 603–616.
- Smith, A., Wall, M., Williams, A., & Singh, K. (2006). Sensitivity to optic flow in human cortical areas MT and MST. *European Journal of Neuroscience*, **423**(2), 561–569.
- Stevens, S. S. (1957). On the psychophysical law. *Psychological Review*, **64**, 153–181.
- Stocker, A., & Simoncelli, E. (2006). Noise characteristics and prior expectations in human visual speed perception. *Nature Neuroscience*, **9**(4), 578–585.
- Stoffregen, T. A., Hettinger, L. J., Haas, M. W., Roe, M. M., & Smart, L. J. (2000). Postural instability and motion sickness in a fixed-base flight simulator. *Human Factors*, **42**(3), 458–469.
- Telban, R. J., & Cardullo, F. M. (2001). An integrated model of human motion perception with visual-vestibular interaction. In *AIAA Modeling and Simulation Technologies Conference and Exhibit*, pp. 1–11 Montreal, Canada. AIAA.
- Telban, R., & Cardullo, F. (2005). Motion Cueing Algorithm Development: Human-Centered Linear and Nonlinear Approaches. Tech. rep. NASA/CR-2005-213747, NASA.
- Thilo, K. V., Guerraz, M., Bronstein, A. M., & Gresty, M. A. (2000). Changes in horizontal oculomotor behaviour coincide with a shift in visual motion perception. *Neuroreport*, **11**(9), 1987–1990.
- Thilo, K. V., Guerraz, M., Bronstein, A. M., & Gresty, M. A. (2002). Percept-related changes in horizontal optokinetic nystagmus at different body orientations in space. *Experimental Brain Research*, **145**(2), 215–221.
- Thilo, K. V., Kleinschmidt, A., & Gresty, M. A. (2003). Perception of self-motion from peripheral optokinetic stimulation suppresses visual evoked responses to central stimuli. *Journal of Neurophysiology*, **90**(2), 723–730.
- Treisman, M. (1977). Motion Sickness: An Evolutionary Hypothesis. *Science*, **197**, 493–495.
- Tschermak, A. (1931). Optischer Raumsinn. In A. Bethe, G. Bergmann, G. Emden, & A. Ellinger (Eds.), *Handbuch der Normalen und Pathologischen Physiologie* (pp. 834–1000). Springer, Berlin.

- Väljamäe, A., Larsson, P., Västfjäll, D., & Kleiner, M. (2004). Auditory Presence, Individualized Head-Related Transfer Functions, and Illusory Ego-Motion in Virtual Environments. In *Proceedings of Seventh Annual Workshop Presence 2004*, pp. 141–147 Valencia, Spain. Available: www.kyb.mpg.de/publication.html?publ=2954.
- Väljamäe, A., Larsson, P., Västfjäll, D., & Kleiner, M. (2005). Sonic self-avatar and self-motion in virtual environments. In *Proceedings of the 8th Annual Workshop of Presence, London 2005* London, England.
- van der Steen, F. A. M. (1998). *Self-motion perception*. Ph.D. thesis, Delft University of Technology.
- van der Steen, F. A. M., & Brockhoff, P. T. M. (2000). Induction and impairment of saturated yaw and surge vection. *Perception & Psychophysics*, **62**(1), 89–99.
- von Helmholtz, H. (1910). *Handbuch der physiologischen Optik*, Vol. 3. Voss.
- von Holst, E., & Mittelstaedt, H. (1950). Das Reafferenzprinzip. *Naturwissenschaften*, **20**, 464–475.
- Wann, J., & Rushton, S. (1994). The illusion of self-motion in virtual-reality environments. *Behavioral and Brain Sciences*, **17**(2), 338–340.
- Warren, R., & Wertheim, A. H. (Eds.). (1990). *Perception & Control of Self-Motion*. New Jersey, London: Erlbaum.
- Warren, W. H. (2006). The dynamics of perception and action. *Psychological Review*, **113**(2), 358–389.
- Warren, W. H., & Hannon, D. J. (1990). Eye-movements and optical-flow. *Journal of the Optical Society of America A-Optics Image Science and Vision*, **7**(1), 160–169.
- Warren, W. H., Kay, B. A., Zosh, W. D., Duchon, A. P., & Sahuc, S. (2001). Optic flow is used to control human walking. *Nature Neuroscience*, **4**(2), 213–216.
- Welch, R. B. (2002). Adapting to Virtual Environments. In K. M. Stanney (Ed.), *Handbook of Virtual Environments* (pp. 619–636). Lawrence Erlbaum.
- Wertheim, A. H. (1994). Motion perception during self-motion - the direct versus inferential controversy revisited. *Behavioral and Brain Sciences*, **17**(2), 293–311.
- Wertheim, A. H., Mesland, B. S., & Bles, W. (2001). Cognitive suppression of tilt sensations during linear horizontal self-motion in the dark. *Perception*, **30**(6), 733–741.
- Wilkie, R., & Wann, J. (2005). The role of visual and nonvisual information in the control of locomotion. *Journal of Experimental Psychology-Human Perception and Performance*, **31**(5), 901–911.
- Wist, E. R., Diener, H. C., Dichgans, J., & Brandt, T. (1975). Perceived distance and perceived speed of self-motion - linear vs angular velocity. *Perception & Psychophysics*, **17**(6), 549–554.
- Witmer, B. G., & Singer, M. J. (1998). Measuring presence in virtual environments: A presence questionnaire. *Presence - Teleoperators and Virtual Environments*, **7**(3), 225–240.

- Wong, S. C. P., & Frost, B. J. (1978). Subjective motion and acceleration induced by the movement of the observer's entire visual field. *Perception & Psychophysics*, **24**(2), 115–120.
- Wong, S. C. P., & Frost, B. J. (1981). The effect of visual-vestibular conflict on the latency of steady-state visually induced subjective rotation. *Perception & Psychophysics*, **30**(3), 228–236.
- 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.
- Wright, W. G., DiZio, P., & Lackner, J. R. (2006). Perceived self-motion in two visual contexts: Dissociable mechanisms underlie perception. *Journal of Vestibular Research*, **16**, 23–28.
- Young, L. R., Dichgans, J., Murphy, R., & Brandt, T. (1973). Interaction of Optokinetic and Vestibular Stimuli in Motion Perception. *Acta Oto-Laryngologica*, **76**(1), 24–31.
- Yuille, A. L., & Bülthoff, H. H. (1996). Bayesian decision theory and psychophysics. In D. C. Knill & W. Richards (Eds.), *Perception as Bayesian Inference* (pp. 123–161). Cambridge University Press.
- 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**, 159–171.
- Zupan, L., Merfeld, D., & 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.

8 Appendix

Acknowledgements

I would like to express my thankfulness to Professor Dr. Heinrich Bülthoff for giving me the great opportunity to conduct this research project at a fantastic research facility, and to Professor Dr. Rainer Guski for kindly offering to be my supervisor. Furthermore, I would like to thank Dr. Bernhard Riecke, Dr. Douglas Cunningham, Dr. Betty Mohler, Dr. Jenny Campos, and Dr. Daniel Berger for fruitful discussions, suggestions and comments on the manuscript, and to Dr. Markus von der Heyde and Franck Caniard for programming many experimental stimuli. I also like to thank all the gentle colleagues at the MPI for Biological Cybernetics for providing such a positive and enjoyable environment. Many thanks to the technical staff for all the incredible technical support, especially to Hans-Günther Nusseck and Michael Weyel.

Thanks to the Max Planck Society and the European Community [FET Proactive Initiative, IST-2001-39223 POEMS Project (Perceptually Oriented Ego-Motion Simulation)] for providing financial support for my research. Many thanks also to Dr. Pontus Larsson, Dr. Daniel Västfjäll, Alexander Våljamäe and Professor Mendel Kleiner at Chalmers University in Gothenburg for fruitful collaboration in the POEMS Project.

I have to thank all the participants of my experiments who patiently took part in many hour-long vection experiments, sitting in darkness and pressing buttons over and over again.

Finally, I'd like to thank my family and my beloved Mine Post for all the invaluable personal support during all this time.

Ehrenwörtliche Erklärung

Hiermit erkläre ich ehrenwörtlich, dass ich diese Dissertation selbständig und ohne unerlaubte fremde Hilfe angefertigt habe. Ich habe außer den im Quellen- und Literaturverzeichnis sowie in den Anmerkungen genannten Hilfsmittel keine weiteren benutzt. Die Arbeit wurde bisher weder im In- oder im Ausland in gleicher oder ähnlicher Form einer anderen Prüfungsbehörde vorgelegt.

Tübingen, im August 2007

(Jörg Schulte-Pelkum)

Bildungsgang von Jörg Schulte-Pelkum

18.05.1973: geb. in Tokyo, Japan

1979: Einschulung Uhlandschule in Hemsbach

1980: Umzug nach Tokyo, Besuch der Deutschen Schule Tokyo

1986: Umzug nach Bochum, Besuch des Goethe-Gymnasiums

1992: Abitur, anschl. Zivildienst

1993: Studium der Psychologie an der Ruhr-Uni-Bochum

2000: Abschluss Diplom

2001: sechsmonatiger Forschungsaufenthalt an der Carlton University, Ottawa, Kanada

2002: Doktorand am Max-Planck-Institut für biologische Kybernetik, Tübingen

2008: Promotion