Perception Coherence Zones in Vehicle Simulation

Ana Rita Valente Pais

Perception Coherence Zones in Vehicle Simulation

Ana Rita Valente Pais

ISBN 978-94-6186-155-9

Printed by Wöhrmann Print Service, Zutphen, The Netherlands.

Cover design by H. J. Damveld and A. R. Valente Pais.
Back cover pictures: Visual Motion Simulator and Cockpit Motion Facility (both, source: NASA).
Front cover pictures: Desdemona Simulator (source: TNO, Soesterberg) and Simona Simulator (source: A. R. Valente Pais).

Perception Coherence Zones in Vehicle Simulation

PROEFSCHRIFT

Ter verkrijging van de graad van doctor
aan de Technische Universiteit Delft,
op gezag van de Rector Magnificus prof.ir. K.Ch.A.M. Luyben,
voorzitter van het College voor Promoties,
in het openbaar te verdedigen
op maandag 27 mei 2013 om 10.00 uur

door

Ana Rita VALENTE PAIS

Ingenieur Luchtvaart en Ruimtevaart
geboren te Vila Franca de Xira, Portugal.

Dit proefschrift is goedgekeurd door de promotor:

Prof.dr.ir. M. Mulder

Copromotor:

Dr.ir. M.M. van Paassen

Samenstelling promotiecommissie:

Rector Magnificus,	voorzitter
Prof.dr.ir. M. Mulder,	Technische Universiteit Delft, promotor
Dr.ir. M.M. van Paassen,	Technische Universiteit Delft, copromotor
Prof.dr. J. Dankelman,	Technische Universiteit Delft
Prof.dr. J.E. Bos,	Vrije Universiteit Amsterdam
Prof.dr. H.H. Bülthoff,	Max Planck Institute for Biological Cybernetics, Duitsland
Dr. P.R. Grant, P. Eng.,	University of Toronto, Canada
Dr.ir. M. Wentink,	Desdemona, B.V.
Prof.dr.ir. J.A. Mulder,	Technische Universiteit Delft, reservelid

Dr.ir. M. Wentink heeft als begeleider in belangrijke mate aan de totstandkoming van het proefschrift bijgedragen.

Dit onderzoek was mogelijk gemaakt door een Toptalent-subsidie van de Nederlandse Organisatie voor Wetenschappelijk Onderzoek (NWO).

Summary

Perception Coherence Zones in Vehicle Simulation

Ana Rita Valente Pais

Motion simulators are used worldwide as pilot and driver training devices. They provide a relatively cheap, safe, accessible, and environmentally-friendly alternative to training in the real vehicle. For these simulators to be effective as training devices, the provided stimuli must allow subjects to recognize and react appropriately to different situations. Studying and monitoring the fidelity of the stimuli provided is therefore essential to guarantee the current safety, and increase the future safety of road and air travel.

One of the main fidelity aspects concerns the simulator motion system. The limited motion space of simulators impedes the exact reproduction of the vehicle motion one-to-one. To provide inertial motion stimuli while maintaining the simulator within its mechanical limits, motion filters are used. These filters, also known as motion cueing algorithms, introduce amplitude attenuation and phase distortion in the inertial motion feedback as compared to the visual motion available through the outside visual displays. This creates a discrepancy between the inertial and the visual stimuli. If this discrepancy, or mismatch, is large enough to be perceived by the subject in the simulator, then the fidelity of the simulation may be impaired.

Knowledge of how those mismatches between the visual and the inertial stimuli affect the human in the simulator is needed for the development of new motion cueing algorithms and simulator designs, and the optimization of motion filters for existing motion platforms. Moreover, a better understanding of human self-motion processes and especially of the

mechanisms involved in the perception of combined visual and inertial motion cues will lead to better and more complete fidelity criteria for the assessment of the quality of inertial motion simulation.

The main goal of this thesis is to contribute to the improvement of inertial motion fidelity in vehicle simulation. From this goal, three more specific objectives are defined.

Objective 1

The first objective is to study the process of motion cueing algorithm design and tuning, and evaluate the consequences of different cueing options, and hence different inertial stimulation, on subjects' perception, control behavior and performance.

Objective 2

The second objective is to focus in more detail on a specific component of simulation fidelity: the perception of combined visual and inertial stimuli. Here, the aim is to extend our knowledge of visual-inertial cue perception in vehicle simulation scenarios.

Objective 3

The third objective is to investigate the possibility of using the knowledge of combined perception of inertial and visual stimuli to derive perception-centered motion cueing guidelines that are independent of a particular motion cueing algorithm structure and motion platform.

To address the first objective, a new motion cueing algorithm for urban curve driving in a large, centrifuge-based motion simulator was designed and evaluated. Three cueing solutions were tested: the new solution, a classical filter normally used with Stewart platforms and a road rumble only algorithm.

To evaluate the effect of inertial cues on subjects' performance and control behavior, very specific performance goals would have had to be set and the type of control task would have had to be very well defined, leaving little room for subjects to choose their own control strategies. To simulate a task as close as possible to actual urban driving, the performance and control behavior constraints were not applied. As a result it was very difficult to compare the three cueing solutions based on performance and control behavior.

A perception analysis based on subjects' answers to questionnaires identified specific inertial cues that were important for either the acceptance or the rejection of specific cueing solutions, but it was not possible to measure how different cues contributed to the overall acceptance.

The process of designing, tuning and evaluating a motion cueing algorithm for an unconventional simulator, for which there are few guidelines and tested solutions, showed that

more platform independent, human-based, perception metrics and guidelines are needed.

The second objective was approached by measuring and evaluating the perception of combined visual and inertial stimuli using the concept of a coherence zone. Since the use of motion filters introduce both amplitude attenuation and phase distortion in the inertial stimulus as compared to the visual stimulus, two types of coherence zones were defined: amplitude coherence zones and phase coherence zones.

An amplitude coherence zone represents the range of inertial motion amplitudes that, although not being a perfect match with the visual motion amplitude, is still perceived by the subjects as being coherent. Similarly, a phase coherence zone represents the values of phase distortion between the inertial and visual stimuli that are still perceived as coherent.

Amplitude coherence zones can be defined by an upper and a lower threshold, or alternatively, by a Point of Mean Coherence (PMC) and a Coherence Zone Width (CZW). The upper threshold represents the highest inertial motion amplitude that is still perceived as coherent with a certain visual stimulus amplitude. The lower threshold represents the lowest inertial motion still perceived as coherent. The PMC is the inertial amplitude level exactly halfway between the upper and the lower threshold and the CZW is the difference between the upper and the lower threshold.

For phase coherence zones only one threshold was measured. Since generic high-pass filters introduce lead in the inertial motion with respect to the visual motion, only thresholds for leading inertial motion were measured. This measured phase-error threshold then represents the largest value of phase lead that can be applied to the inertial motion and is still perceived as coherent. Low-pass filters were not considered.

Several experiments were performed that aimed at investigating either amplitude or phase coherence zones. Amplitude coherence zones were measured for yaw rotation and sway translation. Yaw amplitude coherence zones were measured in two simulators: the Simona simulator and the Desdemona simulator. Phase coherence zones were measured for yaw and pitch rotations. The effects of stimulus amplitude and frequency were measured for both amplitude and phase coherence zones.

Results show that yaw and sway amplitude coherence zones were affected by both the visual stimulus amplitude and frequency. In general, for increasing visual stimulus amplitudes the PMC became larger, as did the CZW. For higher amplitudes of the visual stimulus the upper thresholds tended to become closer to or even lower than the one-to-one line (physical match) and this was identified by a “bending down” of the coherence zones, with the PMC values becoming smaller than the one-to-one line.

The effect of frequency was also similar for yaw and sway amplitude coherence zones. For higher frequencies, both the lower and upper thresholds decreased. The effect of frequency was related to the perception of “motion strength”. It was concluded that perceived motion strength is best explained as a weighted combination of perceived angular accelera-

tion and angular velocity, in the yaw case, and of perceived linear jerk and linear acceleration, in the sway case.

In the comparison of phase coherence zones in yaw and pitch, those were found to be similar and not affected by either the amplitude or the frequency of the stimuli. Subjects behaved like phase-error detectors rather than time-delay detectors and the measured phase-error threshold was approximately 19 deg with the inertial motion leading the visual motion.

In a next step, to approximate more realistic simulation scenarios where subjects are often required to perform flying or driving tasks, the effect of reduced subjects' attention on the perception of amplitude coherence zones was evaluated. Coherence zone measurements were obtained while subjects were performing a manual control boundary-avoidance task. The manual control task was performed based on visual stimuli only and in a different degree-of-freedom than the coherence zone measurements.

The addition of the active task did not significantly affect the perceived coherence of visual and inertial stimuli. A possible explanation is that the perceptual task and the active control task were not performed concurrently, but in sequence, i.e., offset in time. It might take a short period of time to make a decision regarding the perceived coherence, freeing the rest of the time to concentrate on the active task.

These results support the premise that in terms of perceptual fidelity, simulator tests can be performed passively and the results may be generalized for active, pilot-in-control situations. The same might not apply, however, for performance or behavioral fidelity.

As a final step in the study of coherence zones, and as a preparation for addressing the third objective, sway amplitude coherence zones were measured in three hexapod simulators: the Visual Motion Simulator (VMS), the Generic Flight Deck (GFD) and the Integration Flight Deck (IFD), at the NASA Langley Research Center, in Hampton, Virginia, USA. These three simulators had different motion and visual systems characteristics.

The differences between simulator configurations were captured by the differences in the measured upper thresholds of the coherence zone. The lower thresholds were not affected by the simulator configuration nor by the amplitude and frequency of the visual stimuli.

Despite the differences in the upper threshold values, the same trends were observed for the effects of frequency and amplitude in all three simulators. The same was observed for yaw coherence zones when comparing data obtained in the Simona and the Desdemona simulators.

This observation, that the exact values of the coherence zone may vary across simulators, but do maintain the same overall trends with respect to the visual stimulus amplitude and frequency, led to two conclusions. First, despite fundamental differences between simulators, such as the differences between the Simona and the Desdemona simulators (yaw) or the differences between the Visual Motion Simulator (VMS), the Generic Flight Deck (GFD) and the Integration Flight Deck (IFD) (sway), it is still acceptable to test and compare coherence zones measurements made in different apparatus.

Second, it is not only acceptable but actually desirable to do so. Using coherence zones, different simulations can be evaluated on the basis of how well the perception of the combined stimuli fits within the coherent range. This implies that simulators should not be evaluated and compared based on their mechanical performance alone, but on the basis of the combined presentation of visual and inertial stimuli. Coherence zones offer a metric to quantify the adequacy of this combined presentation of stimuli.

To achieve the third objective, the measured data on coherence zones was used to derive motion fidelity criteria and to propose a coherence-zone-based motion fidelity assessment method. The measured lower thresholds were transformed into minimum requirements for motion filter gains. The phase-error thresholds were used to obtain maximum allowed phase distortion criteria. These derived criteria were then compared to established criteria available in literature.

The coherence zones criteria were shown in a modified Sinacori plot. In a Sinacori plot, motion criteria are represented in terms of acceptable motion gain and phase distortion at the frequency of 1 rad/s. Since coherence zones were measured at different frequencies and amplitudes, it was possible to expand that representation to different frequencies and amplitudes.

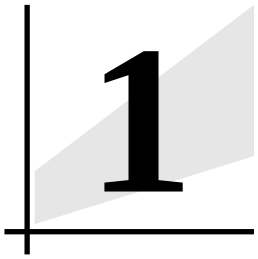
The use of coherence zones as a motion fidelity assessment method relies on the assumption that the perceived coherence between visual and inertial stimuli is indicative of "good" motion. Similarly to other assessment criteria, such as those proposed by Sinacori, Schroeder, and Advani and Hosman, coherence zones only provide a measure for the adequacy of the provided inertial motion with respect to a desired ideal motion. This desired motion corresponds to the simulated vehicle motion and in this thesis, it is assumed to be equivalent to the visual motion. All these assessment methods require vehicle model fidelity and are thus, on their own, not sufficient to guarantee inertial motion fidelity. Nevertheless, they do provide much needed guidance in the design and tuning of motion cueing algorithms.

The proposed coherence zone assessment method and criteria provide three important additions to the already available criteria. First, it presents not only criteria for desirable motion stimuli, but also offers a systematic, objective, human-perception-based method to measure the limits of the criteria.

Second, the coherence zones method and criteria add a third and fourth dimension to the Sinacori plot: frequency and amplitude. By doing so, the coherence zone criteria do not depend on a specific motion filter structure.

Third, by offering a measurement method and allowing different frequencies and amplitudes to be chosen, the coherence zones method can provide simulator-based, task-specific criteria. However, coherence zones as a metric, that is, as a measure of the perceived coherence of the inertial feedback provided, is platform and task independent.

Contents



Introduction

1.1 Motion simulators

On the morning of December 17 of 1903, north of Big Kill Devil Hill, the first powered heavier-than-air flight traveled approximately 36.5 meters in 12 seconds. More than one century later, that first historic flight covered a distance that is still impossible to accomplish with most of the flight simulators of today.

Flight simulators have a limited motion space. Their usefulness, however, lies precisely on that fact. A ground based machine, controlled by layers of software end electronics that keep it from reaching its mechanical limits, provides a safe and easy to access training environment. Simulators allow pilot training under many different scenarios, guaranteeing repeatability of all conditions. Compared to training in a real aircraft, simulators are also a cheaper and more environmentally friendly solution.

Today, the advancement of simulator technology has allowed for a widespread use of simulators as pilot training devices. Simulators are used, alongside with actual flying hours, for the training of commercial pilots. Proficiency checks and type ratings are increasingly relying on simulator training alone and with the exception of maneuvers related to airplane upset prevention and recovery training (Advani et al., 2010b), most flight task training can be performed in simulators (International Civil Aviation Organization, 2009).

Most training flight simulators currently in use consist of a cabin placed on an hexapod, also known as a Stewart platform. Curiously, the hexapod was first developed by Gough (Gough, 1957) for research on car tires.

The automotive industry also relies heavily on simulators. The first driving simulator studies on driving ability and causes of accidents date back to the 1910s and 1920s (Wachtel, 1995). Currently, many car manufactures invest in motion platforms that allow them to research, develop and test new products (Grant et al., 2001; Dagdelen et al., 2004). Road vehicle simulators are also used for research into driver behavior (Godthelp et al., 1984; Van Winsum and Godthelp, 1996; Boer et al., 2000; Reymond et al., 2001; Greenberg et al., 2003; Brünger-Koch et al., 2006) and road safety studies, such as mobile phone usage (Horberry et al., 2006). Recently, car driving simulators are also being used for the training of new drivers.

Motion simulators have also been used for space related research since the 1950s and 1960s. Early centrifuge simulators (Clark and Hardy, 1959) and rotating rooms (Graybiel et al., 1960; Guedry et al., 1962) were used to select and train astronauts and study the physiological impact of space flight. Recently, motion simulators have also become popular as tools for Earth-bound human motricity research (Advani et al., 2010a).

The entertainment industry has been a great catalyst of simulation technology and it is responsible for making motion simulators available to the masses. Motion simulators are now a common sight in innumerous attraction parks and even museums.

Figure 1.1: "Star Tours: The Adventures Continue" ride at Disneyland Park in Anaheim, California, USA, featuring a hydraulic 6 DOF motion platform and a Starspeeder mockup cabin that seats 40 people. (Courtesy of Disney, © Disney/Lucasfilm Ltd, STAR WARS © 2012 Lucasfilm Ltd & TM.)

With the generalized use of motion simulators some concern has arisen about the quality of the stimuli provided. When simulators are used for behavioral studies it is important to elicit similar behavior in the simulator as in the real situation. Especially when the main goal is to train subjects to react appropriately to different situations, the fidelity of the stimuli provided is crucial. In flight simulators, to train pilots to perform a certain task, it is important that all the necessary cues for that task are available and resemble those in the real vehicle.

For this reason, studying and monitoring the fidelity of motion simulators, and in particular of those used as pilot or driver training tools, is essential to guarantee the current, and increase the future safety of road and air travel.

1.2 Vehicle simulators and motion fidelity

Many aspects affect the realism of a vehicle simulation: cockpit and instruments, the control interfaces, the vehicle model, the visual system, the motion base dynamics and the actual motion provided to the subjects.

In flight simulation, considering that simulator pilot training is a standard practice, the evaluation of the quality of simulation has been an important discussion point among the scientific community, flight simulator manufacturers and flight simulator users. Different metrics have been used to assess, standardize and categorize flight simulators and the regu-

lating entities have made a considerable effort to define the desirable characteristics of the separate systems involved (International Civil Aviation Organization, 2009).

One of the main fidelity concerns is the simulator motion system. The limited motion space of simulators impedes the exact reproduction of the aircraft motion one-to-one. Therefore, motion filters are used to translate vehicle motion into simulator motion. These filters, also known as motion cueing algorithms, are used to the emulate the aircraft motion while preventing the simulator from reaching its mechanical limits.

Typical motion filters used in Stewart platforms are high-pass filters, which attenuate the amplitude of the vehicle motion at low frequencies, but also introduce phase distortion (Reid and Nahon, 1985, 1986a,b). This causes a difference between the motion of the vehicle and that of the simulator, which might have an impact on the pilot's perception of motion and control behavior.

The effect of motion filters on pilot's perception of motion and control behavior in the simulator has been extensively studied. Although the results are often platform and task dependent, Schroeder (1999) provides a good summary of many experiments and condenses their results into fidelity criteria, adapted from Sinacori (1977).

These fidelity criteria aim at guiding the process of designing and tuning the motion cueing algorithms to obtain the highest fidelity of motion possible. The criteria defined by Sinacori, and later revised by Schroeder, are based on a so-called Sinacori plot. The Sinacori plot shows the amplitude attenuation versus the phase distortion introduced by motion filters at the frequency of 1 rad/s. Figure 1.2 shows an example of such a plot for angular motion.

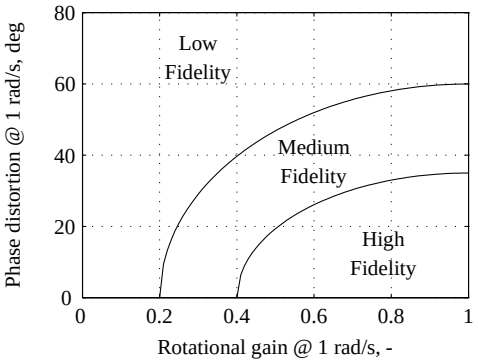


Figure 1.2: Example of a Sinacori plot for angular motion. Adapted by Schroeder (1999).

Ideally, a low attenuation, or high amplitude, and a low phase distortion should be achieved. The fidelity criteria are defined by stipulating boundaries for maximum allowed amplitude attenuation and phase distortion. Three fidelity regions are defined: low, medium and high, with the high fidelity region corresponding to the higher gains and lower phase distortions.

Despite the evidence that motion filters affect pilot's perception and control behavior (Samji and Reid, 1992; Schroeder, 1996, 1999; Grant et al., 2006; Groen et al., 2007; Ellerbroek et al., 2008), the performance of the motion cueing algorithm has not always been

taken into account in the regulations for training devices. Most requirements for the simulators' motion systems used to be limited to the mechanical performance of the motion base alone. However, in the last decade, considerable attention has been given to the fidelity of the inertial motion during actual simulation. As a consequence, the motion requirements now also include the need to present sufficient and adequate inertial cues. The adequacy of inertial cues is evaluated with tests that include not only the motion base mechanics but also the motion cueing software (Joint Aviation Authorities, 2003; Royal Aeronautical Society, 2005; International Civil Aviation Organization, 2009).

The preoccupation with the use of motion in vehicle simulation can also be seen in road vehicle simulation, as some of the most recently built simulators, used for research on car driving, are larger than most research flight simulators. Besides a hexapod motion base, these simulators also include additional translational motion systems, such as lateral rails (Dagdelen et al., 2004; Feenstra et al., 2007; "Mercedes-Benz", 2011) or XY tables ("Toyota", 2007; "NADS-1", 2010; "VTI's", 2012), and added yaw rotation, through gimbals (Feenstra et al., 2007) or rotating platforms ("Toyota", 2007; "NADS-1", 2010).

Nevertheless, despite all efforts to improve the motion fidelity in vehicle simulators, simulator motion will never fully equal the real vehicle motion. There will always be differences and knowledge of how these differences affect the human in the simulator is mandatory for the definition of fidelity criteria, the development of new motion cueing techniques and simulator designs, and the optimization of motion filters for existing motion platforms.

1.3 Perception, Behavior and Performance

For the purpose of training pilots, or drivers, subjects must learn to recognize the important cues, and use them to control the vehicle to attain a desired level safety, comfort and performance. This means there are at least three levels at which we may evaluate the motion in a simulator: perception, behavior and performance.

From a performance point of view, motion in the simulator needs to allow for a similar performance to that achieved in the real vehicle. However, humans are able to adapt to their environment, and for a given task the same level of performance may be attained under different motion conditions.

Despite achieving the same end result, human controllers might have done so by adapting their control strategies. This means that although their performance was the same, their control behavior was different. This indicates that the performance approach alone is not sufficient to understand and evaluate the effect of simulator motion on human controllers and that behavioral metrics should also be considered.

From a behavioral point a view, the simulator motion does not need to equal vehicle motion but it needs to elicit control behavior similar to that observed in the real vehicle. Many studies have focused on the modeling and identification of pilot control behavior in an attempt to capture different control strategies with observed changes in the pilot model (McRuer et al., 1965; Stapleford et al., 1969; Jex et al., 1978; Van der Vaart, 1992; Hosman, 1996; Zaal, 2011; Pool, 2012).

Also in car driving, there has been a considerable effort to evaluate the impact of simulator motion on driver steering and braking behavior (Repa et al., 1982; Boer et al., 2000; Reymond et al., 2001; Siegler et al., 2001; Brünger-Koch et al., 2006). This field of research has contributed to the understanding of how motion and visual cues are used by human controllers and which cues are considered essential for specific manual control tasks. However, behavioral studies, and in particular work on pilot model identification, is currently limited to scenarios where it is possible to model the human controller as a linear system. This requires the use of very specific manual control tasks, with very well defined motion and visual stimuli.

To extend simulator studies to a wider range of scenarios and gather a more complete insight into the effect of motion on humans, it is useful to consider also a perceptual approach.

From a perceptual point of view, the simulator motion does not need to equal vehicle motion fully, but only to the extent to which humans can perceive the difference between both. Many perception studies have, therefore, focused on measuring perception thresholds (Stewart, 1971; Hosman and van der Vaart, 1978, 1980; Benson et al., 1986, 1989; Heerspink et al., 2005).

Measuring human sensitivity to motion in different circumstances helps define the range of possible motion stimuli that can and should be used in simulation. Moreover, for simulation scenarios with above-threshold stimuli, it is also important to understand how motion is perceived and how the combined perception of different types of stimuli, such as visual and inertial, result in perceiving the simulator motion as equivalent to that of the real vehicle, or not.

Considering that humans will base their control actions on the perceived stimuli and that these actions will result in a certain performance level, one might say that, if from a perceptual point of view the stimuli presented to the human are indistinguishable from that in the real vehicle, then the control actions and the resulting performance will also be equivalent to those in the real vehicle. This implies a straightforward, causal relationship between perception, behavior and performance. This would mean that if a simulation has perceptual fidelity, also behavioral and performance fidelity are guaranteed. However, it has not been proven that behavioral fidelity depends exclusively on perceptual fidelity. In fact, recent work has showed that human controllers perceive mismatches between visual and inertial stimuli before these mismatches will start to affect their manual control behavior

(Beckers et al., 2012). This means that having perceptual fidelity might not be necessary to achieve behavioral fidelity.

In this light, it might seem that perceptual fidelity, as a means to achieve behavioral fidelity, is not necessary and that behavioral studies alone would be sufficient to improve motion simulators' usability as training devices. However, independently of the relationship between perception and behavior, perceptual fidelity is fundamental to guarantee immersion and acceptation of the virtual environment, eliminate the possibility of simulator sickness, train subjects for conditions where disorientation might occur, and perhaps the most important reason, to help human controllers recognize and correctly react to important cues during vehicle failure or accident situations.

All in all, it seems fair to say that all three approaches, perceptual, behavioral and performance, are valid approaches and probably all three are necessary for the continuous improvement of motion in vehicle simulation. This thesis focuses on the perceptual approach. This approach starts by understanding the basics of human motion perception and human motion sensors in both real and simulated environments.

1.4 Human self-motion perception

1.4.1 Inertial perception of motion

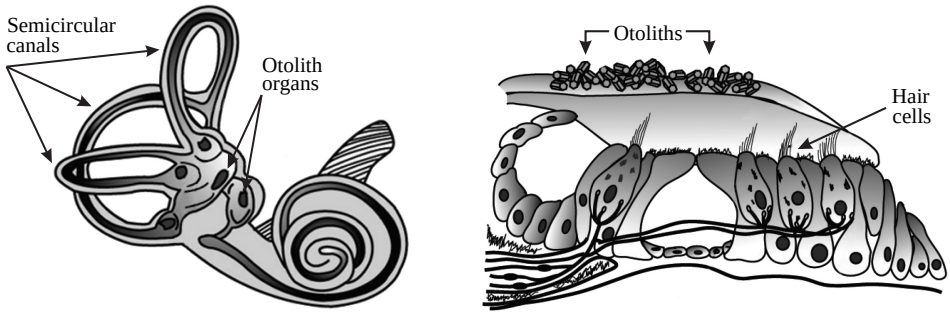
Humans have different systems which are used for the perception of inertial motion: proprioceptive, tactile and vestibular.

The proprioceptive system provides information on muscle contraction in response to external forces and the relative position of our limbs and other body parts with respect to each other. Acceleration externally applied to our bodies, due to vehicle acceleration for example, will have an effect on the necessary forces to keep our body parts in a certain position. This information can be used to make an estimate of one's self-motion.

The tactile system may also be used to derive information about body motion. Changes in skin pressure on the back, buttocks and legs, while sitting in a moving vehicle may indicate the direction and magnitude of the vehicle's acceleration.

The vestibular system is often seen as the main system responsible for sensing self-motion. It is located in the inner ear and consists of the utricle and saccule, which together are called the otolith organs, and the semi-circular canals (SCC) (Howard, 1968). The SCC detect angular motion, whereas the otolith organs are sensitive to specific forces and thus affected by linear acceleration and gravity.

The SCC are three orthogonally oriented circular canals which allow for detection of angular motion in three rotational axis. The canals contain a fluid called endolymph. When the head rotates, the fluid initially remains stationary due to inertia, while the canal moves



(a) Semi-circular canals. **(b)** Otolith organ.

Figure 1.3: Illustration of the vestibular system showing the three semi-circular canals and the otolith organs. (Courtesy of NASA.)

with the head. This causes a membrane, also named cupula, inside the canal to deflect, indicating to the brain that there is angular acceleration (Howard, 1968). If the head continues to rotate at a constant velocity, the endolymph will eventually start to rotate at the same velocity as the canal, and the canal response to angular motion decays. So, the SCC respond to angular acceleration and not angular velocity.

The otolith organs contain a gelatinous layer on top of which there are calcium carbonate crystals known as otoliths. There are two of these structures in each otolith organ. One of these is oriented in the vertical plane and the other in the horizontal plane (Howard, 1968).

When the head is accelerated linearly, the different inertias of the otoliths and the gelatinous layer cause the hair cells inside the gelatinous layer to bend, indicating to the brain that the head is accelerating. If a constant velocity is achieved, the sensation of linear motion stops.

The otolith organs sense specific force and in a constant gravity field this can be used to form a percept of the head orientation relative to the vertical. However, using the otolith organs alone it is not possible to distinguish between linear acceleration and the acceleration of gravity. This differentiation is accomplished by using other sensorial information, such as rotational information from the SCC and visual cues.

1.4.2 Visual perception of motion

The human eye is composed of the cornea, the lens and the retina. The cornea and the lens allow focusing on objects at different distances. The retina is a layer inside of the eye that contains receptor cells called cones and rods. As light reflected from objects enters the eye through the cornea and the lens and reaches the retina, the cones and rods transform that

information into electrical pulses to the brain (Westheimer, 1968; Hood and Finkelstein, 1968).

Images moving across the retina may trigger a perception of movement. However, following a moving object by rotating the head will cause the image of that object to remain fixed with respect to the retina, but the object will still be perceived as moving.

Visual perception of motion depends on other factors than just the eye and the movement of images across the retina. The neuronal paths that transmit the information gathered by the eye receptors and the further processing of this information in the brain, have been shown to have a great importance in visual motion perception (Goldstein, 2002).

The visual system provides information regarding the position and orientation with respect to the outside world, and in addition it also provides velocity information (Van de Grind, 1988; Van Boxtel et al., 2006). For the purpose of studying human self-motion perception in vehicle simulation since the visual perception of motion is done in velocity space (Loose and Probst, 2001), the visual system is often considered to be a velocity sensor (Van der Steen, 1998; Bos et al., 2002, 2004). The visual perception of self-motion is then described by the perception of velocity amplitude and direction.

One important aspect of visual motion perception in vehicle simulation is vection. Vection is the induction of self-motion perception through visual stimulation (Gurnee, 1931; Brandt et al., 1973). A well-know example of vection is the moving train illusion. This illusion occurs when one is sitting in a stationary train and the train on the next track starts moving. The movement of the other train gives the illusion of self-motion in the opposite direction.

In a simulator, the visual system projects moving images representing the outside world which induce the perception of a moving self through a stationary world. Different aspects of the visual system display and the visual scene content have been shown to affect the onset and sustenance of vection, as well as the perceived intensity of the visually-induced self-motion. For a review please refer to Riecke (2010).

The process of vection is often modeled by a low-pass filter, representing the fact that the perception of self-motion after an onset in visual motion stimulation is a slow process (Brandt et al., 1973; Melcher and Henn, 1981; Mergner and Becker, 1990). However, it has been shown that the build-up of vection is faster in the absence of conflicting vestibular stimuli (Young et al., 1973; Berthoz et al., 1975), as for example, when the onset in visual motion is presented simultaneously with a short duration inertial stimulus (Wong and Frost, 1981).

1.5 Motion perception in vehicle simulation

1.5.1 Perception thresholds and perceived motion strength

Human motion perception is obviously a topic that interests fields of research other than just vehicle simulation. A great deal of the existing knowledge was gathered in fields such as medicine and physiology. Examples are the work by Van Egmond et al. (1949) on the mechanics of the SCC, and later the studies of Fernandez and Goldberg (1971) on modeling the SCC of the squirrel monkey. The models developed by Fernandez and Goldberg (1971) formed the basis for one of the models of the human SCC dynamics (Hosman and van der Vaart, 1978, 1980) currently used by the flight simulation community.

Other perception studies relevant for vehicle simulation focus on the limits of human motion perception, the so-called motion perception thresholds. Particularly useful are studies that analyze the effect of stimulus frequency on sensory motion thresholds (Stewart, 1971; Gundry, 1977; Hosman and van der Vaart, 1978; Benson et al., 1986, 1989; Soyka et al., 2011, 2012) or the effect of expectation (Mesland, 1998), studies on motion thresholds measured during concurrent stimulation in other degrees-of-freedom (DOFs) or in the presence of visual cues (Huang and Young, 1981; Zaichik et al., 1999; Rodchenko et al., 2000; Groen et al., 2004; Valente Pais et al., 2006), and studies on the effect on thresholds of performing a manual control task (Roark and Junker, 1978; Hosman and van der Vaart, 1980; Samji and Reid, 1992).

In a vehicle simulation context such studies often aim at solving a specific problem directly related to motion cueing algorithms. For example, the experiments performed by Groen and Bles (2004) determined the pitch motion perception threshold during linear fore-and-aft visual motion stimulation. The threshold value they measured confirmed the empirically determined maximum pitch velocity used during tilt coordination maneuvers.

Although very useful to delimit the ranges of usable motion, these studies offer very little guidance on the use of motion at supra threshold levels. For this reason, it is necessary to also investigate motion perception at higher levels of stimulation.

An example of this type of work is the study by Grant and Haycock (2006) on the relationship between linear acceleration and jerk and the perception of motion strength. Other examples of perception studies with supra threshold stimulation are Just Noticeable Difference (JND) measurements (Mallery et al., 2010; Naseri and Grant, 2011; Dos Santos Buinhas et al., 2013) and sensation estimation (Elsner, 1971) studies.

1.5.2 Visual and inertial stimulation

The physical limitations of the simulator, and the resulting use of motion filters, limit the amplitude and frequency of the motion cues provided in the simulator. Unlike the motion system, the visual system can show any displacement or attitude of the simulated aircraft. This often causes a discrepancy between the inertial motion and the visual motion stimuli.

Understanding the perception of combined visual and inertial stimuli has been, therefore, crucial for the further development of vehicle simulation.

Much of the early research on perception of visual and inertial stimulation was performed in a rotary chair placed inside an also rotating circular screen (Gurnee, 1931; Brandt et al., 1973; Young et al., 1973; Melcher and Henn, 1981; Wertheim and Bles, 1984; Probst et al., 1985; Mergner and Becker, 1990) or other chair-based apparatus (Pavard and Berthoz, 1977; Mesland, 1998; Groen et al., 1999). Later, with the progress of simulator technology and the growing interest of the vehicle simulation community in this topic, other studies appeared performed in motion simulators. In these studies the focus was less on understanding the underlying principles of human-self motion and more directed towards determining which combinations of visual and inertial stimuli resulted in the more comfortable and realistic scenarios (Casali and Wierwille, 1980; Huang and Young, 1981; Schroeder, 1996, 1999; Zaichik et al., 1999; Rodchenko et al., 2000; Groen et al., 2001, 2005; Fortmüller and Meywerk, 2005; Grant and Haycock, 2006; Grant et al., 2006; Groen et al., 2007; Ellerbroek et al., 2008; Fortmüller et al., 2008; Correia Grácio et al., 2010).

A preferred approach to represent the human percept among engineers is to construct mathematical models. Many models have been developed with different levels of complexity varying from models of SCC dynamics (Hosman and van der Vaart, 1978, 1980) to the more complex models that include thresholds (Borah et al., 1988; Greig, 1988; Kamphuis, 1994; Soyka et al., 2011, 2012), visual time delays and vection (Bos and Bles, 2002; Groen et al., 2004), visual and vestibular cues interaction (Zacharias and Young, 1981; Telban and Cardullo, 2001; Bos and Bles, 2002; Reymond et al., 2002; Groen et al., 2004; Naseri et al., 2008) and conflict detection (Zacharias and Young, 1981; Hosman and van der Steen, 1993; Telban and Cardullo, 2001). Outputs of the models include detection of inertial cues, judgment of motion strength, estimation of own velocity, detection of conflicts between visual and inertial cues or an overall sense of self-motion perception expressed in terms of velocities and orientation with respect to gravity.

Most models were successful in explaining or replicating collected experimental data. However, especially for models dealing with combined perception of visual and inertial stimuli, there are none that allow for widespread use. The application of current models is limited to the tasks and scenarios for which they were developed. Nevertheless, summing up all the data collected to develop these models and analyzing the mathematical laws they establish, much can be understood about the integration of visual and inertial cues and the perception of self-motion resulting from this process.

1.5.3 Perception coherence zones

A particular type of research on combined perception of visual and inertial cues focuses on determining the boundaries of accepted visual-motion stimulation. An example is the

work of Van der Steen (1998). Van der Steen measured the mismatch between visual and inertial cues in terms of amplitude. His premise was that during self-motion in the real world, the visual and inertial stimuli are always part of a consistent, coherent movement. In a simulator, the mismatch between the visual and inertial stimuli might reach a point where the motion coherence is lost, or, in other words, a mismatch between the two cues is perceived by the subject.

He used the concept of coherence zone to designate the range of inertial motion amplitudes that, although not being a one-to-one match with the visual motion amplitude, were still considered by the subjects as being part of a coherent movement. Van der Steen considered that coherence was maintained as long as subjects perceived the outside world as stationary. Once the outside world was perceived as moving, coherence was deemed lost.

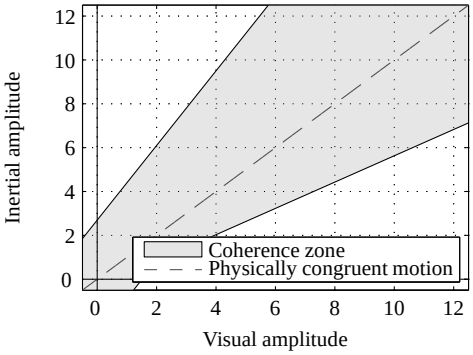


Figure 1.4: Hypothetical amplitude coherence zone.

Van der Steen measured amplitude coherence zones for yaw, roll, swing (sway and roll combined), surge and heave for different visual motion amplitudes and frequencies. Although much relevant data were collected, the amplitude and frequency levels were rather low. Amplitudes ranged from 0 to 12 deg/s for roll and swing, 3 to 18 deg/s for yaw and 0.5 m/s for the linear DOFs. The frequencies varied between 1 and 2 rad/s. For a more direct application to vehicle simulation higher amplitudes and a broader range of frequencies should be tested.

The concept of coherence zone can also be applied to phase differences between visual and inertial stimuli. For example, Grant and Lee (2007) determined the minimum phase lead of the pitch inertial cues relative to the pitch visual cues that could be detected by the subjects in the simulator. They found that the average phase-error detection threshold was 57 deg. This result suggests that for an inertial motion phase lead inferior to the determined threshold, visual and inertial cues are still considered to be coherent. Therefore, the measured threshold can be used to define a phase coherence zone.

1.5.4 Motion assessment and requirements

Despite the large body of data available from many decades of human motion perception research, the available requirements for the combined presentation of visual and inertial stimuli in training simulators is almost non existent.

Generally all requirements are given separately for each system: sound cues, motion cues, visual cues, control feel, etc. Nevertheless, it is possible to derive some constraints for the simulation of combined visual and inertial cues. For example, in terms of timing, the combined fidelity of the motion and visual systems can be assessed based on the transport delay between pilot control input and the motion and visual systems response (Joint Aviation Authorities, 2003; Royal Aeronautical Society, 2005; International Civil Aviation Organization, 2009). By limiting the maximum allowed transport delay for all systems, the appropriate timing between motion and visual cues is also guaranteed.

In terms of amplitude and phase, the matching between inertial and visual cues can be evaluated through the motion filter settings and motion base performance.

According to the International Civil Aviation Organization (ICAO)'s Manual of Criteria for the Qualification of Flight Simulation Training Devices (International Civil Aviation Organization, 2009) the assessment of the combination of motion platform and motion filter can be done by plotting the amplitude attenuation versus the phase distortion, similarly to what is done in a Sinacori plot, but extending the representation to a wider range of frequencies (Advani and Hosman, 2006). This assessment method concerns motion only and it does not include non-linear processes. Considering linear processes only and assuming that the visual motion stimuli, in terms of amplitude, frequency and phase, is equivalent to the motion in the real aircraft, then any difference between the real aircraft motion and the inertial motion cues in the simulator, are also the differences observed between the visual and the motion cues in the simulator. Hence, guaranteeing a desirable simulator motion performance should also lead to a good match between the visual and motion cues.

In the ICAO's Manual (Attachment F) (International Civil Aviation Organization, 2009), it is mentioned that plotting the results in such a manner allows direct comparison of simulators. However, it is also mentioned that the definition of the fidelity boundaries is still an ongoing activity. From this one might conclude that despite the many available studies on human motion perception, to improve the requirements and guidelines for motion cueing, more research is needed. Especially studies that focus on the perception of combined visual and motion cues may help in the definition of boundaries for inertial motion fidelity.

1.6 Research goals

Motion perception studies, whether in a simulation environment or not, have greatly contributed to our understanding of how humans perceive visual and inertial stimuli in a simulator. This knowledge has been used in the development and improvement of motion cueing algorithms currently in use. Nevertheless, concern about the quality of inertial motion dur-

ing vehicle simulation has encouraged the development of new and more complete simulator assessment and comparison methods.

The further development of these methods demands an even better knowledge of human self-motion perception. Moreover, the emergence of new simulator designs calls for new motion cueing techniques for which there are fewer tested and proven solutions. Here as well, a better understanding of human motion perception may help the design and tuning of novel cueing algorithms.

The main goal of this thesis is to contribute to the improvement of inertial motion fidelity in vehicle simulation. From this goal, three more specific objectives are defined.

Objective 1

The first objective is to study the process of motion cueing algorithm design and tuning, and evaluate the consequences of different cueing options, and hence different inertial stimulation, on subjects' perception, control behavior and performance.

Objective 2

The second objective is to focus in more detail on a specific component of simulation fidelity: the perception of combined visual and inertial stimuli. Here, the aim is to extend our knowledge of visual-inertial cue perception in vehicle simulation scenarios.

Objective 3

The third objective is to investigate the possibility of using the knowledge of combined perception of inertial and visual stimuli to derive perception-centered motion cueing guidelines that are independent of a particular motion cueing algorithm structure and motion platform.

1.7 Approach

The three objectives defined above are dealt with in three parts, each with a different level of detail. The first part approaches motion cueing and the role of inertial motion in vehicle simulation from a generic point of view. The second focuses on a more specific component of motion simulation, the perception of combined inertial and visual cues, and analyzes it in more detail. In the third part, the specific knowledge gathered on visual-inertial motion perception is generalized into motion cueing guidelines.

In the first part, the process of design and tuning of motion cueing algorithms is studied by developing a new cueing solution for a centrifuge-based, large motion-space simulator.

For such an unconventional simulator there are fewer guidelines and proven cueing solutions than for a Stewart platform, a fact that exacerbates and highlights the challenges of motion cueing design.

The inertial stimulation attained with the new cueing solution is then evaluated based on perceptual, behavioral and performance metrics. Although there is enough evidence that determining the impact of inertial motion on the fidelity of vehicle simulation is not a trivial problem, by using a large motion space simulator, it may be possible to accentuate the differences between a new algorithm that takes advantage of the large motion space and more conventional approaches that use much less space.

In the second part, the perception of combined visual and inertial stimuli is measured and evaluated using the concept of coherence zone. As motion filters introduce both amplitude attenuation and phase distortion on the inertial motion with respect to the visual motion, both amplitude coherence zones and phase coherence zones are considered.

Compared to subjective motion ratings, coherence zones provide a relatively simple and objective measure of perceived agreement between visual and inertial cues. Moreover, since few assumptions are needed regarding the perceptual mechanisms involved, coherence zone measurements are expected to be valid for more situations than only those in which the measurements were performed.

Amplitude coherence zone measurements are performed mostly in yaw. The large yaw motions present during the car driving scenario from the first part have inspired the choice to investigate yaw coherence zones. Moreover, by choosing yaw, the inertial stimulation is restricted to the SCC, that is, because there are no changes in the orientation with respect to gravity, the stimulation of the otolith organs is negligible. Limiting the stimulation to one vestibular sensor simplifies the analysis of the results and may allow for a better understanding of the influence of one vestibular sensor, in this case the SCC, in the higher level process of perception and integration of multiple cues.

Coherence zones are measured using stimuli with amplitudes and frequencies common to vehicle simulation. The effect of amplitude and frequency on both amplitude and phase coherence zones in yaw and pitch is studied. Afterwards, to approach simulation scenarios where the subject has many tasks to perform, the effect of decreased subject attention on yaw amplitude coherence zones is investigated. Hereafter, amplitude coherence zones in a linear degree-of-freedom (DOF) are addressed. The effect of stimulus frequency and amplitude on sway coherence zones is investigated with measurements in three different simulators. As a preparation for the third part, this step also investigates the possibility of comparing different simulators using coherence zones as a metric.

In the third part, the data on coherence zones collected in all the experimental trials are used to derive motion fidelity criteria. These criteria may be used as guidelines for the

development of new motion cueing algorithms or as metrics for the comparison of different cueing solutions in different simulators.

1.8 Assumptions

This thesis focuses on the perception of combined visual and inertial cues. As mentioned above, the perception of inertial motion relies on different sensors such as the somatosensory system (proprioception and tactile stimuli) and the vestibular system. In this thesis, the relative weight of the sensory signals from both systems on the overall motion percept is not investigated.

Furthermore, mathematical models of the dynamics of the SCC and otolith organs are used to represent the inertial motion sensors as a whole. In these cases, the contribution of the somatosensory system is neglected (Walsh, 1961; Cheung et al., 1989).

In the approach followed in this thesis, the perception of combined visual and inertial stimuli is related to vehicle motion fidelity. As explained before, the focus lies on perceptual fidelity, and not on behavioral or performance fidelity. The motion cueing guidelines that are derived from the experimental findings are perception-based only. Although it is recognized that for a complete understanding of the role of inertial motion on vehicle simulation, also behavioral and performance aspects have to be considered, those are not investigated in this thesis.

Regarding perceptual fidelity, throughout this thesis it is assumed that the visual cues represent the motion of the real vehicle and that any perceived difference between the inertial motion and the visual motion results in a perceived difference between the real vehicle motion and the simulated motion. If no difference is perceived between inertial and visual cues, then there is no perceived difference between the simulated and the real vehicle, and the simulation can be said to be of high quality.

When making this assumption, two important aspects of the simulation are overlooked. The first is the vehicle model. The vehicle motion provided to subjects through the visuals does not always perfectly match the real vehicle motion. The visual motion corresponds to the response of the vehicle model to the control inputs of the pilot. The accuracy of the vehicle model used will determine the quality of the motion shown in the visuals and supplied to the motion cueing algorithm. The second aspect concerns the visual system characteristics. It has been shown that specific aspects of the visual system, such as field-of-view and collimation (Chung et al., 2003), texture (Dearing et al., 2001) and scene content (Sweet and Kaiser, 2005) have an effect on the visually perceived motion. Other aspects, such as focal distance, update rate, contrast and resolution might also influence the perceived visual stimulus.

In other words, the assumption made is that perceived coherence implies motion fidelity, when in fact the fidelity of other simulation components is also necessary. Nevertheless, although coherence does not necessarily imply fidelity, a lack of perceived coherence between the visual and the inertial motion is taken to represent impaired realism of the simulation.

1.9 Outline

Most chapters in this thesis are based on scientific publications that were written independently from each other and can, therefore, be read separately. The publications on which each chapter is based are indicated on the chapter title page.

This thesis is divided in three parts. Figure 1.5 shows a schematic representation of the thesis structure. Each part, I to III, addresses objectives 1 to 3, respectively. The circular representation of the thesis structure highlights the relationship between the three parts. First, in Part I, the study of the process of motion cueing. Then, in Part II, the focus on a specific component of motion simulation: the perception of combined visual and inertial cues. And finally, in Part III, returning to motion cueing by applying the knowledge gathered in Part II to derive motion cueing guidelines.

In Figure 1.5, the dashed arrow between the third and the first parts represents the fact that the developed guidelines have not been fully tested in a motion cueing design process. Doing so would allow us to fully close the circle of Cueing, Perception and Cueing again.

The first part comprises just Chapter 2. In the second part, Chapters 3 to 8 focus on better understanding the combined perception of inertial and visual motion cues, using the concept of coherence zone and Chapter 8 provides a summary of the results and some conclusions. The third part consists of Chapter 9. Below, the contents of each chapter are described in somewhat more detail.

Chapter 2 describes the design and implementation of a car driving motion cueing algorithm for the Desdemona simulator. The developed algorithm is compared to a classical motion cueing algorithm and a road rumble only condition. All three solutions are evaluated using subjective motion ratings and performance metrics.

Chapter 3 presents two experiments. The first experiment uses “Van der Steen-like” motion stimuli, step-like acceleration signals, and measuring method, an adaptive one-up-one-down staircase algorithm, to measure yaw amplitude perception coherence zones. In this experiment, the work of Van der Steen is extended to higher stimulus amplitudes. The second experiment tests a new method for the measurement of perception coherence zones. This new method, named the self-tuning method, aims at a faster and more precise collection of data.

Chapter 4 addresses the effect of frequency on amplitude perception coherence zones. An experiment performed in the Simona Research Simulator is described. Yaw amplitude perception coherence zones are measured using the self-tuning method and employing si-

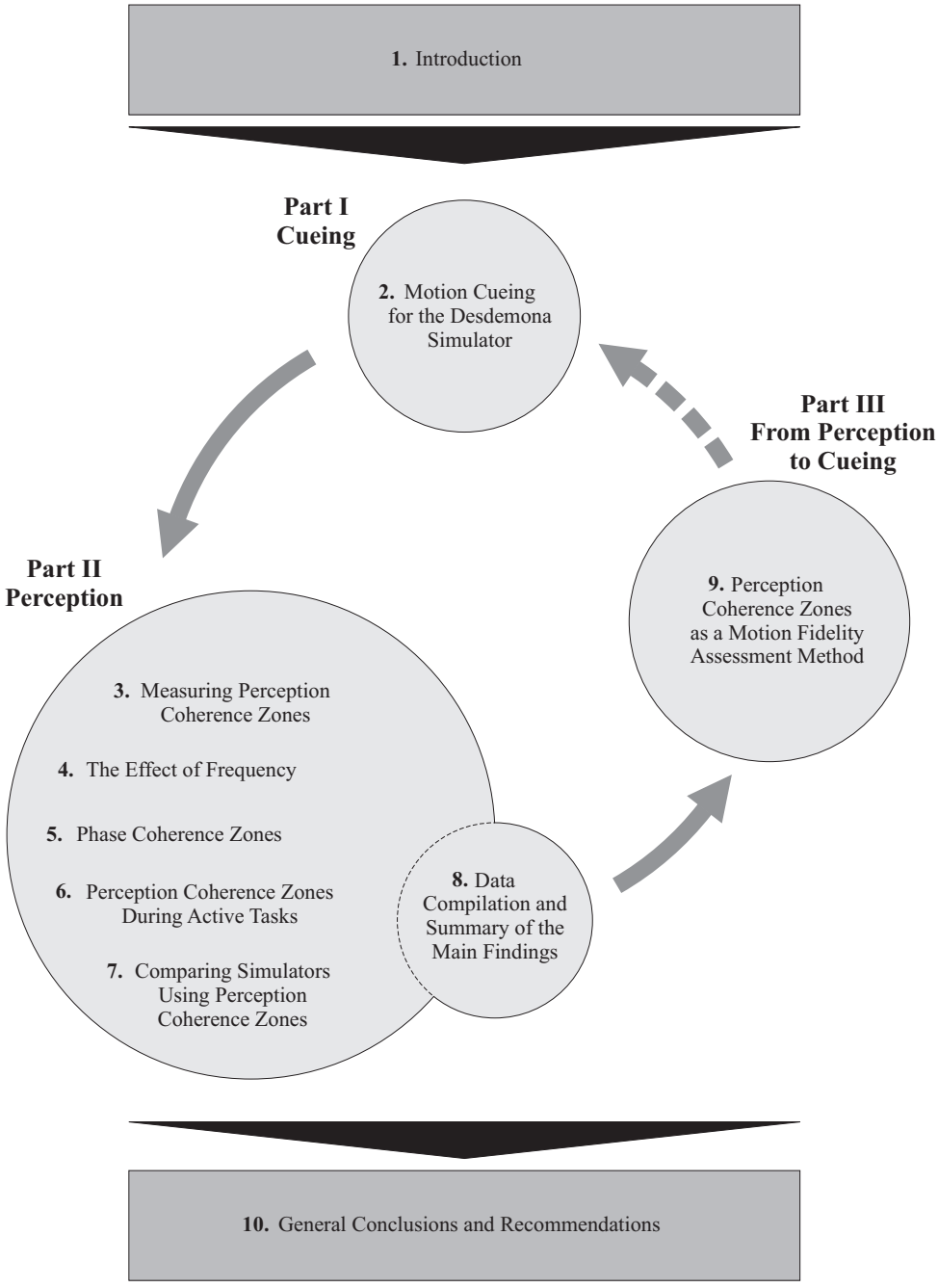


Figure 1.5: Schematic representation of the thesis structure.

nusoidal signals as visual and inertial stimuli. Two frequency values are tested, a medium and a high value. Based on the results of this experiment, a hypothesis is formed that explains the effect of frequency on coherence zones and may be used to extrapolate the results to the whole frequency range. Then, a second experiment is described that tests the posed hypothesis in the medium and low frequency ranges. A low-frequency motion demands a larger motion space. For this reason, the second experiment is performed in the Desdemona simulator, which has an unlimited yaw attitude motion space.

Chapter 5 describes an experiment measuring phase coherence zones in yaw and pitch. The effect of frequency, amplitude and DOF on the perception coherence zones is studied and results are compared to previous research on phase differences and time delays between simulator's visual and inertial motion systems.

Chapter 6 deals with the effect of decreased subject attention on the measured perception amplitude coherence zones. Until now, during a coherence zone measurement subjects are only required to judge the match between inertial and visual motion stimuli and have no other tasks. However, during many simulation scenarios, the pilot will have to provide attention to one or more simultaneous tasks, such as manual control of a vehicle and navigation. In this chapter, an experiment is presented that investigates the effect of performing manual control tasks of different difficulties on the perception coherence zones.

Chapter 7 presents three amplitude coherence zones in sway performed in three simulators at the NASA Langley Research Center. The effect of amplitude, frequency and simulator configuration on the measured coherence zones is investigated.

Chapter 8 summarizes the main findings regarding perception coherence zones and discusses the trends visible in the data across the different experiments.

Chapter 9 uses the perception coherence zone data to construct criteria for the design and tuning of motion filters. The coherence zone criteria are compared to existing fidelity metrics. Considerations are made regarding the potential of coherence zones as perception-based motion fidelity criteria and assessment method.

Part I

Cueing

2

Motion Cueing for the Desdemona Simulator

This chapter is based on the following publication:

Valente Pais, A. R., Wentink, M., Van Paassen, M. M., and Mulder, M. (2009). Comparison of Three Motion Cueing Algorithms for Curve Driving in an Urban Environment. *Presence: Teleoperators & Virtual Environments*, 18(3):200-221.

2.1 Introduction

Throughout the years, research on car driving has been carried out with a multitude of purposes, as for example, understanding and modeling the human driver behavior (Ritchie et al., 1968; Godthelp et al., 1984; Godthelp, 1986; Van Winsum and Godthelp, 1996), assessing potential dangerous driving situations, and studying drivers' reactions to driver assistance systems (Jamson et al., 2007b) or new road designs. The advent of car simulators has improved time and cost effectiveness, while allowing better control and repeatability of the experimental conditions. Furthermore, simulators offer a myriad of possible scenarios while guaranteeing the driver's safety. However, new possibilities bring new questions. The motion and visual stimuli presented to the subject in the simulator are not a replica of the real car situation. Especially regarding motion, many compromises have to be made to be able to maintain the simulator within its physical limits and still provide the driver with the necessary motion cues.

Research has been undertaken to investigate the effect of simulator motion on driving tasks (Repa et al., 1982; Siegler et al., 2001; Greenberg et al., 2003). Others have compared driver motion perception, behavior and performance in a real car and in a simulator (Boer et al., 2000; Panerai et al., 2001; Reymond et al., 2001; Siegler et al., 2001; Hoffman et al., 2002; Brünger-Koch et al., 2006). In these studies behavioral and performance metrics are used to assess the relative and absolute validity of the simulator (Blaauw, 1982). These measurements normally depend on the task at hand and no single metric can be used to summarize the drivers behavior. For braking maneuvers, measures related to the longitudinal control of the car are taken, as for example, maximum deceleration (Brünger-Koch et al., 2006; Hoffman et al., 2002; Siegler et al., 2001; Boer et al., 2000), mean jerk (Siegler et al., 2001), vehicle speed (Brünger-Koch et al., 2006; Panerai et al., 2001), time to collision or time to the stop line when the subject initiates the braking maneuver (Boer et al., 2000; Hoffman et al., 2002; Brünger-Koch et al., 2006). For lateral control maneuvers, such as lane change or cornering tasks, behavior and performance measures performed include root mean square of the heading error and the lateral position error (Repa et al., 1982; Greenberg et al., 2003), steering wheel angle and steering wheel reversal rate (McLean and Hoffmann, 1975; Repa et al., 1982), maximum lane position deviation (Repa et al., 1982), mean trajectory (Siegler et al., 2001), lateral acceleration (Reymond et al., 2001), vehicle angular velocity (Siegler et al., 2001) and curve approach speed (Boer et al., 2000). The choice of objective metrics to be used in a simulator experiment is problematic since it depends on the task difficulty and on pre-determined performance goals. Furthermore, it is difficult to gather sets of studies that have used the same metrics to analyze the same issues. Consequently, the question of which motion cues are necessary for effective driving simulation, is still an open one.

An especially challenging problem is cueing cornering tasks in urban environments. City curves have a smaller radius than highways or country roads. These sharp turns are thought to be more provocative than, for example, highway curves, and can cause disorientation or even simulator sickness (Bertin et al., 2005; Nilsson, 1993). Moreover, when a car enters a curve there is an almost immediate onset of lateral acceleration due to the road curvature. Even at relatively low speeds, small curve radii can cause quite abrupt changes in lateral forces that are difficult to reproduce in a simulator. Both the quick onsets and the sustained forces throughout the curves play an important role in curve driving simulation (Greenberg et al., 2003; Kemeny and Panerai, 2003; Reymond et al., 2001; Siegler et al., 2001; Blaauw, 1982). Furthermore, not only is the curve radius small in urban environments, but also the curve angle is large. It is not uncommon to make 90 deg turns at a crossing or intersection with yaw rates of 30 deg/s (Grant et al., 2004). This type of scenario implies large yaw displacements that are difficult to render in simulators with a limited motion space. The concern about this problem is reflected in the recent development of driving simulators with a much larger motion space, as for example, a tilt platform (like the hexapod) mounted on top of a rail or an XY table (Schwarz et al., 2003; Dagdelen et al., 2004; Jamson et al., 2007a; "Toyota", 2007; Chapron and Colinot, 2007).

The Desdemona simulator (Bles et al., 2000; Feenstra et al., 2007), although it is structurally different from the ones described above, it has a similarly large motion space. This makes it a quite attractive device to be used in road vehicle simulation. Moreover, the Desdemona central yaw axis can be used to simulate both the sustained lateral specific force as well as the yaw rate: the subject sitting in the cabin can actually drive through a curve.

The goal of the present research is twofold. The first part concerns designing and implementing a motion drive algorithm (MDA) for urban curve driving simulation in the Desdemona simulator. The new MDA makes use of the Desdemona centrifuge design to provide high angular rates and sustained accelerations. The second part consists of evaluating the new MDA through experimental comparison with two other motion algorithms. The evaluation of the three motion cueing algorithms will be done based on analysis of motion profiles, scores obtained from questionnaires, and objective measures of drivers' behavior and performance. However, since the task to be performed is relatively easy and no performance goals will be set, there is a limited number of objective metrics that can be used. Signals such as as velocity, acceleration and control inputs will be measured. From these a variety of metrics can be calculated afterwards.

In the following sections we will describe the Desdemona simulator and introduce the concept of motion cueing and motion filters. Then, the design of the new motion drive algorithm is explained, as well as the other two MDA's that were used to create the three experimental motion conditions. Finally, we describe the experimental method, present and discuss the results and draw some conclusions.

2.1.1 The Desdemona simulator

Figure 2.1 is a schematic of Desdemona with indication of its 6 degrees of freedom. Table 2.1 summarizes the Desdemona motion space specifications. The simulator has an 8 meter linear track. This linear track can rotate around its central point, providing a 4 meter centrifuge arm. This DOF is denominated the “central yaw axis”. Motion along the linear track represents displacement along the radius of the centrifuge, so, this DOF is called the “radius” or “radius track”. The structure mounted on the linear track consists of a 2 meter vertical linear track, the “heave track”. A gimbaled structure mounted on the heave track allows the cabin to rotate more than 360 degrees in three orthogonal axes. Using common aeronautical nomenclature we name the rotations around the vertical axis “cabin yaw”, around the lateral axis “cabin pitch”, and around the longitudinal axis “cabin roll”.

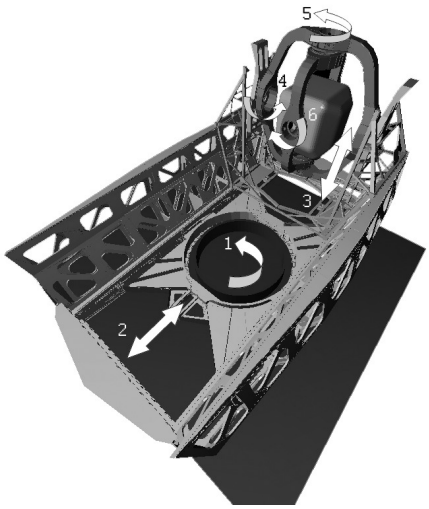


Figure 2.1: Artistic impression of the Desdemona simulator with indication of the degrees of freedom: 1. Central yaw axis. 2. Radius track. 3. Heave track. 4. Cabin roll. 5. Cabin yaw. 6. Cabin pitch.

2.1.2 Motion drive algorithms

Compared to the real vehicle, all simulators have a restricted motion space. This means that a one-to-one replica of the vehicle motion is impossible to accomplish. Therefore, mathematical algorithms are used to transform real vehicle motion into simulator motion. These motion drive algorithms (MDA), or motion filters, serve two purposes. The first is to maintain the simulator within its physical limits, observing not only the maximum

Table 2.1: The Desdemona simulator motion space limits. Pos, Vel and Acc, refer to position, velocity and acceleration limits, respectively.

	Central yaw axis	Radius track	Heave track	Cabin roll	Cabin yaw	Cabin pitch
Pos	> 360 deg	$\pm$ 4.0 m	$\pm$ 1.0 m	> 360 deg	> 360 deg	> 360 deg
Vel	155 deg/s	3.2 m/s	2.2 m/s	180 deg/s	180 deg/s	180 deg/s
Acc	45 deg/s^2	4.9 m/s^2	4.9 m/s^2	90 deg/s^2	90 deg/s^2	90 deg/s^2

displacement, but also velocity and acceleration constraints. The second is to provide the subject in the simulator with sufficient motion cues.

A widely known MDA is the Classical Washout algorithm (Reid and Nahon, 1985, 1986a,b). This algorithm, or variants of it, is mostly used in simulators using Stewart platforms, also known as hexapods. The methods used in this algorithm are common to other MDAs and will be briefly explained here as an introduction to motion cueing techniques.

For the sake of clarity, we define here some terms used later in this chapter. Linear accelerations refer to inertial linear accelerations excluding the gravity component. The combination of both linear acceleration and gravity is called specific force. Linear motion in the vehicle or subject longitudinal, lateral and vertical axes are also referred to as surge, sway and heave, respectively. Rotational movement around the longitudinal, lateral and vertical axes are also denominated roll, pitch and yaw, respectively.

MDAs are a set of motion filters that transform vehicle motion into simulator motion. In the Classical Washout algorithm, sustained linear and angular accelerations are high pass filtered, so as the real vehicle accelerates, the simulator moves in the required direction to render the accelerations (onset cue). In order to prevent the actuators from reaching their limits while the real vehicle continues to accelerate, after the onset cue the simulator moves back to its initial position (washout). The washout creates a simulator motion opposite to the one in the real vehicle. This should optimally be done below the motion perception threshold of the subject. If the washout motion is above the perception threshold, then the return motion will be felt by the subject as a false cue.

In driving simulation, sway motion can be used to provide the subject with the high frequency component of the lateral force present during a curve. When a car enters the curve there is a quick onset of lateral force due to the road curvature. At this point there will be a fast sway movement, followed by a slow washout motion that brings the simulator back to the initial position. When the car leaves the curve there is a sudden decrease in lateral force. This means the simulator will sway in the opposite direction and again washout to the initial position. Both sway motions, the onset-washout when entering the curve and the

onset-washout when exiting the curve, are defined by the same high pass filter, although only the first onset cue is desired. The filter settings should be such that the onset is strong enough to simulate entering the curve but not so strong that it causes a disturbing cue when leaving the curve. Furthermore, the washout motion should be kept below the perception threshold.

In addition to high pass filtering, a technique called tilt coordination is also applied. The linear accelerations of the real vehicle are low pass filtered and coupled to the angular channels. This means that as the real car accelerates forward, for example, the simulator cabin will slowly pitch up. If the tilting of the cabin is done below the rotation perception threshold, then the subject in the simulator will attribute the extra force in its longitudinal axis to a linear forward acceleration. This effect is stronger in the presence of visual cues. The combination of the onset cue from the high pass filter and the tilting of the cabin tries to match the total linear acceleration of the vehicle. It exploits the fact that human sensors cannot distinguish between linear accelerations and gravity (Berthoz and Droulez, 1982). If there is no perception of angular motion, then an increase in specific force can be perceived as an increase in linear acceleration, instead of a different orientation relative to gravity (Groen and Bles, 2004).

This technique can be used in curve driving simulation to provide sustained lateral force using roll tilt. The amplitude of the provided lateral force will depend on the maximum roll angle. The larger the intended lateral force, the larger the maximum roll angle should be. To maintain a subthreshold roll rate the cabin has to rotate slowly causing the lateral force to build up gradually. This low frequency movement complements the quick onset cue provided by the fast sway movement. However, if the cabin takes too long to reach the desired roll angle, there will be a moment when the lateral force provided by the onset cue has passed and the one provided by the roll tilt is not present yet. This may cause a drop in the perceived lateral force. Moreover, when the car leaves the curve, the sustained lateral force decreases abruptly. At this point, the cabin has to quickly rotate back. Again, to maintain the rotation below threshold, it will take some time before the cabin roll washout is finished. This can cause a lateral force false cue at the end of the curve. Thus, the choice of the cabin tilt rate limit is a compromise between, on one hand, a quick build up of the lateral force, a large enough maximum roll angle and a fast roll washout and, on the other hand, a roll rate that is below the perception threshold.

The Classical Washout algorithm, primarily designed for flight simulation in an hexapod, does not make optimal use of the Desdemona motion space. Hexapods can be referred to as parallel simulators (Angeles, 2003, pp.6–10). The design of the motion cueing algorithms for parallel simulators might be considered independent of the motion platform. Although the tuning of the filters must account for the specific limitations and motion space of the platform, the filters' output is expressed in translational and angular motion in an

inertial frame of reference and not in the specific degrees of freedom of the simulator. Conversely, Desdemona may be considered a serial simulator, since each DOF is connected to the next, forming an open-loop kinematic structure. In Desdemona, the motion cueing strategy is closely related to the design of the simulator and its separate degrees of freedom. A general motion cueing strategy for Desdemona has been designed before, the spherical washout algorithm (Wentink et al., 2005). Although this filter makes better use of Desdemona's motion space than the Classical Washout algorithm, it is difficult to tune. We believe that an MDA specifically designed for Desdemona, with a defined task in sight, can make a much more effective use of the Desdemona motion space and motion characteristics.

2.2 Three motion filters

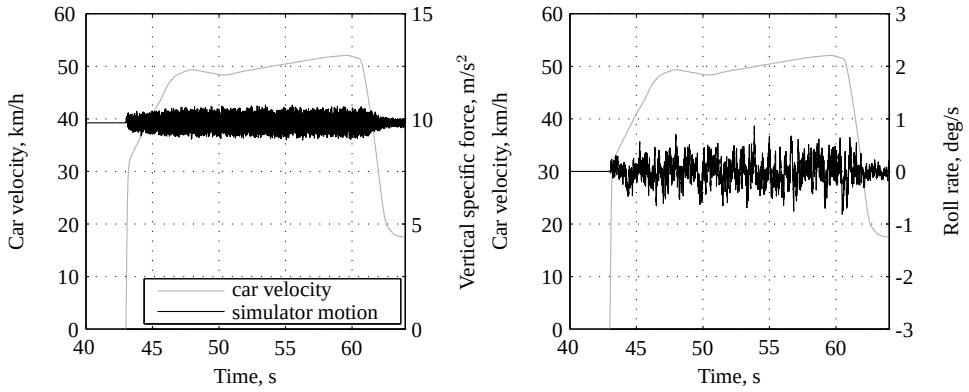
Three motion drive algorithms (MDA) were implemented in the Desdemona simulator: the Rumble filter, the Classical filter and the One-to-one yaw filter. The Rumble filter consisted of only road rumble motion. The Classical filter was a Classical Washout algorithm, similar to the one described in Section 2.1.2, adapted to Desdemona's motion space. The One-to-one yaw filter was especially designed for curve driving in Desdemona, and as the name indicates, it provided one-to-one yaw rate, with no need to washout lateral position.

2.2.1 The Rumble algorithm

The first filter was designed to provide a control "no motion" condition. However, not moving the simulator at all would allow subjects to recognize this condition too easily, potentially biasing the results. Therefore, the "no motion" condition was changed into a rumble only condition, i.e., no car accelerations were cued but there was motion in heave and roll that mimicked the vibrations and oscillations due to the car engine and the road irregularities. The road rumble algorithm was developed at TNO, Soesterberg and had been in use in TNO's small hexapod simulator. This motion condition provided the subjects with roll and heave motion with frequencies and amplitudes varying with car longitudinal velocity, but unrelated to the accelerations of the simulated vehicle (see Figure 2.2).

2.2.2 The Classical algorithm

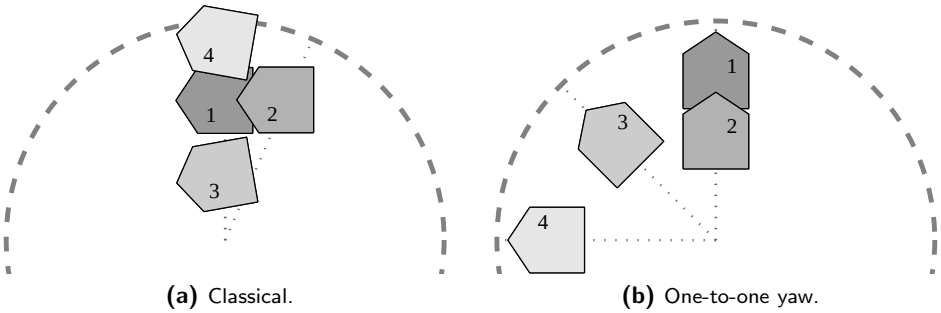
The second motion filter was a Classical Washout algorithm adapted to the Desdemona simulator motion space. The cabin initial or neutral position was halfway along the heave track and at 1 meter from the end of the radius track. The cabin was oriented perpendicular to the radius track. Figure 2.3a shows the motion of the cabin in the horizontal plane when the simulated car makes a left turn. As the simulated car approached the turn (1) and braked (2), the cabin moved backwards. The rotation of the central yaw axis was used to provide onset longitudinal acceleration. Small displacements in the radius and cabin yaw were used



(a) Heave. **(b)** Roll.

Figure 2.2: Heave and roll simulator motion varying with the car model longitudinal velocity.

to maintain the specific force in the driver's longitudinal axis, as the central yaw axis rotated. When the car entered the curve (3), the cabin moved along the radius and the yaw gimbal turned, providing onset yaw and lateral acceleration. Coming out of the curve (4) generated a similar response as entering the curve, but with the cabin yaw and the displacement along the radius in the opposite direction. Between each movement, the cabin was washed out back to the neutral position. Cabin pitch and cabin roll were used throughout the experiment for tilt coordination and for onset cues in roll and pitch. The road rumble was simulated using the same algorithm as in the Rumble filter.



(a) Classical. **(b)** One-to-one yaw.

Figure 2.3: Schematic of the cabin motion during one simulated left turn with the Classical and the One-to-one yaw motion filter. Cabin position as the simulated car approaches the turn (1), brakes (2), turns left into the curve (3) and leaves the curve (4).

2.2.3 The One-to-one yaw algorithm

The neutral position of the cabin was at 1.25 meter from the end of the radius track and halfway along the heave structure. It was oriented radially, i.e., the cabin's longitudinal axis was parallel to the radius track, with the subjects facing outward. The two outer rings were in the horizontal plane, causing a gimbal lock. However, the impossible rotation was yaw, which could be done using the central yaw axis. The cabin could still roll, using the yaw gimbal, and pitch, using either the pitch or the roll gimbals. Figure 2.3b shows the position of the cabin throughout a simulated left turn. When the simulated vehicle approached the turn (1), it decelerated (2). The cabin moved backwards along the radius track providing the subject with onset longitudinal acceleration cues. When the vehicle entered the curve (3), the central yaw axis rotated at the same yaw rate as the car. The tangential acceleration from the rotation of the yaw axis, and not the centripetal acceleration, as one would expect, was used to simulate the lateral forces. When the car left the curve (4), the yaw axis decelerated. Since the central yaw axis did not have limited displacement, there was no need to bring it to the neutral position, so there was no lateral position or yaw angle washout. Pitch motion was used to provide sustained longitudinal specific forces and to compensate for the centripetal acceleration generated by the rotation of the central yaw axis. Roll motion was also used to provide sustained lateral specific forces. For the road rumble we used the same algorithm as in the Rumble filter.

Figure 2.4 shows a block diagram of the developed motion cueing algorithm. The main elements will be analyzed in more detail in the following sections.

2.2.3.1 Longitudinal and vertical motion

In Figure 2.4, HP_{radius} and HP_{heave} were composed of a first order high-pass filter HP_{1st} followed by a second order high-pass, limiting filter ($HP_{2nd\,lim}$). The specific forces at the driver's head (f_{car}) were transformed to the subject's longitudinal and vertical axis, using the cabin orientation (Θ). The calculated x and z components of the specific force were then filtered and coupled to the radius and heave DOFs, respectively. The output of the filters H_{radius} and HP_{heave} were the commanded motion of the radius and heave DOFs in terms of position, velocity and acceleration. The limiting filter ($HP_{2nd\,lim}$) had two working modes: high-pass filter mode or limiting mode. After high-pass filtering the signal, the limiting algorithm looked at the current position, velocity and acceleration to predict the position in the near future. Depending on the calculated future position, one of the two working modes was chosen. If the future position was within the position limits, no limiting was necessary. This meant that the output of the total limiting filter ($HP_{2nd\,lim}$) was simply the input signal after a second order high-pass filter (cueing motion). On the other hand, if the future position exceeded the position limits, the limiting mode took over by braking the simulator and repositioning it at a safety distance (limiting motion). The simulator maximum velocity and acceleration were limited for both working modes, so both the cueing

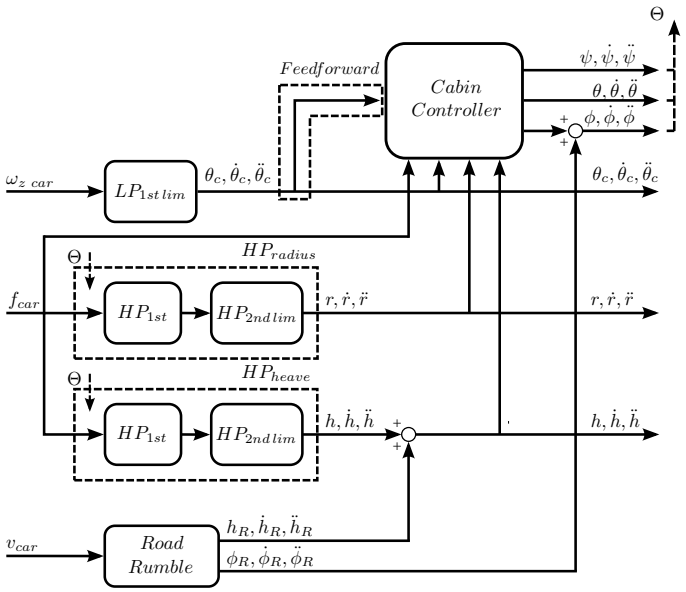


Figure 2.4: Block diagram of the One-to-one yaw motion cueing algorithm.

motion as well as the limiting motion had limited velocities and accelerations, although different limits were used for the two modes. The first order filter, HP_{1st}, prevented the limiting filter, $HP_{2nd\,lim}$, from alternating between the two modes continuously, when the input was a high sustained acceleration signal. This would cause the simulator to oscillate between the safety distance and the position limit.

2.2.3.2 Lateral motion

The central yaw axis was used to provide both the lateral specific forces and the yaw rotation. The car yaw velocity (ω_z) was coupled almost directly to the central yaw axis rotational velocity. In Figure 2.4, $LP_{1st\,lim}$ was a low pass filter with velocity and acceleration limiting. However, the cutoff frequency was sufficiently high that the filter approximated unity at the frequencies of interest, resulting in a one-to-one yaw rate. The output of this filter was the motion of the central yaw axis in terms of position, velocity and acceleration. The tangential acceleration generated by the acceleration and deceleration of the central yaw axis simulated the lateral forces through the curve. However, the accelerations generated were not large enough, so to increase the lateral specific forces during the curves the cabin was tilted in roll. Increasing the rotational acceleration of the central yaw axis could provide higher tangential accelerations, thus decreasing or even omitting the roll rotation. However, this would also increase the resultant centripetal acceleration. Accordingly,

there would be higher specific forces in the subjects' longitudinal axis that would require faster pitch rotation. Thus, the choice of the central yaw axis rotational acceleration was actually a compromise between adding some roll rotation or increasing the pitch rotational acceleration to eventually supra-threshold levels.

2.2.3.3 Tilt coordination and the cabin controller

In addition to roll, pitch rotation was also used to provide sustained specific forces. Pitch tilt complemented the radius track onset cues and compensated for the centripetal force associated with the central yaw axis rotation. The tilt coordination algorithm was implemented in the block $Cabin\ Controller$ in Figure 2.4. The inputs for the $Cabin\ Controller$ were the motion of the first three DOFs of the simulator (central yaw axis, radius and heave) and the desired specific forces at the subject's head (f_{car}). From the inputs, the specific force vector generated by the motion of the first three DOFs (f_3) was computed. Then, the $Cabin\ Controller$, using only the pitch and roll DOFs, oriented the cabin so that the direction of f_3 coincided with the direction of f_{car}.

To improve the timing of the cabin roll rotations with the desired lateral forces, we fed forward the car yaw rate (ω_z) to the roll channel of the $Cabin\ Controller$. The car yaw rate signal was used, instead of the lateral specific force signal, for two reasons. First, in the type of curves we used, the shape of the two signals was generally the same. Second, the car yaw rate signal was much smoother than the lateral specific force signal.

2.3 The experiment

2.3.1 Hypotheses

From the three implemented motion filters, we expected the One-to-one yaw filter to be rated best by subjects. The One-to-one yaw filter provided more motion than the Rumble filter, which we expected to be favorable to the realism of the simulation. Compared to the Classical algorithm, the One-to-one yaw filter did not need to washout lateral position nor yaw angle and provided a one-to-one yaw, instead of only yaw onset cues. We hypothesized that these features would improve the realism of lateral motion during turns.

With respect to motion sickness, we expected the Classical filter to be the most provocative due to the existence of false cues during the washout of the roll angle. The Rumble filter, since it provides very little motion, was hypothesized to be the least provocative. We also expected that if subjects did get motion sick, then throughout the experiment that condition would worsen, i.e., motion sickness would tend to increase throughout the experiment.

2.3.2 Method

2.3.2.1 Apparatus

The experiment was performed in the Desdemona simulator. The cabin was equipped with a generic car cockpit, see Figure 2.5.

Figure 2.5: The interior of the Desdemona cabin with the car cockpit installed. Outside world view from the driver's perspective.

The visual database was built using StRoadDesign ("StRoadDesign", 2008) and OpenSceneGraph (Burns and Osfield, 2004). A PC-based computer generated image system was used to render the outside world. In the cabin, three computers generated real-time images with an update rate of 60 Hz. Three projectors (resolution: 1024×768 px) projected the image on a three part flat screen, placed at approximately 1.5 m from the driver's eyes, creating an out-of-the-window field-of-view of 120 degrees horizontal and 32 degrees vertical. Blending and image distortion was also computed in the three computers in the cabin.

The dashboard consisted of a speedometer displayed on an LCD screen, placed behind the steering wheel and connected to the "on board" I/O computer. The sound system was developed at TNO. It reproduced wind and engine sound depending on vehicle velocity and engine RPMs. Direct drive electrical motors placed inside the cabin provided the control loading for the steering wheel, the gas and the brake pedals. Pedals and steering wheel position and velocity were read by the "on board" I/O computer with a sampling frequency of 1 MHz for the pedals and 100 Hz for the steering wheel. The I/O computer connected the controls, the dashboard and the audio system to the vehicle model at a frequency of 400 Hz.

Two computers ran the car model and the motion filters. The car model was implemented as an s-function generated by CarSim, running on Matlab Simulink at 400 Hz. The motion filters were also implemented in Matlab Simulink and ran at a frequency of 200 Hz.

One supervisor computer hosted the operator interface and logged data. The commanded motion was sent from the motion filters to the Desdemona computer via a bridge computer with a frequency of 200 Hz.

2.3.2.2 Experimental design and procedure

An experiment was performed using three different motion filters: the Rumble, the Classical and the One-to-one yaw filters, described in Section 2.2. Using each motion filter subjects drove two times around a square city block, performing only left turns. Figure 2.6 shows the top view of the circuit. Two of the turns were 20 meter radius curves and the other two, in diagonally opposite corners, were perpendicular crossings, or intersections, with rounded shoulders. The radius of the rounded shoulders was 8.5 meters.

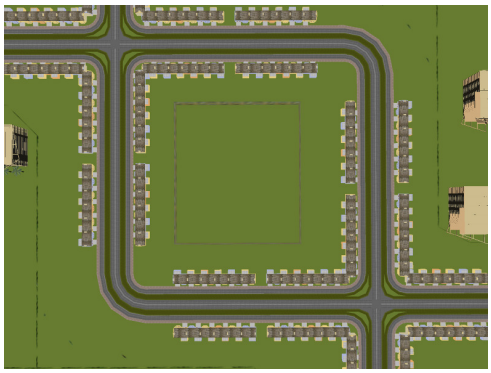


Figure 2.6: Top view of the driving circuit: a square city block with 150 meter straight segments, two 20 meter radius curves and two intersections.

After each run (8 left turns), the subjects answered a questionnaire. After the first and last run they also filled in a motion sickness scale. At the end of the experiment they were asked to rank the three filters according to their preference. The presentation order of the motion filters was randomized and balanced for all subjects. Each subject performed four runs: three experimental runs and one trial run. The trial run was performed with the same motion filter as the first experimental run. This means that the trial run was not the same for all subjects, but it was balanced for all subjects. Between each run there was time for the subjects to fill in the questionnaire and the motion sickness scale. Subjects indicated when they were ready for the next run. The total time inside the simulator, per subject, was between 20 and 30 minutes.

2.3.2.3 Questionnaires and motion sickness scale

The questionnaires consisted of seven questions to be answered by placing a mark on an analog scale, one mark per scale. The analog scales were represented by a horizontal line of 10 cm with beginning, middle and end markings. The extremes of the scale were from "totally unrealistic motion", on the left, to "just like a real car", on the right. Different aspects of the motion could contribute to the perception of unrealistic motion, such as: motion in the wrong DOF (false cues), motion that is either too strong or too weak (mismatch in amplitude), poor timing (mismatch in phase or time delays), or a mismatch in frequency. Subjects were asked to answer the following questions:

1. How realistic or unrealistic was the overall motion while driving, specially focusing on the curved segments (curves and intersections)?
2. How easy or difficult was it to steer the car (staying on the lane)?
3. How realistic or unrealistic did the road rumble feel?
4. How realistic or unrealistic did entering the curves feel?
5. How realistic or unrealistic did leaving the curves feel?
6. How realistic or unrealistic did accelerating feel?
7. How realistic or unrealistic did braking feel?

There are several motion sickness scales available (Kennedy et al., 1989, 1993). The motions sickness scale used was the Misery Scale (MISC), developed and validated at TNO Human Factors (Wertheim et al., 1992; Bos et al., 2005), see Table 2.2.

2.3.2.4 Subjects and subjects' instructions

24 volunteer subjects participated in the experiment, 16 male and 8 female subjects. Subjects were aged from 23 to 58 (the mean was 32 years, the median was 31 years), with a driving experience between 2 to 39 years (the mean was 13 years, the median was 12 years). All except two subjects had experience with automatic gear shifted cars. All but six subjects had previous experience with some sort of vehicle simulator.

Subjects did not see the simulator move in any of the motion conditions before they went in for the experiment. We instructed subjects to drive like they normally would in their cars, trying to keep the car in the center of the right lane and keeping an acceptable velocity. They were reminded that it was a city environment and the speed limit was 50 km/h. With respect to the questionnaires, we asked subjects to consider the simulator motion to answer the questions, i.e., not to focus too much on other simulation features like the visuals, the dashboard or the lack of mirrors and car frame. Also, we advised them to take as a reference

Table 2.2: The MISC: the rating scale used to evaluate motion sickness.

Symptom		Score
No problems		0
Slight discomfort but no specific symptoms		1
Dizziness, warm, headache, stomach awareness, sweating, etc.	vague	2
	some	3
	medium	4
	severe	5
Nausea	some	6
	medium	7
	severe	8
	retching	9
Vomiting		10

a rental car, a small family car with automatic gear shift. By doing so, we tried to establish an absolute reference, common for all subjects.

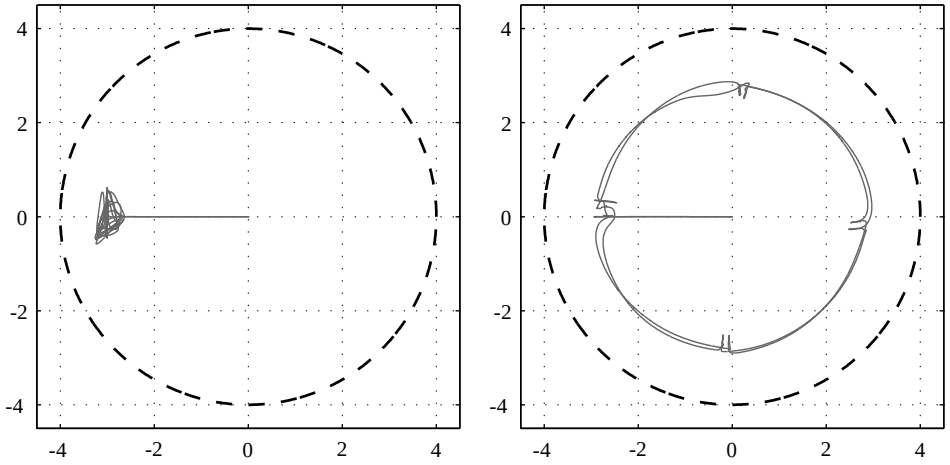
2.4 Results

2.4.1 Simulator motion

The three motion filters resulted in three different simulator motion profiles. For the Rumble filter the resulting motion was quite trivial and consisted of high frequency motion in roll and heave. For the Classical and One-to-one yaw filters it is interesting to look at the motion space used by the two filters, see Figure 2.7.

In terms of longitudinal motion the motion space used was equivalent for both motion filters. The tuning of the longitudinal channel was quite conservative and that can be seen in the limited motion space used, approximately half a meter. The lateral cueing in the One-to-one yaw filters was done using the central yaw axis, which lead to the 360 deg foot print shown in Figure 2.7b.

For the One-to-one yaw and Classical filters also the motion provided to the subject in the simulator and in the car was compared. The specific forces and yaw rate at the subject's head for one subject during four curves are displayed in Figure 2.8. The "car" signals were computed from the output of the car model and the "simulator" signals from the output of the motion filters.



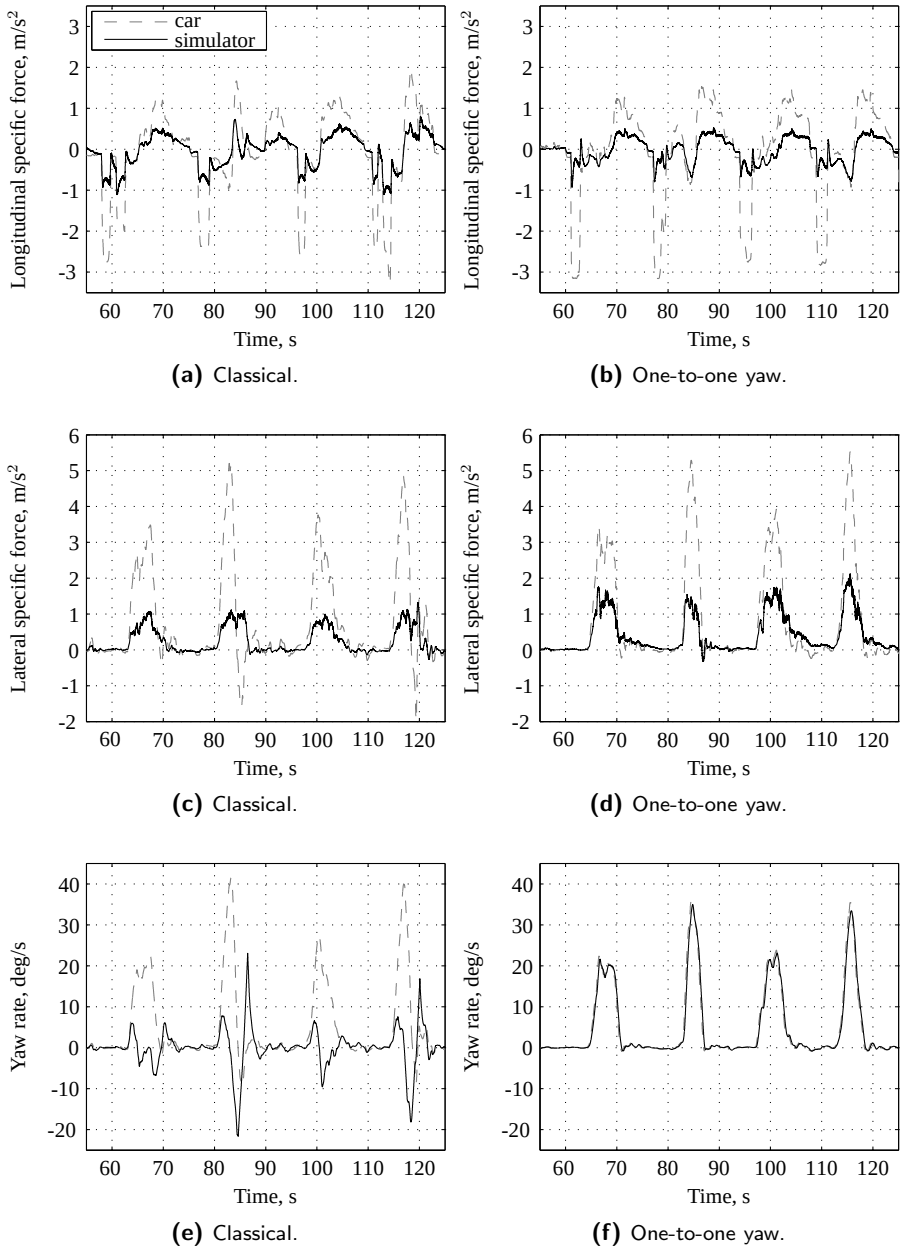
(a) Classical. **(b)** One-to-one yaw.

Figure 2.7: Motion space used by the Classical and the One-to-one yaw filters: foot print of the cabin motion (gray line) and simulator maximum radius (dashed black line).

The longitudinal cueing of both algorithms have similar characteristics, so the differences in the resulting longitudinal specific forces are minimal.

With respect to the lateral specific forces, the One-to-one yaw filter provided higher magnitudes than the Classical. Tuning the Classical filter to provide higher amplitudes of lateral specific force would lead to higher roll angles that would also take longer to washout, increasing the occurrence of situations like the one depicted in Figure 2.9. In this case, the washout of the roll angle in the Classical filter was too slow causing a peak in lateral force at the end of the curve. We relate this artifact to the many subjects' reports of feeling tilted sideways when coming out of the curve. The different behavior of the Classical motion filter shown in Figure 2.9 was related to the subjects' driving strategies: a combination of chosen velocity and trajectory. The One-to-one yaw filter allowed roll rates of 6 deg/s in tilt coordination, twice as high as in the Classical algorithm. Nevertheless, there were no complaints from subjects regarding false roll cues in the One-to-one yaw filter.

The major difference between the two motion filters, One-to-one yaw and Classical, was the yaw motion, see Figure 2.8e and Figure 2.8f. In a real car, driving into a curve results in an initial angular acceleration in the direction of the curve (positive) and then, leaving the curve, there will be an angular acceleration in the opposite direction (negative). With the Classical filter, only onset cues were provided. This means that the cabin first turned in the positive direction (onset) and then immediately returned to the initial orientation (washout).



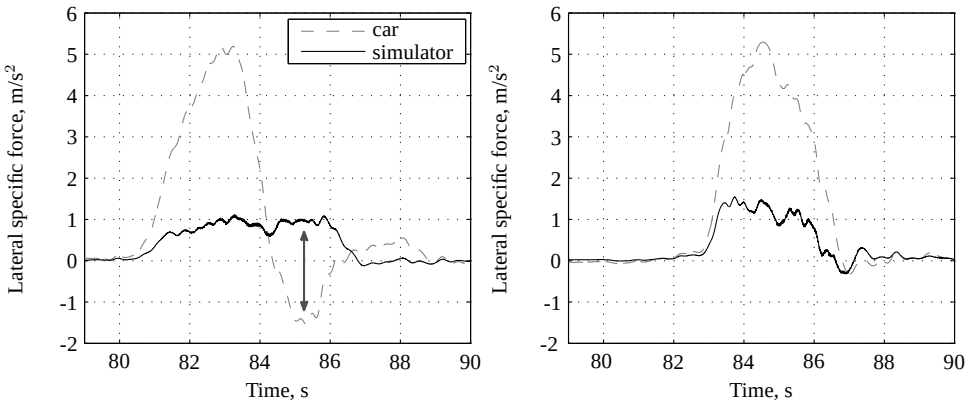
(a) Classical. **(b)** One-to-one yaw.

(c) Classical. **(d)** One-to-one yaw.

(e) Classical. **(f)** One-to-one yaw.

Figure 2.8: Longitudinal and lateral specific forces and yaw rate at the subject's head from the car model and in the simulator with the Classical and One-to-one yaw motion filters during four curves.

When leaving the curve, a similar behavior occurred: the cabin turned in the negative direction (onset) and then immediately returned to the initial orientation (washout). This behavior resulted in the yaw rate depicted in Figure 2.8e. The One-to-one yaw algorithm, on the other hand, did not have a washout, since the yaw cue was provided one-to-one using the central yaw axis.



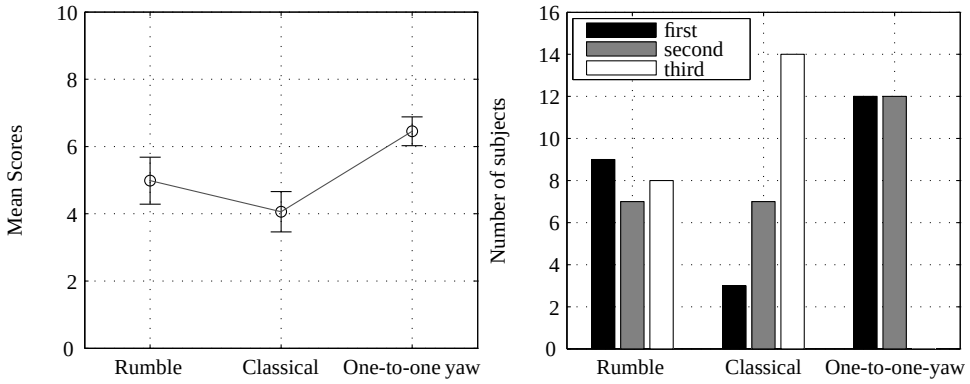
(a) Classical. **(b)** One-to-one yaw.

Figure 2.9: Lateral specific force at the subject's head from the car model and in the simulator for one subject driving the same curve with two different motion conditions: the Classical and the One-to-one yaw motion filters. At the end of the curve, the slow washout of the roll angle with the Classical algorithm causes a false cue in lateral force.

2.4.2 Questionnaires: drivers' ratings

The answers to the questionnaires were converted from the analog scale to a numerical value from 0 to 10. This numerical value was taken as the score on each of the seven questions in the questionnaire. The means of the scores, adjusted for all subjects (Field, 2005, pp. 279–285) and the 95% confidence interval of the means are shown in Figure 2.10a and in Figure 2.11. Each plot corresponds to a question on the questionnaire: overall score, easiness of driving, road feel, entering the curves, leaving the curves, accelerating and braking.

Figure 2.10 shows that with respect to the overall realism of the motion, the One-to-one yaw filter scored best, Rumble second and Classical last. The results of the ranking question are shown in Figure 2.10b. The One-to-one yaw filter was voted as the best by half of the subjects and the other half placed it in second place. The Classical filter was classified third by more than half the subjects and the Rumble filter was classified first, second and third almost the same number of times.



(a) Overall score. **(b)** Ranking.

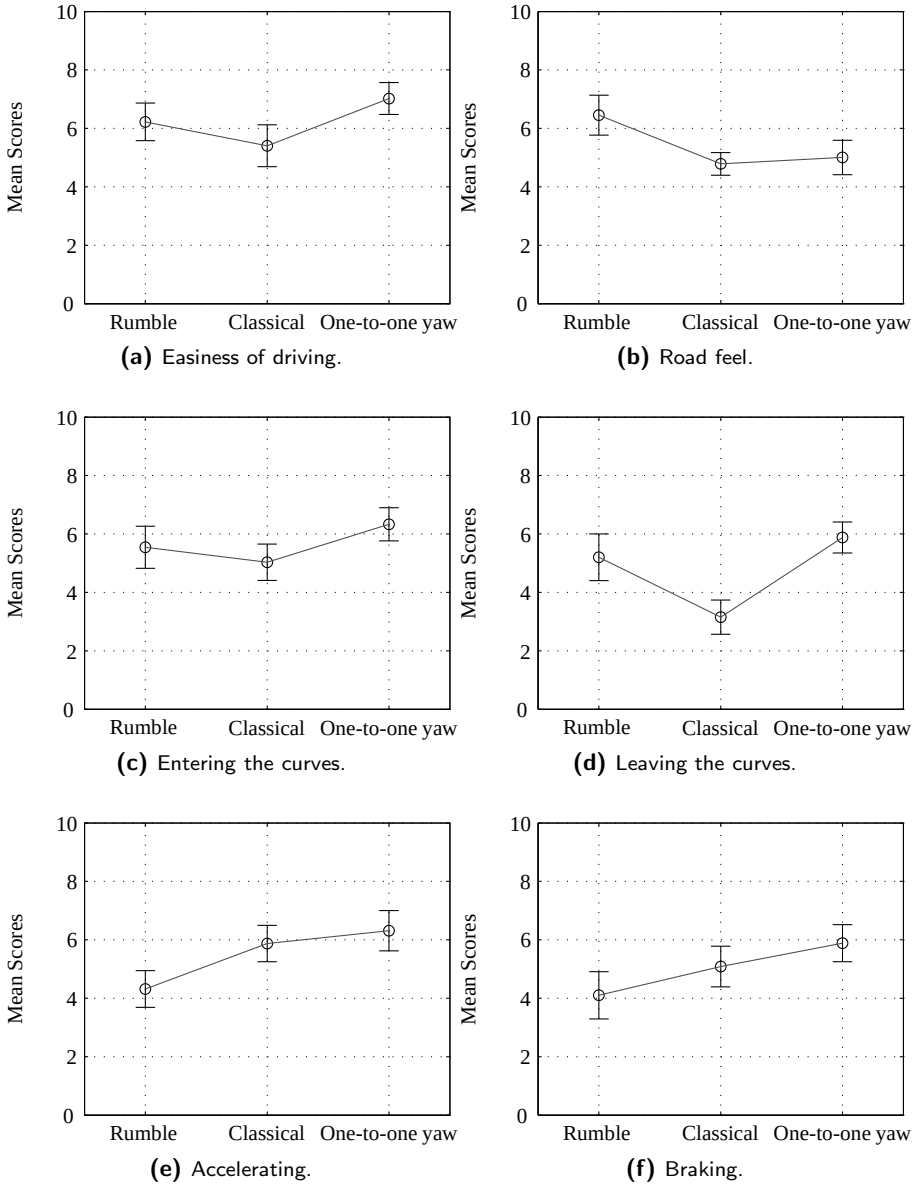
Figure 2.10: Realism of the simulator motion for the three motion conditions from the scores on the questionnaire and the results of the ranking question. The bars in (a) represent the 95% confidence interval of the means.

A repeated measures Analysis of Variance (ANOVA) was performed on the scores on each question. The independent variable was the motion filter (Rumble (R), Classical (C) and One-to-one yaw (O)) and the dependent measures were the scores on each of the questions in the questionnaire. There was a significant effect of the motion filter on the scores for all questions except question 4: realism of the motion entering curves. For all other questions post hoc pairwise comparisons were performed using Bonferroni correction for the level of significance (Field, 2005, pp. 339–341). Table 2.3 shows the results of the ANOVA and the post hoc tests.

On the question about the overall realism, the One-to-one yaw filter had a significantly higher score than Rumble and Classical. Regarding the easiness of driving, the One-to-one yaw filter had also the highest score, significantly higher than the Classical but not significantly higher than the Rumble filter. On question 3, realism of the road feel, the Rumble filter scored best, significantly higher than the Classical but not significantly higher than the One-to-one yaw filter.

Lateral motion was evaluated by questions 4 and 5. On question 4, realism of the motion while entering curves, the One-to-one yaw filter scored best, Rumble second and Classical last. Also leaving the curves, the Classical filter was considered the worse of the three. However, whereas the differences in scores while entering the curves are not statistically different, leaving the curves, the Classical condition shows a significantly lower score.

Regarding longitudinal motion, accelerating was considered significantly more realistic in the conditions with motion than with the Rumble condition. Similarly, braking with the One-to-one yaw condition was significantly better than the Rumble condition but was not



(a) Easiness of driving. **(b)** Road feel.

(c) Entering the curves. **(d)** Leaving the curves.

(e) Accelerating. **(f)** Braking.

Figure 2.11: Answers to the questionnaire. Mean scores and the 95% confidence interval of the means (from 0, not realistic at all, to 10, just like a real car).

Table 2.3: ANOVA and post hoc tests results of the answers to the questionnaire, where ** is highly significant ($p < 0.01$), * is significant ($0.01 \leq p < 0.05$), - is not significant ($p \geq 0.05$), and $n.a.$: is not applicable.

	ANOVA			Pairwise comparison		
Question	df	F	sig.	R-C	R-O	C-O
Total score	1.58, 36.30[a]	12.08	**	–	*	**
Easiness of driving	2, 46	4.55	*	–	–	*
Road feel	1.45, 33.45[a]	7.77	**	**	–	–
Entering curves	2, 46	2.80	–	$n.a.$	$n.a.$	$n.a.$
Leaving curves	1.56, 35.96[a]	13.65	**	*	–	**
Accelerating	2, 46	7.49	**	*	**	–
Braking	2, 46	4.42	*	–	*	–

[a] Greenhouse-Geisser sphericity correction applied.

statistically different from the Classical algorithm. However, for the braking maneuvers there was no statistical difference between the Classical and the Rumble filters' scores.

2.4.3 Motion sickness scales

Of all the subjects, only one was unable to perform the experiment due to motion sickness. The subject expressed a wish to stop after the trial run. We then asked one more subject to perform the experiment, to keep a balanced design. 24 subjects finished the experiment, from which, 8 started with the Rumble filter, 8 with the Classical filter and 8 with the One-to-one yaw filter. To evaluate the scores on the MISC, the subjects who did not get sick at all (subjects who scored zero twice on the MISC) were excluded from the statistical analysis. In total 8 subjects were excluded from the data set, 3 had started with the Rumble filter, 2 with the Classical and 3 with the One-to-one yaw filter. The MISC scores of the 16 subjects after the first and after the third runs were normally distributed.

An independent one-way ANOVA was performed to evaluate the difference in motion sickness scores after the first run. The subjects who started with the Classical filter presented the highest scores on the MISC and the ones who started with the Rumble filter presented the lowest scores. However, the ANOVA showed that the motion filter did not have a significant effect on the MISC scores ($F\,(2, 13) = 0.72$, $p > 0.05$).

The cumulative trait of motion sickness was evaluated by performing a repeated measures ANOVA to compare the MISC scores after the first run and at the end of the experiment. There was indeed a significant increase ($F\,(1, 15) = 11.52$, $p < 0.01$) from the scores after the first run ($M = 1.5$, $SE = 0.34$) to the scores at the end of the experiment

($M = 3.3$, $SE = 0.66$). Figure 2.12 shows the MISC scores after the first run and at the end of the experiment for the motion filter presented first to the subjects.

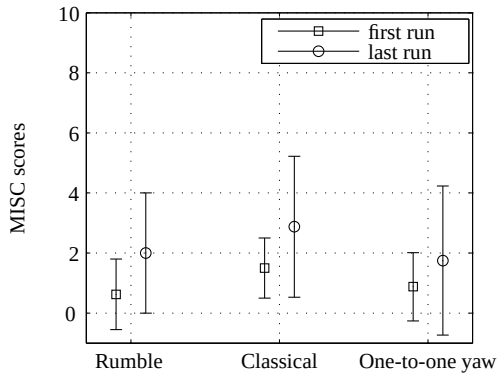


Figure 2.12: Scores on the MISC. Mean and the 95% confidence interval of the mean of the MISC scores after the first and last run, for the motion filter presented first to the subjects.

2.4.4 Objective measures

We expected the subjects to adopt a different driving behavior or have a different performance depending on the motion filter. We measured different signals: steering wheel angle, brake and gas pedal deflection, lateral and longitudinal acceleration, and velocity. The only metric for which we found a relevant and statistically significant effect of the motion filter was on the maximum deceleration.

The majority of times subjects pressed the brake pedal at the final part of a straight segment and continued pressing it in the beginning of the curve or intersection. So, the two laps around the square city block were divided in two sections: the intersections (Intersection), including the straight segment before it and the 20 meter radius curves (Curve), also including the straight segment before the curve. For each filter we computed the maximum deceleration in each of these sections. Each driver, in each motion condition, drove each section of the trajectory four times. For the statistical analysis we used the average of the values calculated for the second and third runs. Figure 2.13 shows the average maximum deceleration in each section of the trajectory. The values displayed are the adjusted means for all subjects (Field, 2005, pp. 279–285), for each motion filter and section of the road.

The data were analyzed using a two-way, repeated measures ANOVA. The independent variables were the motion filter (Rumble, Classical and One-to-one yaw) and the road section (Intersection and Curve). The dependent measure was the maximum deceleration. Mauchly's tests indicated that the assumption of sphericity was violated for some of the

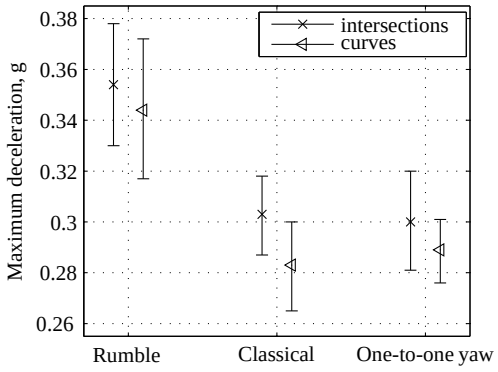


Figure 2.13: Mean maximum deceleration for each motion filter and section of the trajectory. The error bars represent the 95% confidence interval of the means.

effects. In these cases, we used Greenhouse-Geisser estimates of sphericity to correct the degrees of freedom. We also performed post hoc pairwise comparisons using Bonferroni correction for the level of significance (Field, 2005, pp. 339–341). The ANOVA showed a significant effect of the motion filter ($F\,(1.36, 31.27) = 9.996$, $p < 0.01$) and of the section of the trajectory ($F\,(1, 23) = 5.23$, $p < 0.05$) on the maximum deceleration. The post hoc tests indicated that the maximum deceleration was significantly lower ($p < 0.05$) with the Classical ($M = 0.29$, $SE = 0.009$) and One-to-one yaw ($M = 0.30$, $SE = 0.009$) filters than with the Rumble filter ($M = 0.35, SE = 0.014$).

2.5 Discussion

In the present study experiments with three motion drive algorithms for car driving have been performed. This study did not include "real" car driving and even in the literature not much data from actual car experiments were found. Therefore, the motion drive algorithms are mainly compared against each other in terms of their use of motion space and with respect to the subjective ratings of driving realism and motion sickness. The maximum brake acceleration could be compared to values found in literature from both simulator and "real" car studies.

2.5.1 Simulator motion

The longitudinal motion cueing of the One-to-one yaw filter was very similar to the Classical washout. However, since this was the first experiment to run in the Desdemona simulator, extra care was taken for the motion constraints. To prevent the simulator from reaching

the radius arm limits during any part of the experiment, the longitudinal motion tuning was very conservative. Less conservative filter parameters would allow better use of the motion space.

In the One-to-one yaw algorithm the maximum roll rates were twice as high as the typical 3 deg/s used in tilt coordination, without complaints from the subjects. In our opinion two things explain this result. First, the feedforward loop provided a roll rate that was much better timed to the onset of lateral force. Second, in a real vehicle, there is also a roll onset while entering and leaving a curve. The provided roll rate fit the expected vehicle roll, making it easier to accept it as a "good cue". We think this technique can be used in other types of vehicle simulation, like flight simulation. The success of the tilt coordination using the feedforward loop lies on the choice of the driving signal. In this case, we used the vehicle yaw rate, which had the same shape as the desired roll angle.

2.5.2 Questionnaires: drivers' ratings

The One-to-one yaw filter scored best in terms of realism of the simulation. Both the scores on the question about the overall realism and the ranking results confirm that the order of preference was the One-to-one yaw filter first, second the Rumble and third the Classical. Also the question about the easiness of driving showed the same order of preference. The road feel was less realistic in the conditions with motion (Classical and One-to-one yaw filters). Some subjects reported that the feel of the road in these conditions was too strong and it sometimes felt like they were driving a small truck. This result probably reflects an interaction between the road feel cueing and the vehicle motion cueing, which amplified the simulator accelerations.

Looking at the questions about the lateral motion, entering and leaving curves, the One-to-one yaw filter scored best and the Classical worst. When leaving the curve with the Classical filter, many subjects reported feeling tilted when coming out of the curve. Some of the subjects added that it was sickening, disorienting or simply unpleasant. This lead us to believe that the lower scores of the Classical filter in this question were due to this artifact. With the Rumble condition there were no lateral cues, except for roll vibrations, and still the score was just slightly lower than with the One-to-one yaw filter. More obvious differences on paper, like the one-to-one yaw rate provided in the One-to-one yaw filter, were not positively noted by the subjects, not even in the experimental debriefing. These observations indicate that the scores reflect not so much the good cueing as they do the "bad cueing" or "bad motion". The presence of false cues, even brief ones, are strongly penalized, whereas the absence of both good cues and false cues, like in the Rumble condition, seems to be tolerated quite well.

Regarding the longitudinal motion, the One-to-one yaw filter scored best and the Rumble condition worst. It seems that here, unlike in the lateral case, there was a recognition of the lack of "good cues". The Rumble filter was not reported to be disturbing or disori-

enting, but some subjects did report that something was missing, although others also said that the acceleration and braking feeling was the best in this condition. Nevertheless, the scores clearly show that, although there were no false cues, the Rumble filter was indeed considered less realistic than the conditions with motion. The score difference between the Classical and the One-to-one yaw filters was not large, which was to be expected, since both filters used similar algorithms to cue longitudinal motion. However, for the braking maneuvers the Classical algorithm did not show an improvement with respect to the Rumble condition, whereas the One-to-one yaw clearly did. Tentatively, this may be due to the fact that, although both algorithms cued longitudinal motion similarly, they were coupled to different degrees of freedom. The Classical algorithm used the central yaw axis and the cabin yaw, whereas the One-to-one yaw used the radius. Moreover, the two filters were dramatically different in the other degrees of freedom which implies different interactions with the longitudinal motion channel. This cross-talk between motion in different degrees of freedom might also be the cause for the small difference between the scores of the two algorithms. Another possibility is that the subjects were slightly biased in their ratings of the longitudinal motion due to differences in perceived lateral motion.

In general, the assessment of new motion drive algorithms is a difficult task, since there is no standard method and it always relies on one specific set of tuned parameters. The comparison with other MDA's rests on the assumption that all motion filters were tuned equally well. In this experiment we tuned all the motion filters using the authors and a few others as test subjects. The tuning of the Classical algorithm was especially difficult. Higher gains, and hence more motion, caused the already mentioned lagging of the roll angle when leaving the curves. Lower gains provided smaller motion cues and it became difficult to distinguish the motion with the Classical filter from the motion with the Rumble filter. On the whole, the questionnaire method and the breakup of the questionnaire into questions about the longitudinal and lateral motion seems to be a very successful way of understanding the strong and weak points of new concepts for motion cueing.

2.5.3 Motion sickness scales

The low scores obtained on the MISC and the fact that only one subject was unable to finish the experiment, was quite satisfactory. The scores on the MISC were lower with the Rumble filter and higher with the Classical filter. No statistical difference was found though, possibly due to the short duration of each run. The higher scores with the Classical filter are probably an effect of the false cues present at the end of the curve.

The scores at the end of the three runs were on average higher than the scores after the first run, supporting the hypothesis that motion sickness is cumulative.

2.5.4 Objective measures

The only relevant metric that showed a significant effect of the motion filter was the maximum deceleration before and in the beginning of the curves and intersections. Maximum deceleration values from real car experiments are not abundant. Brünger-Koch et al. (2006) reported a mean maximum deceleration value of 0.2 g, when subjects were asked to make a full stop at a specified point, with an approaching speed of 50 km/h. Boer et al. (2000) reported mean values between 0.2 g and 0.33 g for approaching speeds between 52 to 65 km/h. Boer et al. (2000) also showed results for a 40 meter radius, left curve negotiation. With approaching speeds between 72 and 76 km/h, the maximum deceleration values were between 0.23 g and 0.36 g (means per subject). Although the experimental settings were different, the values found in the present study fit well within the ones from tests in a real car.

The mean maximum deceleration values obtained show that in the conditions with motion (Classical and One-to-one yaw filters), the maximum deceleration was lower than in the condition without motion (Rumble filter). This is in agreement with the findings of Siegler et al. (2001). Siegler et al. (2001) performed an experiment on braking behavior in a driving simulator with and without motion. The driving speed was 80 km/h. They reported higher deceleration values in the "no motion" condition than in the "motion" condition. For self initiated braking, the maximum deceleration with the "no motion" condition was 0.54 g whereas with the "motion" condition it was 0.44 g. For braking triggered by signposts, with the "no motion" condition the maximum deceleration was 0.48 g and with "motion" it was 0.43 g. Although these values are larger than the ones obtained in the present study, a few points should be taken into account. First, these values refer to maneuvers where subjects were asked to reach a full stop. Second, the nominal driving speeds were larger than the 50 km/h maximum set for the present task. Third, the subjects were asked to reach a full stop at a certain pre-determined line. In order to meet the performance goal and decelerate the car to a full stop, the braking maneuver was probably more aggressive than the one needed in the present experiment, to approach a curve or intersection. Nevertheless, the maximum deceleration with the Rumble condition was significantly higher than with the conditions with motion, indicating that for the braking maneuver, motion was indeed relevant.

2.6 Conclusions

The designed motion drive algorithm showed potential for urban curve driving simulation. The use of the central yaw axis of the Desdemona simulator allows left and right curves with no need to washout the lateral position. Lateral motion was considered most realistic with the One-to-one yaw filter, although the Rumble filter was a very close second. The lower scores of the Classical filter, when leaving the curves, were related to the washout of

the roll angle at the end of the curve. The results on the lateral motion support the idea that "no motion is better than bad motion".

For the longitudinal motion the "no motion" condition (Rumble) was considered less realistic than the conditions with motion. The only performance metric that was affected by the different motion conditions was the maximum deceleration. Similar to other studies, in the "no motion" condition the average maximum deceleration was higher than in the conditions with motion. Comparison with values from real car experiments lead us to believe that the addition of motion contributed positively to the realism of the braking maneuver.

Since no specific task or performance goal was set for the experiment, not many behavioral and performance metrics could be used to compare the different filters. The simple task of driving around the city block, keeping in the lane, was not difficult enough to push subjects to their limits. A more demanding task or a very specific performance goal would perhaps force subjects to search for an optimal control behavior. By doing so, the motion cues in the simulator could become crucial in correctly assessing the vehicle state. Further investigation and improvement of the One-to-one yaw motion drive algorithm should include a comparative study between real car driving and simulator driving, setting a clear performance goal and assigning tasks of increased difficulty.

Part II

Perception

3

Measuring Perception Coherence Zones

This chapter is based on the following publication:

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

3.1 Introduction

Flight simulators have a limited motion space and it is impossible to reproduce aircraft motion one-to-one. Motion filters, which are used to transform aircraft motion into simulator motion, introduce phase and amplitude distortions. Whereas the motion cues are constrained to the simulator mechanical limits, the visual cues can represent the modeled aircraft true displacement and attitude. This situation causes the visual and the inertial cues provided to the pilot in the simulator to differ. It is important for the fidelity of the simulation that the pilot does not perceive this discrepancy between cues. While designing and tuning simulator motion filters, an attempt is made to minimize the difference between the pilot perceived motion cues and the aircraft actual motion that is shown in the visuals. For this reason, knowledge on human motion perception thresholds and perception and integration of visual and inertial cues is of crucial importance for flight simulation.

Throughout the past decades much research has been done on sensory motion thresholds (Stewart, 1971; Hosman and van der Vaart, 1978, 1980; Benson et al., 1986, 1989; Heerspink et al., 2005) and motion thresholds in the presence of other visual and motion cues (Zaichik et al., 1999; Rodchenko et al., 2000; Groen et al., 2004; Valente Pais et al., 2006, 2012; Pool et al., 2012). One step toward application of human perception knowledge in flight simulation are experiments with supra threshold visual and inertial stimulation. Initially, many experiments on perception of combined visual and inertial cues were carried out on rotating chairs enclosed by a rotating drum (Young et al., 1973; Huang and Young, 1981; Melcher and Henn, 1981; Wertheim and Bles, 1984; Probst et al., 1985; Mergner and Becker, 1990). These studies mainly focused on inducing and measuring self-motion perception during visual and inertial stimulation but failed to present a clear measure of to what extent the mismatch between visual and inertial cues was perceived by the subjects. Visual and inertial cues might differ in terms of amplitude, frequency, phase, timing or direction. If the differences are detectable then the realism of the simulation might be impaired. Conversely, if no mismatch between the cues is perceived, then they are interpreted as consistent with each other and as belonging to one realistic scenario, that is, the cues are perceived as being “coherent”.

More recent studies, performed in flight simulators, have measured such mismatches in terms of phase and amplitude. For example, Grant and Lee (2007) investigated the maximum phase difference between visual and motion that could go undetected by subjects in a simulator. Van der Steen (1998) researched amplitude differences between visual and motion that still resulted in a realistic simulation. He called the coherent set of values between visual and inertial motion amplitude, the coherence zone. Although not explicitly stated by the authors, the study of Lee and Grant, in fact, also measured a coherence zone. Their measured threshold, the maximum motion phase lead for which visual and inertial cues were still considered to be coherent, defines a “phase coherence zone”.

The studies by Grant and Lee (2007) and Van der Steen (1998), like most of the available work on perception, has been done during passive tasks, that is, there was no pilot in control. During a pilot-in-the-loop simulation it is difficult to have control over the precise visual and motion cues provided which complicates the design and analysis of experiments. More specifically, the frequency content and amplitude of the stimuli greatly depend on the pilot's control strategy (Zaal et al., 2009; Pool et al., 2010). The precise influence of the amplitude and frequency of the stimuli on the perception of coherent visual and vestibular motion is still unknown.

In this chapter, although still with passive tasks, steps are given in the direction of measuring perceived coherence between visual and inertial cues during a pilot-in-the-loop simulation. The effects of stimulus amplitude on perception coherence zones for yaw motion are investigated in two experiments in the Simona Research Simulator of the Delft University of Technology.

The first experiment extends the coherence zones measured by Van der Steen (1998) to amplitudes closer to the ones used in vehicle simulation. The higher amplitudes are chosen based on data from helicopter yaw capture tasks (Schroeder, 1999; Grant et al., 2006; Ellerbroek et al., 2008). The experimental procedure is similar to the one used by Van der Steen (1998). That is, for a certain visual motion amplitude, the inertial motion amplitude is varied throughout a set of trials. Then, after each trial, subjects are asked about the coherence between the visual and the inertial cues. The subject's answer determines the inertial amplitude of the following trial according to a simple up-down staircase procedure (Levitt, 1971).

The second experiment is designed to validate a new measuring method that gives the subjects a more active role. Using the same setup as in the first experiment, subjects are asked to tune the amplitude of the motion cue by increasing or decreasing its amplitude across trials. This self-tuning method aims to be faster and less tiring for subjects than the staircase method.

The chapter is structured as follows. Section 3.2 defines the concept of coherence zone and summarizes the work of Van der Steen (1998) on this topic. Afterwards, Sections 3.3 and 3.4 describe the two experiments, including the corresponding results. At the end, a general discussion of all the experimental results is presented in Section 3.5 and the final conclusions are drawn in Section 3.6.

3.2 Coherence zones

The term coherence zone was first introduced by Van der Steen (1998). However, many others have studied the influence of combined visual and inertial cues in self-motion perception (Young et al., 1973; Pavard and Berthoz, 1977; Zacharias and Young, 1981; Probst et al.,

1985; Wertheim, 1994; Mesland, 1998; Zaichik et al., 1999; Rodchenko et al., 2000; Groen et al., 2004; Valente Pais et al., 2006) and similar concepts that used different terminology can be found elsewhere (Groen et al., 2001; Fortmüller and Meywerk, 2005; Fortmüller et al., 2008; Grant and Lee, 2007). Since the current work was based on the experimental methods of Van der Steen (1998), this section provides a description of his work on coherence zones. Also, the definition of coherence zone and related concepts as presented by Van der Steen (1998) and as used in this thesis are briefly discussed.

A coherence zone is formed by the various combinations of visual and inertial cues that are perceived by the pilot as being part of a coherent, realistic simulation. The limits of such a coherence zone are thus dependent on the definition of a realistic simulation. Van der Steen (1998) defined a coherence zone as a combination of visual and inertial stimuli that, although being physically incongruent, still "provides the perception of an earth stationary visual scene". This is based on the fact that during locomotion in the real world, the visual scene is normally perceived as stationary with respect to an earth-fixed reference frame. Since the perceived self-velocity and the perceived velocity of the visual scene may be assumed to be inaccurate to a certain extent, there will be a range of values for which self-velocity and visual scene velocity will be perceived as matching although they are physically different. A large enough mismatch between these two would cause the perception of a non earth-stationary visual scene. In a virtual environment, the perception of a moving outside world signifies that the simulation is impaired.

Van der Steen (1998) performed two studies in flight simulators. In the first study the perceived coherence between the inertial and visual stimuli was used as a measure for the uncertainty in perceived inertial self-motion. For that purpose, the visual stimuli consisted of checkerboard patterns displayed very shortly on monitors located in the subject's peripheral field of vision. Van der Steen (1998) varied the visual stimulus amplitude for given amplitudes of the inertial motion. Both the visual and the inertial stimuli were sinusoidal. For each inertial amplitude, he measured the maximum and minimum visual cue amplitudes that still resulted in the perception of an earth-stationary visual scene. When the visual cue amplitude was excessively large with respect to the inertial cue, subjects perceived the visual scene to move against their perceived self-motion direction. On the other hand, if the amplitude of the visual cue was too small, subjects perceived the visual scene to move in the same direction as their perceived self-motion, that is, as moving with them. These two scenarios represented two velocity amplitudes that defined the coherence zone (CZ): a "fast" threshold (V_{FAST}) and a "slow" threshold (V_{SLOW}):

$$CZ = V_{FAST} - V_{SLOW} \tag{3.1}$$

Using a staircase method (Levitt, 1971), Van der Steen (1998) measured these thresholds and for each pair of thresholds he calculated the point of mean coherence (PMC) and the gain of the PMC (GMC):

$$PMC = 0.5\,(V_{FAST} + V_{SLOW}) \tag{3.2}$$

$$GMC = \frac{PMC}{V_{SELF}} \tag{3.3}$$

The GMC is a measure for the position of the coherence zone with respect to the inertial amplitude (V_{SELF}). A GMC less than one signifies that the measured PMC is less than the provided inertial velocity, indicating that subjects are underestimating their inertial amplitude or overestimating their visually perceived velocity. If the GMC is larger than one, then the PMC is larger than the inertial velocity, indicating an overestimation of the inertial velocity or an underestimation of the visually perceived velocity.

Van der Steen (1998) measured thresholds for surge, heave, roll, swing (combined sway and roll) and yaw using different combinations of inertial amplitudes and motion frequencies. For surge and heave one motion amplitude was tested: 0.5 m/s. For roll and swing, motion amplitudes varied from 3.7 to 12.4 deg/s and for yaw, they varied from 2.9 to 17.8 deg/s. The profiles frequencies were 1.2, 1.5 and 2 rad/s for all degrees of freedom and for roll and swing an extra condition with frequency of 1 rad/s was tested.

The calculated GMCs were in general below unity for surge and heave motion. Comparing his results with previous studies (Zeppenfeldt, 1991; Wertheim, 1994; Mesland and Wertheim, 1995) Van der Steen (1998) concluded that for increasing inertial motion amplitudes the GMCs slightly decreased. For roll and swing motions the GMCs were above one and decreased with increasing inertial motion amplitude. For the yaw motion the GMCs varied from 2 at the lowest inertial motion amplitude to slightly below one for inertial amplitudes of 9.7 deg/s and higher. This indicates that subjects overestimated their inertial velocity or underestimated their visually perceived velocity at the lower amplitudes. Van der Steen (1998) concluded that these results were in agreement with previous studies by Wertheim and Bles (Wertheim and Bles, 1984), but only for the larger amplitudes. A particularly interesting finding was that the yaw coherence zone did not increase with increasing amplitude as was hypothesized based on findings by Wertheim (1994).

In a second study, “the outside world is being perceived as stationary” was taken as a measure for the realism of the simulation. Van der Steen (1998) varied the amplitude of the inertial motion for fixed amplitudes of the visual motion. The visual scene consisted of a hilly grass field with some houses and trees, displayed on a dome with a field of view

of 142 deg horizontal and 110 deg vertical (partially occluded by the cockpit). The visual and inertial motion profiles consisted of acceleration steps. The amplitude of the inertial cue was varied, again using a staircase procedure. The highest and lowest inertial motion amplitudes that still resulted in the perception of a stationary outside world were called the upper and lower threshold, respectively. Van der Steen (1998) measured these thresholds in roll and yaw for visual amplitudes of 0, 2, 4, 8 and 12 deg/s^2.

The resulting coherence zones for both roll and yaw became wider for larger visual motion amplitudes. For yaw visual amplitudes higher than 4 deg/s^2 the coherence zones were fairly symmetric, that is, the upper and lower thresholds were at similar distances from the one-to-one line. For roll this symmetry was not present: the upper thresholds were at a greater distance from the one-to-one line than the lower thresholds. For both roll and yaw motion the upper and lower thresholds could be linearly fitted. If the linear fit is extrapolated to higher visual motion amplitudes, however, the coherence zones became quite large. Figure 3.1 shows Van der Steen's (1998) upper and lower thresholds for yaw motion and the corresponding linear regression lines.

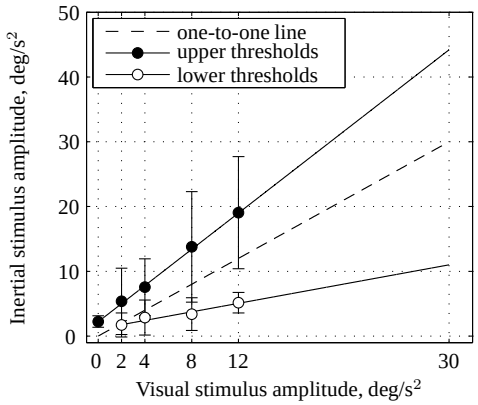


Figure 3.1: Mean upper and lower yaw thresholds measured by Van der Steen (1998) with corresponding regression lines. Error bars indicate the standard deviations.

As can be observed, for yaw visual motion with an amplitude of 30 deg/s^2 the coherence zone predicted by the linear regression would be between 10 and 45 deg/s^2. This implies that for vehicle simulation around this amplitude the motion gain can be decreased to as little as 0.3.

Van der Steen (1998) also collected verbal remarks from his subjects and divided their answers into three categories: "the motion was not realistic", "something was out of the ordinary" and "the motion was realistic". He reported that answers of the first category were given when the visual scene was perceived as non stationary and answers of the last

category when the scene was perceived stationary. Reports that something was out of the ordinary occurred around the measured thresholds.

The verbal reports collected by Van der Steen (1998) seem to indicate that around the threshold value the amplitude mismatch was detected but the visual scene was still perceived stationary. If we consider that the realism of the simulation is hindered when an amplitude mismatch is perceived, then this should be the criterion used to delimit the perception coherence zone. For this reason, the definition of simulation realism adopted in the present work was different from the one Van der Steen (1998) used. In this thesis, the simulation will be considered to be impaired when the inertial motion is perceived as "too strong" or "too weak" with respect to the presented visual scene.

3.3 Experiment 1

The goal of the first experiment was to extend the yaw coherence zones measured by Van der Steen (1998) to higher amplitudes. The coherence zone was measured in terms of an upper and a lower threshold. The upper threshold was defined to be the maximum inertial motion amplitude that was still considered by subjects to be coherent with the provided visual scene. The lower threshold was the smallest inertial motion still accepted by the subjects as coherent with the visual motion.

3.3.1 Method

The experimental method used was based on the one used by Van der Steen (1998), with a few adaptations. Also, a different simulator and a different visual database were used. The following sections describe the method, indicating were it differs from the original experiment.

3.3.1.1 Apparatus

The experiment was conducted in the Simona Research Simulator, shown in Figure 3.2. The Simona simulator has a hydraulic 6 DOF motion base which allows for a maximum displacement of $\pm$ 41.6 deg in yaw. The visual system consists of three LCD projectors, with a resolution of 1280 $\times$ 1024 pixels per projector, and a collimating mirror that provides a field of view of 180 deg $\times$ 40 deg. The visual update and refresh rates are 60 Hz. For a more detailed description of the Simona simulator motion and visual systems capabilities and the computer architecture and software used, please refer to Van Paassen and Stroosma (2000), Stroosma et al. (2003) and Berkouwer et al. (2005).

For the visual scene, Van der Steen (1998) used a grass field with hills and trees. In this experiment, a different database was used but an attempt was made to preserve the same number of visual objects. Figure 3.3 shows the outside visual scene, which consisted of a

Figure 3.2: The Simona simulator.

view of the Amsterdam Schiphol airport including the control tower, some lower buildings, part of a runway and some grass fields. The viewpoint height was the same as in the original experiment: 5 meters.

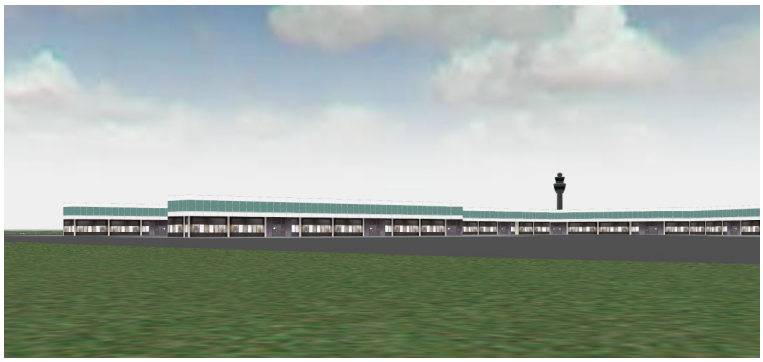


Figure 3.3: Central part of the outside visual scene showing a view over Schiphol airport.

3.3.1.2 Experimental design

The experiment had a one-way repeated measures design, with the amplitude of the visual stimulus being the independent factor. As explained above, the goal was to extend the co-

herence zones measured by Van der Steen (1998) to higher amplitudes. Based on data from helicopter tracking tasks (Schroeder, 1999; Grant et al., 2006; Ellerbroek et al., 2008) and taking into account the available motion space of the flight simulator, a maximum visual motion amplitude of 30 deg/s^2 was chosen. Also, to allow for a comparison with the results obtained by Van der Steen, some of his conditions were repeated. A balance had to be found between maintaining the number of conditions as low as possible while still obtaining enough data points along the amplitude axis. Coherence zones were measured for amplitudes of the visual scene motion of 0, 4, 12, 18, 22, 26 and 30 deg/s^2. The conditions with 0, 4 and 12 deg/s^2 stimuli matched the lowest, median and highest amplitudes tested by Van der Steen. To extend the amplitude levels up to 30 deg/s^2, four evenly space amplitude levels were chosen: 18, 22, 26 and 30 deg/s^2.

For all visual motion conditions a lower and an upper threshold for inertial motion were determined. Per subject the 14 threshold conditions were measured twice. This resulted in 28 experimental trials per subject.

3.3.1.3 Motion and visual signals

The stimulus profile used for visual and motion consisted of a sequence of smoothed steps in acceleration. Figure 3.4 shows an example of the position, velocity and acceleration time histories. The smoothing was done using a quarter of a period of a squared cosine function with a frequency of 10.5 rad/s. The longest acceleration plateau had a duration of 1.5 seconds.

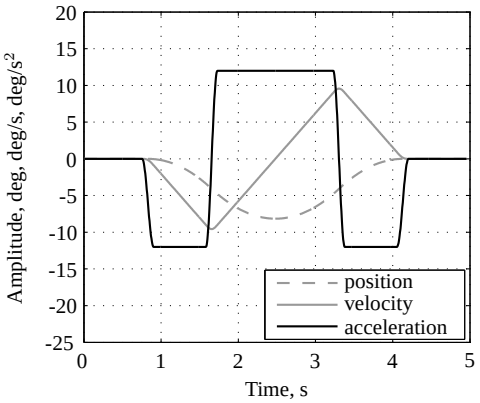


Figure 3.4: Example of the step-like motion profile for an acceleration plateau of 12 deg/s^2.

3.3.1.4 Procedure

Subjects were seated in the left-hand chair of the simulator cabin. Motion was applied so that the subject's head was in the center of rotation. The subject wore a headset with active noise cancellation on which engine noise was played. Three buttons located in the control column in front of the participants were used to record their answers throughout the experimental runs.

The order of the experimental conditions was randomized for every subject. For each experimental trial, the visual motion amplitude was kept constant and the inertial motion amplitude was varied through a set of runs. After each run, subjects were asked to indicate whether or not the inertial motion amplitude was perceived as coherent with the visual scene motion. The subjects' answer determined the inertial amplitude of the following trial according to first, a search algorithm to guarantee that subjects were within their coherence zone, and afterwards, a simple up-down staircase algorithm (Levitt, 1971).

Figure 3.5 shows an example of a sequence of trials starting with the search algorithm and making the transition to the staircase algorithm after the first positive answer. The procedure described by Van der Steen (1998) did not include a search algorithm.

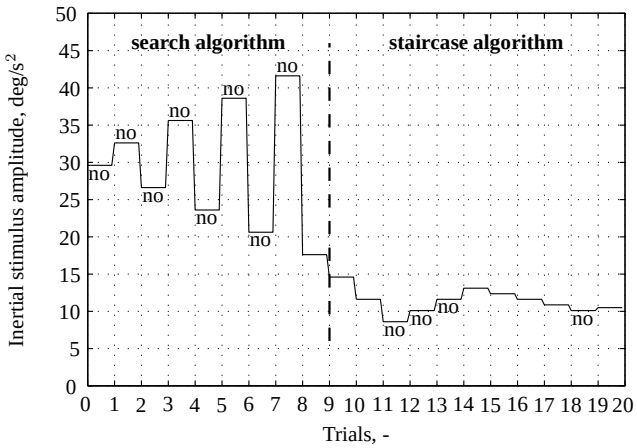


Figure 3.5: Example of a sequence of trials illustrating the algorithms used, for a lower threshold measurement at a visual motion amplitude of 30 deg/s^2. The negative answers are indicated above or below the corresponding amplitude level.

For both algorithms, after each answer, the following run was calculated based on a step size and a step sign. A positive step sign meant that the following run would have a higher inertial motion amplitude than the previous one. A negative sign meant that the next run would have a weaker motion.

The search algorithm started with an inertial motion between 0.9 and 1.1 times the visual scene and a step size of 10% of the visual motion amplitude. For the 0 deg/s^2 amplitude case, the initial step size was a random number between 0.2 and 0.3 deg/s^2 and the first trial amplitude equaled the initial step size. After each negative answer from the subjects, the step size was doubled and the step sign was inverted. After the first positive answer, indicating that the subjects were within their coherence zone, the staircase algorithm started.

In the staircase algorithm, the initial step sign was negative for lower threshold measurements and positive for upper threshold measurements. The initial step size was chosen to be 0.2 times the visual amplitude. For the 0 deg/s^2 amplitude case, the initial step size was a random number between 0.2 and 0.3 deg/s^2. After a reversal in answers, that is, a "no" answer after a "yes" answer or vice-versa, the step sign was reversed and the step size was halved. After four consecutive answers of the same type, always yes or always no, the step size was doubled. The staircase ended when the step size reached 1/8th of the initial step size or at the end of 30 trials.

Before the actual experimental session started, subjects did two test trials to get acquainted with the procedure.

3.3.1.5 Subjects and subjects' instructions

In total 8 male subjects aged between 23 and 28 years (mean age of 25 years) participated in the experiment.

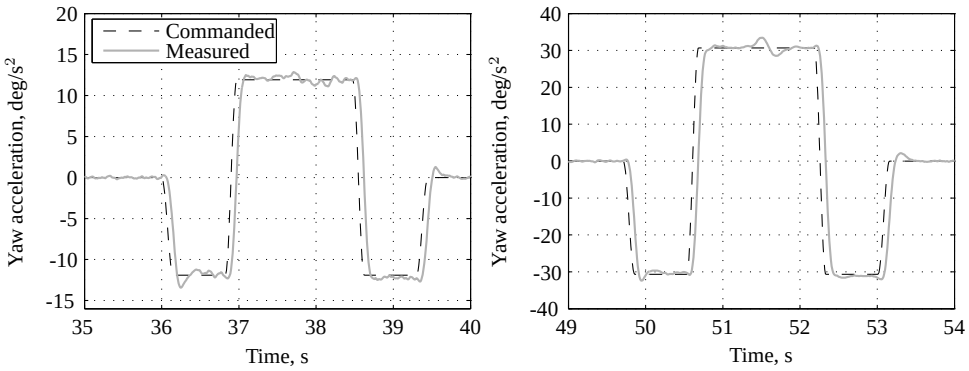
The participants were instructed to sit upright and refrain from making head movements throughout the experiment. They were, however, allowed to gaze over the visual scene at will.

Subjects were told they were to perform a series of runs divided in blocks. In each block of runs the visual scene would move the same way but the amplitude of the simulator motion would vary. Subjects were asked to answer the question "Did the amplitude of the visual movement correspond with the magnitude of the motion?" at the end of each trial. A positive answer would mean that they perceived the amplitude of the visual and of the motion to match, and a negative answer would mean that they considered the simulator inertial motion to be either too strong or too weak with respect to the visual motion.

Further, subjects were told that in each block of runs a sequence of amplitude levels would be tested that was chosen by an algorithm. No further information was given on the algorithm, nor were they informed that their answers had an effect on the next amplitude level.

3.3.2 Results

An example of the time histories of the commanded and measured motion signals for low and high amplitude values, approximately 12 deg/s^2 and 30 deg/s^2, respectively, is shown in Figure 3.6.



(a) 12 deg/s^2. **(b)** 30 deg/s^2.

Figure 3.6: Time histories of the commanded and the measured yaw acceleration for amplitudes of approximately 12 and 30 deg/s^2.

The small peaks observed in the measured signal halfway the acceleration step correspond to the so-called "turnaround bump". In percentage, this acceleration peak is around 8% of the commanded amplitude in the 12 deg/s^2 signal and 9% in the 30 deg/s^2 signal.

For the determination of the upper and lower thresholds, the amplitude levels tested during the search algorithm part were not taken into account. Only the data collected from the staircase algorithm were used. For each of the amplitude levels tested in one staircase, the corresponding answers were converted into a value. Negative answers were attributed a value of 0 and positive answers a value of 1. To each of these data sets a psychometric curve was fit using a least squares method. The curves were defined in terms of a mean and a standard deviation as described in Equation (3.4), for the lower thresholds, and Equation (3.5), for the upper thresholds. P_{lo} and P_{up} represent the probability of a certain motion amplitude (A) to be perceived as coherent with the visual amplitude. The estimated value of μ, which is the amplitude level at 50% probability, was taken as the threshold of the coherence zone.

$$P_{lo}\left(A\right) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{A} e^{\frac{-1}{2\sigma^2}(x-\mu)^2} dx \tag{3.4}$$

$$P_{up}\left(A\right) = 1 - \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{A} e^{\frac{-1}{2\sigma^2}(x-\mu)^2} dx \tag{3.5}$$

The estimated upper and lower thresholds were averaged for every repetition of every subject. The subjects' mean values are displayed in Figure 3.7 together with data from Van der Steen (1998). It should be noted that the standard deviations presented correspond to

the deviation of the estimated thresholds from the overall mean, and not to the estimated standard deviations of the psychometric curves.

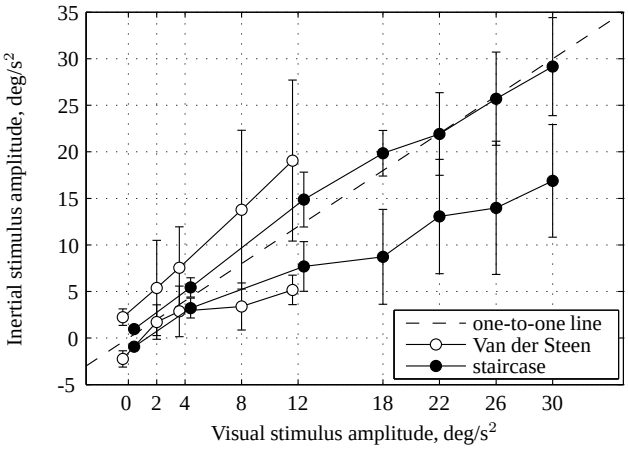


Figure 3.7: Mean estimated upper and lower thresholds plotted together with data from Van der Steen (1998). The error bars indicate the standard deviation.

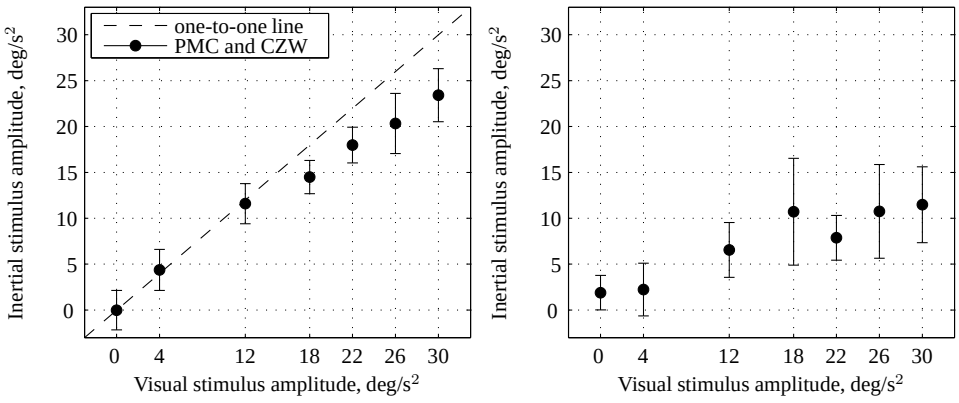
From Figure 3.7 it can be seen that the measured coherence zone is narrower than the coherence zone measured by Van der Steen (1998). This might be a result of the different questions posed to the subjects. The participants in his experiment indicated when the outside world was perceived to move whereas in the present experiment subjects signaled an amplitude mismatch between visual and inertial motion. Perhaps the perception of an amplitude mismatch precedes the perception of a non-stationary outside world.

For a clearer visualization of the coherence zone width and symmetry, a coherence zone width (CZW) and a point of mean coherence (PMC) were calculated from the upper (th_{up}) and lower (th_{lo}) threshold values using Equations 3.6 and 3.7. The PMC and CZW for all visual amplitudes are shown in Figure 3.8.

$$CZW = th_{up} - th_{lo} \tag{3.6}$$

$$PMC = th_{lo} + \frac{CZW}{2} \tag{3.7}$$

Up to the amplitude of 12 deg/s^2 the present results are comparable to the ones from Van der Steen (1998). The coherence zone is fairly symmetric, as indicated by the PMCs very close to the corresponding visual amplitudes, and the coherence zone width increases with increasing visual amplitude stimulus.



(a) Point of mean coherence. **(b)** Coherence zone width.

Figure 3.8: Measured coherence zones. Error bars indicate the 95% confidence interval of the mean.

For amplitudes higher than 12 deg/s^2 the coherence zone is no longer symmetric. The upper thresholds align with the one-to-one line causing the PMCs to deviate from the one-to-one-line. This indicates that for the higher amplitudes of the visual stimulus, subjects overestimated their inertial amplitude or underestimated their visually perceived velocity. It is also interesting to observe that for the visual amplitudes above 18 deg/s^2 the CZW hardly increases.

3.4 Experiment 2

Although the experimental method used in the first experiment seemed suitable, it demanded a relatively large effort from the part of the subjects. All subjects reported the task to be difficult, as it required a constant high level of concentration for a long time. In an attempt to make the task easier and also accelerate the experimental data collection, a new method was designed, a self-tuning method that gave participants the control over the motion amplitude of each of the runs. It was expected that this would lead subjects to converge to a threshold value in fewer runs than with the staircase procedure. To validate this new method, a small experiment was performed. Lower and upper thresholds were measured for two visual amplitudes with the staircase method used in Experiment 1 and with the new self-tuning method.

3.4.1 Method

3.4.1.1 Apparatus

As in the first experiment, all trials were conducted in the Simona Research Simulator, using the same visual scene.

3.4.1.2 Experimental design

A two-way repeated measures design was chosen. Upper and lower thresholds were measured for amplitudes of the visual scene motion of 12 and 30 deg/s^2, with both the staircase and the self-tuning method. This resulted in 4 experimental conditions that participants repeated 6 times, 3 times for measurements of the lower threshold and 3 times for measurements of the upper threshold. In total, each subject performed 24 experimental trials.

3.4.1.3 Motion and visual signals

The visual and motion profiles were the same as the ones used in Experiment 1.

3.4.1.4 Procedure

The experimental trials were divided in two blocks, one using the staircase method and the other using the self-tuning method. From a total of 5 subjects, 2 started with the staircase method and 3 with the self-tuning method. Within each block, the presentation order of the experimental conditions was randomized for every subject.

The staircase method was already described in Section 3.3.1.4. For the self-tuning method, in each experimental condition the visual amplitude was kept constant while the motion amplitude was varied throughout the runs of one trial. At the beginning of the trial, subjects were informed whether that trial corresponded to a lower or an upper threshold measurement.

In each trial, the amplitude of the first run was randomly selected between 1.1 and 0.9 times the visual amplitude. At the end of each run subjects could change the motion of the next run. They did this by pushing a switch button multiple times up or down until they reached a certain number of increments or decrements. The chosen number was displayed on the outside visual. A positive number meant the next run would have a higher amplitude motion, and a negative number meant a lower amplitude motion. After giving their answer, subjects pressed a second button to signal that they were ready for the next run. The trial ended when subjects' answers had two consecutive reversals of one increment or decrement, i.e., a sequence of 1, -1, 1, or -1, 1, -1. This indicated that subjects converged to a certain amplitude of motion that could not be increased or decreased anymore. To guarantee that this method had a similar measurement resolution to the staircase method, when subjects reached the conversion point, the difference between the last two amplitude levels tested

should be the same as in the staircase method. Recalling that the staircase method started with a step size of 20% of the visual amplitude and stopped when the step size was 1/8th of the initial step size, then the logical size of the increment or decrement in the self-tuning method is 1/8th of 20% of the visual amplitude.

At the end of all trials, subjects were asked which method they preferred and why.

3.4.1.5 Subjects and subjects' instructions

All 5 subjects had participated in Experiment 1. They were aged between 25 and 28 years with a mean of 25.6 years.

Subjects were informed that the experiment consisted of two parts, each with a different experimental procedure. The two procedures were labeled "Method 1" and "Method 2", corresponding to the staircase procedure and the self-tuning procedure, respectively. The instructions for the staircase part of the experiment were the same as the ones provided in Experiment 1. For the self-tuning part, participants were instructed to start their tuning procedure by finding a motion amplitude that matched the visual. From that point on, they should tune the motion up or down, dependent on the threshold measurement of that trial. For example, when measuring an upper threshold they should increment the motion until it was perceived as too strong. Then, they were told to decrease it and increase it as many times as needed until they found the strongest motion condition that was still perceived as coherent with the visual motion. Subjects were advised to start with increments of 8, 10 or more and decrease the number of increments or decrements at every direction reversal. They were informed of the stopping criteria of the trials.

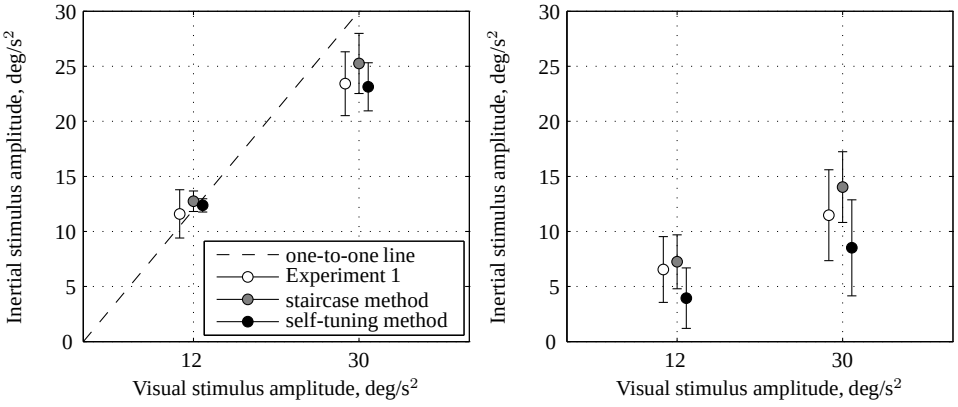
Before each part of the experiment subjects performed 2 to 4 test trials.

3.4.2 Results

The upper and lower thresholds measured with the staircase method were calculated as described in Section 3.3.2. For the self-tuning method, the threshold for each run was calculated by averaging the amplitude of the last two runs. For every subject, the threshold value for a certain experimental condition was averaged for all repetitions of that condition.

As in Experiment 1, the obtained upper and lower thresholds were used to calculate the point of mean coherence (PMC) and the coherence zone width (CZW). Figure 3.9 shows these values for both methods tested. For reference purpose, the PMC and CZW from Experiment 1 for the amplitudes of 12 and 30 deg/s^2 are also presented.

The PMC and CZW values from Experiment 1 and those from Experiment 2 with the two different methods were quite similar. In Experiment 2, the mean PMC at the visual amplitude of 30 deg/s^2 was around 24 deg/s^2 whereas at the visual amplitude of 12 deg/s^2 was quite close to the one-to-one point (12.5 deg/s^2). Similar to the first experiment, the coherence zone appears to bend below the one-to-one line at higher visual amplitudes. As also seen in Experiment 1, as the visual amplitude grows from 12 to 30 deg/s^2, the CZW



(a) Point of mean coherence. **(b)** Coherence zone width.

Figure 3.9: Measured coherence zones for two visual amplitudes using two different measuring methods. Error bars indicate the 95% confidence interval of the mean.

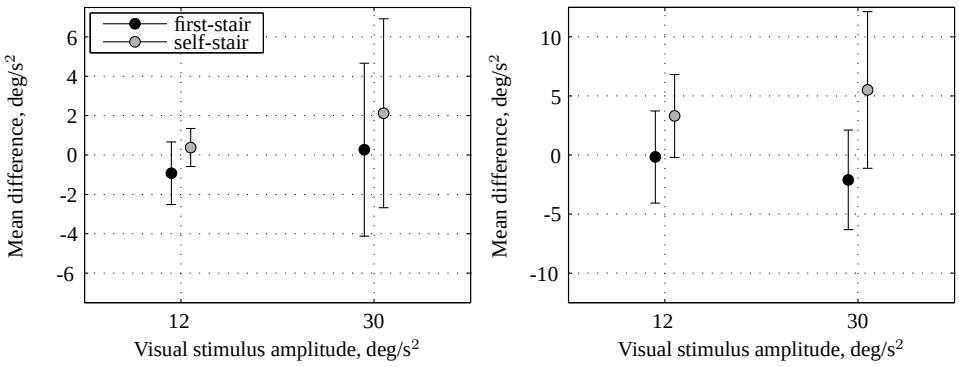
also increases. The CZW measured with the self-tuning method was slightly lower than the CZW measured with the staircase method.

To assess whether or not the self-tuning method was equivalent to the staircase method, a two one-sided t-test (TOST) was performed. To consider the two methods equivalent within a certain margin, the difference in values obtained with each method should be smaller than a chosen value δ. Two null hypotheses are made, that the difference between the two means ($M1 - M2$) is larger than δ and smaller than $-\delta$. If both null hypotheses are rejected, than it can be said that the absolute value of the difference in means is smaller than delta ($|M1 - M2| < \delta$).

In Experiment 1, thresholds were measured with the staircase method. In Experiment 2, thresholds were measured using both the staircase method and the self-tuning method. For the TOST it is assumed that the difference between the values obtained in Experiment 2 with the two different methods should not be larger than the differences obtained in Experiment 1 and Experiment 2 using the same method. Hence, the value of δ can be chosen based on the the values obtained in Experiment 1 and in Experiment 2 using the staircase method. For this purpose, the mean differences between the values obtained using the staircase method in Experiment 1 and in Experiment 2 and the corresponding 95% confidence intervals were calculated.

Figure 3.10 shows the mean differences between the values obtained in Experiment 1 (first) and Experiment 2 using the staircase method (stair), and the mean differences between the values obtained in Experiment 2 using the staircase method and the self-tuning method

(self). All confidence intervals include the zero, which indicates that the data are consistent with the means obtained in each experiment being equivalent.



(a) Point of mean coherence. **(b)** Coherence zone width.

Figure 3.10: Mean differences between the PMC and CZW values measured in Experiment 1 (first), in Experiment 2 using the staircase method (stair) and in Experiment 2 using the self-tuning method (self). Error bars indicate the 95% confidence interval of the mean.

The width of the confidence intervals for the mean difference between PMC values obtained in Experiment 1 and Experiment 2, both using the staircase method, were 3.18 and 8.79 deg/s^2 for visual amplitudes of 12 and 30 deg/s^2, respectively. For the CZWs these values were 7.81 and 8.44 deg/s^2 for visual stimulus of 12 and 30 deg/s^2, respectively.

The confidence intervals were normalized with the amplitude of the visual stimulus and then averaged across amplitudes. The average confidence interval widths in percentage of the visual stimulus were 28% and 47% for the PMC and CZW, respectively. These values represent the mean difference between the two experiments using the same method and were chosen as the δ values to be used in the TOST.

Table 3.1 shows the results of the TOST for the normalized difference in PMCs and CZWs. The TOST indicates that both null hypotheses can be rejected and thus both methods may be considered equivalent for a difference of 28% and 47% in the values of the PMC and CZW, respectively.

The new method was neither faster nor slower than the staircase method in terms of the number of runs needed to converge. So, although from the experimenter point of view there was no advantage in using this method, when asked about it, all participants preferred the self-tuning method over the staircase method. Subjects answered that the new method was nicer and more motivating.

Table 3.1: TOST results for the comparison of the staircase method with the self-tuning method, where ** is highly significant ($p < 0.01$), * is significant ($0.01 \leq p < 0.05$), and - is not significant ($p \geq 0.05$).

		$M1 - M2 > \delta$			$M1 - M2 < -\delta$		
	δ (%)	df	t	sig.	df	t	sig.
normalized diff. PMC	28	4	-2.2122	*	4	5.5073	**
normalized diff. CZW	47	4	-2.1697	*	4	7.4078	**

3.5 Discussion

Two experiments were performed to measure perception coherence zones in yaw for different amplitudes and using two different measuring methods for each of the conditions. From the measured data a point of mean coherence and a coherence zone width were calculated.

3.5.1 The experimental method

The method used in the first experiment was quite similar to the one used in the original experiment of Van der Steen (1998), with a few changes, among them, the question posed to the subjects. In the present experiment the participants had to make a judgment about the relative magnitudes of the visual and the inertial cue, whereas in Van der Steen's (1998) experiment they were asked to indicate a non-stationary outside world. The assumption that perceiving a mismatch in cues amplitude precedes the perception of a non-stationary outside world may explain why the measured coherence zones were narrower than the ones measured by Van der Steen (1998). The fact that the coherence zone measurements may vary according to the question posed does not harm the validity of the results or the method. The formulation of the question should follow from the definition of coherence zone. Using different definitions implies that different thresholds are being measured, which justifies the use of distinct questions. Despite all these considerations, no definite conclusions can be drawn as Van der Steen (1998) also used a different visual system that might also have influenced the visually perceived yaw velocity and consequently the width of the coherence zone.

In the second experiment a self-tuning method to measure coherence zones was validated. The intention was to create a procedure that was faster and easier for the participants. By avoiding subjects to become tired, distracted or bored, the data collection could become more reliable. The new method was neither faster nor slower than the staircase method in terms of number of trials per run. Nevertheless, all participants were positive about the self-tuning method, reporting that it was more motivating. This increased motivation for the subjects can help prevent boredom during the experiments.

3.5.2 The effect of amplitude

In both experiments the PMCs decreased below the one-to-one line for amplitudes of the visual cue above 12 deg/s^2. This indicates that at high visual amplitudes the participants preferred relatively lower inertial amplitudes. This result could arise from an artifact introduced by the simulator's motion system. However, after examining the data from the inertial measurement unit (IMU) mounted on the Simona simulator, very small differences were observed in the simulator motion performance at higher and lower amplitudes. For the higher amplitude conditions the turnaround bump was larger in absolute terms, but relative to the commanded acceleration amplitude, it was approximately the same at higher and lower amplitudes.

Another explanation could be that at higher amplitudes subjects were underestimating the visually perceived velocity. The GMCs presented by Van der Steen (1998), which were in general above one for angular motion, also indicate an underestimation of the visual velocity. However, for the particular case of yaw motion, Van der Steen (1998) registered GMCs above one only for amplitudes below 9.7 deg/s, which is not in agreement with what was found here. Conversely, Melcher and Henn (1981) have demonstrated that at high accelerations a visual scene is less optokinetic. Their results show that during visual stimulation only, the latency for the detection of self-motion increased for accelerations above 10 deg/s^2. They also reported that for increasing visual stimuli velocities, the self-velocity was underestimated. An underestimation of the visually perceived velocity could lead to the down tuning of the inertial motion. For further investigation into this effect several visual related aspects should be taken into account, such as the characteristics of the visual system, the information content of the visual scene and the process of vection.

The visual cue amplitude also had a significant effect on the CZW. In general, the CZW increased with increasing visual cue amplitude. Based on Van der Steen's (1998) work, this was an expected result. In the first experiment, this trend was observed only for visual cue amplitudes up to 18 deg/s^2. For higher amplitudes the CZW remained fairly constant. Apparently, for higher amplitudes of the visual stimulus, subjects become less accurate in their estimation of self-velocity but at the same time they are relatively more sensitive to deviations from the visually perceived velocity. No explanation for this result was found.

The present results suggest that signals with high amplitudes or high frequencies can be more attenuated than low amplitude profiles. This finding favors the efficient use of motion space. More specifically, high amplitude yaw motion can be tuned down to as much as 0.5, if considering the lower threshold of the coherence zone, or approximately 0.8, if considering the PMC.

3.6 Conclusions

The work of Van der Steen (1998) on coherence zones was extended in this chapter to better approximate realistic flight simulation scenarios. First, amplitudes closer to the ones found in vehicle motion were tested. Second, a measuring method which gave participants a more active role in the measurement procedure was validated.

The measured coherence zones were narrower than the ones measured by Van der Steen (1998) which might be due to the different question posed to the subjects and consequently the different definition of coherence zone used. The perception of a mismatch between inertial and visual cue, as used in this work to define the limits of the coherence zone, might precede the perception of a non-stationary outside world.

The self-tuning method used in Experiment 2 was not faster than the staircase method used in Experiment 1, but it was considered by the subjects as more motivating and a nicer task to perform. The results obtained with both methods were equivalent.

In the two experiments it was observed that the amplitude of the visual cue affected both the PMC and the CZW. The PMC and CZW increased with increasing amplitude of the visual cue, although for amplitudes above 18 deg/s^2 the PMC decreases below the one-to-one line and the CZW remains fairly constant.

Based on these results, and considering the lower threshold of the coherence zones, in a flight simulation scenario, high amplitude yaw motions might be down tuned to 0.5 of the visual cue.

4

The Effect of Frequency

This chapter is based on the following publications:

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

Valente Pais, A. R., Van Paassen, M. M., Mulder, M., and Wentink, M. (2010). Perception of Combined Visual and Inertial Low-Frequency Yaw Motion. In *Proceedings of the AIAA Modeling and Simulation Technologies Conference, Toronto, ON, Canada, August 2-5*, AIAA 2010-8093.

4.1 Introduction

In the previous chapter (Chapter 3) Van der Steen's work (Van der Steen, 1998) was extended by measuring yaw coherence zones at higher amplitudes, between 0 and 30 deg/s^2, using a step-like profile. To further advance our knowledge of coherence zones and to be able to apply it in a flight simulation context, also the effect of frequency on the coherence zones should be studied.

Van der Steen (1998) measured coherence zones in yaw using sinusoidal signals with frequencies between 1.2 and 2 rad/s and found that the stimulus frequency had no significant effect on the measured thresholds. However, as also remarked by Van der Steen (1998), the range of frequencies tested was rather small. For a better understanding of the effect of frequency on the combined perception of visual and inertial cues, more data should be collected for a larger range of frequencies.

This chapter describes two experiments where coherence zones are measured for frequencies as low as 0.2 rad/s and as high as 10 rad/s.

The first experiment investigates the effect of the stimulus frequency on the yaw coherence zones using the self-tuning method validated in Chapter 3. Measurements are made with three different motion profiles, a step-like profile and sinusoidal profiles with frequencies of 2 rad/s and 10 rad/s. The results are used to formulate an hypothesis regarding the influence of the SCC dynamics on the coherence zones.

The second experiment focuses on the low frequency range, using sinusoidal signals of 0.2 rad/s and 2 rad/s, and tries to verify the hypothesis posed based on the high frequency results. The results are analyzed in velocity and acceleration space.

This chapter is organized as follows. First, the method and results of the first experiment are described followed by the hypothesis posed based on these results. Afterwards, the second experiment method and results are presented and compared with those from the first experiment. At the end, a general discussion is presented, followed by conclusions.

4.2 Experiment 1

4.2.1 Method

4.2.1.1 Apparatus

All trials were conducted in the Simona Research Simulator and the same visual scene as described in both experiments of Chapter 3 was used.

4.2.1.2 Experimental design

The experiment had a two-way repeated measures design. The experimental conditions were defined by two visual amplitudes, 12 and 30 deg/s^2, and three stimulus profiles. Two of the profiles were sinusoids with frequencies of 2 and 10 rad/s. The third profile, included as a reference profile, was the step-like signal used in the previous two experiments. Both upper and lower thresholds were measured and each subject repeated each experimental condition 3 times. Every subject performed 36 experimental trials.

4.2.1.3 Motion and visual signals

The step-like profile was described before in Section 3.3.1.3. The other two profiles were sinusoids with frequency ω and maximum amplitude A defined by the experimental condition. The sinusoidal signals had a duration of 4 periods. The first and last periods were used to fade in and out. The whole acceleration signal, including the fade-in and fade-out phases, is defined in Equation (4.1), where $T = 2\pi/\omega$ and $\omega_s = \omega/2$.

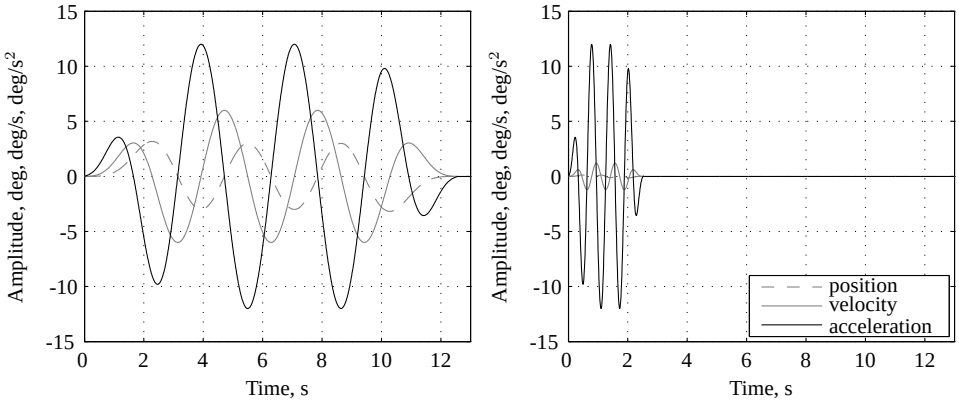
$$a(t) = \begin{cases} A\sin(\omega t)\left(0.5 - 0.5\cos(\omega_s t)\right) + \underbrace{A_c \sin(\omega_c t)}_{\text{compensation}} & , \quad \begin{array}{c} 0 < t \leq T \\ and \\ 3T < t \leq 4T \end{array} \\ A\sin(\omega t) & , \quad T < t \leq 3T \end{cases} \tag{4.1}$$

If the acceleration signal without the compensation term is integrated, a velocity signal is obtained that does not start at zero. A constant could be added to the velocity signal to compensate for the velocity initial value but that would result in a position signal that diverges with time. To prevent this situation, the compensation term was added. The amplitude $A_c = A/12$ and frequency $\omega_c = \omega/2$ were chosen such that the velocity signal starts at zero and is continuous at $t = T$.

Figure 4.1 shows an example of the signals for the two frequencies tested and an amplitude of 12 deg/s^2.

4.2.1.4 Procedure

The procedure followed was the same as the one used in Experiment 2 in Chapter 3 regarding the self-tuning method part. For a description of the procedure see Section 3.4.1.4. The order of the experimental runs was randomized per subject.



(a) Frequency of 2 rad/s. **(b)** Frequency of 10 rad/s.

Figure 4.1: Example of two motion profiles of different frequencies with a maximum acceleration amplitude of 12 deg/s^2.

4.2.1.5 Subjects and subjects' instructions

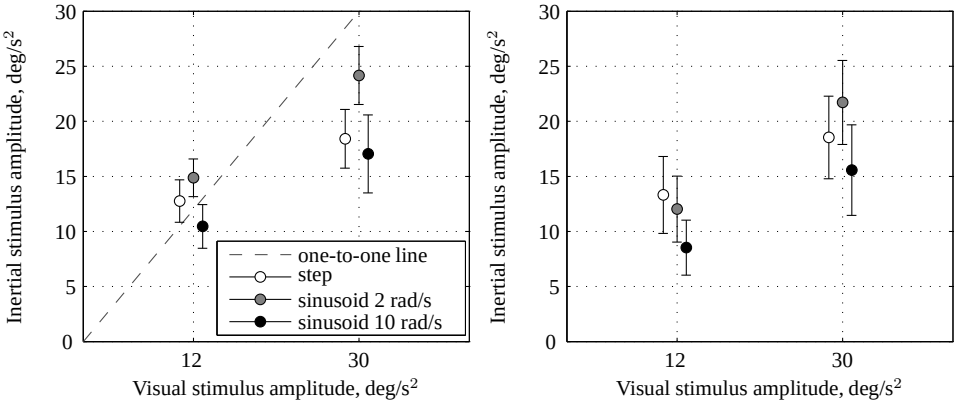
There were 8 participants, 6 males and 2 females, with ages between 23 and 34 years (mean age of 25.9 years).

The instructions given to the subjects were the same as for the self-tuning method part of Experiment 2 in Chapter 3. Subjects were told there would be different profiles of motion, some faster and some slower. Before the experiment, subjects performed 4 to 6 test runs.

4.2.2 Results

Again, as in Experiment 2 in Chapter 3, the obtained upper and lower thresholds were used to calculate the PMC and the CZW. They are presented in Figure 4.2 for the three motion profiles tested. In this and all subsequent figures the error bars indicate the 95% confidence interval of the adjusted means.

As in the experiments in Chapter 3, the PMC for the higher visual amplitude was significantly lower than the one-to-one point. Comparing both sinusoidal signals, the PMC was clearly lower for the higher frequency signal than for the lower frequency signal. The same trend can be observed in the CZW. With respect to the step-like signal, since this profile contains both high and low frequencies it was not obvious where the results with this profile would be with respect to the other two profiles. The PMC was somewhat in between both sinusoidal signals and the CZW for the higher amplitude was also in between the two sinusoids. The CZW for the low amplitude was higher than for the two sinusoidal profiles.



(a) Point of mean coherence. **(b)** Coherence zone width.

Figure 4.2: Measured coherence zones for two visual amplitudes and two stimulus frequencies.

An ANOVA was performed on the effect of the amplitude of the visual cue and the stimuli profile on the PMC and CZW. The results are summarized in Table 4.1.

Table 4.1: ANOVA results for the point of mean coherence (PMC) and the coherence zone width (CZW), where ** is highly significant ($p < 0.01$), * is significant ($0.01 \leq p < 0.05$), and - is not significant ($p \geq 0.05$).

Independent variables	Dependent measures					
	PMC			CZW		
	df	F	sig.	df	F	sig.
Amplitude	1, 7	132.87	**	1, 7	44.49	**
Profile	2, 14	7.25	**	2, 14	4.57	*
Amplitude × Profile	2, 14	2.78	-	2, 14	1.05	-

The amplitude had a highly significant effect on both the PMC and the CZW. The effect of the profile on the PMC was also highly significant. Post-hoc pairwise comparisons, using Bonferroni correction for the level of significance, show that the PMC for the 2 rad/s signal was significantly higher ($p < 0.05$) than the other two profiles.

The main effect of the profile on the CZW was also significant. However, the pairwise comparisons did not show any significant differences between the profiles. The effect of the frequency on the CZW may be related to the effect of the frequency on the PMC. It was

observed that for higher visual amplitudes, both the PMC and the CZW were higher. If for the 2 rad/s signal the PMC was higher, perhaps a higher CZW resulted only from a higher PMC and not necessarily from the effect of the stimulus frequency.

4.3 Hypotheses

An explanation for the different PMCs with different profiles can be found in the dynamics of the SCC. The hypothesis is made that the dynamics of the SCC are not taken into account in the internal comparison between perceived visual and perceived inertial motion, which implies that the perceived coherence zones will vary depending on the frequency of the stimuli used. To explain this dependency, the dynamics of the SCC are shown in Figure 4.3 together with three motion stimuli with frequencies of 0.2, 2 and 10 rad/s (black circles). The SCC dynamics are represented in terms of velocity input and acceleration input.

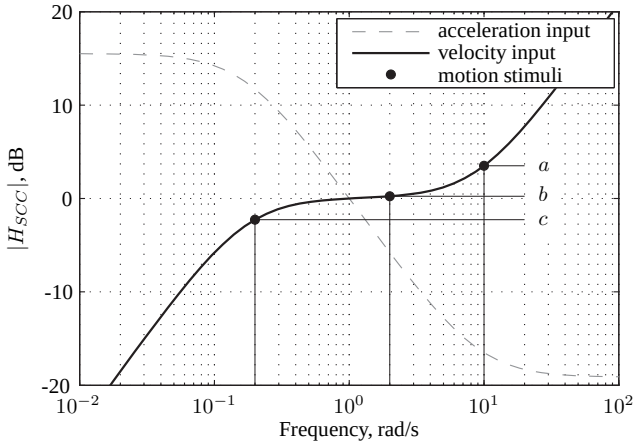


Figure 4.3: Semi-circular canals dynamics represented in terms of velocity input and motion stimuli with frequencies of 0.2, 2 and 10 rad/s (black circles).

The SCC model used here is the one fitted by Hosman and Van der Vaart (1980) through threshold measurements. The gain was set such that the model has a gain of one at the frequency of 1 rad/s. The SCC model for an angular velocity input and an output that is proportional to angular velocity is given by:

$$H_{SCC}(j\omega) = 5.972j\omega\frac{0.1097j\omega + 1}{5.924j\omega + 1} \tag{4.2}$$

For a frequency range between 1 and 9 rad/s the SCC dynamics equal approximately a gain of one. For frequencies higher than 9 rad/s the SCC introduce a gain higher than one

in the vestibular path signal. If the visual and vestibular signals are directly compared and there is no internal compensation for the gain introduced by the SCC dynamics, then the vestibularly perceived motion amplitude will be higher than the visually perceived motion amplitude. As a result of this perceived difference, subjects will require lower amplitudes of inertial motion than the amplitude that physically corresponds with the visual cue.

This reasoning is supported by the results of the previous study. The PMCs at 10 rad/s were 1.42 times higher than at 2 rad/s, which corresponds quite well with the ratio of 1.45 between the SCC gains at 10 rad/s and at 2 rad/s (in Figure 4.3, $a/b = 1.45$).

For the lower frequency range, since the SCC gain is lower, subjects will prefer higher amplitudes of inertial motion than for stimuli in the range of 1 to 9 rad/s. For the frequency of 0.2 rad/s the gain introduced by the SCC is lower than one so, the hypothesis is posed that subjects will prefer inertial motion amplitudes that are higher than the one-to-one match with the visual stimulus.

Although this line of reasoning seems to explain the previously obtained results, it depends on the assumption that the dynamics of the SCC are not being taken into account during the internal comparison of the visual and inertial signals. This would imply that there is no internal representation of the SCC dynamics or that this representation is not accurate for too high or too low frequencies.

4.4 Experiment 2

4.4.1 Method

4.4.1.1 Apparatus

To measure coherence zones in yaw with a low frequency stimulus, a very large position motion space is necessary. For this reason, experiment 2 was conducted in the Desdemona simulator at TNO Defence, Security and Safety in Soesterberg, The Netherlands, shown in Figure 4.4.

The simulator has an 8 meter linear track that can rotate around its central point, providing a 4 meter centrifuge arm. The structure mounted on the linear track consists of a 2 meter vertical linear track. A gimbaled structure mounted on the heave track allows the cabin to rotate more than 360 degrees in three orthogonal axes. Using common aeronautical nomenclature we name the rotations around the vertical axis “cabin yaw”. In this experiment only the cabin yaw was used. The cabin yaw can has an acceleration limit of 90 deg/s^2 and a velocity limit of 180 deg/s^2.

A PC-based computer generated image system was used to render the outside world. In the cabin, three computers generated real-time images with an update rate of 60 Hz. Three projectors (resolution: 1024 $\times$ 768 px) projected the image on a three-part flat screen, placed at approximately 1.5 m from the participants’ eyes, creating an out-of-the-window

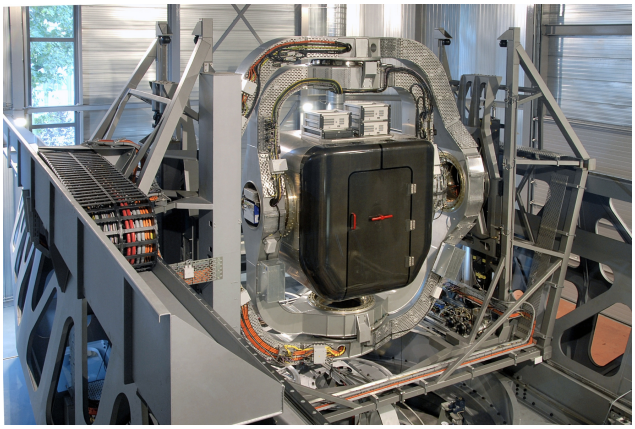


Figure 4.4: The Desdemona simulator. (Courtesy of TNO, Soesterberg.)

field-of-view of 120 degrees horizontal and 32 degrees vertical. The outside visual scene consisted of a view of an airport including some buildings, part of a runway and some grass fields. The viewpoint height was 5 meters.

4.4.1.2 Experimental design

In order to be able to compare the results of the present experiment with the ones from experiment 1 (Section 4.2), performed in the Simona simulator, some experimental conditions were chosen to match exactly the ones also used there. So, as in the Simona study, two experimental conditions were defined by two visual amplitudes: 12 and 30 deg/s^2, and one stimulus frequency: 2 rad/s.

Furthermore, to investigate the effect of frequency on the coherence zones in both velocity and acceleration space a combination of frequency and amplitude of the stimuli had to be found that allowed for a comparison between signals with two different frequencies and the same acceleration amplitude and two different frequencies and the same velocity amplitude. For the low frequency stimulus a frequency of 0.2 rad/s was chosen. Since due to simulator velocity limits it was not possible to perform the condition defined by amplitude 30 deg/s^2 and frequency 0.2 rad/s, one other amplitude value was defined: 3 deg/s^2. Using this amplitude at a frequency of 0.2 rad/s allowed for a direct comparison, in velocity space, with the condition with visual amplitude of 30 deg/s^2 and frequency of 2 rad/s, since both result in a velocity amplitude of 15 deg/s.

The visual scene was also chosen to provide similar references as in the Simona study. However, since the visual databases used were different and the outside scene was not exactly the same, it was thought that some investigation into the effect of the content of the

visual scene would be interesting as well. For this reason and with an exploratory character, a condition was added with a different visual scene consisting of a star field. At an amplitude of 12 deg/s^2 and a frequency of 2 rad/s, measurements were made under two different outside visual conditions: the airport scene and the star field, which consisted of white dots on a black background. In this way only visual flow was provided and there was no other elements that could give a real measure of distance or attitude. The lack of concrete references is thought to influence subject's estimate of visually perceived attitude(Sweet and Kaiser, 2005).

Table 4.2 summarizes the visual motion conditions. Note that only the condition with amplitude of 12 deg/s^2 and frequency of 2 rad/s was tested with both the airport scene and the star field.

Table 4.2: Experimental conditions defined in terms of acceleration amplitude and frequency of the visual stimuli and the resulting visual amplitude in deg/s.

Frequency, rad/s	Amplitude, deg/s^2		
	3	12	30
0.2	15	60	-
2	1.5	6	15

In total there were six different experimental conditions. For each condition both upper and lower thresholds were measured and each subject repeated each measurement 3 times which resulted in a total of 36 trials per subject.

4.4.1.3 Procedure

Thresholds were measured using an in-the-loop self-tuning procedure also used in reference (Correia Grácio et al., 2010). While being subject to a continuous sinusoidal yaw motion both visually and inertially, subjects were asked to use a joystick to change the inertial motion gain.

The joystick gain was different for the two frequencies and it was chosen as a compromise between control sensitivity and tuning procedure duration. A high joystick gain would cause abrupt or fast changes in the inertial motion amplitude, possibly resulting in discomfort or making it very difficult for subjects to fine-tune the motion gain. A too low gain would result in very long experimental trials. The joystick gain was set such that at the maximum positive deflection, the joystick signal would cause the velocity signal amplitude to double in approximately 84% of the stimulus period. The joystick signal was integrated and to further smooth the transition between inertial motion gains a second-order low-pass filter ($\omega_n = 5, \zeta = 1$) was applied. The output of the second-order filter was then applied to

the inertial velocity signal. Additionally, some logic was applied to the motion gain signal to prevent negative gains or too high gains that would cause the simulator to reach its motion limits.

For measurements of the upper threshold, participants were asked to increase the motion gain as much as possible while still perceiving the inertial motion amplitude to match the visually perceived motion. For measurements of the lower threshold, they were asked to decrease the motion gain as much as possible while still perceiving matching inertial and visual motion amplitudes. When the desired motion gain was achieved, subjects pressed the trigger button on the joystick to end the trial. A second press on the trigger indicated they were ready for the next trial. The initial motion gain was a random value between 1.15 and 0.85.

4.4.1.4 Subjects and subjects' instructions

Seven male subjects and one female subject participated in the experiment. Their ages were between 21 and 28 years, with an average of 24.63 years.

Participants were seated and held the joystick in their right hand. They were asked to keep their head as much as possible in the seat's head-rest and refrain from making head movements, but they were free to scan the outside view at will. They wore a head-set on which white noise was played to mask simulator noise. The head-set also allowed them to communicate with the experiment supervisor at all times.

Subjects were instructed to make small adjustments to the motion gain and after every adjustment let the motion profile run for at least one full period. They were also advised to adjust the motion gain until they were outside their coherence zone, that is, while measuring an upper threshold they should increase the motion gain until they felt the inertial motion was too strong. From that point they should slowly decrease the motion gain until it felt congruent again.

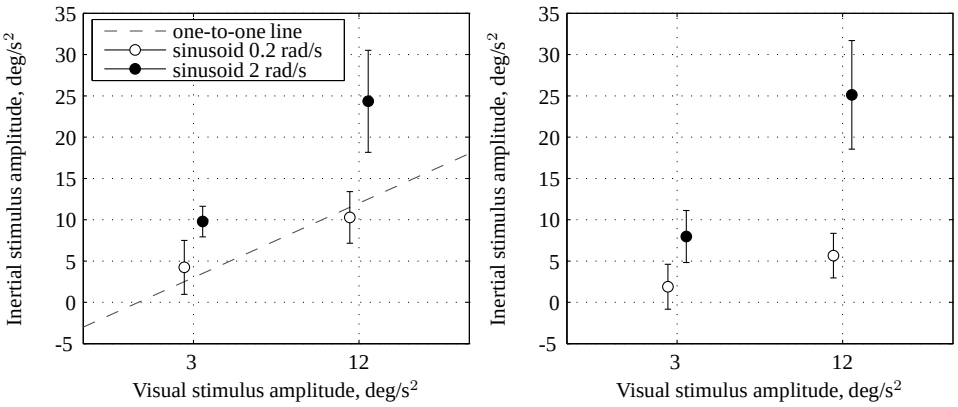
4.4.2 Results

Threshold values were obtained by computing the maximum inertial motion amplitude during the last period of the sinusoidal signal in each trial. From the upper and lower threshold values a coherence zone width (CZW) and a point of mean coherence (PMC) were calculated using Equation (4.3) and Equation (4.4), respectively.

$$CZW = th_{up} - th_{lo} \tag{4.3}$$

$$PMC = th_{lo} + \frac{CZW}{2} \tag{4.4}$$

Figure 4.5 shows the PMC and CZW for two visual amplitudes and two frequencies. Values are presented in acceleration. In this and all subsequent figures the error bars indicate the 95% confidence interval of the adjusted means.



(a) Point of mean coherence. **(b)** Coherence zone width.

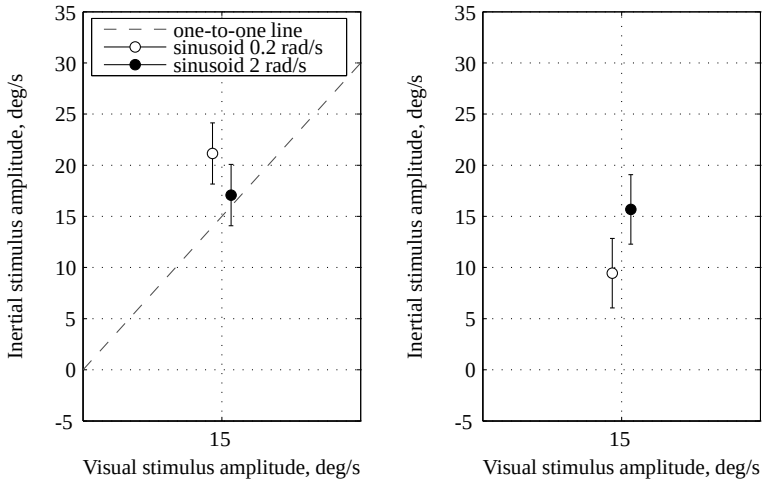
Figure 4.5: Measured coherence zones for two visual amplitudes and two stimulus frequencies.

An ANOVA was performed on the effect of the amplitude of the visual cue and the stimulus frequency on the PMC and CZW. The results are listed in Table 4.3. Similar to previous experiments, both amplitude and frequency had a significant effect on the PMC and CZW. However, a significant interaction between frequency and amplitude was found. At the amplitude of 12 deg/s^2 and frequency of 2 rad/s the PMC increases with respect to the one-to-one line, instead of slightly decreasing, as observed for the lower frequency. The same trend can be observed for the CZW.

From Figure 4.5 it can also be observed that the PMCs with the lower frequency stimulus were not larger than the PMCs with the higher frequency as it was expected. However, if the same results are represented in velocity space, as shown in Figure 4.6, the reverse situation is observed. Here only the two conditions with a velocity amplitude of 15 deg/s are shown.

Although the PMC with the lower frequency stimulus is higher than with the higher frequency stimulus, comparison using a T-test showed no significant difference between both conditions, $t(7) = 1.605, p > 0.05$. The CZW values show the inverse trend, with the higher frequency condition having a larger CZW than the lower frequency condition. However, also here the T-test showed that the difference is not significant, $t(7) = -2.166, p > 0.05$.

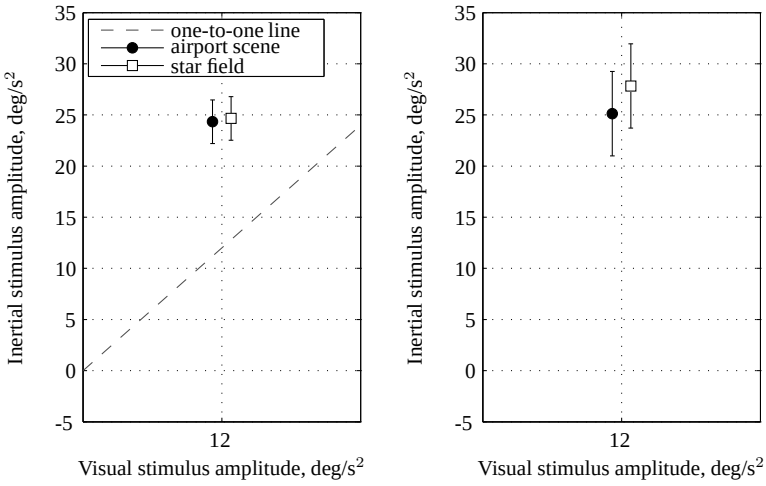
Figure 4.7 displays the PMCs and CZWs for the two conditions with different outside visuals.



(a) Point of mean coherence.

(b) Coherence zone width.

Figure 4.6: Measured coherence zones for two conditions with the same velocity amplitude and different frequencies.



(a) Point of mean coherence.

(b) Coherence zone width.

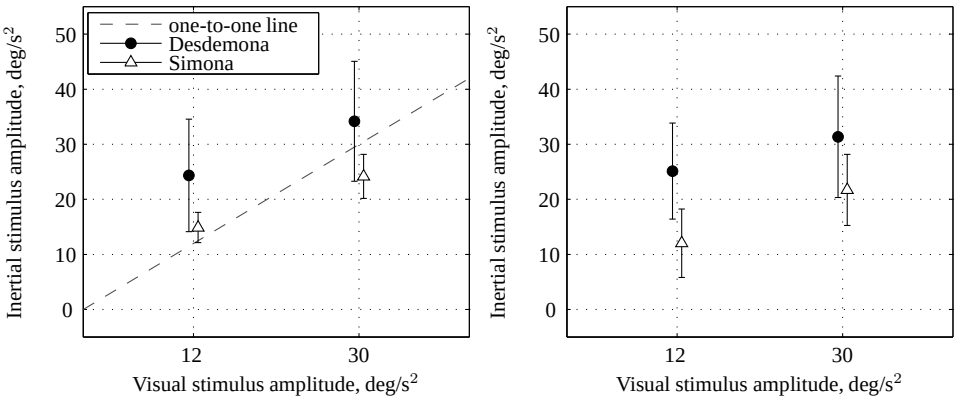
Figure 4.7: Measured coherence zones for a stimulus frequency of 2 rad/s and two different outside visual scenes.

Table 4.3: ANOVA results for the point of mean coherence (PMC) and the coherence zone width (CZW), where ** is highly significant ($p < 0.01$), * is significant ($0.01 \leq p < 0.05$), and - is not significant ($p \geq 0.05$).

Independent variables	Dependent measures					
	PMC			CZW		
	df	F	sig.	df	F	sig.
Amplitude	1, 7	33.13	**	1, 7	36.26	**
Frequency	1, 7	15.50	**	1, 7	33.80	**
Amplitude × Frequency	1, 7	10.73	*	1, 7	10.46	*

The PMC and CZW for the two visual scenes were fairly similar and the T-tests did not indicate a significant difference between the two conditions: $t(7) = -0.178, p > 0.05$ for the PMC and $t(7) = -0.776, p > 0.05$ for the CZW.

The conditions with visual cue amplitude of 12 and 30 deg/s^2 and frequency of 2 rad/s were compared to the ones from the previous study in the Simona simulator. In the latter, the same stimulus characteristics were used, although a slightly different tuning method was used and, as mentioned before, different visual display characteristics. Figure 4.8 shows the means and 95% confidence interval of the PMC and CZW for the experimental results in the two simulators, Desdemona and Simona.



(a) Point of mean coherence. **(b)** Coherence zone width.

Figure 4.8: Measured coherence zones for a stimulus frequency of 2 rad/s and two visual amplitudes in two different experiments in the Desdemona and the Simona simulators.

The differences between results on both simulators are quite large, with the PMC and CZW measured in Simona lower than the ones measured in Desdemona. Also, the 95% confidence intervals are considerably larger for the experiment in Desdemona. Table 4.4 shows the results of the ANOVA, where the amplitude of the visual cues was considered a within-subjects effect and the simulator used as a between-subjects effect. To guarantee homogeneity of variance, the PMC data were transformed to the reciprocal and the ANOVA was performed on the transformed data.

Table 4.4: ANOVA results for the point of mean coherence (PMC) and the coherence zone width (CZW), where ** is highly significant ($p < 0.01$), * is significant ($0.01 \leq p < 0.05$), and - is not significant ($p \geq 0.05$).

Independent variables	Dependent measures					
	PMC			CZW		
	df	F	sig.	df	F	sig.
Amplitude	1, 14	39.37	**	1, 14	14.13	**
Simulator	1, 14	3.77	-	1, 14	6.33	*
Amplitude $\times$ Simulator	1, 14	1.92	-	1, 14	0.67	-

The statistical test revealed a significant difference between the two simulators for the CZW but not for the PMC. As expected, the amplitude had a significant effect on both the PMC and CZW and there were no interaction effects between amplitude and simulator.

4.5 Discussion

4.5.1 The imperfect internal representation hypothesis

The first experiment investigated the effect of the stimulus frequency on the coherence zones. The PMCs for the 10 rad/s signal were significantly lower than for the 2 rad/s signal. At the higher frequency subjects preferred inertial motion amplitudes that are lower than the physically correspondent visual cue amplitudes. This finding was explained using the dynamics of the SCC: at higher frequencies, above 9 rad/s the velocity signal is more amplified than at the middle frequencies (from 1 to 9 rad/s). At higher frequencies a higher gain is introduced in the vestibular sensory path but not in the visual path. When the visual and vestibular signals are compared internally, the vestibular signal will be perceived as having higher amplitude than the visual signal. When asked to match the amplitude of the inertial signal with that of the visual, subjects will then tend to tune the inertial signal down.

It is important to note that we are assuming that the comparison between the visual and the vestibular paths is done in terms of velocity. Although there is no clear evidence to support or reject this assumption, such a representation is compelling since visually we can perceive velocity (Hosman, 1996) and the output of the SCC, for a frequency range between approximately 1 and 9 rad/s, is also proportional to velocity (Groen and Bles, 2004).

The experimental design, however, considered the acceleration amplitude to be an independent factor. The choice to keep the acceleration amplitude, and not the velocity amplitude, constant across different frequencies was based on the fact that, although the SCC output is proportional to velocity, they are sensitive to accelerations (Van Egmond et al., 1949). The 2 and 10 rad/s signals had equivalent amplitudes in acceleration. In terms of velocity, the amplitudes have to be divided by the signals' frequencies. This resulted in different velocity amplitudes depending on the frequency of the profile, making it difficult to draw definite conclusions over the effect of velocity on the coherence zones. Perhaps in future work it would be useful to consider different frequency stimuli at fixed velocity amplitudes.

Qualitatively, the gain of the SCC at the two frequencies could explain the differences found in the PMCs. From a modeling point of view, this implies that the conflict detection mechanism does not take into account the SCC dynamics in its totality. The increased gain for higher frequencies could be accounted for, by for example, passing the visual signal through an internal model of the SCC canals, as proposed by Zacharias and Young (Zacharias and Young, 1981) and later also modeled by Telban and Cardullo (Telban and Cardullo, 2001). In that case, the conflict detection between the signals from the visual and from the vestibular path would be independent of the stimulus frequency, with the exception maybe for a frequency dependency of vection.

Since there seems to be a frequency dependency on the visual-vestibular conflict detection, it can be suggested that the internal representation of the SCC dynamics is not perfect. For frequencies between 1 and 9 rad/s the SCC function approximately like a velocity feedthrough, or, if an acceleration input is considered, like an integrator. It could be argued that for the lower and higher frequencies, which are maybe not that common in normal head movements, the internal representation of the SCC is still just an integrator, or a velocity feedthrough. That is to say that for frequencies less common in natural head movements the SCC internal representation is simply extended from the dynamics at more "common" frequency ranges.

One way to test these assumptions would be to pre-filter the inertial motion signal with the inverse dynamics of the SCC, so that the total system of pre-filter and SCC dynamics would become the perfect velocity sensor at all frequencies, for an example, see Wentink et al. (Wentink et al., 2009). In such a case, the PMC should not be influenced by the stimulus frequency.

The effect of the stimulus frequency on coherence zones was also tested by Van der Steen (Van der Steen, 1998). Although he found no significant results, it should be reminded that the range of frequencies tested, from 1 to 2 rad/s, was within the range in which the SCC function as a velocity feedthrough. Besides the work of Van der Steen, no other study could be found that directly showed the effect of the stimulus frequency on the perception of coherent visual and inertial cues. However, it has been shown (Melcher and Henn, 1981; Mergner and Becker, 1990) that vection is a low frequency process that complements the high frequency characteristics of the vestibular system. As a consequence, at the higher frequency profile vection might not be occurring, making the judgment of coherent visual and inertial cues more difficult.

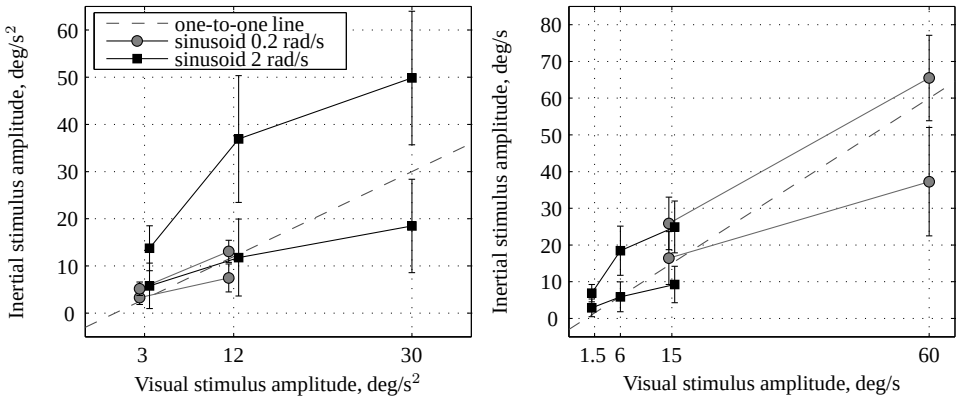
4.5.2 Testing the hypothesis at low frequencies

In the second experiment, like in previous studies the perceived coherence zones were influenced by the amplitude of the visual cue and the frequency of the stimuli. However, unlike expected, at the higher amplitude and higher frequency condition, the PMC and CZW values increased with respect to the one-to-one line. This condition also presented the larger 95% confidence interval, indicating perhaps that subjects had more difficulty or were less consistent in determining the limits of the coherence zone for this condition.

The comparison between the two frequencies in terms of acceleration did not show an increase in PMC for the conditions with lower frequency, as was hypothesized. However, when representing the conditions and respective PMC and CZW values in velocity, the expected trend is visible. For the lower frequency, the PMC was higher than for the higher condition. Although this is in agreement with the assumption that the gain imposed by the dynamics of the SCC is not taken into account during the internal comparison of the visual and inertial paths, definite conclusions can not be made, since the difference between the PMCs proved to be statistically not significant.

Also interesting to note is the inverse trend of the CZW which shows a lower value for the lower frequency than for the higher frequency. It has been observed in Chapter 3 and in Experiment 1 that the CZW increase is roughly proportional to the increase in PMC, similar to what can be seen in the present results when PMC and CZW values are shown in acceleration space. When represented in velocity space, however, the CZW seems to be less closely related to the PMC. This result might suggest that the perceived inertial velocity amplitude is being taken into account for the internal comparison with the visual signal and the determination of a subjective one-to-one match, but that the perceived acceleration level influences how much subjects allow a deviation from that one-to-one match. One other possibility would be that for the lower frequency condition, which has also the lowest amplitude in terms of acceleration, 3 deg/s^2, the absolute motion perception threshold is influencing the lower bound of the coherence zone. For a better insight into this effect, in

Figure 4.9, the coherence zones are plotted in terms of upper and lower threshold for both acceleration and velocity amplitudes.



(a) Acceleration. **(b)** Velocity.

Figure 4.9: Upper and lower thresholds for all conditions displayed in both acceleration and velocity units.

As can be seen in Figure 4.9b, the upper thresholds for both conditions are quite similar whereas the lower thresholds differ, with the condition with the lowest frequency showing the highest value. Figure 4.9a shows that the lower threshold for the lower frequency condition approaches 3 deg/s^2. Although this value is probably still well above the sensory threshold for this frequency (Heerspink et al., 2005; Valente Pais et al., 2006), it might be too low to allow subjects to make a fair comparison with the perceived visual motion. If the lower bound of the coherence zone is in fact being influenced by the inertial sensory threshold, then, not only the decreased CZW but also the increased PMC for the lower frequency condition might be a direct consequence of this. In that case, the perceived one-to-one match might be lower than what the PMC indicates, but the upper and lower thresholds are not symmetrically distributed around the subjective one-to-one point. This would refute the hypothesis that the coherence zones are influenced by the SCC dynamics.

To avoid a biased measurement of the coherence zone lower threshold, higher acceleration amplitudes should be chosen. However, it is difficult to find a suitable velocity signal amplitude for two such different frequencies as 0.2 and 2 rad/s that still results in acceleration signals that are comfortable for the subjects and fit within the simulator's velocity and acceleration motion space. One other possibility would be to choose frequency values closer to each other, for example, 0.3 and 1 rad/s. At closer frequencies it is easier to find an appropriate velocity amplitude, although the gains introduced by the SCC are also closer. If the

gains are similar, then any existing difference between the coherence zones measured at the two frequencies will be more difficult to observe due to intra- and inter-subject variability.

4.5.3 Comparison of the high and low frequency studies

The comparison of the results obtained in the Simona and in the Desdemona simulators shows a considerable difference between the two simulators. Although the values of the PMC and CZW obtained in Desdemona were much higher that the ones from Simona, statistically only the CZW values were considered to be significantly affected by the simulator factor. The data from Desdemona had also a much higher variance resulting in a larger confidence interval.

The values for the PMC found in the Desdemona simulator were considerably higher than the one-to-one line, indicating an overestimation of the visually perceived self-velocity or an underestimation of the inertially perceived self-motion. As the visual display in the Desdemona simulator is much closer to the observer than in the Simona simulator, and there is no collimation, the objects in the visual scene might be perceived as closer to the observer then in Simona. If that is the case, than the displacement of these objects at a certain velocity might be interpreted as a higher self-velocity than if the objects are perceived as being further away from the observer.

Moreover, the proximity of the display in the Desdemona simulator will result in a higher sensitivity to head movements. For a shorter focal distance, each head movement will result in a displacement of the objects displayed with respect to the subject's head, when in fact, these objects are represented as being far away, and hence, should move much less with each head movement. This motion of the represented objects with each head movement might also lead to the perception that objects are closer to the subject. When displacements due to head movements are combined with the motion of the visual stimulus, objects might be perceived as moving at a higher velocity.

This overestimation of the visually perceived motion would then lead to a higher matching inertial motion, explaining the higher values found for the PMC.

Although the two visual display systems differed also in other aspects, collimation might be the one playing the most important role in the subjective estimation of self-velocity. Chung et al. (Chung et al., 2003) investigated the effect of collimation, field-of-view and resolution on the pilot's control performance and subjective ratings during a hover task. They concluded that collimation had a positive influence on lateral velocity control and pilot's perceived sense of motion. A wider field-of-view showed added value for position and attitude control and resolution had no significant effect. So, whereas the field-of-view might be important for the perception of orientation with respect to the world, absence or presence of collimation might influence subjects' perception of self-velocity and overall sense of motion.

Besides the difference in visual system, also the visual scene was slightly different, although if the results from the comparison between the star field and airport scene are to be trusted, the specific visual scene might have less influence than the visual system characteristics.

One other important difference between the two studies was the experimental method used. Both relied on a self-tuning method, although in Simona that was accomplished by letting subjects increment and decrement the inertial motion amplitude in consecutive trials, whereas in Desdemona tuning was done in-the-loop and the inertial motion gain could be changed continuously. Nevertheless, the effect of amplitude on the coherence zones showed the same trend in both simulator studies and was independent of the simulator used.

4.5.4 Application to flight simulation

The present results suggest that to obtain the same level of perceived coherence, signals with high amplitudes or high frequencies can be more attenuated than low amplitude, low frequency profiles. Attenuating high frequencies leads to a low-pass filtering action, which is contrary to what is done in most motion cueing algorithms. Since low frequency signals use more motion space than high frequency ones, most algorithms high-pass filter inertial accelerations and use tilt-coordination and visual stimuli for low frequency cueing. Thus, the presented trends are not in line with what is currently done in flight simulation.

Conversely, with respect to the amplitude of the inertial cue, the findings do favor the efficient use of motion space. More specifically, high amplitude, high frequency yaw motion can be tuned down to as much as 0.3, if considering the lower threshold of the coherence zone, or approximately 0.6, if considering the PMC. However, care should be taken when extrapolating the trends found here to other motion profiles. The step-like profile tested, for example, although containing frequencies above the 10 rad/s, resulted in PMCs and CZWs somewhat in between the two sinusoidal signals. This indicates that when moving from motion profiles with one frequency to signals with a more complex frequency content, as it happens in flight simulation, then the perception of coherent visual and inertial cues might not follow the amplitude and frequency trends found with simple sinusoids. Even assuming that the effects of frequency and amplitude remain the same for more complex signals, then still the phase of the different frequency components has to be taken into account. By solely changing the phase values one can modify the time signal such that, for example, peak amplitudes vary considerably.

When transferring the current findings to motion cueing applications, not only the above mentioned effects should be considered but also the degree of freedom. In pitch and roll and during linear accelerations dedicated measurements should be made, since other inertial sensors play a role. Moreover, the combined stimulation in more than one degree of freedom might also influence the perception of coherent visual and inertial cues.

4.6 Conclusions

The effect of stimulus frequency on the coherence zones was investigated in two experiments performed in the Simona and the Desdemona simulators.

In both experiments it was observed that the amplitude of the visual cue influenced both the PMC and the CZW.

In the first experiment, the frequency of the stimuli had a significant effect on the PMC and the CZW. For a medium frequency stimulus (2 rad/s), the PMC and the CZW were higher than for the high frequency stimulus. A model of the SCC was used to qualitatively explain these trends and an hypothesis was posed regarding an imperfect internal representation of the SCC dynamics.

In the second experiment this hypothesis was tested. If the coherence zone measurements are observed in velocity space, then the trend of the PMC for lower stimulus frequency seems to support the hypothesis that the SCC dynamics influence the perceived coherence between visual and inertial cues. However, the values found for the CZW and the individual lower and upper thresholds indicate that there might be an effect of the inertial sensory thresholds for the condition with the lowest acceleration amplitude. For this reason the main hypothesis could not be confirmed.

For further research into this effect, care should be taken in the choice of suitable stimulus frequencies and amplitudes. At low frequencies higher acceleration profiles will have high velocity amplitudes which might be uncomfortable for subjects and/or reach simulator limits. The effect of frequency on perception coherence zones is, for this reason, easier to investigate at the middle frequency range (1 to 9 rad/s) and high frequency range ($>$ 9rad/s).

The comparison of the two studies in different simulators showed large differences in the PMC and CZW. This might seem discouraging for motion perception research in motion simulators, since results in one simulator can apparently not be easily transposed to other simulators. Nevertheless, if the focus remains on the trends rather than on the absolute values, then the conclusions drawn are similar for both studies: coherence zone widths increase with increasing visual cue amplitude, and at higher amplitudes the point of mean coherence tends to lower with respect to the one-to-one line.

The visual scene content had no influence on the measured coherence zones. However, the particular characteristics of the visual system might have a large influence on the exact values of the measured thresholds, and, as mentioned before, might explain the large differences found between the coherence zones measured in the two simulators.

These findings imply that in a vehicle simulation scenario high amplitude yaw motions might be tuned down to approximately 0.6 of the visual cue and that high frequencies can be attenuated more than low frequencies. However, in a full motion scenario with cues that occur simultaneously in more than one degree of freedom, this value might change. Further

research on this topic should extend the range of frequencies and amplitudes tested as well as add motion cues in other degrees of freedom.

Moreover, to move towards a pilot-in-control situation, the effect of workload on the perception coherence zones should also be investigated. During a vehicle simulation scenario, subjects attention is generally dedicated to manual or supervisory control tasks, or both. In such a scenario, the attention given to the perception task of matching perceived visual and inertial cues decreases, which might influence the boundaries of the perception coherence zone.

5

Phase Coherence Zones

This chapter is partially based on the following publication:

Jonik, P. M., Valente Pais, A. R., Van Paassen, M. M., and Mulder, M. (2011). Phase Coherence Zones in Flight Simulation. In *Proceedings of the AIAA Modeling and Simulation Technologies Conference, Portland, OR, USA, August 8-11*, AIAA 2011-6555.

5.1 Introduction

Motion filters distort not only the amplitude of the inertial cues but also the phase. For a complete assessment of the effect of motion filtering on pilot motion perception, and more specifically, on the perception of combined visual and inertial cues, both amplitude and phase should be taken into account.

Several studies have investigated the effect of motion filtering on pilot performance (Shirachi and Shirley, 1977; White and Rodchenko, 1999; Atencio, Jr., 1993; Schroeder, 1996; Grant et al., 2006), pilot manual control behavior (Pool, 2012), handling qualities (Chung and Schroeder, 1997; Damveld, 2009) and pilot motion acceptance and perception (Schroeder, 1996; Grant et al., 2006). Using different metrics, they related filter gain and break-frequency to a desired level of motion fidelity. These gains and break-frequencies can be related to the amplitude attenuation and phase distortion of the inertial cue with respect to the visual cue. However, a specific filter gain will correspond to different levels of attenuation at different frequencies. The same way, a specific break-frequency value will lead to different levels of phase distortion depending on the stimulus frequency that is being considered.

When measuring pilot motion acceptance with different motion filter settings, both amplitude and phase mismatches occur and it is not clear which parameter will cause a specific setting to be rated as good or bad motion. Moreover, similarly to what was found for amplitude mismatches(Valente Pais et al., 2010a), the perception of phase distortion between cues may depend on the frequency of the stimuli. This frequency dependence behavior cannot be captured with an approach that lumps together different phase distortion values under one break-frequency value. To better understand the mechanisms behind perception of phase mismatches between inertial and visual cues, and as it has been done for amplitude mismatches, the phase coherence zone should be measured at different stimulus frequencies.

There are few studies that directly measured phase mismatches between inertial and visual cues. Earlier works mainly focus on time-delays resulting from technical limitations of the visual and inertial systems rather than phase mismatches due to the effect of motion filters (Frank et al., 1988). For example, Miller and Riley (1977) measured the effect of time delays between the inertial and visual cues on the performance of one subject during a tracking task. They concluded that performance was significantly decreased when the inertial cue lagged the visual cue by 300 ms and their subjects tended to become nauseated for time delays larger than 200 ms. They related the measured time-delay values to phase differences based on the short period frequency of the controlled aircraft model, which was 2.38 rad/s. However, it is not clear that the controlled aircraft, and hence also the simulator inertial motion, was in fact sinusoidal.

The tracking task consisted of following a target aircraft which moved vertically with a frequency of 0.21 rad/s and the subject had control of the own aircraft in pitch and roll.

The frequency of motion in the vertical plane was thus very dependent on the pilot's control inputs and the response of the aircraft model. Moreover, they tested mostly conditions with the inertial motion lagging the visual motion, and not, as it could be expected by the use of a high-pass motion filter on the inertial channel, the inertial cue leading the visual cue. Furthermore, since they only used one subject and there was motion in more than one degree of freedom, it is difficult to generalize and use these results to determine a phase coherence zone.

Grant and Lee (2007) directly measured detection of phase-error thresholds for sinusoidal pitch motion. Although not explicitly stated by the authors, they have measured phase coherence zones. They have determined a phase-error threshold below which the visual and inertial cues were perceived as synchronized, although there was a phase difference between them. Correspondingly, phase differences above the phase-error threshold lead to the perception of incoherent, unsynchronized visual and inertial cues. Since generally motion filters provide phase lead to the inertial cue, they have considered only the cases where the inertial cue leads the visual cue. This lead to only one phase-error threshold, as opposed to other coherence zone measurements which have a lower and an upper threshold. Nevertheless, to all phase values between 0 deg and the measured phase-error threshold may be considered a phase coherence zone. Grant and Lee (2007) measured phase-error thresholds for pitch motion with frequencies of 1.257 rad/s and 6.283 rad/s and amplitudes of 2.29 deg/s and 5.73 deg/s. They used a flight simulator for the inertial motion and provided the visual cue through a head-mounted display. The mean phase-error threshold found was 57 deg.

The goal of the present experiment is to extend Grant and Lee's work (2007) to higher amplitudes and frequencies and establish a basis for a theoretical model which could be extrapolated to other conditions. Although they used pitch motion, in the present work yaw is used. By using yaw motion instead of pitch, the effect of gravitational cues is avoided, which may simplify the interpretation of the results. Nevertheless, in order to have a direct comparison with Grant and Lee's data (2007), two conditions are performed with pitch motion.

This chapter first presents the hypotheses posed with respect to the effect of amplitude and frequency on the phase-error thresholds, followed by a description of the experimental method used. The results are shown and compared with the data from Grant and Lee (2007). The chapter ends with a general discussion and conclusions.

5.2 Hypotheses

Human perception of stimuli synchronization can be regarded as a time dependent or a phase dependent mechanism. One might argue that the detection of unsynchronized visual

and inertial cues is simply a matter of identifying the occurrence of two events in time. If that is so, there would be no reason to believe that the frequency of the stimuli could change that time-based threshold. That is to say, the perception threshold is independent of the frequency of the motion signals. However, both visual and inertial stimuli are not sharp events that occur at a specific point in time. Rather, they should be considered as stimuli that may continuously vary in time. In fact, the judgment of how well both stimuli are synchronized should occur during stimulation. For this reason, it is hypothesized that subjects will act like phase detectors, rather than time-delay detectors.

This hypothesis states that subjects are sensitive to phase differences between two sinusoidal motion stimuli, that is, the measured thresholds will be phase-error thresholds. These phase-error thresholds may vary depending on the frequency and amplitude of the stimuli. The phase-error thresholds obtained by Grant and Lee (2007) will be used to make further considerations on these effects.

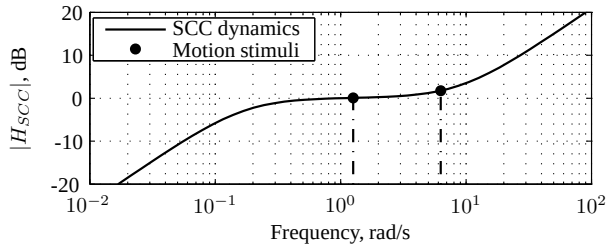
Grant and Lee (2007) measured phase-error thresholds for pitch motion with the inertial motion leading the visual motion. They used sinusoidal motion signals with two different frequencies, 1.257 rad/s and 6.283 rad/s, and two different amplitudes, 2.29 deg/s and 5.73 deg/s. This resulted in four experimental conditions which they tested twice, once using a motion gain of 1 and once using a motion gain of 0.5 applied to the inertial motion. They also investigated the effect of visual scene complexity by testing all the conditions with a simple visual consisting of a horizon line and a more complex visual scene which included houses and trees. In total, they measured phase-error thresholds for 16 conditions. A summary of their results is given in Table 5.1.

Table 5.1: Average phase-error thresholds obtained by Grant and Lee (2007).

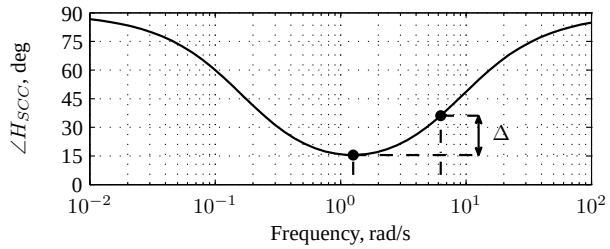
Motion gain	Visual complexity	Frequency, rad/s	Amplitude, deg/s	
			2.29	5.73
0.5	Low	1.257	87	50
0.5	Low	6.283	68	46
0.5	High	1.257	82	45
0.5	High	6.283	60	41
1.0	Low	1.257	70	38
1.0	Low	6.283	46	32
1.0	High	1.257	54	40
1.0	High	6.283	66	38

Based on these results two opposing hypotheses may be posed on the effect of frequency on the phase-error thresholds.

The first hypothesis is based on the results from the conditions with the lowest amplitude (2.29 deg/s). For these conditions, the phase-error thresholds were approximately 20 deg higher for the lower frequency than for the higher frequency (except for the two conditions with a motion gain of 1 and high visual complexity scene). Similar to what was done for amplitude coherence zones, this result can be explained by the dynamic characteristics of the SCC. Figure 5.1 shows the SCC dynamics in terms of magnitude and phase, as described by the model in Equation (5.1) (Hosman and van der Vaart, 1980).



(a) Magnitude.



(b) Phase.

Figure 5.1: Semi-circular canals dynamics with indication of the experimental stimulus frequencies in Grant and Lee (2007).

$$H_{SCC}(j\omega) = 5.972j\omega\frac{0.1097j\omega + 1}{5.924j\omega + 1} \tag{5.1}$$

As can be seen in the phase plot, Figure 5.1b, the SCC introduce an additional phase lead (indicated by Δ) at the frequency of 6.283 rad/s when compared to the frequency of 1.257 rad/s. Since this extra phase lead is provided by the SCC, it will only affect the perception of the inertial cue. Due to this phase lead, during the internal comparison of the visual and inertial cues, the inertial stimulus might be perceived as being ahead of the visual stimulus. This effect is based on the assumption that humans do not have an accurate internal representation of their SCC, and therefore, cannot internally compensate for it in their comparison of the two stimuli. During the experiment, an artificial phase lead is introduced in the inertial signal, which is added to the lead provided by the SCC, resulting in a larger

phase difference between cues than the difference accounted for by the experimental procedure. Assuming a constant phase-error threshold at the moment the perceived visual and inertial cues are internally compared, the conditions with a higher frequency should present a lower phase-error threshold, since the SCC are already introducing phase lead. The difference in the measured thresholds at the two different frequencies should be entirely explained by the difference in phase lead introduced by the SCC. In fact, the difference in phase lead introduced by the SCC is 21 deg, which corresponds quite well with the approximately 20 deg difference observed in the measured thresholds.

The second hypothesis is based on the results of Grant and Lee (2007) for the higher amplitude conditions. Here, the phase-error threshold remained fairly constant across frequencies. It could be hypothesized that at the lower amplitude conditions subjects had difficulties perceiving the motion stimuli and the apparent effect of frequency is in fact an effect of amplitude. Although the lower amplitude signals are well above threshold, they might still hinder subjects' judgment when they need to simultaneously evaluate both inertial and visual cues. This fact seems to be supported by the fact that thresholds are slightly higher for the conditions with a motion gain of 0.5. Furthermore, at the higher frequencies, for the same velocity amplitude, the acceleration amplitude is higher. This could explain why the thresholds found were lower for the high frequency conditions. Increasing the acceleration amplitude of the stimuli, either by increasing the frequency or by increasing the velocity amplitude, will perhaps result in stimuli which are easier to perceive and judge. Therefore, if large enough stimuli are used, the phase-error thresholds should be independent of the frequency.

This hypothesis also implies that there is an effect of amplitude on the phase-error thresholds. It could be said that as the amplitude of the stimuli increases, the phase-error threshold will tend to decrease. It is thought, though, that this effect saturates above a certain amplitude.

5.3 Method

5.3.1 Apparatus

The experiment was conducted in the Simona Research Simulator which has an hydraulic 6 degree-of-freedom motion base which allows for a maximum displacement of $\pm$ 41.6 deg in yaw and -23.7 to 24.3 deg in pitch. The visual system consists of three LCD projectors, with a resolution of 1280 $\times$ 1024 pixels per projector, and a collimating mirror that provides a field of view of 180 deg $\times$ 40 deg. The visual update and refresh rates are 60 Hz. For a more detailed description of the Simona simulator motion and visual systems capabilities and the computer architecture and software used, please refer to Van Paassen and Stroosma (2000); Stroosma et al. (2003); Berkouwer et al. (2005). The outside visual scene consisted of a

view of the Amsterdam Schiphol airport including the control tower, some lower buildings, part of a runway and some grass fields, from a viewpoint height of 5 meters.

5.3.2 Experimental Conditions

Phase-error thresholds were measured for yaw motion signals with three frequencies and two amplitudes. Thresholds were measured only for the case where the inertial motion leads the visual motion, since generally motion filters provide phase lead in the inertial cue. To allow for a direct comparison with the data from Grant and Lee (2007), phase-error thresholds were also measured in pitch, for two frequencies and one amplitude. The two lower frequencies chosen were the same as the ones used in the their study, 1.257 and 6.283 rad/s. The highest frequency, 10.053 rad/s was chosen to be in the range above 10 rad/s where the semicircular canals dynamics introduce a larger phase lead, but it is not a too high frequency that it would become uncomfortable for the subjects in the simulator. The low amplitude chosen of 5.73 deg/s was also used by Grant and Lee (2007). The other amplitude of 9.17 deg/s was chosen to extend their data to higher amplitudes. All experimental conditions are listed in Table 5.2.

Table 5.2: Experimental conditions.

Amplitude, deg/s	Frequency, rad/s		
	1.257	6.283	10.053
5.73	Yaw/Pitch	Yaw/Pitch	Yaw
9.17	Yaw	Yaw	Yaw

The specific values of 9.17 deg/s for the amplitude and 10.053 rad/s for the frequency were selected such that two of the conditions would also produce identical amplitudes in acceleration. Initially, this was thought to provide an interesting comparison, although the idea was later discarded.

Table 5.3 shows the resulting amplitudes in acceleration, velocity and attitude for all conditions. The amplitude defining a condition is denoted in velocity. All conditions were chosen to produce reasonable values of acceleration and stay within the limits of the Simona simulator.

5.3.3 Motion Profile

Both the visual and inertial motion signals were sinusoidal, with equal amplitude and frequency but different phases. To guarantee that the acceleration, velocity and attitude of the simulator always started and ended at zero in each run, the inertial motion signals were

Table 5.3: Attitude, velocity and acceleration amplitudes for each of the experimental conditions.

	Frequency, rad/s					
	1.257	6.283	10.051	1.257	6.283	10.051
Attitude, deg	4.56	0.91	0.57	7.30	1.46	0.91
Velocity, deg/s	5.73	5.73	5.73	9.17	9.17	9.17
Acceleration, deg/s^2	7.20	36.00	57.60	11.52	57.60	92.17

faded in and out. To avoid any unwanted discrepancies between the visual and the inertial signals, the fade-in and fade-out were also applied to the visual signal. Equation (5.2) describes the fade-in and fade-out parts of the motion profile in terms of acceleration. A and ω are the amplitude and frequency of the acceleration signal, as given by the experimental conditions, in deg/s^2 and rad/s, respectively. The amplitude $A_c = A/12$ and the frequency $\omega_c = \omega/2$ were chosen such that the velocity signal started at zero and was continuous at $t = 2\pi/\omega = T$.

$$f(t) = \frac{1}{2}A\sin(\omega t) - \frac{1}{2}A\sin(\omega t)\cos(\omega_s t) + A_c\sin(\omega_c t) \tag{5.2}$$

Equation (5.3) shows the complete motion signal in terms of acceleration. T is the period of the signal and N is the number of periods in each experimental run, excluding the fade-in and fade-out periods. The value of N depended on the frequency of the motion signal. A minimum of two periods and 10 seconds of duration was considered to be necessary to allow for a good judgment of the match between the visual and the inertial signals. For the lowest frequency, the value of N was 2, resulting in experimental runs of approximately 20 seconds. For the middle and highest frequencies of 6.283 and 10.051 rad/s, N was set at 8 and 14, respectively, resulting in experimental runs of approximately 10 seconds.

$$a(t) = \begin{cases} f(t), & 0 < t \leq T \\ A\sin(\omega t), & T < t \leq (N+1)T \\ f(t-T), & (N+1)T < t \leq (N+2)T \end{cases} \tag{5.3}$$

For the both the yaw and pitch conditions, the center of rotation was the subject's head. During pitch motion, and to maintain as much as possible similar conditions to the ones in Grant and Lee's study (2007), there was no compensation for the specific forces arising from tilting with respect to gravity.

In each run, the inertial motion lead the visual motion by a constant phase. By keeping a constant phase difference also during the fade-in, the inertial motion effectively started while the visual scene was still stationary, and during fade-out the visual motion kept on

moving, while the inertial motion had already stopped. Although the inertial motion was considered to be sub-threshold during these short time frames at the beginning and end of each run, it could not be excluded that this might aid subjects in the detection of a time delay between the onset and stopping of both cues. To avoid this additional clue, the visual scene was also faded-in. A blank screen was shown at the beginning of inertial motion, and only after a half period the visual scene would appear. To assure a soft transition between the blank screen and the visual scene, the visual scene was faded in during 500 ms.

The synchronization of the inertial and visual cues was adjusted according to known intrinsic time delays of the visual (Stroosma et al., 2007) and motion systems (Berkouwer et al., 2005) of the simulator. To compensate for these time delays, the visual scene was artificially delayed by 6 ms.

5.3.4 Procedure

Subjects were seated in the right-hand-side chair of the simulator cabin. Subjects wore a headset with active noise cancellation where engine noise was played to further prevent any noise cues. The trim button on the sidestick, on the right side of the participants, was used to record their answers throughout the experimental runs. Communication with the experiment supervisor was possible at all times.

The experiment was divided in trials. Each experimental trial consisted of several runs. In each run, the visual and inertial motion signals had the same amplitude and frequency but a different phase. Throughout one trial the visual signal was kept constant whereas the inertial signal phase was varied according to a so-called one-up-one-down adaptive staircase algorithm, similar to the one used in Experiment 1 in Chapter 3. The choice of measuring method was determined by the fact that, although subjects can easily identify a phase mismatch, they are rarely able to determine which of the signals (visual or inertial) was leading and which one was lagging. This precludes the use of the self-tuning method.

The staircase algorithm determined the phase-error to be applied to each next run depending on the phase-error of the previous run and on the subjects' judgment of the visual and inertial motion signals. After each run subjects were asked if the inertial and visual cues were synchronized. They provided their answer by pushing the trim button up, signifying "yes", or down, signifying "no". Depending on the subject's answer the phase-error in the next run was either increased (if the answer was "yes") or decreased (if the answer was "no"). The amount by which the phase-error was increased or decreased was called the step size. After each reversal in the subjects' answer the step size was halved. If the same answer was given more than four times in a row the step size was doubled to assure a faster approach to the threshold. The initial phase-error was randomized between 0 and 5 deg and the initial step size was 8 deg. An increase in phase-error meant that the inertial motion would lead the visual motion by a larger value.

While there was no upper limit, the lower limit for the phase error was zero degrees. If the value reached zero and the subject answered "no" for two times in a row, the phase error in the next run was randomized between 50 and 70 deg. This assured that subjects who, due to tiredness or lack of attention, perceived phase-errors although there were none, could "restart" their judgment. The experimental trial finished when either the step size reached 1 deg or the number of reversals reached ten.

For every subject the order of conditions was randomized. Every subject repeated each condition three times, which resulted in a total of 24 experimental trials per subject. Each subject spent approximately 2 hours in the simulator with a 15 minute break halfway.

Before the experiment started, subjects were given three trials to get used to the procedure.

5.3.5 Subjects and subjects' instructions

Eight subjects participated in the experiment, seven males and one female. Their ages were between 23 and 26 years with a mean of 24.75 years. Five subjects indicated having previous experience with full motion simulators.

Subjects were informed that the experiment was divided in a number of trials. Each trial consisted of a random number of runs, of which some contained a phase-error and others not. The number of runs with or without phase error was also stated to be randomized. Although not explicitly mentioned, it was suggested that their answers had no influence on the phase-error of the following run. Subjects were asked to judge whether the visual and the inertial stimuli were synchronized in time. They were instructed to concentrate solely on this synchronization detection and to avoid getting distracted by other factors in the experiment, such as possibly perceived amplitude mismatches. They were, however, encouraged to report any kind of cues or strategies used to detect the synchronization errors, or any other impressions about the experimental conditions and procedure.

5.4 Results

For each experimental trial a phase-error threshold was determined using the subjects' answers in each run. A "yes" answer was assigned a value of zero and a "no" answer a value of one. A psychometric curve based on a normally distributed cumulative probability function was then fitted to the answers using a least squares method. The psychometric curves were defined by a mean (μ) and a standard deviation (σ) as described in Equation (5.4).

$$P(A) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{A} e^{\frac{-1}{2\sigma^2}(x-\mu)^2} dx \tag{5.4}$$

Here, the variable x represents the phase-error or phase difference between the visual and the inertial motion signals. P is the probability of a phase-error with the value A to be perceived as unsynchronized visual and inertial stimuli.

The estimated value of μ, which is the phase-error at the 50% probability level, was taken as the phase-error threshold. The estimated phase-error thresholds were then averaged across repetitions of the same experimental condition. The averaged phase-error thresholds for all subject are shown in Figure 5.4 for pitch and yaw motion. The bars represent the 95% confidence interval.

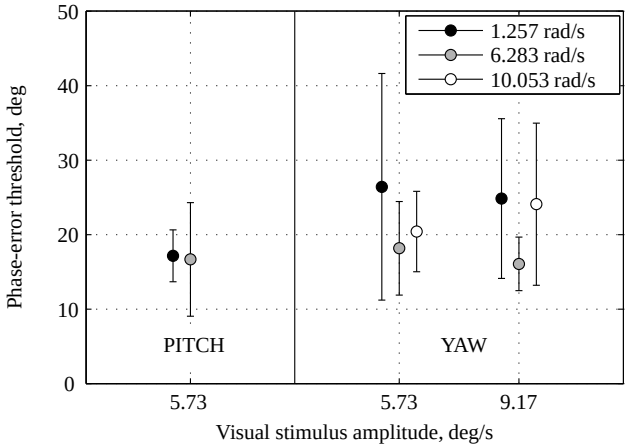


Figure 5.2: Phase error thresholds for both yaw and pitch motion.

The pitch phase-error threshold was approximately the same for both frequencies. The yaw phase-error thresholds did not appear to be influenced by the amplitude of the cue, but there was a slight variation with frequency. The lowest and highest frequencies showed higher thresholds than the middle frequency. A remarkable difference between conditions seems to be the spread in values, with the lowest frequency conditions in yaw and the highest amplitude and highest frequency condition showing a much larger 95% confidence interval.

To further investigate the spread in the measured thresholds, the data was organized in box plots, as seen in Figure 5.3. In this figure, the central line represents the median and the edges of the boxes are the 25th and the 75th percentiles. The whiskers extend to all data points that are not outliers. Outliers are considered data points that are larger than $q_3 + 1.5(q_3 - q_1)$ and smaller than $q_1 - 1.5(q_3 - q_1)$, where q_1 and q_3 are the 25th and the 75th percentiles, respectively. Outliers are given together with the indication of the subject number.

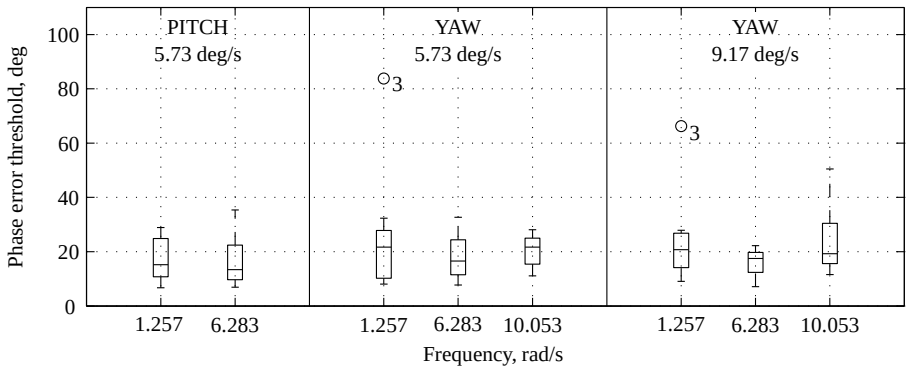
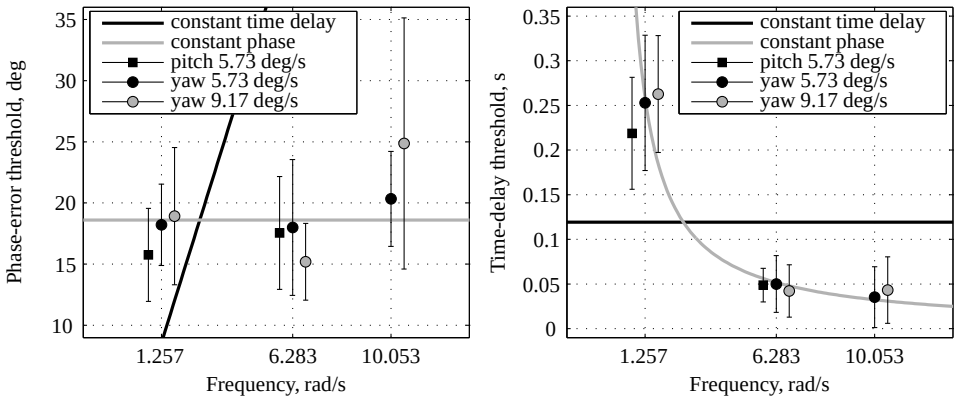


Figure 5.3: Box plot of all conditions showing outliers.

As can be observed from Figure 5.3, the large spread in the yaw conditions with the lowest frequencies can be explained by the fact that subject number 3 presented phase-error thresholds that were much higher than the mean for those conditions. Subject number 3 reported having difficulty perceiving the inertial motion for these conditions. In fact, these conditions result in the lowest acceleration amplitudes (see Table 5.3) and difficulty in perceiving the cue might have strongly impaired the subject's ability to perform the task. The data from subject number 3 was removed from the data set and it is not included in the results shown from this point on.

The averaged thresholds excluding subject number 3 are shown in Figure 5.4, where the bars indicate the 95 % confidence interval. Here, the thresholds are presented in terms of phase error and time delay. The time delay values were calculated from the phase-error thresholds measured and the frequency of the signal. Since the stimuli used were single sine waves, each phase-error value corresponds to one specific time delay between the inertial and the visual motion. In this figure two hypothetical lines are also shown that represent the trend the thresholds should follow in the case that subjects behaved like phase-error detectors and in the case they behaved like time-delay detectors. These lines were calculated by considering the constant phase and constant time-delay to be the average of the measured thresholds for all conditions.

In Figure 5.4a the measured phase-error thresholds seem to follow the constant phase line much more than the constant time-delay line. This result can be better observed in Figure 5.4b, where the thresholds follow the constant phase line quite closely. The time-delay threshold is much higher for the low frequency conditions, which is to be expected since for lower frequencies, the same phase error corresponds to a larger time delay than for the higher frequencies.



(a) Phase error. **(b)** Time delay.

Figure 5.4: Thresholds expressed in terms of phase error and time delay.

An ANOVA was performed on the data. For the yaw phase-error thresholds, neither the amplitude ($F(1,6) = 0.118$, $p > 0.05$) nor the frequency ($F(2,12) = 3.111$, $p > 0.05$) had a significant effect on the thresholds. There were also no interaction effects ($F(2,12) = 0.903$, $p > 0.05$). For the pitch thresholds, there was no significant difference between frequencies ($F(1,6) = 0.132$, $p > 0.05$). Between yaw and pitch thresholds there was also no significant differences ($F(1,6) = 0.527$, $p > 0.05$).

For comparison purposes, the measured thresholds are shown together with the data from Grant and Lee 2007 in Figure 5.5. Their thresholds were 20 deg to 50 deg larger than the ones measured in the present experiment. These differences might be due to the different visual system and experimental method used.

The phase-error thresholds were determined assuming that the time delay between the visual and the inertial stimuli was solely determined by the phase-error introduced by the experimental procedure. Any other synchronization issues were taken into account by artificially delaying the visual scene by 6 ms to match the known intrinsic time delay of the motion system. However, both the visual system time delay and the motion system time delay were measured for specific configurations of visual graphics and motion profiles. It could be argued that for the specific stimuli of the present experiment these time delays would change.

To investigate the impact that unaccounted for timing differences might have on the measured thresholds, modified phase-error thresholds were calculated. These modified thresholds show the subjects' average thresholds for hypothetical timing differences between the

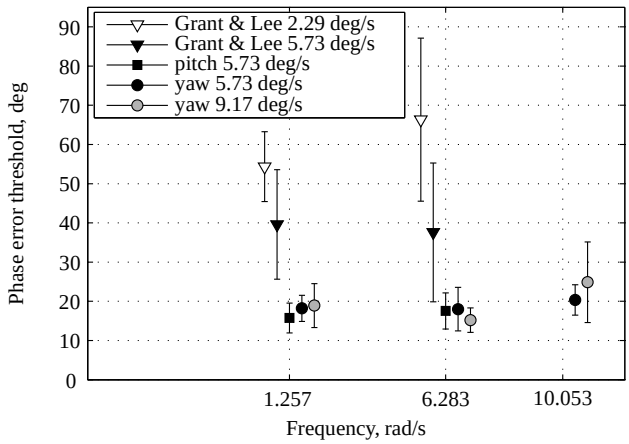


Figure 5.5: Comparison of the measured phase-error thresholds with the data from Grant and Lee (2007).

inertial and visual signals. Figure 5.6 shows these thresholds for timing differences from -10 to 10 ms. A positive timing difference indicates that either the inertial motion system had a smaller time delay than what was compensated for, or the visual system had a larger time delay, or both. A negative value shows the opposite situation, either the inertial motion system was slower or the visual system was faster than initially considered.

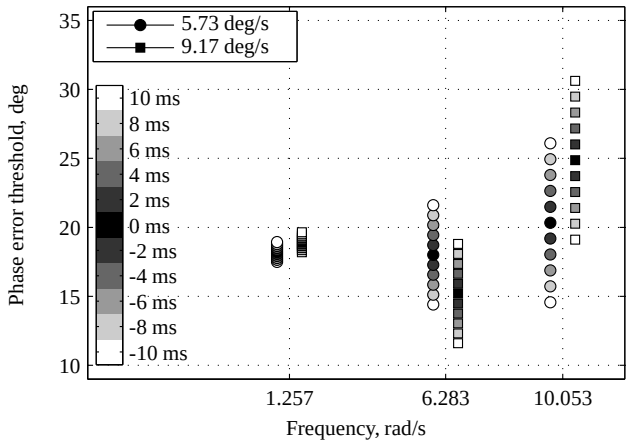


Figure 5.6: Impact of unaccounted for timing differences between visual and inertial signals on the phase-error thresholds.

As expected, any small timing difference that was unaccounted for would have a larger impact on the phase-error at the higher frequencies than at the lower ones. Despite the fact that there seems to be quite a difference between the measured thresholds and the thresholds for the maximum timing differences at the highest frequencies, most values are still within the 95% confidence interval calculated for the measured data.

5.5 Discussion

An experiment was performed to measure the maximum phase lead of the inertial cue that was still perceived by subjects as being synchronized with the visual cue. This maximum value, also denominated phase-error threshold, was measured for pitch and yaw motion at different frequencies and amplitudes of the stimuli.

The measured phase-error thresholds did not show significant difference between yaw and pitch thresholds. The specific forces arising from the pitch motion were either not relevant for the phase-error detection task or perhaps, although they were above the surge threshold, not large enough to be useful.

The yaw phase-error thresholds were also not affected by the amplitude of the stimuli. This is not in agreement with the values recorded by Grant and Lee (2007), where the thresholds at the lower amplitude were between 15 deg and 40 deg higher than for the higher amplitude. That effect was even more pronounced for the conditions with a motion gain of 0.5. Their results resemble the ones of subject 3 in the present experiment, whose phase-error thresholds for the lower frequency yaw conditions were around 50 deg to 60 deg higher than at other frequencies. The subject had stated, during these low frequency conditions, that it was difficult to perceive the inertial motion. The low frequency conditions have, in fact, the lowest acceleration amplitudes (see Table 5.3). These low values, although well above threshold, might still be insufficient to allow for a good comparison between the visual and inertial stimuli. Curiously, the subject's low frequency pitch threshold was in the same range as the other values. However, during pitch motion an extra gravitational cue can be used to detect the sinusoidal inertial motion, facilitating the comparison of both stimuli. For all the other subjects the amplitude of the stimuli was apparently large enough to allow for a good detection of mismatches between inertial and visual cues.

The stimulus frequency did not significantly affect either the yaw or the pitch phase-error thresholds. This result is in agreement with what was observed for the higher amplitude conditions in Grant and Lee's study (2007). In their study, frequency had an evident effect only for the low amplitude conditions. As discussed above, the combination of low amplitude and low frequency leads to very low acceleration amplitudes. As the stimuli acceleration amplitude decreases, subjects seem to become more lenient in their judgment of matching visual and inertial cues, resulting in higher phase-error thresholds. The amplitudes and frequencies chosen in the present experiment resulted in apparently large enough

acceleration amplitudes, which allowed all subjects, with the exception of subject number 3, to make a good evaluation of both inertial and visual cues. Looking at all the results, the ones presented here and the ones from Grant and Lee (2007), it could be said that an effect of amplitude and frequency is only observable for low amplitude stimuli.

The constant phase-error thresholds measured across frequencies also indicates that human subjects are sensitive to phase differences between visual and inertial cues, rather than time delays. If subjects would be time-delay detectors one would expect that as frequencies increase, the phase-thresholds would also increase. The results found for both yaw and pitch motion showed that as the frequency increases, the corresponding time-delay threshold greatly decreases, following the expected trend of a constant phase-error threshold.

Comparing the actual threshold values found in this study with the ones from Grant and Lee (2007) it is clear that the pitch phase-error thresholds were much lower for the present work. Despite the fact that the higher values they found at the lower amplitude conditions may be explained by an insufficient inertial cue amplitude, for the higher amplitude conditions, there is still a 20 deg difference. Grant and Lee (2007) performed measurements of their visual and motion systems inherent time delays specifically for their experiment whereas in the present experiment, previously measured values were used. The time delays inherent to the systems were not measured for the purpose of the present experiment and might have gradually changed with time or be slightly off due to specific characteristics of the experimental setup, such as visual scene complexity and the frequency of the stimuli. To investigate whether such timing differences might explain the differences between the results in both studies, modified thresholds were calculated based on hypothetical timing differences between visual and inertial signals. It was thought that the modified thresholds could also reveal a small effect of frequency that was masked by timing differences that were not accounted for in the experimental setup.

As expected, the modified thresholds at higher frequencies differed the most from the measured thresholds. Nevertheless, for unaccounted timing differences of as much as 10 ms, at the highest frequency, the phase thresholds would increase or decrease by a maximum of around 5 deg. On the one hand, a higher phase-threshold at higher frequencies could be expected if humans behaved like time-delay detectors. However, although 5 deg represents a 20 % to 25 % increase in the threshold value, it is still not as much as it would be expected from a constant time-delay threshold. On the other hand, a lower phase-error threshold at higher frequencies is in agreement with the hypothesis that the phase lead introduced by the SCC is not compensated for in the internal comparison of visual and inertial cues. Nevertheless, a decrease of only 5 deg does correspond well with the 30 deg phase lead introduced by the SCC at the highest frequency with respect to the lowest frequency. It is thus unlikely that there is an effect of frequency masked by small timing mismatches between the visual and inertial systems.

The study of the modified thresholds also did not help explain the 20 deg difference found between Grant and Lee's (2007) study and the present one. Perhaps more straightforward reasons for the discrepancy can be pointed out, such as the fact that both studies were performed in different simulators and the visual systems used differed considerably. The influence of visual system characteristics, such as field-of-view and collimation, have been shown to influence subjects' sense of motion (Chung et al., 2003). Nevertheless, to be able to understand the effect of these characteristics on the perception of cue synchronization, further investigations are needed.

Differences in the experimental methods used may also contribute to the difference in the measured thresholds. In this study the threshold value was defined to be the phase-error at the 50% probability level in the psychometric curve. Grant and Lee (2007) used a 79% probability level. For a higher detection probability a higher phase-error value would be expected. As perhaps expected, the phase-error threshold is not a hard boundary, but more a gradual change, where the higher the phase-error, the higher is the probability that it is detected.

For a better direct comparison with the work from Grant and Lee (2007), it would be necessary to calculate the 79% probability level from the measured thresholds. However, the staircase procedure used in the present experiment provides a good estimate of the psychometric curve mean (the 50% probability level), but not of its inclination. For this reason it is not possible to obtain a reliable estimation of the complete psychometric curve that would allow us to estimate the 79% probability level.

The presented results have two implications for vehicle simulation. First, the idea of different threshold levels, or different probabilities of detection, agrees well with criteria for simulator motion which consider different levels of fidelity. An example of such criteria is the Sinacori plot (Sinacori, 1977), where the phase distortion introduced by a motion filter at 1 rad/s is used to delimit high, medium and low fidelity regions.

Second, a phase-error threshold that remains constant with frequency indicates that at the lower frequencies, where high pass filters introduce more phase lead than at higher frequencies, there is a higher risk of disrupting the realism of the simulation. However, how different phase distortions at different frequencies might influence the lumped phase-error threshold is not yet known.

5.6 Conclusions

An experiment was conducted to measure the influence of stimulus frequency and amplitude on the visual-vestibular phase-error detection thresholds. The measurements were made in yaw and pitch.

The measured thresholds suggest that humans act like phase-error detectors, rather than like time delay sensors and that for the range of amplitudes and frequencies tested, the mean phase-error thresholds is 19 degrees.

The frequency of the stimulus did not affect phase error detection. These results are partially in agreement with the findings of Grant and Lee (2007). The mean phase-threshold value, however, was much lower than the one found in that study. This difference might be attributed to the experimental methodologies use. The staircase method used here was designed to obtain a reliable threshold value at the 50% probability level of the psychometric curve, whereas in Grant and Lee (2007) the probability level was 79%. Differences in the visual systems characteristics might have also played a role, but without further investigation no decisive conclusions can be drawn.

The measured thresholds for yaw were similar to the ones found in pitch. The stimulus amplitude was probably high enough above the angular motion sensory threshold to render the extra cue, the otolith stimulation, unnecessary.

From the combined analysis of the present results and the ones from Grant and Lee (2007), it appears that the frequency and amplitude of the stimuli only have an influence on the phase-error threshold for low angular acceleration amplitudes.

If the results found here are to be directly applied to motion filter tuning, then care should be taken that the angular inertial motion does not lead the visual cues by more than 19 degrees.

6

Perception Coherence Zones During Active Tasks

This chapter is based on the following publication:

Valente Pais, A. R., Van Paassen, M. M., Mulder, M., and Wentink, M. (2011). Effect of Performing a Boundary-Avoidance Tracking Task on the Perception of Coherence Between Visual and Inertial Cues. In *Proceedings of the AIAA Modeling and Simulation Technologies Conference, Portland, OR, USA, August 8-11*, AIAA-2011-6324.

6.1 Introduction

During flight simulation, the inertial and visual stimuli provided to the pilot differ considerably. The motion cueing algorithms (MCA) transform the inertial motion amplitude and phase differently for different signal frequencies. In order to quantify the effect of these MCAs on the realism of the simulation, it is necessary to evaluate how much the differences between inertial and visual cues affect human perception and behavior.

Many studies have been performed on the integration of inertial and visual stimuli in various simulation scenarios (Groen et al., 2001; Groen and Bles, 2004; Grant et al., 2006; Grant and Lee, 2007). A specific set of studies has concentrated on the concept of coherence zones (Van der Steen, 1998; Valente Pais et al., 2010b,a). A coherence zone represents a range of inertial motion levels, either amplitude or phase levels, which although not being a match with the visual motion, are still perceived by humans as one realistic, coherent movement. These coherence zones determine the boundaries within which the simulator inertial motion can be adjusted or modified to fit the simulator motion limits without hindering the quality of the simulation.

Thus far, coherence zones have been measured for different amplitudes and frequencies of angular motion. All experimental setups had subjects judge whether or not visual and inertial cues matched in terms of amplitude (Chapters 3 and 4; Van der Steen, 1998) or phase (Chapter 5; Grant and Lee, 2007). Although much more data can be collected, there is now a basic understanding of how coherence zones are affected by stimulus amplitude and frequency. However, before this knowledge can be applied to pilot-in-the-loop simulations, more knowledge must be gained as to how the coherence zones change when the pilot is not fully concentrated on the visual and inertial cues, but actually has a task to perform, as for example, a manual control task.

Introducing a manual control task in a perceptual experiment is not without its challenges. A pilot-in-the-loop situation implies that there is less control over the amplitude and frequency of the visual and inertial stimuli that subjects are exposed to, making it difficult to analyze and compare results.

Probably for this reason, the few studies that have been done, measuring the effect of performing a manual control task on perception thresholds (Roark and Junker, 1978; Hosman and van der Vaart, 1978; Samji and Reid, 1992), have maintained the inertial motion relevant for the threshold measurement separated from the inertial motion relevant for the control task. This was accomplished by using separate degrees-of-freedom (Hosman and van der Vaart, 1978; Samji and Reid, 1992) or measuring the indifference thresholds by superimposing the inertial stimulus used to measure the threshold on the inertial stimulus resulting from the control task (Roark and Junker, 1978).

These studies have measured the effect of decreased attention or increased mental load, through the addition of a manual control task, on the motion perception thresholds. Al-

though both inertial and visual stimuli were used, the goal was to measure the inertial perception threshold only. The effect of increased mental load on the combined perception of visual and inertial cues has not been investigated.

These studies have shown that the perception thresholds increase considerably when the control task is added. However, it is not known whether the same holds for coherence zones. Although the coherence zone boundaries may also depend on the same mechanisms that determine the sensory motion perception thresholds, other higher level brain processes are probably involved in the integration of inertial and visual stimuli.

In order to measure the effect of decreased attention on the perception coherence zones, an experiment was designed where subjects' yaw amplitude coherence zones were measured while performing a pitch boundary-avoidance tracking task at different difficulty levels. For comparison purposes, also the passive coherence zones (without the control task) were measured.

It is hypothesized that, similar to the trends observed for inertial motion thresholds, when measuring coherence zones while performing a control task, subjects spend less attention on the perception task and become more lenient. They will allow for wider coherence zones due to a decrease in the lower threshold and an increase in the upper threshold.

This chapter describes the experimental method for both the passive and the active coherence zone measurements and presents the results. In the Discussion, the experimental findings for the active part are compared to the ones for the passive part and to other work found in literature. The chapter ends with the conclusions.

6.2 Method

6.2.1 Apparatus

The experiment was conducted in the Simona Research Simulator which has an hydraulic 6 degree-of-freedom motion base which allows for a maximum displacement of $\pm$ 41.6 deg in yaw. The visual system consists of three LCD projectors, with a resolution of 1280 $\times$ 1024 pixels per projector, and a collimating mirror that provides a field of view of 180 deg $\times$ 40 deg. The visual update and refresh rates are 60 Hz. For a more detailed description of the Simona simulator motion and visual systems capabilities and the computer architecture and software used, please refer to Van Paassen and Stroosma (2000); Stroosma et al. (2003); Berkouwer et al. (2005). The outside visual scene consisted of a view of the Amsterdam Schiphol airport including the control tower, some lower buildings, part of a runway and some grass fields, from a viewpoint height of 5 meters. A Head-Up-Display (HUD), described further in Section 6.2.4, was also shown throughout all experimental conditions.

6.2.2 Experimental design

The experiment was divided in two parts. The first part consisted of measuring coherence zones without any control task, so a passive perception task only. In the second part coherence zones were measured for the same conditions as in the first part, although this time subjects had to perform an active control task simultaneously with the perception task. Both parts considered, the experiment had a two-way repeated measures design, with the two independent variables being the frequency of the motion signal and the task difficulty. Two frequencies and three task difficulty levels were tested, resulting in a total of 6 conditions. The frequencies tested were 2 and 5 rad/s and the difficulty levels were perception task only, perception task and active control task of low difficulty, and perception task and active control task of increased difficulty. For all conditions a lower and an upper threshold for inertial motion were determined and each measurement was made 3 times, which resulted in 36 experimental trials per subject. Table 6.1 summarizes the experimental conditions.

Table 6.1: Threshold measurements performed at each experimental condition.

Frequency, rad/s	Part I	Part II	
	Passive	Active easy	Active difficult
2	Upper/Lower	Upper/Lower	Upper/Lower
5	Upper/Lower	Upper/Lower	Upper/Lower

6.2.3 Motion and visual signals

The choice of frequencies and amplitudes was based on values previously used (Chapter 4), to allow for comparison with earlier studies. However, one additional factor was taken into consideration. It was chosen to test signals of different frequencies but with the same velocity amplitude, and not acceleration amplitude, as it has been done until now. It was thought that by doing so, it would be possible to further investigate the effect of frequency and amplitude on perception coherence zones.

The amplitude of the visual in acceleration (A) was 12 deg/s^2 for the 2 rad/s signal and 30 deg/s^2 for the 5 rad/s signal. These acceleration amplitudes have been used in previous chapters, and so has the 2 rad/s frequency. The 5 rad/s frequency was chosen such that the two signals with different frequencies would have the same velocity amplitude: 6 deg/s.

The first and last period of the sinusoidal signals were used to fade in and out.

Equation (6.1) describes the fade-in and fade-out parts of the motion profile in terms of acceleration, where ω is the frequency of the acceleration signal, as given by the experimen-

tal conditions. The amplitude $A_c = A/12$ and the frequency $\omega_c = \omega/2$ were chosen such that the velocity signal started at zero and was continuous at $t = 2\pi/\omega = T$.

$$f(t) = \frac{1}{2}A\sin(\omega t) - \frac{1}{2}A\sin(\omega t)\cos(\omega_s t) + A_c\sin(\omega_c t) \tag{6.1}$$

Equation (6.2) shows the complete motion signal in terms of acceleration. T is the period of the signal and N is the number of periods of each motion signal, excluding the fade-in and fade-out periods. The value of N depended on the frequency of the motion signal: for the 2 rad/s signal, N was 5, and for the 5 rad/s signal, N was 2. This resulted in the same duration for both conditions: 15.71 seconds.

$$a(t) = \begin{cases} f(t), & 0 < t \leq T \\ A\sin(\omega t), & T < t \leq (N+1)T \\ f(t-T), & (N+1)T < t \leq (N+2)T \end{cases} \tag{6.2}$$

6.2.4 Control Task

The control task consisted of a boundary-avoidance pitch tracking task. The tracking error and boundaries were shown on a compensatory display superimposed on the outside visual scene. This HUD consisted of an horizontal gray bar and the aircraft symbol, as shown in Figure 6.1. The gray bar's width corresponded to a pitch angle of 3 deg.

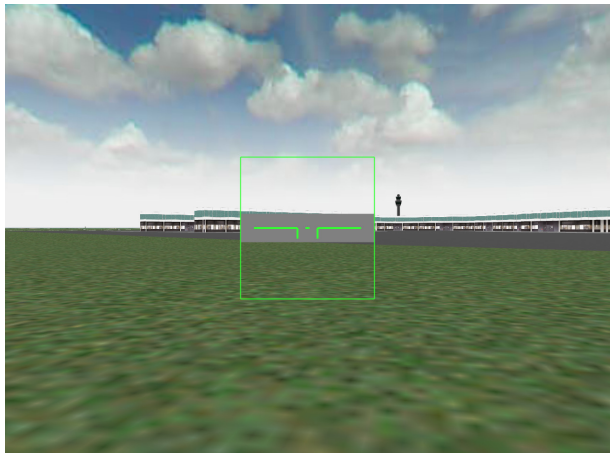


Figure 6.1: Representation of the center part of the outside visual with the HUD.

The outside visual scene was dedicated to the perception task and showed only yaw motion. Information on the pitch boundary-avoidance task was only available through the HUD. As the outside visual scene moved, the aircraft symbol and the surrounding frame

remained fixed with respect to the pilot's view point, such that the HUD was always directly in front of the pilot. The vertical motion of the gray bar showed the tracking error between the desired pitch angle and the controlled element pitch angle. During the perception task only conditions, the HUD was shown with a stationary gray bar.

There were two difficulty levels for the control task, which were achieved by using two different system dynamics. The controlled system dynamics for the easy task was a single integrator; for the difficult task it was a second order system with a positive pole, that is, an unstable system. The transfer functions for both systems, from stick deflection (δ_s) to pitch angle (θ) are shown in Equations 6.3 and 6.4. Both elements were tuned such that they had the same gain ($K = 0.8594$) at a frequency of 1 rad/s.

$$H(j\omega) = K\frac{1}{j\omega} \tag{6.3}$$

$$H(j\omega) = K\frac{2.6926}{2.5(j\omega)^2 - j\omega} \tag{6.4}$$

The tracking signal, or forcing function, used was based on a pitch tracking task study reported in literature (Damveld, 2009) and simplified to a sum of 5 sines with amplitudes, frequencies and phases as listed in Table 6.2.

Table 6.2: Characteristics of the sine signals used to compose the forcing function. k_i indicates the multiple of the fundamental frequency, ω_i the corresponding frequency in rad/s, A_i the amplitude is radians and ϕ_i the phase in radians.

k_i	ω_i	A_i	ϕ_i
7	0.41	0.0173	2.99
13	0.76	0.0173	2.65
17	1.00	0.0173	-3.12
24	1.41	0.0173	-0.56
31	1.82	0.0173	3.81

The forcing function was chosen to be relatively simple, when compared to other manual tracking control studies (Zaal, 2011; Pool, 2012). As these studies focused on pilot model identification, the tracking signal frequency content was of crucial importance. Since that is not the case in the present study, the higher frequency content was not necessary. Moreover, this simple forcing function allowed for a very simple task when using the single integrator element, and still a considerable control challenge when using the second order unstable system. For the latter system, a more difficult forcing function would result in a constant loss of control. Using these two tasks with very contrasting difficulty levels was thought to deliver larger, and hence easier to observe, differences in the coherence thresholds.

Since the forcing function was simple, to avoid memorization of the tracking signal, the time signal was mirrored around the amplitude axis, around the time axis and around both, creating four different time signals with equivalent difficulty. These time signals were used sequentially in each run.

The boundary avoidance task was preferred over a pure tracking task because it was thought that for the latter, despite the two levels of difficulty, subjects would deliver the same amount of attention and effort to the control task. The only difference would then be on the performance attained. With a boundary avoidance task there can be periods of time when no pilot input is needed to keep acceptable performance, in this case, to stay within the gray area. These periods can be made more or less frequent by changing the controlled element. By using an unstable system, distractions and lack of attention were heavily punished, making this a much more attention demanding task than the single integrator.

6.2.5 Procedure

The experiment was performed in two parts, a passive and an active part. To avoid subject fatigue, each subject performed each part on a different day.

The perception task procedure was similar to the experiments described in Chapters 3 and 4. In each experimental condition the visual amplitude was kept constant while the motion amplitude was varied throughout the runs of one trial. At the beginning of the trial, subjects were informed whether that trial corresponded to a lower or an upper threshold measurement.

In each trial, the amplitude of the first run was randomly selected between 1.1 and 0.9 times the visual amplitude. At the end of each run subjects could change the motion of the next run. They did this by pushing a switch button multiple times up or down until they reached a certain number of increments or decrements. The chosen number was displayed on the outside visual. A positive number meant that the next run would have a higher amplitude motion, and a negative number resulted in a lower amplitude motion. After giving their answer, subjects pressed a second button to signal that they were ready for the next run. The trial ended when subjects' answers had two consecutive reversals of one increment or decrement, i.e., a sequence of 1, -1, 1, or -1, 1, -1. This indicated that subjects converged to a certain amplitude of motion that could not be increased or decreased anymore. The size of one increment or decrement was 0.025 of the visual amplitude.

For the part of the experiment where subjects had to perform both the perception task as well as an active control task, this same procedure was used for the perception task. However, during each run subjects had not only to judge the amplitude of the inertial motion with respect to the visual amplitude, but they also had to perform the pitch tracking task.

At the beginning of the second part and before the actual measurements were made, subjects performed training sessions to reach a constant performance on the tracking task.

These sessions consisted of a minimum of 20 runs for each tracking task difficulty, per subject. Training runs were done until the subject showed a steady performance for at least 10 runs. If at the end of 20 runs there was evidence that a participant was still improving or adjusting his control strategy, more runs would be performed. The root-mean-square (RMS) of the subjects' control input and total error signals were monitored to guarantee a consistent tracking.

To assess whether or not the two active task difficulties were indeed perceived as less and more demanding by the subjects, an effort scale was used, from "No effort at all" to "Maximum effort". After each of the runs subjects were asked to fill in the effort scale on paper, with a vertical mark placed on a 10 cm line.

After training for both control task difficulties, the measurement part started. The experimental trials were divided in 6 blocks of 4 conditions each. In each block always the same task difficulty was used. Each block corresponded to measurements of the upper and lower thresholds for the two signal frequencies. In each pair of blocks all experimental conditions were tested. The presentation order of each pair of blocks was balanced across all 8 subjects. Before each block, if the task difficulty changed from the previous block (so, a different controlled element was used), subjects performed one test trial to adapt to the new task.

In the beginning of each run subjects had 3 seconds to get acquainted with the control task. After these 3 seconds the inertial and visual motion would start. The tracking task stopped at the same time the inertial and visual motion stopped, indicating the end of a run. At the end of one trial, subjects were asked to fill in the average effort scale for the tracking task across all runs in that trial.

6.2.6 Subjects and subjects' instructions

There were 8 male participants with ages between 23 and 35 years (mean of 27.5 years).

For the perception task, participants were instructed to tune the motion up or down, depending on the threshold measurement of that trial. For example, when measuring an upper threshold they should increment the motion until it was perceived as too strong. Then, they were told to decrease it and increase it as many times as needed until they found the strongest motion condition that was still perceived as coherent with the visual motion. Subjects were advised to start with increments of 10 or more and decrease the number of increments or decrements at every direction reversal. They were informed of the stopping criteria of the trials.

With respect to the active task, subjects were instructed during training that they should keep the horizontal green line displayed in the HUD within the moving gray area, with the least effort possible. It was explained to them that after the training they would be performing both the perception task as well as the control task. During the measurement

part subjects were encouraged to maintain the same level of control input and performance as they did during training.

After each trial during the active part of the experiment subjects filled in the effort scale. They were instructed to judge only the effort and attention they had to spend on the control task to keep acceptable performance. Their effort score should reflect the averaged effort across all the runs in one trial. For the training phase the instructions were the same but subjects were asked to fill in the scale after each run.

6.3 Results

During one of the training sessions of one subject there was a problem with the data recording. Since no reliable information was collected, this subject's data were removed from all the analysis concerning training sessions. However, the recorded threshold data during the passive and active part of the experiment were still used.

6.3.1 Effort scores

To assess whether or not both control task difficulties were indeed perceived as requiring different levels of effort and attention, subjects were asked to fill in an effort scale. The marks made on paper were converted to a value with two decimal places, from 0, representing "No effort at all", to 10, "Maximum effort". For the training sessions, the mean score was calculated from the average scores over the last 10 runs, where subjects had showed similar performance. For the active part, the mean scores were the averaged values recorded in each of the three repetitions of one condition.

Figure 6.2 shows the mean scores for the easy and difficult control tasks during the training and active parts. The error bars in this and all following figures in this chapter represent the 95% confidence interval of the mean.

The task difficulty was clearly reflected in the effort scores, with higher scores for the more difficult task. To investigate whether this effect was significant an ANOVA was performed and the results are shown in Table 6.3.

The task difficulty had indeed a significant effect on the effort scores and there was no difference in perceived effort between the training sessions and the active part of the experiment.

Additional statistical tests have shown that except for the task difficulty, there were no other statistically significant differences found in the effort scores between experimental conditions in the active part. The stimulus frequency and threshold measured (upper or lower) did not affect the mean effort scores.

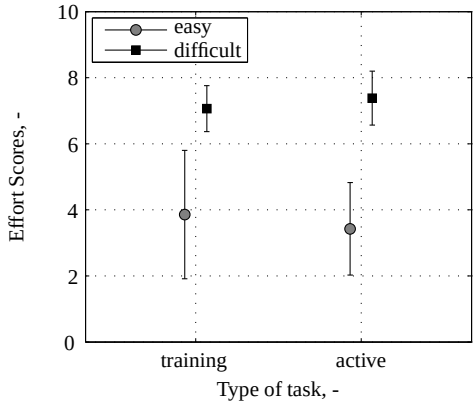


Figure 6.2: Mean effort scores for both tasks, during training and actual experiment.

Table 6.3: ANOVA results for effort scores, where ** is highly significant ($p < 0.01$), * is significant ($0.01 \leq p < 0.05$), and - is not significant ($p \geq 0.05$).

Independent variables	Dependent measures		
	Effort scores		
	df	F	sig.
Type	1, 6	0.25	-
Difficulty	1, 6	28.01	**
Type $\times$ Difficulty	1, 6	1.02	-

6.3.2 Control input and performance measures

As a complementary measure to assess the effort and attention demand of the task difficulty levels, the average duty cycles were calculated. The duty cycle is defined as the percentage of time during which subjects are acting on the control stick and thus changing their control input. In this analysis, these percentages were calculated by using the rate of change of the stick deflection signal.

A rate of change different than zero meant subjects were acting on the stick. However, in practice the rate of change is never exactly zero, since subjects can not steady their hand so perfectly as to immobilize the stick completely. For the duty cycle calculations the rate of change was considered small enough as to indicate a “no input change” period when it was between $\pm$ 0.015 rad/s ($\pm$ 0.86 deg/s). This value was chosen after visual inspection of the stick deflection signal. Figure 6.3 shows part of a time trace of the rate of change of

the stick deflection with indication of the $\pm$ 0.86 deg/s limits. The insert shows the original signal and the signal used for the duty cycle calculations.

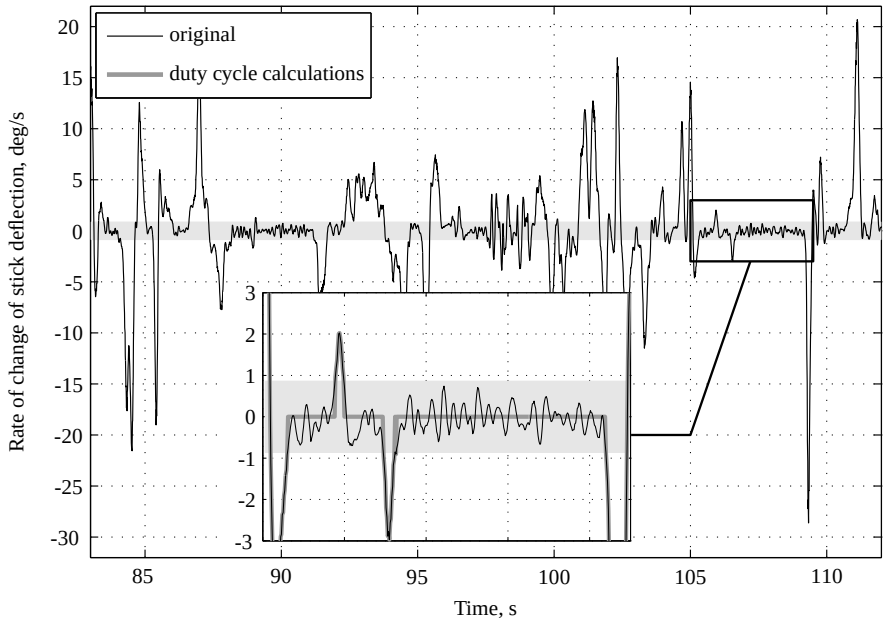


Figure 6.3: Example of a time trace of the rate of change of the stick deflection. The gray region represents the $\pm$ 0.86 deg/s limits.

The calculated duty cycle values were averaged across conditions and subjects to investigate any change of control activity between the training sessions and the active part of the experiment. For the training sessions only the values of the last 10 runs were used. No significant differences were found between the duty cycle values during the training sessions and during the active part of the experiment ($F(1,6) = 0.24, p > 0.05$). The task difficulty, on the other hand, had a very clear effect on the duty cycle values both during the training and the active experiment ($F(1,6) = 26.91$, $p < 0.01$), with the duty cycle values being much higher for the difficult task than for the easy task.

To investigate a possible interaction between the control activity and the stimuli provided for the perception task, the duty cycle values were also analyzed throughout the different conditions of the active part of the experiment, that is, for trials with both upper and lower threshold measurements and the two different stimulus frequencies. Figure 6.4 shows the average values for all the active conditions and the training sessions.

The root mean square (RMS) of the input and error signals were also calculated to get some insight on subjects' control activity and performance throughout the experiment. The error signal was zero whenever the system's pitch angle was within the boundaries defined

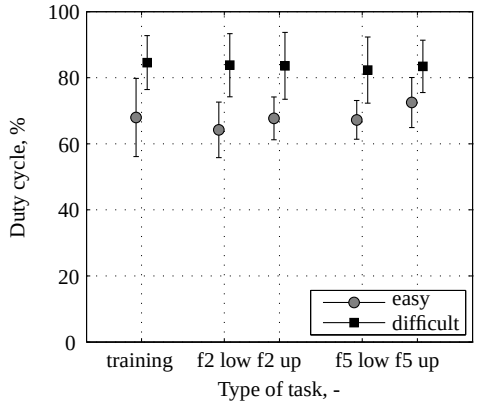
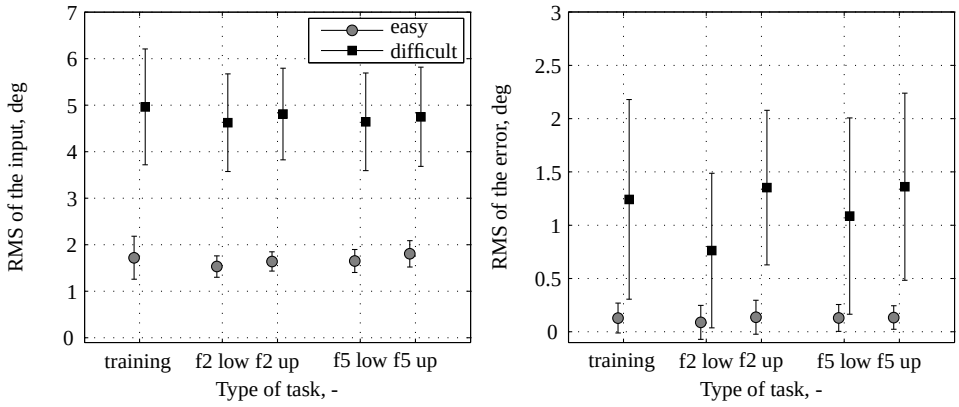


Figure 6.4: Mean duty cycle values for both difficulty levels, during the training and active parts of the experiment. f2 and f5 refer to stimulus frequencies of 2 and 5 rad/s, respectively. low and up indicate lower and upper threshold measurements.

by the gray area in Figure 6.1. When the pitch angle exceeded these boundaries, the error was the difference between the system's pitch angle and the boundary of the gray bar. For the training sessions only the last 10 runs were considered. The averaged values for the RMS of the input and error signals for all conditions are shown in Figure 6.5.



(a) RMS of the input signal. **(b)** RMS of the error signal.

Figure 6.5: Mean RMS values of the input and error signals for both difficulty levels, during the training and active parts of the experiment. f2 and f5 refer to stimulus frequencies of 2 and 5 rad/s, respectively. low and up indicate lower and upper threshold measurements.

The RMS of the input signal shows the same trends as the duty cycle, with significantly higher values for the difficult task than for the easy task ($F(1,6) = 48.13$, $p > 0.05$). There were no significant differences between the mean RMS values during training and during the active part of the experiment ($F(1,6) = 0.54$, $p > 0.05$). The RMS of the error signal shows slightly different results, with higher values for the active part of the experiment than for the training sessions ($F(1,6) = 18.42$, $p < 0.01$). This degradation in performance was more pronounced in the task with a higher difficulty level, which can be confirmed by a significant effect of the interaction term Type (training or active) $\times$ Difficulty ($F(1,6) = 14.49$, $p < 0.01$). The difficulty level also affected the RMS of the error significantly ($F(1,6) = 9.07$, $p < 0.05$).

Across the different conditions of the active part of the experiment also differences were observed between the control activity indicators (duty cycle and RMS of the control input signal) and performance indicators (RMS of the error signal signal). The ANOVA results for the duty cycle, and the RMS of the input and error signals across all conditions in the active part are shown in Table 6.4.

Table 6.4: ANOVA results for the duty cycle and the RMS of the input and error signals. Only the main factors and the interactions for which there were statistical significant effects are shown. ** is highly significant ($p < 0.01$), * is significant ($0.01 \leq p < 0.05$), and - is not significant ($p \geq 0.05$).

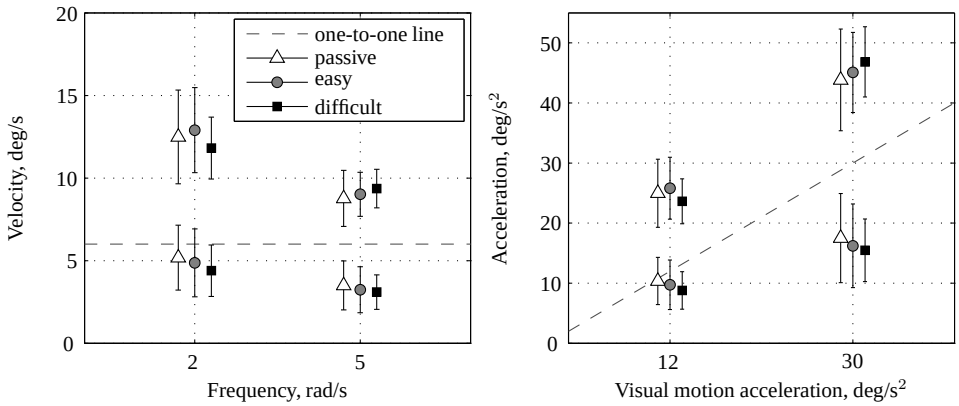
Independent variables	Dependent measures								
	Duty cycle			RMS input			RMS error		
	df	F	sig.	df	F	sig.	df	F	sig.
Threshold	1, 7	24.49	**	1, 7	17.54	**	1, 7	4.12	-
Frequency	1, 7	4.40	-	1, 7	1.97	-	1, 7	0.43	-
Difficulty	1, 7	42.48	**	1, 7	54.13	**	1, 7	11.54	*
Threshold $\times$ Difficulty	1, 7	11.35	*	1, 7	0.08	-	1, 7	6.58	*
Frequency $\times$ Difficulty	1, 7	18.72	**	1, 7	4.61	-	1, 7	0.42	-

The task difficulty had a significant effect on all three metrics, showing a higher control activity and worse performance for the difficult task. The threshold measurement had an effect on the duty cycle and the RMS of the input signal, with higher values for the upper threshold measurement, but it did not affect the RMS of the error signal.

The interaction terms Threshold $\times$ Difficulty and Frequency $\times$ Difficulty also showed an effect on the duty cycle values. In Figure 6.4 it can be seen that for the more difficult task, the duty cycle values remain fairly constant, whereas for the easy task the values increase with increasing frequency and going from lower to upper threshold measurements. For the RMS of the error the opposite seems to happen, that is, the RMS the values for the easy task remaining fairly constant across conditions and show some changes only in the more difficult task. The performance seems to be worse in the upper threshold measurement runs.

6.3.3 Thresholds and Coherence Zones

The inertial amplitudes selected by the subjects in each condition were converted to upper and lower thresholds by averaging the last two amplitudes of every trial. The mean thresholds values are shown shown in Figure 6.6 in velocity (Figure 6.6a) and acceleration units (Figure 6.6b).



(a) Velocity representation. **(b)** Acceleration representation.

Figure 6.6: Mean thresholds and standard deviations represented in terms of acceleration and velocity units.

In general terms the upper thresholds increase and the lower thresholds decrease as we go from the passive (perception only) task to the easy active task and then to the difficult active task. One exception is the upper threshold for the 2 rad/s stimulus and the difficult task, which is lower than the corresponding values for the other task difficulty levels. To investigate if these trends are significant, an ANOVA was performed and the results are shown in Table 6.5.

For both the upper and the lower threshold measurements only the frequency had a significant effect, with lower values for the higher frequency, similar to what has been seen in Chapter 4. The task difficulty did not have a significant effect on the measured thresholds.

Although there was no significant effect of the task difficulty, it is still interesting to analyze the threshold results in terms of a coherence zone. A CZW and a PMC were calculated according to Equations 6.5 and 6.6.

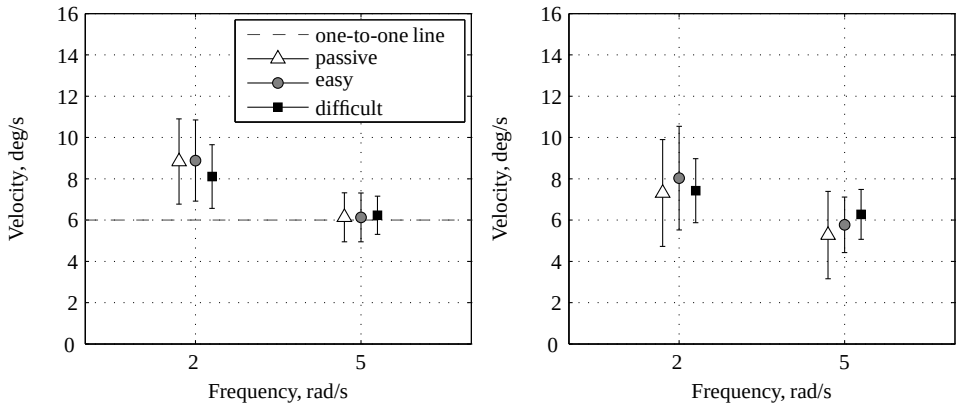
$$CZW = th_{up} - th_{lo} \tag{6.5}$$

$$PMC = th_{lo} + \frac{CZW}{2} \tag{6.6}$$

Table 6.5: ANOVA results for the upper and lower threshold measurements, where ** is highly significant ($p < 0.01$), * is significant ($0.01 \leq p < 0.05$), and - is not significant ($p \geq 0.05$).

Independent variables	Dependent measures					
	Lower threshold			Upper threshold		
	df	F	sig.	df	F	sig.
Frequency	1, 7	29.86	**	1, 7	32.80	**
Difficulty	2, 14	2.31	-	2, 14	0.25	-
Frequency × Difficulty	2, 14	0.89	-	2, 14	1.92	-

The PMC and CZW in velocity and acceleration units are shown in Figure 6.7a and Figure 6.7b, respectively.



(a) Point of Mean Coherence. **(b)** Coherence Zone Width.

Figure 6.7: Coherence zones represented in terms of velocity units.

The PMC remains fairly constant whereas the CZW slightly increases with increasing task difficulty. The exception to this trend is the condition with frequency 2 rad/s and the difficult task, as also seen in the threshold results. Since the effect of the task difficulty was not statistical significant in the threshold data, it is not expected that this is different for the coherence zone data. However, for the sake of completeness, also an ANOVA was performed on these values and the results are presented in Table 6.6.

Table 6.6: ANOVA results for the Point of Mean Coherence (PMC) and the Coherence Zone Width (CZW), where ** is highly significant ($p < 0.01$), * is significant ($0.01 \leq p < 0.05$), and - is not significant ($p \geq 0.05$).

Independent variables	Dependent measures					
	PMC			CZW		
	df	F	sig.	df	F	sig.
Frequency	1, 7	36.45	**	1, 7	17.58	**
Difficulty	2, 14	0.57	-	2, 14	0.78	-
Frequency $\times$ Difficulty	2, 14	2.88	-	2, 14	0.86	-

6.4 Discussion

Coherence zone thresholds were measured during a perception only task, an easy control task and a difficult control task. As a means to monitor the attention demanded by the control task, a subjective effort rating was used, as well as control activity and performance metrics.

It was important to analyze these metrics to investigate whether or not the two levels of difficulty of the control task were indeed any different. The duty cycle values and the RMS of the control input were much higher for the difficult task than for the easy task. This result is consistent with what was expected, since for the difficult task, the unstable controlled element required larger deflections of the control stick to attain acceptable performance. The confirmation that the two control tasks were of different enough difficulty levels comes from the large difference found between the effort scores for both tasks.

One other important goal of recording effort scores and control activity metrics was to guarantee that the attention effort dedicated to the control task was constant across the training and the active parts of the experiment. The training sessions set the base line for how much effort should be put into the control task to attain acceptable performance. If subjects maintained this level of attention, then it could be assumed that there would be reduced attention for the perception task.

The analysis of the subjective effort scores, duty cycle values and RMS of the input signal show that, in general, subjects were successful in maintaining a constant level of attention dedicated to the control task even when they were required to perform a perception task simultaneously. The effort scores and the RMS of the control input showed no difference in the perceived effort between the training and the active parts of the experiment.

The duty cycle values, on the other hand, despite similar values between training and active runs in general, show small but significant changes throughout the conditions of the

active part. Namely, the threshold measurement had a significant effect and there were interaction effects for the task difficulty and threshold measurement and task difficulty and frequency. Looking at the average duty cycle values, it seems that for the more difficult task the duty cycle values were fairly constant, but for the easy task they increase with frequency and going from lower to upper threshold. It could be that the simulator inertial motion affected the subjects interaction with the control stick, artificially raising the time subjects were acting on the controls. However, if that would be the case, the same trends should be observed for the difficult task.

The different results found for the duty cycle values and the RMS of the control input are not contradictory. The duty cycle indicates the time is spent changing the input, but it does not characterize how that change is made. The RMS of the control input is one way of doing that characterization. Gray (2009) lists the RMS of the control input rate as an average measure of aggressiveness. He also mentions that a better measured of aggressiveness is the total work applied on the stick (stick displacement times the force applied). He defends that the combination of a high duty cycle and a high aggressiveness result in high pilot inceptor workload. Perhaps analyzing the control effort in these terms would provide a clearer view of the differences across conditions. Nevertheless, the currently used metrics seem sufficient to argue that the control activity was constant between training and active sessions.

The performance, on the other hand, was not constant. The RMS of the error signal was higher for the active part of the experiment than for the training part. In the active part, the upper threshold measurements also resulted in worse performance, that is, larger errors. Despite that the perceived effort scores and the control activity metrics indicate that the control effort was maintained from training to active runs, performance was degraded.

This may imply that although subjects dedicated the same effort to the control task, their control inputs were less accurate, or perhaps delayed in time, in such a way as to result in larger excursions outside the tracking boundaries. One other hypothesis is that subjects were bad judges of their own effort and in fact, despite the effort scores indicating otherwise, they spent less attention on the control task when they were to simultaneously perform the perception task.

Contrary to what was expected and is found in literature (Roark and Junker, 1978; Hosman and van der Vaart, 1980; Samji and Reid, 1992) the coherence thresholds did not increase with the addition of the control task. The general trend of wider coherence zones for more difficult tasks is present, but small and definitely not significant. Furthermore, there was a decrease in the upper threshold during the difficult task at the lowest frequency, which contradicts the expected results. In this case it might be that the task became too difficult and any kind of inertial motion was considered just a distraction from the control task. This could lead some subjects to tune down the inertial motion. However, if that would have been

the case, the upper threshold measurement for the higher frequency should show the same trend. The fact that it only happens at the lowest frequency seems to dismiss this hypothesis, though care should be taken while interpreting trends in the data when there is no statistical significance.

The indifference threshold increase so clearly found in other studies (Roark and Junker, 1978; Hosman and van der Vaart, 1980; Samji and Reid, 1992) was not encountered in the coherence zone thresholds. Although intuitively one might expect the same trend, it may be that the perceptual mechanism involved in a coherence threshold cannot be directly related to the perception mechanism underlying sensory inertial thresholds in the presence of other cues, or during higher workload conditions.

Inertial motion perception thresholds are often referred to as a signal-to-noise ratio mechanism (Greig, 1988). When sensor output rises above neuronal noise, the inertial stimulus is detected. For thresholds in the presence of other cues or during a control task, the signal-to-noise ratio has to be higher than normal before it is detected. For coherence zone measurements, the inertial motion is supra-threshold and the signal-to-noise ratio is high enough so that the individual signals are always detected. The coherence zone threshold is the outcome of the comparison mechanism between inertial and visual stimuli, which may not be so quickly affected by the addition of a control task.

One other explanation for the obtained results is that this comparison is not something subjects perform simultaneously with the control task, but in distinct periods of time. The inertial and visual stimulus comparison can be made during a short period of time within a run, leaving the rest of the time to perform the control task. Although control activity metrics and effort scores were recorded, these values refer to mean values, averaged across runs.

It might be that what is being assessed is then not the effect of attention demanded by the control task, but available time to decide whether or not two cues match. For the perception task subjects had the entire run to decide. For the easy task they could dedicate a fair amount of time for the comparison while still maintaining the controlled element between boundaries. For the difficult task, since the controlled element was unstable, any loss of attention was heavily punished, so subjects had less time for the comparison. Nevertheless, it may be that even for the most difficult condition subjects had enough time to make a decision and in the other conditions they had time to spare. If that is the case, then no difference should be expected in the measured coherence zones.

This being true, for vehicle simulation application, the subjects' acceptation of simulated inertial cues might be measured passively and directly applied to active situations, such as procedural or supervisory tasks. For manual control tasks with inertial motion feedback, if the control task is in a different degree-of-freedom of the one being assessed, the

same might hold, although for definite conclusions more information is needed on how coherence zones change in the presence of inertial cues unrelated to the perception task.

If the manual control task is in-the-loop, that is, if there is inertial feedback relevant to the control task in the same degree-of-freedom as the one being assessed, the passively measured coherence zones will probably not suffice to decide on the inertial motion feedback characteristics. In this case, subjects need the inertial cues for performance, and inertial motion affects not only perception but also behavior (Shirley and Young, 1968; Pool et al., 2011b).

The frequency had a significant effect on the measured thresholds. As also seen in the previous chapters, for higher frequencies subjects prefer less motion. The PMC decreases from above the one-to-one line at 2 rad/s to being close to a physical match with the visual stimulus at 5 rad/s. The CZW also decreases with the decrease in the PMC. The effect of frequency on the PMC has been qualitatively explained in Chapter 4 with the fact that at higher frequencies the dynamics of the SCC have a higher gain in velocity which is not accounted for in the internal comparison of visual and inertial cues. This effect is present in all the threshold measurements indicating that not only for the perception task only, but also for the active tasks, this relationship remains valid.

6.5 Conclusions

Perception coherence zones for yaw motion were measured during passive and active situations. During the active part of the experiment, subjects were required to perform an out-of-the-loop pitch boundary-avoidance task. Two different levels of difficulty for the control task were designed.

The perception coherence zones were not significantly affected by the addition of the control task. It seems that the decision whether or not the inertial and visual stimuli were a match could be done in a short period of time, leaving the rest of each run to perform the control task. This means that there is a time separation of the perception task and the active control task.

This being the case, in flight simulation applications, the acceptable range of inertial motion amplitudes for a certain visual stimulus may be determined in a passive manner and directly used in situations were the pilot has a task to perform, such as procedural training and supervisory tasks. Also for manual control tasks, the passively determined coherence zones might apply, if the control task is in a different degree-of-freedom of the one being assessed. The same does not apply, however, for control tasks in-the-loop, where inertial motion feedback is important not only for perceptual fidelity but also to attain adequate performance and control behavior.

7

Comparing Simulators Using Perception Coherence Zones

This chapter is partially based on the following publication:

Correia Grácio, B. J., Valente Pais, A. R., Van Paassen, M. M., Mulder, M., Kelly, L. C., and Houck, J. A. (2011). Relationship Between Optimal Gain and Coherence Zone in Flight Simulation. In *Proceedings of the AIAA Modeling and Simulation Technologies Conference, Portland, OR, USA, August 8-11*, AIAA 2011-6556.

7.1 Introduction

The use of motion filters in flight simulators ensures that the mechanical limits of the simulator are not reached while providing inertial cues to the pilots. These motion filters introduce amplitude and phase distortions with respect to the real aircraft motion and to the displayed visual cues. Despite the fact that motion simulators have been used commercially since the 1930s, only in the last decade some serious effort has been done to include the characteristics of the motion filters in the quality assessment of flight simulators (Royal Aeronautical Society, 2005; International Civil Aviation Organization, 2009). Although only recently the motion filters are being taken into account by regulatory authorities, their influence on the quality of the simulation and the impact on pilot perception and behavior has been an important topic of research in the scientific community (Samji and Reid, 1992; Schroeder, 1999).

Many studies have focused on understanding pilot perception in simulator environments (Roark and Junker, 1978; Hosman and van der Vaart, 1978, 1980; Samji and Reid, 1992; Zaichik et al., 1999; Groen et al., 2001; Bos et al., 2002; Greenberg et al., 2003; Groen and Bles, 2004; Heerspink et al., 2005; Fortmüller and Meywerk, 2005; Grant et al., 2006; Naseri and Grant, 2011) and others have investigated the effect of motion cues and motion filter settings on pilot behavior (Shirley and Young, 1968; Jex et al., 1978; Greig, 1988; Grant et al., 2005; Pool et al., 2010, 2011a). There also has been some effort in categorizing different filter settings in equivalent fidelity regions (Mudd, 1968; Schroeder, 1996; White and Rodchenko, 1999; Schroeder, 1999).

The so-called Sinacori plots relate motion filter gain and phase distortion at the frequency of 1 rad/s to three levels of simulation fidelity: low, medium and high. These fidelity levels have been determined using the judgment of expert pilots during specific tasks in the simulator. They also assumed a fixed structure for the motion filter, such that by knowing the gain and phase distortion at one frequency, the other properties of the filter are also known. This approach starts with constraints regarding motion filter design and then derives fidelity levels based on pilot judgment and performance.

A different approach is to start by determining what is considered acceptable and desirable in terms of human motion perception, or using the current terminology, what is high or low fidelity, then design motion filters that fit those requirements. This second approach starts with constraints regarding human motion perception and then derives fidelity levels for motion filter design.

Research on human motion perception in flight simulation scenarios can be considered essential for the second approach. Especially studies that focus on the combined perception of visual and inertial stimuli (Hosman and van der Vaart, 1981; Grant and Lee, 2007) are of importance, since the differences in amplitude and phase introduced by the motion filters

result in different visual and inertial stimuli being provided to the pilots. The coherence zone research is one of these studies.

Amplitude coherence zones are influenced by a number of factors such as the amplitude (Chapter 3) and frequency (Chapter 4) of the visual stimuli. Coherence zones seem to also be significantly affected by the type of visual and motion system used. In Chapter 4, data from two different simulators was compared and it was observed that the measured coherence zones were different between simulators. Although the differences were statistically significant only for the CZW and not for the PMC, such a result indicates that when transforming coherence zone knowledge to motion filter design and tuning, attention should be given to the simulator specific configuration in terms of motion base and visual system.

From this, one can say that the evaluation of the motion base performance and motion filter settings as the sole basis for assessing the motion cueing quality of a simulator might be incomplete, as it does not consider the combination of visual and motion systems as a whole. For different visual systems, the same motion filter settings may result in different perceptual fidelity levels, and vice-versa.

To investigate whether coherence zones could be used as a metric to complete the current fidelity assessment methods, it is necessary to test how much of the differences among simulators can be captured by coherence zones. To this purpose, an experiment was designed that measured coherence zones in three different simulators.

This experiment was part of a larger project and since many of the decisions regarding the experimental design can only be fully understood when considering the whole project, the following paragraphs briefly explain the the project.

In a recent study, the perception of inertial motion strength as compared to visual motion has been investigated using a different concept than the coherence zones. In a study by Correia Grácio et al. (Correia Grácio et al., 2010) the matching of inertial and visual linear lateral motion has been studied using the concept of optimal gain. Using sinusoidal signals as stimuli, they asked their subjects to tune the inertial motion amplitude such that it would optimally match the amplitude of the presented visual scene motion.

To the ratio of the amplitude value chosen by subjects over the amplitude of the visual signal they called optimal gain. This gain represents the factor by which the vehicle motion should be filtered to obtain inertial motion that matches the visually perceived motion, that is, the optimal gain corresponds to a possible motion filter gain.

One interesting finding from this study was that the measured optimal gain values depended on the initial inertial amplitude provided to the subjects. When the amplitude of the inertial motion of their first run was well above the amplitude of the visual stimulus, subjects tended to choose higher optimal gains than when having a first run with a lower inertial amplitude. As a consequence, the optimal gain value became an optimal gain zone, delimited by the higher and lower values found.

From this result, the question arose to whether the measured optimal gains were not in fact, a coherence zone. To investigate what was the relationship between optimal gain and coherence zone, and whether these two metrics could indeed differentiate between simulator configurations, a study was designed where optimal gain measurements and coherence zone measurements in sway were made using the same subjects, stimuli and identical measurement methods in three different simulators (Correia Grácio et al., 2013).

This chapter describes the part of the experiment concerning coherence zones only.

7.2 Method

7.2.1 Apparatus

The experiment was conducted in the Visual Motion Simulator (VMS) and the Cockpit Motion Facility (CMF), located at NASA Langley Research Center (LaRC) in Hampton, Virginia, USA. The CMF consists of one motion base and three interchangeable simulator cabins. For this experiment, two different cabins were used on the CMF, the Generic Flight Deck (GFD) and the Integration Flight Deck (IFD). Figure 7.1 shows the VMS and CMF motion bases.

(a) VMS. **(b)** CMF.

Figure 7.1: The VMS and CMF motion bases. (Courtesy of NASA.)

7.2.1.1 VMS

The VMS has a six DOF, hexapod type, motion base with an actuator stroke of 1.5 m. The maximum displacement, velocity and acceleration in the lateral axis are ± 1.2 m, ± 0.6 m/s and ± 5.9 m/s^2, respectively. For motion control, commands were sent to the motion base at 50 Hz.

The visual system consists of four Wide Angle Collimated (WAC) windows of 1024 by 944 pixels with an update rate of 60 Hz. The front windows provide a field-of-view (FoV) of 65.93 deg horizontal by 45.23 deg vertical. The lateral window on the left-hand side has a FoV of 48.50 deg horizontal by 35.50 deg vertical.

The lateral acceleration of the motion base was measured using Sundstrand QA-900 accelerometers (serial number 1271). These have been tested and calibrated to $\pm$2g at up to 100 Hz, and have deviations of less than 1000 μg.

7.2.1.2 GFD

The motion base of the CMF is also a 6 DOF, hexapod motion base, although it is larger and newer than the VMS. The actuator stroke is 1.9 m and the lateral motion limits are $\pm$1.4 m, $\pm$1.0 m/s, and $\pm$6.9 m/s^2 in position, velocity and acceleration, respectively. The motion base was controlled at 50 Hz.

The GFD visual system consists of four WAC windows with an update rate of 60 Hz. The front windows have a horizontal FoV of 46 deg and a vertical FoV of 34 deg. The lateral windows have a FoV of 49 deg horizontal by 37.5 deg vertical. Although the FoV is smaller in this cockpit than in the VMS, the resolution of the screens is higher, with each window having 1280 by 1024 pixels.

The lateral acceleration was recorded with a Honeywell Q-Flex(R) accelerometer (Model QA-700) which was placed under the dynamic platform at the centroid position.

7.2.1.3 IFD

The IFD cockpit is also part of the CMF, so the motion base description given for the GFD also apply to this simulator.

The visual system is quite different from the GFD and VMS. It consists of a collimated panoramic display, with a horizontal FoV of 200 deg and a vertical FoV of 40 deg. Four projectors are used, each with 1440 by 1024 pixels, and an update rate of 60 Hz.

The interior of the three simulators are shown in Figure 7.2.

7.2.2 Experimental design

The experiment had a three way repeated measures design. The independent factors considered were the three simulators described above, two visual stimulus amplitudes, and two stimulus frequencies.

The visual stimulus amplitudes used were 0.5 and 1 m/s^2. These amplitudes were chosen such that the results could be directly compared to previous studies on optimal gain tuning performed in other simulators.

The choice of frequencies was less straightforward. Initially three frequencies of 2, 3, and 5 rad/s were chosen for the GFD and IFD part of the experiment. The lowest frequency

(a) VMS. **(b)** GFD. **(c)** IFD.

Figure 7.2: View of the interior of the three cabins. (Courtesy of NASA.)

of 2 rad/s was the lowest possible frequency to be tested while still remaining within the motion base limits. For the VMS the minimum frequency was 3 rad/s, however, so only two frequencies would be tested in this simulator.

During preliminary tests it became clear that because both coherence zones and optimal gain measurements were being performed, the experimental sessions were too long and there was the risk that subjects would become too tired. For this reason, one of the frequencies was eliminated from the tests in the GFD and IFD. To maintain symmetry with respect to the tests in the VMS, it would have been better to eliminate the 2 rad/s condition. However, it was thought that maintaining this low frequency would allow a more direct comparison to results from other studies that used the same stimulus frequency. Moreover, the larger the differences between tested frequencies the easier it would be to observe the effect of frequency on the coherence zones. It was then decided to maintain the 2 and 5 rad/s conditions for the GFD and IFD and test the 3 and 5 rad/s conditions in the VMS. With this design, comparison among the three simulators can be done only at the frequency of 5 rad/s.

For each condition two measurements were taken, one for the upper threshold and one for the lower threshold. For each of these measurements three repetitions were made, resulting in a total of 24 experimental trials in each simulator.

7.2.3 Motion and visual signals

The visual and inertial motion stimuli consisted of sinusoidal signals with amplitude and frequency defined by the experimental conditions described above. The signals were designed such that experimental runs of different frequencies would have the same duration. The length of the motion signals were 2, 4 and 8 periods for the conditions with frequencies of 2 rad/s, 3 rad/s and 5 rad/s, respectively.

These sinusoidal signals were faded in and out to guarantee that the acceleration, velocity and position signals always started and ended at zero. The fade in and fade out parts of the signal are described by Equation (7.1), where A is the amplitude in m/s^2, ω is the

signal frequency in rad/s, and ω_s and ω_c are the smoothing and compensation frequencies, respectively, also in rad/s. Both the smoothing and the compensation frequencies equaled half of the signal frequency.

$$f(t) = \frac{1}{2}A\sin(\omega t) - \frac{1}{2}A\sin(\omega t)\cos(\omega_s t) + A_c\sin(\omega_c t) \tag{7.1}$$

The complete motion signal is given by Equation (7.2) where T is the period of the signal and N is the number of periods in one run. The number of periods does not include the two periods that are necessary to perform the fade in and fade out. Including the fade in and the fade out, the total length of one run was 12.57 seconds.

$$a(t) = \begin{cases} f(t), & 0 < t \leq T \\ A\sin(\omega t), & T < t \leq (N+1)T \\ f(t-T), & (N+1)T < t \leq (N+2)T \end{cases} \tag{7.2}$$

Figure 7.3 shows examples of complete runs for all three frequencies and an amplitude of 1 m/s^2.

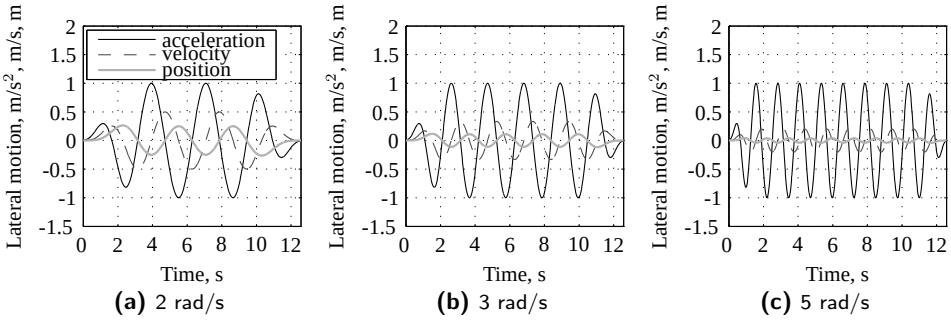


Figure 7.3: Example of the motion signals during one run for the three different frequencies and an amplitude of 1 m/s^2.

7.2.4 Procedure

The experiment was divided in three parts. The first part was conducted in the GFD, the second part in the VMS and the third part in the IFD. Due to simulator scheduling it was not possible to have all simulators available at the same time, so randomization between simulators was not possible. There was one month separating the tests in the GFD and in the VMS and four months between the VMS and the IFD parts. The same subjects were used in all three simulators. The order of the 24 experimental trials performed in each

simulator was randomized for every subject. For each subject, the trial order was the same on all three simulators.

Subjects were seated in the left-hand chair of the simulator cabin. The subject wore a headset with active noise cancellation which allowed to communicate with the experiment supervisor. Three buttons located in the sidestick, on the left side of the participants were used to record their answers throughout the experimental runs.

For each experimental trial, the visual motion amplitude was kept constant and the inertial motion amplitude was varied through a set of runs. In each trial, the inertial motion amplitude of the first run was randomly selected between 1.1 and 0.9 of the visual amplitude. Before each trial started, subjects were informed whether that trial corresponded to a lower or an upper threshold measurement.

At the end of each run within a trial subjects could change the motion of the next run. They did this by pushing a switch button multiple times up or down until they reached a certain number of increments or decrements. The chosen number was shown on a head-down display placed directly in front of the subjects. A positive number meant the next run would have a higher amplitude motion, and a negative number meant a lower amplitude motion. After giving their answer, subjects pressed a second button to signal that they were ready for the next run.

The trial ended when subjects' answers had two consecutive reversals of one increment or decrement, i.e., a sequence of 1, -1, 1, or -1, 1, -1. This indicated that subjects converged to a certain amplitude of motion that could not be increased or decreased anymore. To avoid fatigue, trials were also stopped if subjects reached 30 runs. The size of the increment or decrement was 0.025 of the visual amplitude, which corresponded to 0.0125 m/s^2 for the lowest amplitude condition and 0.050 m/s^2 for the highest amplitude condition.

Before starting the experiment, subjects performed three randomly chosen experimental trials for training purposes.

7.2.5 Subjects and subjects' instructions

Eight subjects were selected from the employees of the LaRC Flight Simulation Facility. There were seven male participants and one female participant. The subjects' average age was 49 years, ranging between 31 and 64 years old. All eight subjects were able to complete the experiment, and there were no complaints of motion sickness.

The participants were instructed to sit upright and refrain from making head movements throughout the experiment. They were, however, allowed to gaze over the visual scene at will.

Participants were told they were to perform a series of experimental trials which consisted of several runs. In each trial the visual scene would move the same way but the amplitude of the simulator inertial motion would vary between runs depending on their input.

Subjects were told at the beginning of the trial whether an upper or a lower threshold measurement was being performed. For an upper threshold measurement subjects were asked to find the strongest inertial motion amplitude that was still perceived as coherent with the visual cue. For a lower threshold measurement they were asked to find the weakest inertial motion amplitude that was still coherent with the visual cue. Subjects were instructed to decrease and increase the inertial motion amplitude as many times as needed until they were satisfied with their choice. Subjects were advised to start with increments of 10 or more and decrease the number of increments or decrements at every direction reversal. They were informed of the stopping criteria of the trials.

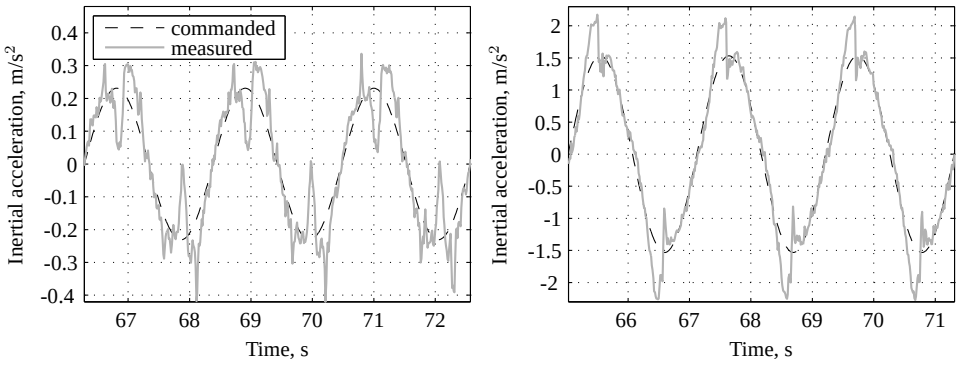
7.3 Results

To observe the differences in motion base performance between the three simulators, the commanded lateral acceleration was compared to the measured signals for each of the three simulators. Figure 7.4 and Figure 7.5 show time histories of the commanded and measured signals for all frequencies at high and low amplitudes. The amplitudes chosen to be plotted are around 0.2 m/s^2 and 1.5 m/s^2 for the commanded signal, and are representative of the limits of the search interval for most subjects.

As can be seen, there were quite some differences in performance between the VMS and the CMF motion bases. The VMS showed a more pronounced “turn around bump” at the low amplitude and low frequency conditions and a clear overshoot at the high amplitude conditions. The low amplitude conditions at a frequency of 5 rad/s seemed to be equally demanding for both motion bases. In these conditions the time histories clearly showed an additional oscillation around 15 rad/s for the VMS and 20 rad/s for the CMF. Although these oscillations were more pronounced at the low amplitude and high frequency conditions, a spectral analysis of the time histories showed that they were present in all other conditions as well.

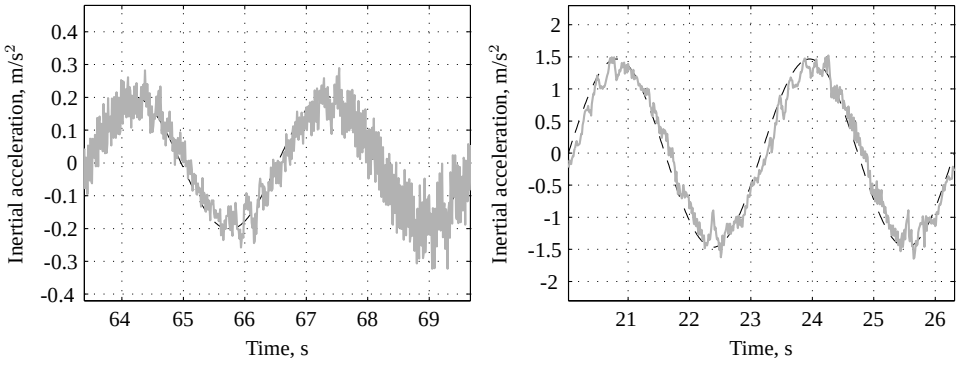
The data from the CMF motion base also showed oscillations around 180 rad/s, which were stronger with the GFD than with the IFD cabin. The oscillations can clearly be seen in the time histories in Figure 7.4c and Figure 7.5c. Although both cabins were mounted on the same motion base, differences in mass of the cabin, if not fully taken into account by the motion base controller, may have affected the overall performance.

Since there were considerable differences between the commanded and the measured motion, the thresholds should be based not on the amplitudes of the commanded signals, but on the amplitudes of the measured signals. However, some care should be taken, since some of the peak amplitudes of the measured signal might have been caused by noise inherent to the accelerometer and not motion of the platform.



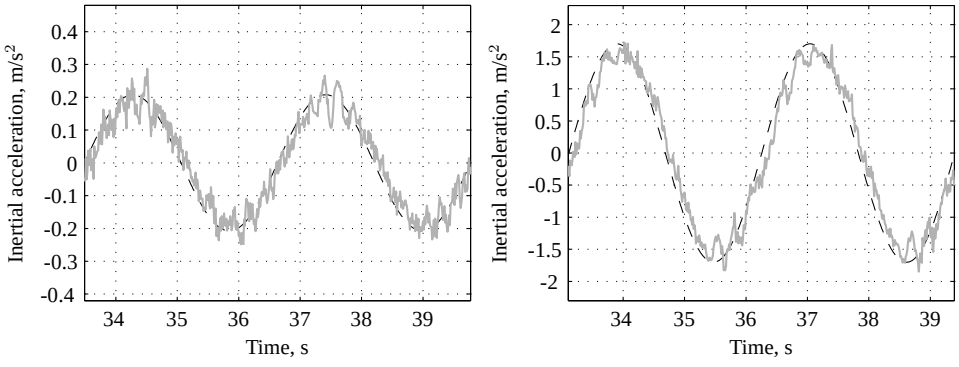
(a) VMS, 3 rad/s, low amplitude.

(b) VMS, 3 rad/s, high amplitude.



(c) GFD, 2 rad/s, low amplitude.

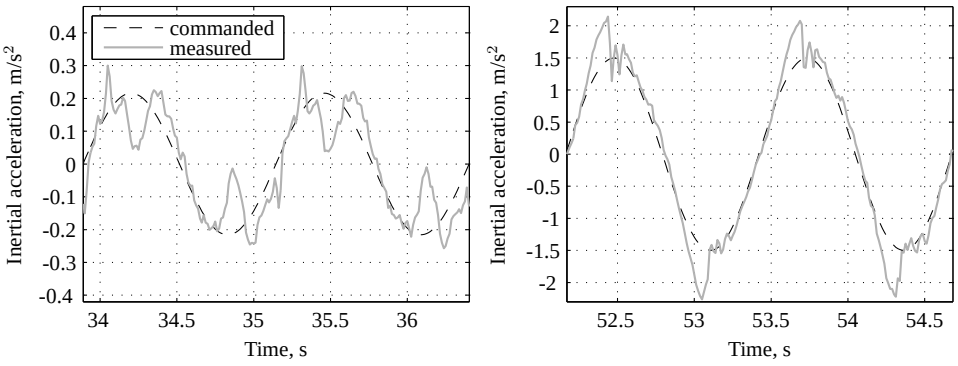
(d) GFD, 2 rad/s, high amplitude.



(e) IFD, 2 rad/s, low amplitude.

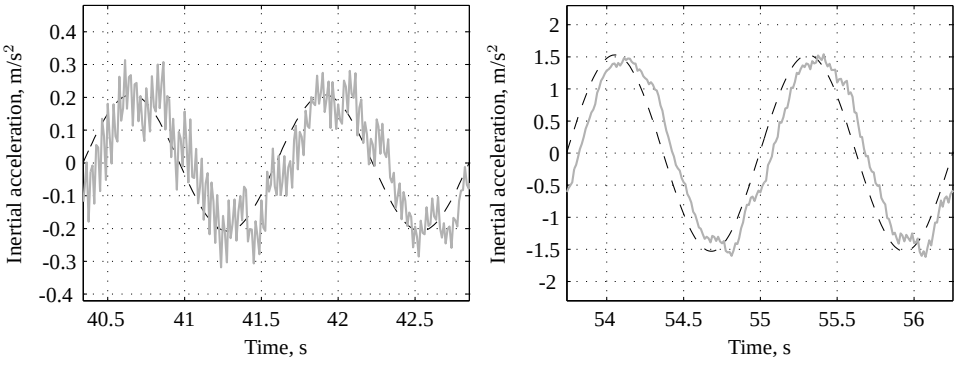
(f) IFD, 2 rad/s, high amplitude.

Figure 7.4: Time histories of the commanded and the measured lateral acceleration for frequencies of 3 and 2 rad/s.



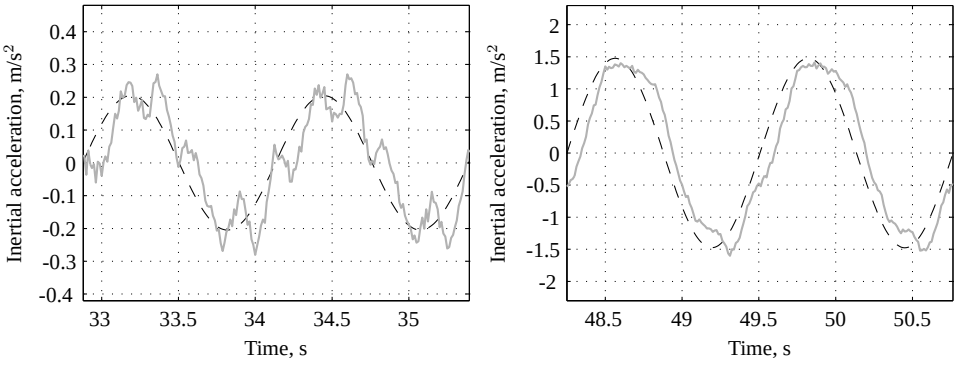
(a) VMS, 5 rad/s, low amplitude.

(b) VMS, 5 rad/s, high amplitude.



(c) GFD, 5 rad/s, low amplitude.

(d) GFD, 5 rad/s, high amplitude.



(e) IFD, 5 rad/s, low amplitude.

(f) IFD, 5 rad/s, high amplitude.

Figure 7.5: Time histories of the commanded and the measured lateral acceleration for the 5 rad/s frequency.

To preserve the measured accelerations and eliminate high frequency noise that is unlikely to have resulted from the inertial motion of the platform, the measured signals were filtered with a second order filter with a cutoff frequency of 50 rad/s. Examples of the measured signals before and after the filtering are shown in Figure 7.6. Only examples of the low amplitude runs are plotted, since in these the differences between the measured and filtered signals are more noticeable.

The threshold values were determined by averaging the peak amplitude of the last two runs in each trial. The so determined thresholds were then averaged across the three repetitions of each condition. Especially for the VMS, the differences in peak amplitudes between the commanded and the filtered signals were quite large, so a large difference can be expected between the thresholds if they are determined from the commanded signals or from the filtered signals. For the tests in the VMS, the thresholds determined from the commanded signals are shown together with the thresholds determined from the filtered signals in Figure 7.7.

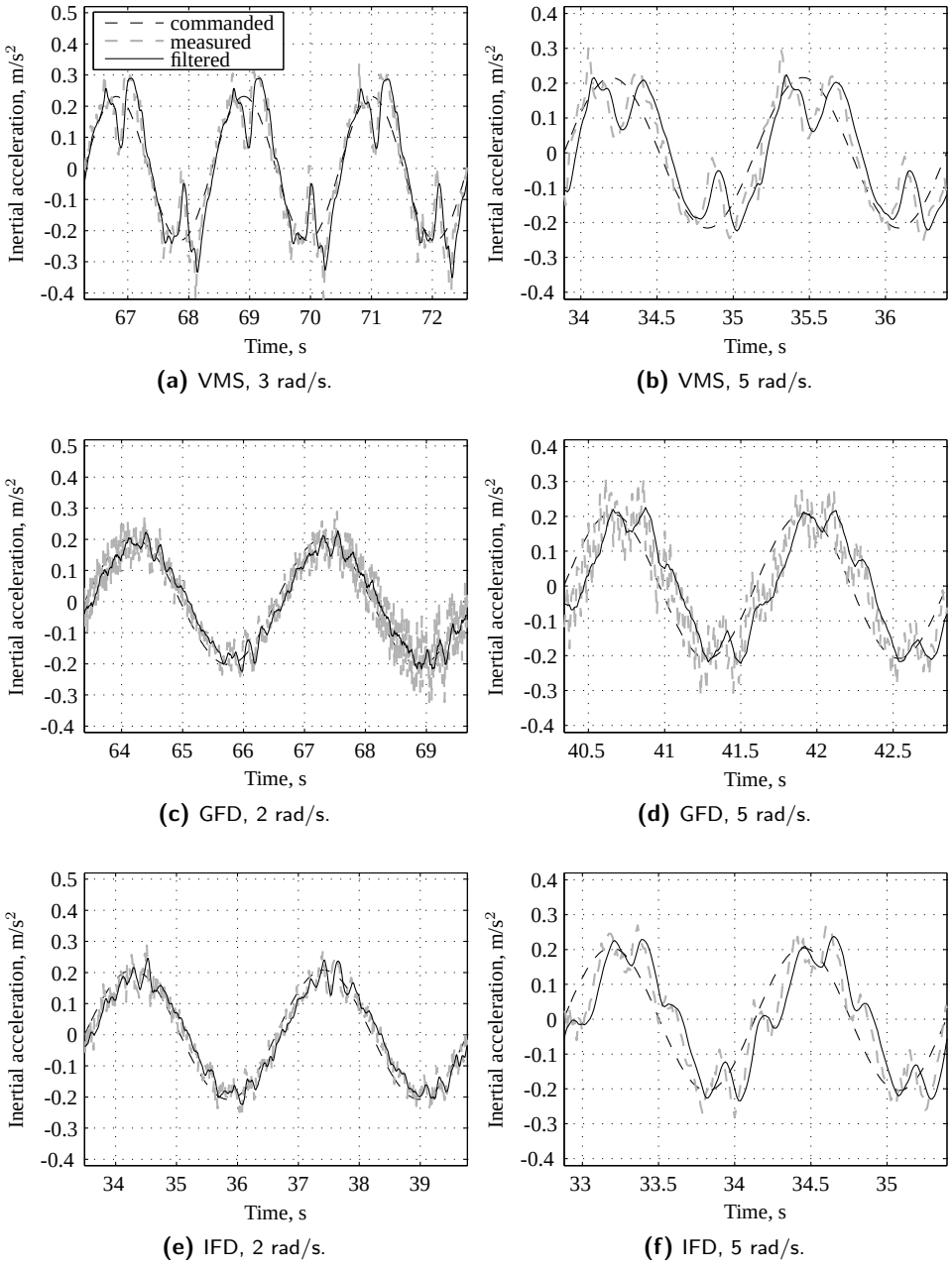
As can be seen, there was indeed a considerable difference between the commanded signal thresholds and the filtered signals thresholds, especially for the upper thresholds of the high amplitude conditions. Although the difference between commanded and filtered thresholds for the other simulators was negligible, for the sake of consistency, all thresholds were determined from the filtered signals. Therefore, except where explicitly mentioned otherwise, all threshold values presented here were determined from the filtered signals.

Figure 7.8 shows the determined thresholds for all conditions in the VMS. In this and all following figures, the error bars indicate the 95% confidence interval of the mean.

For both frequencies and amplitudes of the visual stimulus, the upper and lower thresholds defined a coherence zone that included the one-to-one line. The variation of the lower threshold values across conditions was small, whereas the upper thresholds clearly increased with the visual stimulus amplitude and decreased with frequency.

A repeated measures ANOVA was performed to investigate whether the observed differences were significant. The results of the analysis are shown in Table 7.1. All data were normally distributed. Both the frequency and the amplitude had a significant effect on the upper thresholds, but not on the lower thresholds. There was no significant interaction effects.

The thresholds determined in the GFD and IFD are shown in Figure 7.9. Similarly to the results from the VMS, the visual stimulus amplitude and frequency had a more noticeable effect on the upper thresholds than on the lower thresholds. For both simulators, there was a very small increase in the lower thresholds for the higher amplitude and a slight decrease for the higher frequency conditions. The same trend can be seen for the upper thresholds, but here the differences are much larger. The upper thresholds measured in the GFD were slightly lower than the ones measured in the IFD, and the lower thresholds were slightly higher.

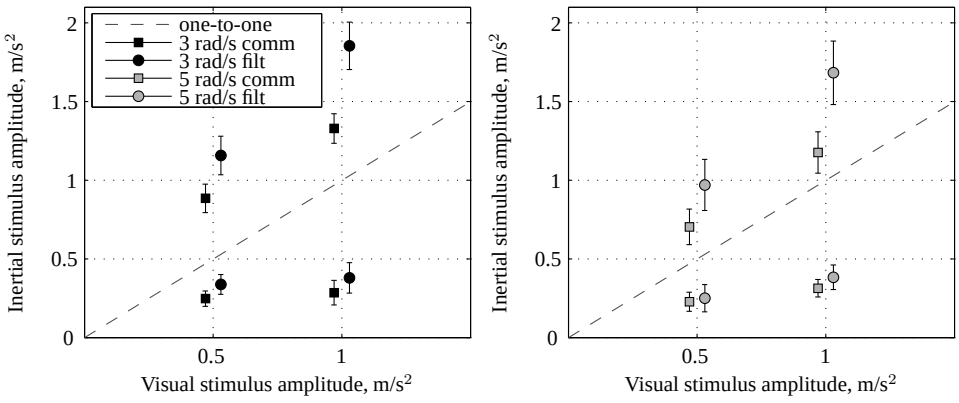


(a) VMS, 3 rad/s. **(b)** VMS, 5 rad/s.

(c) GFD, 2 rad/s. **(d)** GFD, 5 rad/s.

(e) IFD, 2 rad/s. **(f)** IFD, 5 rad/s.

Figure 7.6: Time histories of the measured and the filtered lateral acceleration signals for all frequencies and low amplitude inertial motion.



(a) 3 rad/s. **(b)** 5 rad/s.

Figure 7.7: Upper and lower thresholds based on commanded (comm) and filtered (filt) signals in the VMS. Error bars indicate the 95% confidence interval of the mean.

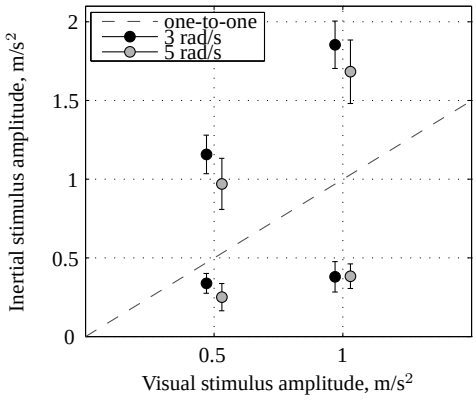
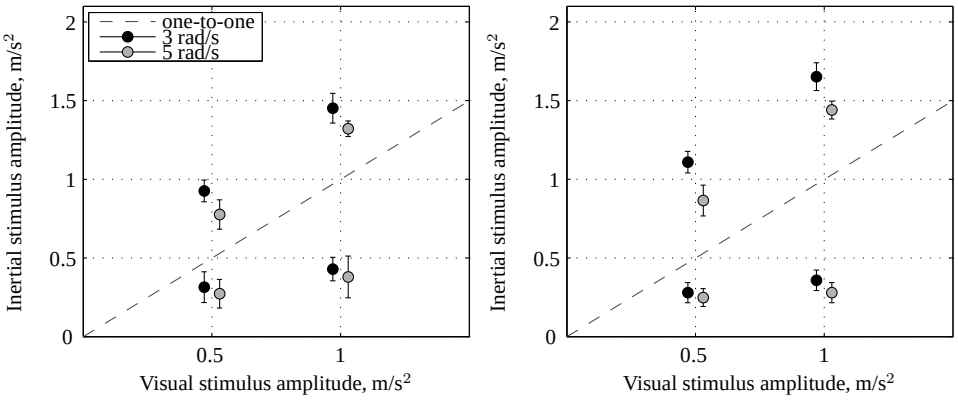


Figure 7.8: Upper and lower thresholds for the VMS.

To investigate whether these effects were significant, an ANOVA was performed on the data. Table 7.2 shows the results of this analysis. Only the main effects and significant interactions are shown. The effect of amplitude and frequency on the lower threshold values was not significant. Moreover, the lower thresholds from the two simulators were not significantly different. The upper threshold values were significantly affected by the simulator, amplitude and frequency factors. In addition, the effect of frequency on the upper thresholds

Table 7.1: Results of the ANOVA for the measured thresholds in the VMS, where ** is highly significant ($p \leq 0.01$), * is marginally significant ($p \leq 0.05$), and - is not significant ($p > 0.05$).

Independent variables	Dependent measures					
	Lower threshold			Upper threshold		
	df	F	sig.	df	F	sig.
Amplitude	1,7	2.47	-	1,7	48.36	**
Frequency	1,7	1.60	-	1,7	5.59	*
Amplitude × Frequency	1,7	3.07	-	1,7	0.02	-



(a) GFD. **(b)** IFD.

Figure 7.9: Upper and lower thresholds for the GFD and the IFD.

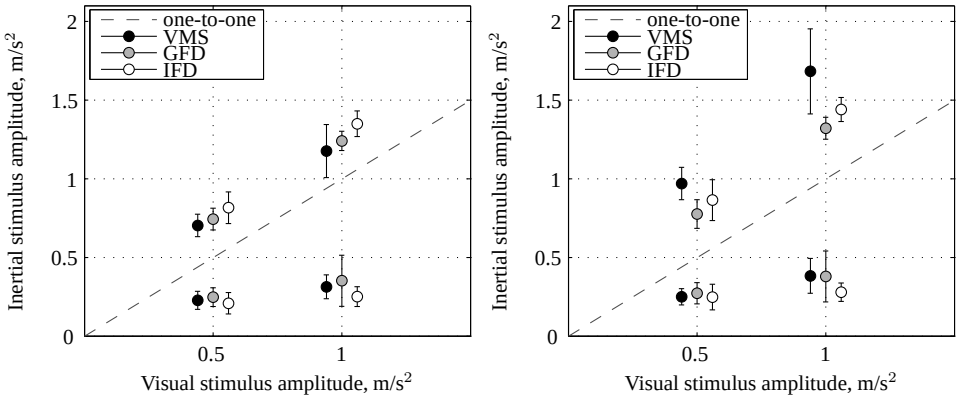
was larger on the IFD than on the GFD, as is confirmed by the significant interaction effect found between simulator and frequency.

The measured thresholds from all three simulators can be compared only for the highest frequency, since in the VMS the lower frequency was 3 rad/s and not 2 rad/s as in the GFD and IFD. The upper and lower thresholds at 5 rad/s and at the two amplitudes for all simulators are shown in Figure 7.10. For comparison purposes both the thresholds calculated from the commanded and from the measured signals are shown.

The lower thresholds did not vary much from one simulator to the other. The largest differences were seen in the upper threshold values. The VMS showed the highest values, especially for the conditions with the 1 m/s^2 amplitude. The GFD showed the lowest upper thresholds. An ANOVA was performed on the data. The results are shown in Table 7.3.

Table 7.2: Results of the ANOVA for the measured thresholds in the GFD and IFD, where ** is highly significant ($p \leq 0.01$), * is marginally significant ($p \leq 0.05$), and - is not significant ($p > 0.05$).

Independent variables	Dependent measures					
	Lower threshold			Upper threshold		
	df	F	sig.	df	F	sig.
Simulator	1,7	2.48	-	1,7	24.84	**
Amplitude	1,7	2.56	-	1,7	156.07	**
Frequency	1,7	2.96	-	1,7	21.77	**
Simulator × Frequency	1,7	0.05	-	1,7	7.37	*



(a) Commanded signals. **(b)** Measured signals.

Figure 7.10: Upper and lower thresholds, for the 5 rad/s frequency conditions, in the VMS, GFD and IFD, obtained from commanded and measured signals.

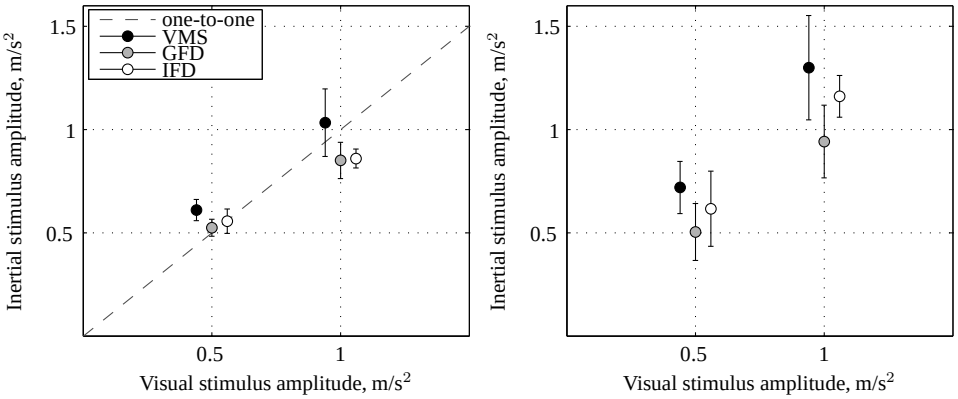
The lower thresholds were not significantly different across conditions and simulators. The simulator and the amplitude had a significant effect on the upper thresholds. Although the larger difference seemed to be between the VMS and the other two simulators, a post-hoc pairwise comparison, using Bonferroni adjustments for multiple comparisons, showed that the VMS and the IFD were not significantly different, but the GFD was significantly different from the VMS and the IFD.

The coherence zones limited by the upper and lower thresholds can also be represented in terms of a PMC and a CZW. Figure 7.11 shows the results in the three simulators in terms of PMC and CZW.

Table 7.3: Results of the repeated measures ANOVA for the measured thresholds in all three simulators for a stimulus frequency of 5 rad/s, where ** is highly significant ($p \leq 0.01$), * is marginally significant ($p \leq 0.05$), and - is not significant ($p > 0.05$).

Independent variables	Dependent measures					
	Lower threshold			Upper threshold		
	df	F	sig.	df	F	sig.
Simulator	2,14	1.32	-	1.19,10.18[a]	7.96	*
Amplitude	1,7	3.15	-	1,7	72.64	**
Simulator × Amplitude	2,14	1.03	-	12,14	2.03	-

[a] Greenhouse-Geisser sphericity correction applied.



(a) Point of mean coherence. **(b)** Coherence zone width.

Figure 7.11: Measured coherence zones.

The PMCs increased with amplitude, approximately following the one-to-one line, although for the higher amplitude they were slightly lower with respect to this line than for the lower amplitude conditions. The VMS had the largest PMCs, which is to be expected, given that the upper thresholds were also much higher in this simulator. The CZWs were also larger for the higher amplitude conditions. The VMS had the wider coherence zones, followed by the IFD. The GFD presented the narrowest zones.

The ANOVA results, shown in Table 7.4, indicate that the effect of amplitude was significant on both the PMC and the CZW. The simulator factor was also significant on both metrics, although it was not possible to statistically differentiate the CZWs in the three sim-

ulators using a post-hoc pairwise comparison (again using Bonferroni adjustment). For the PMC the post-hoc tests showed a significant difference between the VMS and the GFD, but not between the VMS and the IFD and between the GFD and the IFD.

Table 7.4: Results of the repeated measures ANOVA for the PMC and CZW in all three simulators for a stimulus frequency of 5 rad/s, where ** is highly significant ($p \leq 0.01$), * is marginally significant ($p \leq 0.05$), and $-$ is not significant ($p > 0.05$).

Independent variables	Dependent measures					
	PMC			CZW		
	df	F	sig.	df	F	sig.
Simulator	1.20,8.41[a]	5.39	*	2,14	7.44	**
Amplitude	1,7	108.49	**	1,7	24.82	**
Simulator $\times$ Amplitude	1.95,8.37[a]	1.67	$-$	2,14	1.51	$-$

[a] Greenhouse-Geisser sphericity correction applied.

7.4 Discussion

The three motion simulators used in this experiment combine two types of visual systems, WAC windows and collimated panoramic display, and two different motion bases. Both motion bases are hexapods, but the VMS is a considerably older platform than the motion base of the GFD and IFD. The performance of the motion bases were monitored by looking at the commanded and measured motion signal.

7.4.1 VMS

As was to be expected, the older motion platform, the VMS, showed larger turnaround "bumps" that were relatively worse at the lower amplitudes. This situation was known a priori and the worse performance of the motion base was thought to have an influence especially on the lower threshold values, since here the amplitudes of the inertial motion were lower. It was thought that the large turnaround bump could increase the overall perception of strength of the inertial motion, leading the subjects to tune down the motion for both high and low amplitudes.

However, at the very low amplitudes, possible crosstalk (motion in other DOFs) might mask the inertial lateral motion, which could lead subjects to tune the motion up, such that motion in the lateral DOF could be distinguished from the crosstalk. The lower thresholds results show that if any of these issues played a role, they either canceled each other out, or

the effect was too small to be measurable.

The turnaround bump might also have influenced the measured upper thresholds in the VMS. The threshold values obtained from the commanded signals where clearly lower than the ones computed from the measured and filtered signals. The upper thresholds from the commanded signals were in fact lower than the ones obtained in the other two simulators, but when using the measured signals, the peak amplitudes due to the turnaround bump result in higher threshold values.

On the one hand, one might say that including the peak amplitude of the turnaround bump into the calculation is wrong because it artificially raises the threshold value. On the other hand, it is also not correct to base the threshold values on the peak amplitude of the commanded signals instead of the measured signals when the difference is so large. It was considered that presenting the results in terms of actual measured amplitudes was the most correct. Nevertheless, both the commanded and measured threshold values in the VMS were computed for comparison purposes.

When comparing the commanded upper threshold values in the VMS, to the ones in the GFD, which has the same type of visual system, one might argue that if the turnaround was the cause of lower values then considering the peak amplitude of the turnaround bump in the threshold computation should result in similar thresholds. The fact that the measured thresholds are in fact higher than the ones in the GFD, indicate that although the turnaround bump influences subjects' perceived motion strength, it does not account for the whole difference between the VMS and the GFD.

The turnaround bump represents a high-frequency component of the motion signal. Subjects perhaps based their judgment on the lower frequency component of the signal and took the turnaround bump only partially in consideration.

The measured coherence zones in the VMS show a significant effect of frequency, with higher frequencies leading to lower upper thresholds. This result is similar to what was found for yaw motion in Chapters 4 and 6 and it is in agreement with previous studies that showed that subjects judge motion strength not only on the basis of acceleration but also of jerk (Grant and Haycock, 2006).

In Chapter 4 it was argued that the influence of frequency on the yaw coherence zone might be related to the dynamics of the SCC. The gains applied to the inertial stimulus by the SCC might not be taken into account during the internal comparison of visual and inertial stimuli resulting in a deviation of the PMC from the one-to-one line. If the same rationalization was to be applied to the lateral motion, one would expect a similar but weaker effect, since the gain of the otoliths also increases with frequency (Heerspink et al., 2005), but only slightly, for a frequency range between 1 and 10 rad/s. This may help explain the variation in the upper threshold values with increasing frequency, but it is not agreement with what was found for the lower threshold values.

Surprisingly, the lower thresholds were not affected by the stimulus frequency. One explanation for this fact might be the performance of the motion platform at the low amplitudes. The crosstalk might have masked the lateral motion to such an extent that it hindered subjects in their task. One other explanation could be the relatively lower resolution of the measurement method at the lower amplitudes. In the experimental setup, the minimum amplitude intervals that are tested depend on the amplitude of the visual stimulus, so they are the same for the lower and the upper threshold trials. A small enough amplitude interval at the higher amplitudes runs may be too large to capture the slight differences in lower threshold measurements. Of course, it may happen that for these specific combinations of lateral motion amplitudes and frequencies the lower thresholds of the perceived coherence zone are indeed very similar.

The amplitude of the visual stimulus also had a significant effect on the upper thresholds but not on the lower thresholds. Again, the lower threshold runs might have been affected by the resolution of the experimental method. Although not immediately observable from the upper and lower threshold plots, very similar to what happens for yaw coherence zones, the coherence zone bends down with respect to the one-to-one line. This effect is easier to see by looking at the PMC values. For the higher amplitude the PMC value is still higher than 1 m/s^2, but it is much closer to the one-to-one line than for the lower amplitude.

7.4.2 GDF and IFD

The thresholds measured in the GFD and IFD were also based on the measured signals. Unlike the VMS, the difference between the commanded and measured signals peak amplitudes in these two simulators was small, which resulted in very similar commanded and measured threshold values.

The differences in weight of the two cabins, with the IFD weighting 10% more than the GFD, was expected to slightly influence the motion base performance. However, apart from a high frequency oscillation that is stronger in the GFD than in the IFD, the lateral motion performance of the motion platform was very similar. It is not clear what are the causes of the high frequency oscillation. One hypothesis would be that it is not actually an oscillation of the motion base, but of the measurement unit. The frequency of this oscillation was around 180 rad/s or 28 Hz, which may be too low for electrical noise, but it could be explained by some type of mechanical vibration at the attachment point between the measurement unit and the platform.

The influence of frequency and amplitude on the thresholds followed the same trends as the ones seen in the VMS. The higher frequency conditions resulted in lower values, although that effect was only statistically significant for the upper thresholds. The fact that

also in these two simulators the different conditions did not significantly affect the lower thresholds indicates that either the amplitudes and frequencies chosen were not different enough, or indeed that the measuring method lacks resolution at the lower amplitudes. Since this is a result that is constant across simulators, the probability that it is related to a specific simulator configuration, is small.

Contrary to the lower thresholds, the upper thresholds were different between the two simulators. The IFD presented higher upper thresholds than the GFD. When looking at the motion platform performance, the larger differences between simulators are observed at the lower amplitudes. Any differences in the results that would derive from the motion base would then be expected to be more accentuated on the lower thresholds. However, it is precisely in the upper thresholds that significant differences are seen. This indicates that the differences found between the threshold values were probably due to the visual system.

Chung et al. (2003) showed that collimation greatly influenced subjects' perception of lateral velocity and a larger field-of-view improved position control during a hover task. Both the IFD and the GFD have collimated displays, but the IFD has a panoramic display whereas the GFD has four WAC windows. For a subject sitting on the left-hand seat of the simulator only two of those windows provide visual stimulation. This greatly decreases the field-of-view which might lead to an underestimation of the visual cue amplitude in the GFD relative to the IFD. The effect of this underestimation is not observed on the lower thresholds, which, as mentioned above, might be related to a lack of resolution of the measuring method at low amplitudes.

The thresholds measured in the IFD resulted in a wider coherence zone, although the point of mean coherence was not significantly different between simulators. The larger panoramic display of the IFD favors the onset and strength of vection (Brandt et al., 1973; Held et al., 1975; Kawakita et al., 2000; Nakamura, 2001) which can lead to a stronger sense of self motion, which in turn allows larger inertial amplitudes before subjects detect some mismatch between the visual and the inertial cue. One can even say that the IFD has a more convincing display, that allows the inertial cue amplitude to divert further from the visual cue before it is noticed by subjects.

7.4.3 VMS, GDF and IFD

Comparing all three simulators for the conditions with a stimulus frequency of 5 rad/s, again the lower thresholds show little variation, whereas the upper thresholds were significantly affected by the amplitude and the simulator. Based on the thresholds calculated from the measured signals, the post-hoc tests showed that statistically, the upper thresholds in the VMS and the IFD were not different from each other but were significantly different from the values in the GFD.

It should be noted that since all subjects performed the experiments in each of the simulators in the same order, an effect of learning should not be disregarded. However, there is some evidence that in this type of task, the effect of learning is negligible. In this experiment, as well as in the other experiments described in this thesis, each condition is always repeated by the same subject two or three times. Although this is not explicitly reported, an effect of the repetition has never been found. Moreover, in Chapter 3, two Experiments are described. A small subset of subjects performed both experiments, with a few weeks of time in between. Also here, no effect of learning or habituation was found.

The differences found between the upper thresholds in the VMS and IFD as compared to the GFD are somewhat non intuitive. Similarities could be expected between simulators that have similar visual systems, such as the VMS and the GFD, or simulators with the same motion base, such as the GFD and IFD. In this respect, the GFD could be expected to be in between the VMS and the IFD. The GFD has a newer motion platform than the VMS, with improved performance, but it has a visual system that is not as sophisticated as the one in the IFD.

When comparing the upper thresholds based on the commanded signals, indeed the results fit the expectations, with the VMS showing the lowest thresholds and the IFD the highest. As noted before, calculating the thresholds from commanded signals might not be entirely correct but does help in understanding how the turnaround bump might influence subjects' judgment of motion. The lower upper thresholds in the VMS are in agreement with the hypothesis that subjects tuned the motion down to compensate for the turnaround bump.

The comparison of the three simulators might be summarized in a very generic way by saying that better simulators allow for larger differences between the inertial and visual cues. That is, for higher quality systems, subjects are more lenient in their judgment of matching visual and inertial cues. Moving from the VMS to the IFD, the quality of the inertial motion feedback and the visual systems could be said to improve. This improvement lead to higher upper thresholds, with the exception of the thresholds in the VMS that include the peaks caused by the turnaround bump. However, the same effect was not observed for the lower thresholds during the tested conditions and these are perhaps the most interesting for flight simulation applications.

From a simulation point of view it is much more interesting to look at the lower thresholds, since the goal is to provide enough inertial feedback using the smallest motion space. If motion filter gains are set at, for example 0.7 of the visual cue, then based on these tests, subjects' perception of combined visual and inertial cues will be equivalent in all three simulators. However, the same might not apply for other degrees of freedom or other combinations of amplitudes and frequencies.

On the other hand, the results for the upper thresholds might be interesting for repositioning motion of the simulator. In certain motion profiles, the amplitude of the inertial

motion might be exaggerated with respect to the visual cue such that it is used to cue the vehicle motion but also to allow faster movements when pre-positioning the simulator or bringing it back to an initial position. According to these results, the increase in inertial motion with respect to the visual cue will not be noticed by the subjects and more even for higher quality simulators.

7.5 Conclusions

The coherence zone for lateral acceleration motion was measured in three different simulators for two different visual stimulus amplitudes and two stimulus frequencies. The effect of amplitude and frequency on the upper thresholds of the coherence zones were similar to what has been previously found for yaw motion. The upper thresholds decreased with frequency and increased with amplitude. However, for the higher amplitude the point of mean coherence decreased with respect to the one-to-one line.

The lower thresholds were not significantly affected by either amplitude or frequency. This result was unexpected and since it was consistent across simulators it is thought to be a particular case of the chosen motion conditions. Nevertheless, the comparison of the three simulators using coherence zones was successful, in the sense that differences in simulator configuration resulted in differences in the measured upper thresholds.

To further investigate the potential of coherence zones as a perceptual metric to quantify differences in simulator and motion cueing performances, more degrees of freedom, amplitudes and frequencies should be tested across different simulators. One practical way of bypassing the complicated logistics of running multiple simulator studies is to use one simulator and artificially degrade both the motion base performance and the visual system quality (see Nieuwenhuizen, 2012).

8

Data Compilation and Summary of the Main Findings

8.1 Introduction

The effects of amplitude, frequency, subject's attention and simulator characteristics on amplitude coherence zones have been studied in different experiments using different experimental methods. Throughout the thesis, each chapter described one or two experiments that focused on investigating a particular aspect of coherence zones. In each chapter, the results were discussed and conclusions were drawn that applied to the studied scenarios.

As explained in the Introduction Chapter, each chapter can be read almost independently of the others. Although this allows for an easier access to the present thesis, the fragmented nature of such a presentation might hinder the understanding of the general principles regarding coherence zones. An attempt is made in this chapter to approach all the experimental data as a whole.

All the data from the previous experiments has been gathered and will be presented side-by-side. In doing so, it becomes easier to draw general conclusions and, although the important details of each experiment are temporarily overlooked, it also provides for a summary of the most important findings.

8.2 Yaw amplitude coherence zones

Chapters 3, 4 and 6 described a total of five experiments where amplitude perception coherence zones were measured in yaw. In total, five different motion profiles were used: one was an acceleration step-like motion signal and the other four were sinusoidal signals with frequencies of 0.2, 2, 5 and 10 rad/s. Four experiments were performed in the Simona simulator and one in the Desdemona simulator; three different measuring methods were used. Despite all the variations in the different experimental setups, combining all the data in one plot would allow the main trends to be observed.

The mean data across all subjects from each of the experiments was averaged for each of the conditions, determined by motion profile and amplitude, to obtain overall mean coherence zone data. Figure 8.1 shows the data of all five experiments as a function of the motion profile and the amplitudes of the visual motion stimuli. The perception coherence zones' limits, the lower and upper thresholds, are represented by the lower and upper limits of the colored bars. Each bar represents an interval, that is, a pair of lower and upper thresholds. Each color represents a different motion profile. For clarity purposes, whenever more than one pair of thresholds were measured at the same visual amplitude, the bars are plotted next to each other surrounding the corresponding visual amplitude. Horizontal black lines, placed at the inertial amplitude that physically matches the visual amplitude, connect those pairs of thresholds.

Looking at Figure 8.1 it becomes clear that a large part of the data was collected for stimuli with amplitudes of 12 and 30 deg/s^2. For these conditions, which were performed with different experimental setups, the spread in threshold values seems quite large. This is

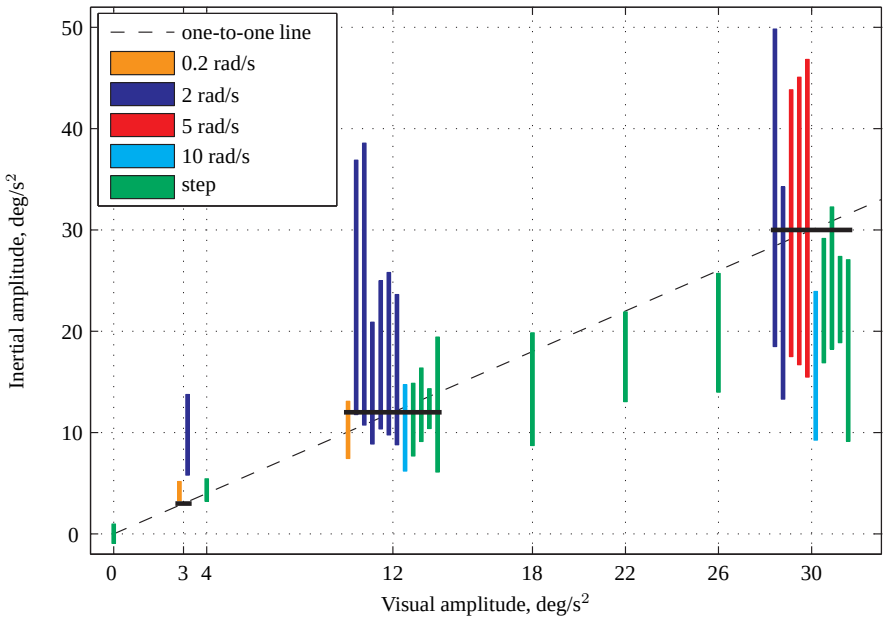


Figure 8.1: Coherence zones represented as the values between the measured lower and upper threshold values, across all amplitudes and frequencies, for five different experiments. The horizontal black lines span different measurements made at the same visual amplitude.

a direct result of using data from both the Simona and the Desdemona simulators, which, as discussed in Chapter 4, can have a very large influence on the exact values of the measured thresholds. Although the thresholds' values are dependent on the specific characteristics of the simulators' motion and visual systems, the trends observed for the effect of amplitude and frequency have remained constant across platforms. Namely, it has been shown that with increasing amplitude of the visual stimulus, coherence zones tend to bend down with respect to the one-to-one line and the coherence zone width increases. Also, an increase in frequency generally led to lower threshold values.

In Figure 8.1, the trends observed for each of the experiments separately can still be observed. The coherence zones at higher amplitudes seem to be lower, relative to the one-to-one line, than the coherence zones at lower amplitudes. Moreover, for a given visual stimulus amplitude, with the exception of the 0.2 rad/s stimuli, coherence zone limits tend to decrease as the stimulus frequency increases. It is remarkable that even across different experiments and simulators, it is possible to observe such trends.

8.3 Frequency

In Chapter 4 the effect of frequency on yaw coherence zones was related to the dynamics of the SCC. There, it was argued that although the SCC are stimulated by rotational acceleration, its dynamics resemble that of an integrator for frequencies between approximately 1 and 9 rad/s (Hosman and van der Vaart, 1978). For this range of frequencies, where normal head motion occurs, the output of the SCC is proportional to velocity and, outside this range, it is proportional to acceleration. In this case, if subjects use a simplified internal representation of the SCC that extends the integrator dynamics to all frequencies, high frequency and low frequency acceleration stimuli would be over-and underestimated, respectively, see Figure 8.2.

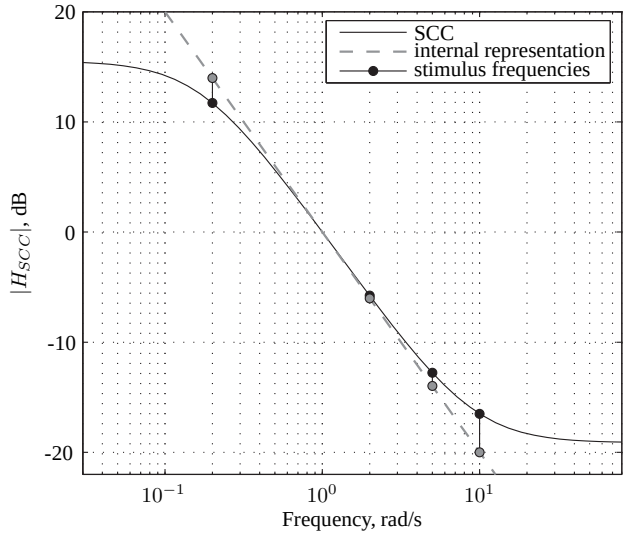


Figure 8.2: Difference in gain between the SCC dynamics and the hypothesized internal representation at the frequencies of the stimuli used.

For frequencies within the 1 to 9 rad/s range, a large effect of frequency would not be expected. In fact, comparing the results of coherence zone measurements made using 5 rad/s stimuli with the ones using 2 rad/s (visual amplitude of 30 deg/s^2) it is difficult to see a clear effect of frequency. However, when measuring coherence zones for very low frequency stimuli (0.2 rad/s) this hypothesis did not seem to hold. As can be seen in Figure 8.1, for the visual amplitudes of 3 and 12 deg/s^2, the thresholds measured for the 0.2 rad/s stimuli present lower values than that of the 2 rad/s stimuli.

Because the output of the SCC is proportional to velocity between 1 and 9 rad/s and since visually perceived self-motion is also based on velocity, in Chapter 4 it was argued

that it is reasonable to assume that the internal comparison of inertial and visual stimuli is done in "velocity space".

Figure 8.3 shows the same data as in Figure 8.1, but now both the visual amplitudes as well as the inertial amplitudes are represented in velocity units. As before, horizontal black lines, placed at the inertial amplitude that physically matches the visual amplitude, connect the different pairs of thresholds corresponding to the same visual amplitude.

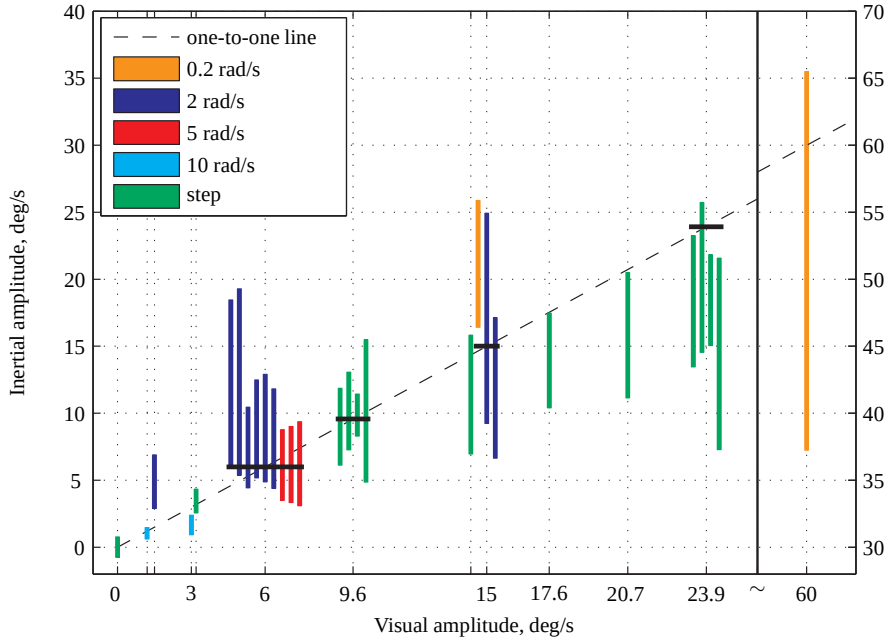


Figure 8.3: Coherence zones represented as the values between the measured lower and upper threshold values in velocity units, for five experiments. The horizontal black lines span different measurements made at the same visual amplitude.

When comparing the threshold results in velocity space, the same reasoning as above could be applied. Due to a simplified internal representation of the SCC, for very high frequencies ($>$ 9 rad/s) the inertial motion may be overestimated and for very low frequencies ($<$ 1 rad/s) it may be underestimated.

The step profile contains several frequencies, from below 2 rad/s to above 10 rad/s. The thresholds determined using this profile are likely to depend on the relative amplitude of each frequency composing the signal and on how the combined perception of different frequencies affects the overall perception of inertial motion strength.

Comparing the thresholds measured at the visual amplitude of 15 deg/s, the 0.2 rad/s stimulus presents higher thresholds values than the 2 rad/s. For the visual amplitude of 6

deg/s, the 2 rad/s stimuli result in higher threshold values than the ones from the 5 rad/s stimuli. Here, the trend of decreasing coherence zone thresholds with increasing stimulus frequency is more clearly shown than in Figure 8.1. This seems to confirm the hypothesis of the simplified internal representation (Chapter 4). However, based on the dynamics of the SCC, the differences in the threshold values between the 0.2 rad/s and 2 rad/s stimuli were expected to be larger than the differences between the 2 rad/s and the 5 rad/s. The latter pair of frequencies are within the region in which the SCC approximates an integrator and thus also, the simplified internal representation.

Moreover, if the differences between thresholds measured at different frequencies could be fully explained by the simplified internal representation hypothesis, then the expected trends should be observable not only in velocity but also in acceleration units. Figure 8.4 zooms in on Figure 8.1 and shows the thresholds measured for a visual amplitude of 12 deg/s^2. For simplification, the results from the step profile have been omitted. The first three bars on the left-hand side in the figure represent data collected in the Desdemona simulator and the other, data collected in the Simona simulator. One can see that the threshold values at 0.2 rad/s are still much lower than the ones at 2 rad/s. This cannot be explained using only the internal representation hypothesis.

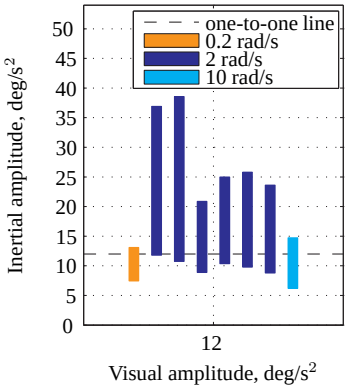
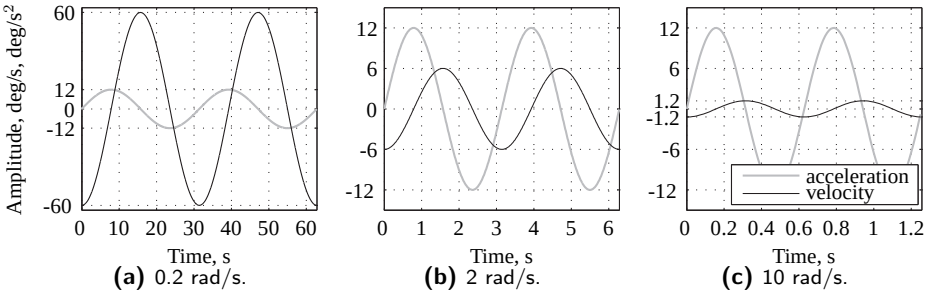


Figure 8.4: Coherence zones measured in acceleration units at a visual stimulus amplitude of 12 deg/s.

8.4 Acceleration and velocity

A different explanation for the effect of frequency was also offered in Chapter 4. There it was concluded that subjects seem to base their judgment on a *combination* of both velocity and acceleration amplitudes. This is similar to what has been discussed in the work of Grant and Haycock (2006) and Soyka et al. (2009), where it has been shown that perception of linear inertial motion strength depends not only on the acceleration amplitude but also on the jerk amplitude. The otolith organ's models presented by Hosman and Van der Vaart (1978) show that the response of these organs is proportional to both acceleration and jerk. Transposing this to the angular motion case, where the output of the SCC is proportional to velocity and acceleration, one would say that perceived angular motion strength depends on both velocity amplitude and acceleration amplitude.

That being the case, the effect of frequency could in fact be an effect of velocity and acceleration amplitudes. That is, for the same acceleration amplitude, a high frequency acceleration stimulus results in a smaller velocity amplitude than a lower frequency one. For stimuli with frequencies as different as 0.2 and 10 rad/s the differences in velocity amplitudes are very large. As an illustration, Figure 8.5 shows three sinusoidal signals with the same acceleration amplitude and different velocity amplitudes.



(a) 0.2 rad/s. **(b)** 2 rad/s. **(c)** 10 rad/s.

Figure 8.5: Example of three motion profiles with a maximum acceleration amplitude of 30 deg/s^2 and different frequencies, resulting in different maximum velocity amplitudes.

For the same acceleration amplitude of 12 deg/s^2, the lowest frequency signal has a velocity amplitude of 60 deg/s. As mentioned above, for large visual motion amplitudes, coherence zones tend to bend below the one-to-one line. This means that at high amplitude movements subjects have an increasing tendency to down tune the inertial motion. So, comparing the thresholds measured at 0.2 rad/s with the ones measured at 2 rad/s, with a visual amplitude of 12 deg/s^2, two factors are playing a role: frequency and amplitude. As a general trend, lower frequencies lead to higher acceleration thresholds. On the other hand, high motion amplitude leads to a down tuning of the inertial motion, leading to lower thresholds. The 0.2 rad/s profile is a much lower frequency than the 2 rad/s profile but has a very large velocity amplitude.

The relationship between velocity and acceleration amplitudes and frequency can be better observed by showing the coherence zone data in both velocity and acceleration space, while retaining the information about the stimulus frequency. Figure 8.6 shows the coherence zones measured in terms of visual velocity and inertial acceleration, for each motion profile.

In this figure there is a different one-to-one line for each motion profile, since, depending on the stimulus frequency, the same acceleration amplitude corresponds to different velocity amplitudes. For the same stimulus frequency, higher visual velocity amplitudes result in coherence zones that are bent down with respect to the one-to-one line. Concurrently, for the

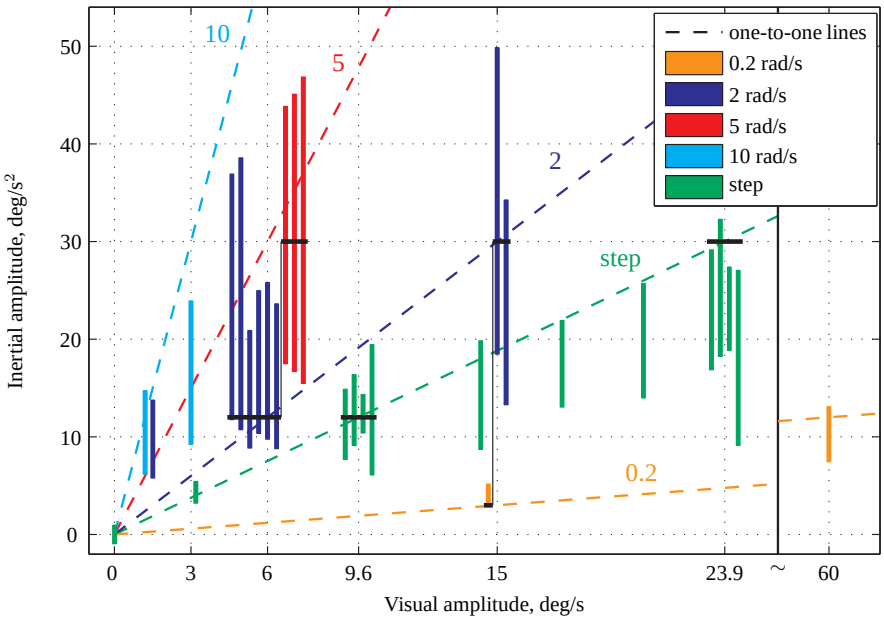


Figure 8.6: Coherence zones represented as the values between the measured lower and upper threshold values in acceleration units, across all visual velocity amplitudes, for five different experiments. The horizontal black lines span different measurements made at the same visual amplitude.

same visual amplitude, higher frequency stimuli also lead to lower thresholds with respect to the physical match.

From this it seems possible to conclude that coherence zones are affected by both the velocity and the acceleration amplitude of the motion stimulus. Subjects seem to rely on a weighted combination of both velocity and acceleration estimates to match the perceived inertial motion to the perceived visual motion. The effect of frequency that is observed when representing the coherence zones as visual versus inertial amplitudes might just be a consequence of this weighing between velocity and acceleration.

In general, for high frequency and high amplitude motion, which might be considered outside the normal range of head and body movement, subjects tend to down tune the inertial motion with respect to the physical match.

8.5 Yaw motion gains

The effect of frequency is especially important when applying knowledge of perception coherence zones to the tuning of motion filters. Depending on the available motion space, high-pass filters might introduce more or less attenuation at different frequencies, depending on the filter's gain and cutoff frequency. In order to be able to use coherence zone measurements in the optimization of such filters, thresholds can also be represented in terms of gains at different profile frequencies.

Figure 8.7 shows the coherence zones in terms of motion gains, that is, the upper and lower thresholds were divided by the corresponding visual acceleration amplitude for each profile frequency. The data was divided in three different visual acceleration amplitudes: 3, 12 and 30 deg/s^2. The step-like profile was a periodic signal with a fundamental frequency of 1.8 rad/s. However, it contained other higher frequency components and so, it can not correctly be plotted at only one frequency. For this reason, the step-like profile is plotted separately on the right hand side of the figure. For clarity, the step profile data corresponding to amplitudes other than 12 or 30 deg/s^2 are not shown.

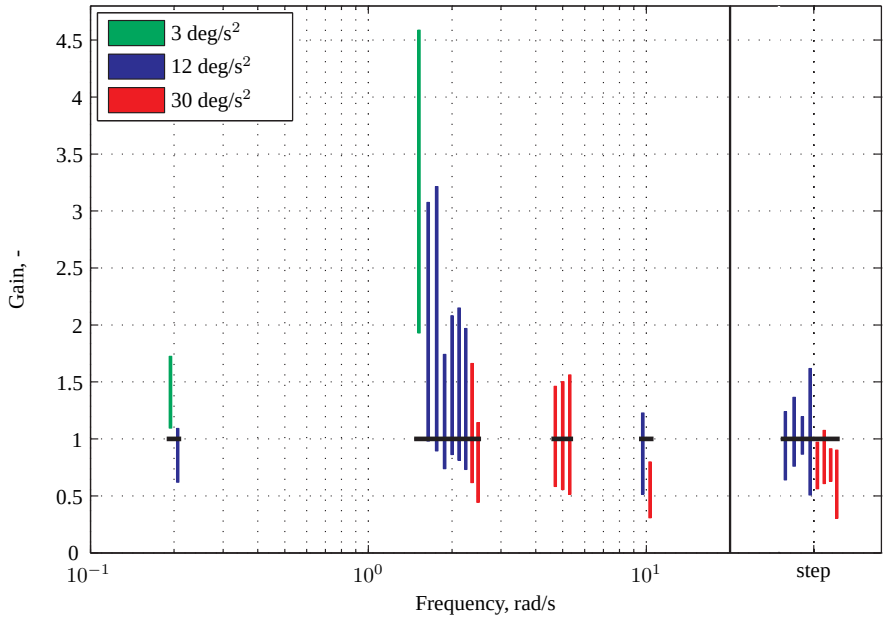


Figure 8.7: Coherence zones represented as the maximum and minimum motion gains obtained from the threshold values, across three amplitudes for all yaw motion experiments. The horizontal black lines span different measurements made at the same frequency.

The upper threshold gains are clearly higher for the middle frequencies of 2 and 5 rad/s. The differences in the lower threshold gains are not as pronounced but they also present higher values for the middle frequencies. The values obtained using the step profile seem to be somewhere between what one could expect for stimulus frequencies between 2 and 10 rad/s. The effect of the visual cue amplitude can also be seen, with higher amplitudes leading to lower gains. This corresponds to the previously mentioned "bending" of the coherence zone with respect to the one-to-one line at higher amplitudes.

The large upper threshold value for the 3 deg/s^2 stimulus corresponds to measurements performed in the Desdemona simulator. As discussed in Chapter 4, for the same experimental conditions, the thresholds values obtained in the Desdemona simulator were generally larger than the ones obtained in the Simona simulator, probably due to the specific characteristics of Desdemona's visual system, such as the proximity of the display and the absence of collimation.

8.6 Sway motion gains

For the experiments in sway described in Chapter 7, the same data treatment can be applied to obtain threshold gains. The measured lower and upper thresholds for sway amplitude coherence zones were divided by the corresponding visual amplitude. Figure 8.8 shows these threshold gains for three stimulus frequencies and two amplitudes, collected in three different simulators.

The trends in the data are similar to what was observed for yaw coherence zones. Both the upper and lower threshold gains decrease slightly with increasing frequency and the gains are lower for the highest amplitude of the visual cue. There is one exception at the frequency of 3 rad/s, where the upper threshold gains are larger than the ones at 2 rad/s. The thresholds at 3 rad/s correspond to measurements made in the Visual Motion Simulator (VMS), whereas the ones at 2 rad/s were performed in the Generic Flight Deck (GFD) and the Integration Flight Deck (IFD). As discussed in Chapter 7, the VMS presents considerable overshoot in the motion response, which results in higher threshold values.

8.7 Coherence zones and optimal gain

What is perhaps more striking in Figure 8.7 and Figure 8.8 is that the upper threshold gains can be very high. For sway, upper gains are above 2 and for yaw, at the frequency of 2 rad/s and a visual cue amplitude of 12 deg/s^2 upper threshold gains can be above 3. This is a remarkable result, since previous research has shown that subjects in the simulator prefer motion gains below one (Groen et al., 2007; Correia Grácio et al., 2010).

Perhaps a crucial difference between these studies and the present one is what exactly is asked of subjects during the experiment. In this study, subjects are requested to find the

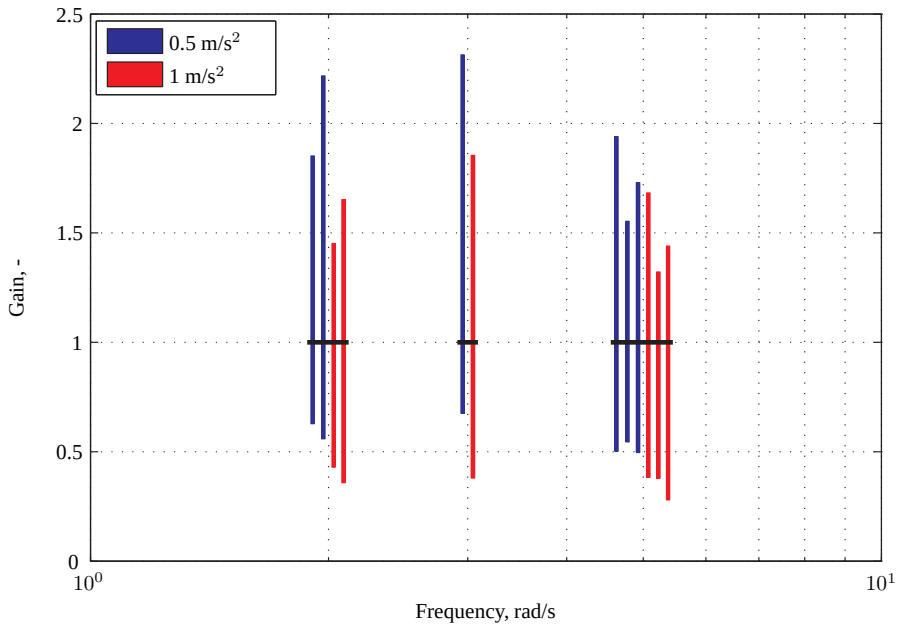


Figure 8.8: Coherence zones represented as the maximum and minimum motion gains obtained from the threshold values, across all amplitudes and all sway motion experiments.

boundaries of their coherence zones. They are invited to determine what is the strongest and weakest inertial motion amplitude that still matches the visual cue amplitude. In the cited studies, subjects are required to either rate the motion or indicate what is the inertial amplitude that is the best possible match to the perceived amplitude of the visual stimulus. It is quite intuitive to think that this optimal inertial amplitude is contained within the coherence zone. In fact, a study on both perception coherence zones and optimal gain showed that the optimal region was contained within the lower part of the coherence zone (Correia Grácio et al., 2013). What exactly motivates subjects to choose a certain range of values within the coherence zones as being "optimal" is still not known.

Nevertheless, when intending to optimize the motion filtering process, it is perhaps more useful to look at the lower part of the coherence zones. This, not only because of the optimal region location, but also because lower motion gains will lead to a more efficient and still effective use of the motion space of the simulator. In this sense, the lower thresholds of the coherence zone provide a good indication for the motion gains to be used. However, measured upper and lower thresholds showed a decreasing trend with increasing frequency. This preference for weaker inertial motion amplitude at the higher frequencies is at odds

with the use of high-pass filters, which show a higher gain at higher frequencies. The relationship between coherence zone measurements, especially with respect to the lower thresholds, and motion filtering optimization is further discussed in Chapter 9.

Part III

From Perception to Cueing

9

Perception Coherence Zones as a Motion Fidelity Assessment Method

9.1 Introduction

The lower thresholds of the coherence zones represent the minimum inertial motion amplitudes that are perceived by subjects in the simulator as still being coherent with the visual stimuli. If these values are used as the minimum allowed motion gain at a specific frequency, then it is likely that the inertial and visual cues at that frequency will be perceived as having matching amplitudes. The same reasoning can be applied to the phase coherence zone measurements described in Chapter 5. The phase-thresholds measured can represent the maximum phase value at which the inertial cue may lead the visual cue and still be perceived as having a matching phase.

It is assumed that since the visual stimulus represents the vehicle motion, for a realistic model of the vehicle dynamics, any perceived difference between the visual and the inertial cue will correspond to a perceived difference between the vehicle motion and the simulator motion. If the simulator motion and the vehicle motion are perceived as different, then it can be said that the simulation is impaired.

Taking the perception of coherent or incoherent visual and inertial cues as a measure of simulation fidelity, the coherence zone measurements can be used as criteria for motion filter design and tuning. Although ideally much more data would be needed to draw up motion fidelity criteria that extend to different simulator configurations, types of task and degrees-of-freedom, the currently available data can still be used to demonstrate the potential of coherence zones measurements as guidelines for the development and tuning of motion filters.

In Chapter 8 the measured coherence zone thresholds across different experiments were averaged for different frequencies and amplitudes of the visual stimulus. By dividing these threshold values by the corresponding amplitude of the visual stimulus, threshold gains were obtained. These gains represent the coherence zones' limits expressed as a ratio between the visual amplitude and the inertial motion amplitude.

In this chapter the calculated average threshold gains, and also the phase-error thresholds measured in Chapter 4, will be used as coherence-zone based motion fidelity criteria.

9.2 Sinacori plot and Schroeder's revised criteria

One widely used criterion in flight simulation is the so-called Sinacori plot with fidelity regions (Sinacori, 1977). The fidelity regions limits correspond to criteria for gain and phase distortion introduced by motion filters at the frequency of 1 rad/s. Schroeder (1999) performed a careful review of motion filter assessment studies and, based on experimental data, suggested revised criteria for the rotational and translational axes. The revised criteria by Schroeder are shown in Figure 9.1.

Two lines delimit three regions of fidelity. Along the line separating the "High" from the "Medium" fidelity regions, four points were marked, represented by letters A to D.

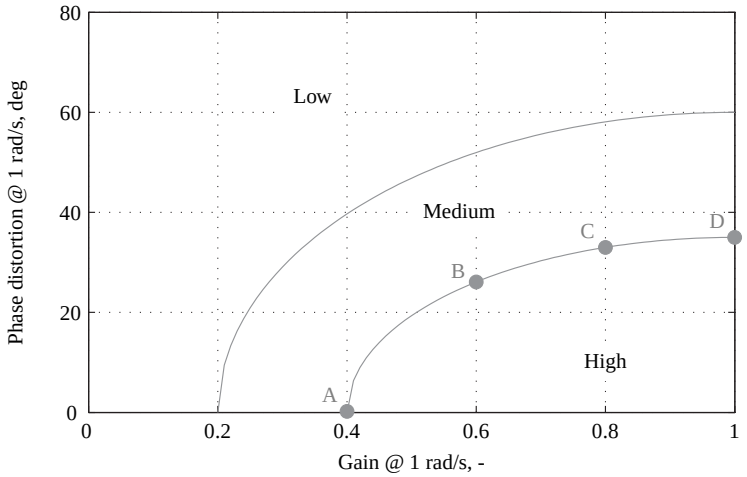


Figure 9.1: Sinacori plot with motion fidelity criteria for the rotational axes revised by Schroeder (1999).

Each of these points represent the amplitude attenuation and phase distortion introduced by a given motion filter at 1 rad/s. There are several motion filter settings that may result in the same amplitude attenuation and phase distortion values at 1 rad/s. However, for most flight simulators with a Stewart platform, the most widely used method to cue vehicle accelerations in the limited space of a simulator is to use a high-pass filter. Moreover, to make sure that after each initial motion of the simulator, the cabin always returns to a neutral position, motion filters need to be of second order or higher. If the motion filter structure is defined as a second order high-pass filter with a fixed damping value of 0.7, each pair of gain and phase-distortion values in the Sinacori plot will correspond to one unique set of motion filter parameters (scale factor and break frequency). The modulus across frequencies for each of these filters are shown in Figure 9.2. Here, the lower threshold gains for four frequencies and two visual signal amplitudes are shown.

The lower threshold gains represent the lower limit of the inertial motion gain for which the combined inertial and visual cues are perceived as coherent. Similarly, for fixed values of phase-distortion at 1 rad/s at each of the points A to D, the motion filter lines represent the lower limit of the inertial motion gain for which, according to Schroeder's criteria, the motion cueing is considered to be of high fidelity. It should be noted though, that the criteria derived by Schroeder for the rotational axes are mainly aimed at roll and pitch motion, whereas the coherence zone measurements were made in yaw. Nevertheless, there are a few observations that are interesting to make.

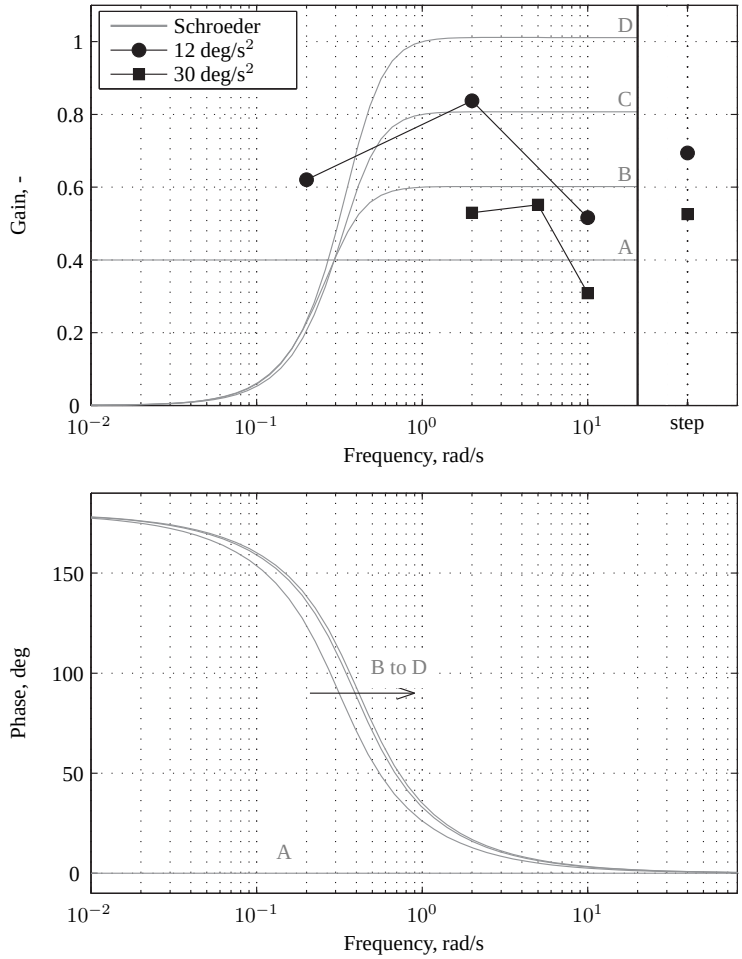


Figure 9.2: Lower threshold gains for five motion profiles and two visual stimulus amplitudes and motion filters derived from Schroeder's revised criteria.

The coherence zone measurements were made with a zero phase-distortion, that is, only the gain was varied, whereas Schroeder's criteria were based on experimental data collected using motion filters that introduce both amplitude attenuation and phase distortion. At a zero phase-distortion level (filter A), Schroeder's criteria are, in general, less strict than the coherence zone criteria, with the former setting a minimum motion gain of 0.4 whereas the latter demands gains as high as 0.8.

Since the depiction of Schroeder's criteria is made based on high-pass filters, it is to be expected that the gain increases from low to high frequencies and remains constant for

frequencies higher than the filter break-frequency. One exception is filter A that corresponds to a single scale factors (the break frequency is zero). In contrast, the limits based on the coherence zones seem to indicate that high gains are needed at the middle frequencies (around 2 rad/s) and they can be lower for high and low frequencies. Furthermore, the criteria based on coherence zones also allow for different values at different amplitudes of the visual cue. This is only possible since measurements were made using single sinusoid signals with a specific amplitude.

The coherence zone gains were derived from human self-motion perception metrics and, apart from choosing the lower thresholds as the lower limits for motion filter scale factors, no other considerations have been made regarding the current simulation technology. In contrast, the initial Sinacori plot and Schroeder's revised criteria are centered on typical motion filtering techniques and aim at achieving the best possible result within the limitations of the simulator. The measurement of coherence zones may deliver criteria that are impossible to meet with a specific simulator or during a specific type of task, rendering the criteria useless for direct application. On the other hand, Schroeder's criteria or any type of criteria that assumes a specific motion filtering technique might be limiting the development of novel cueing methods.

9.3 Advani-Hosman criteria

The proposed use of coherence zones as fidelity criteria while still considering simulator motion space issues can be accomplished by expanding the existing Sinacori plots with criteria defined for both frequency and amplitude. One step towards this concept has been made by Advani and Hosman.

Advani and Hosman (2006) have proposed a joint representation of the motion base and the motion filter performance. To this purpose, they consider the total transfer function from the simulated vehicle motion to the inertial stimuli provided to the subject, including the dynamics of the motion base hardware, the characteristics of the motion base controller and the effects of the motion filter. The transfer function can then be plotted in a modulus versus phase plot for all frequencies of interest. In such a plot, a region of gain and phase-distortion values can be defined which will lead to acceptable motion fidelity within a certain range of frequencies.

Figure 9.3 shows the limits for acceptable motion for the rotational axes that were proposed by Advani and Hosman on a temporary basis, i.e., until more experimental data is collected. An example of a transfer function of the combined hexapod motion base and motion filter is also plotted. The motion base characteristics used were those of the Simona simulator for yaw motion, with a second order high-pass motion filter with a scale factor of 0.55 and a cutoff frequency of 0.35 rad/s. The phase values presented are absolute values,

so they may indicate either inertial motion phase lead (up to 4 rad/s) or phase lag (from 4 to 10 rad/s).

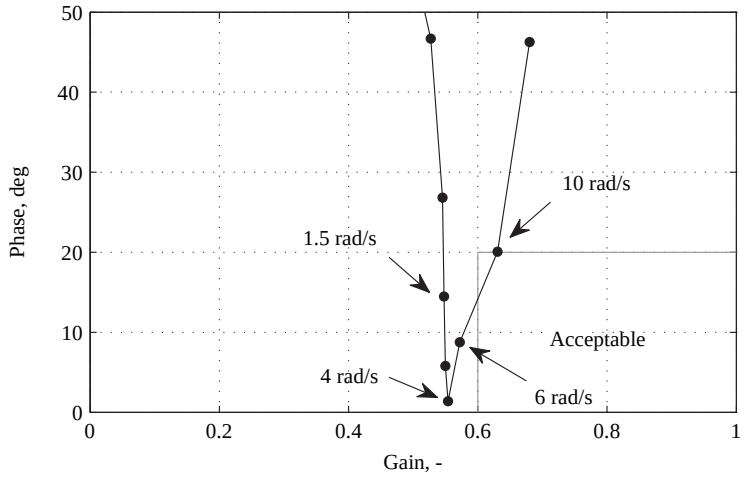


Figure 9.3: Advani-Hosman proposed criteria for the rotational axes and an example of a combined transfer function.

As can be seen, the simulator motion can be described at different frequencies and the fidelity of the motion provided can be assessed independently of the type of motion filter used. From the figure, one might conclude that the inertial motion, at almost all frequencies, is not acceptable. However, below, it will be shown that when using criteria that distinguish between stimulus frequencies and amplitudes, the provided inertial motion, at high amplitudes, is in fact acceptable for most frequencies.

9.4 Coherence zone criteria

It is now proposed that coherence zone measurements be used to calibrate the acceptable motion region, or better said, regions. Not only one region can be defined, but for the same transfer function, different regions can be created that correspond to different frequency ranges and amplitudes.

As an example, the lower threshold gains shown in Figure 9.2 are used to determine the left boundary of the acceptable motion region. For the upper boundary, corresponding to the phase-distortion limit, the thresholds measured in Chapter 5 are used. The phase-threshold values obtained were not measured at the same frequencies and amplitudes as the amplitude coherence zones, so some approximations will be made. The phase-thresholds were measured in yaw for amplitudes of 5.73 and 9.17 deg/s and frequencies of 1.257, 6.283

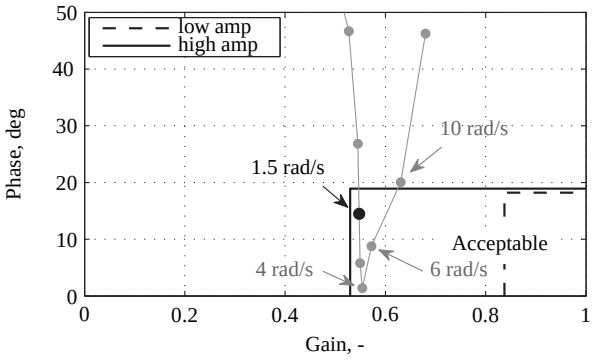
and 10.051 rad/s (see Chapter 5 for the values in acceleration units), whereas the lower thresholds gains from Figure 9.2 correspond to amplitudes of 12 and 30 deg/s^2 and frequencies of 0.2, 2, 5 and 10 rad/s. For reasons of simplification, the amplitudes of the visual stimulus will be considered either high, 9.17 deg/s in the phase-threshold experiment and 30 deg/s^2 in the amplitude coherence zone experiments, or low, 5.73 deg/s and 12 deg/s^2. The phase-threshold values at the frequencies of 1.257, 6.283 and 10.051 will be taken as corresponding to the frequencies of 2, 5 and 10 rad/s, respectively. It should be noted that the phase-thresholds measured showed no significant effect of either the amplitude nor of the frequency of the visual stimulus. Hence, using an averaged phase-error threshold for all amplitudes and frequencies would also be an acceptable approximation.

Figure 9.4 shows the different fidelity regions derived from the lower thresholds of the amplitude coherence zones measurements and the phase coherence zone measurements. The same transfer function from Figure 9.3 is also shown.

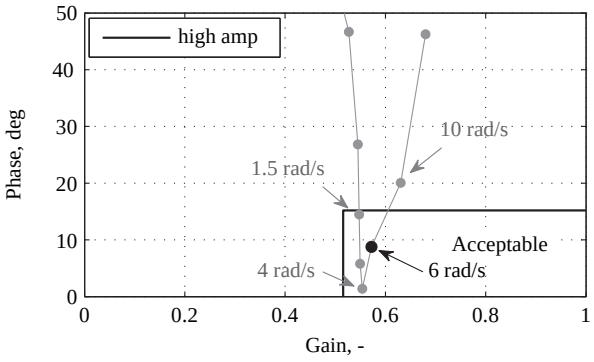
As already mentioned before, the phase coherence zones and the amplitude coherence zones were obtained independently, so there is no certainty that the combined coherence region is rectangular. A rectangular region assumes that for varying motion gain values, the phase-threshold is constant and vice-versa. The rounded regions proposed by Schroeder were based on experimental work that used motion filters for the generation of inertial motion cues. Typically, when the cutoff frequency of a high-pass filter is decreased, the gain has to be decreased as well to maintain the same usage of the motion space. Generally, a decrease in the cutoff frequency leads to a reduction of the phase-distortion at 1 rad/s. So, for a typical hexapod simulator the trade-off to be made is between high gain and high phase distortion or low phase distortion and low gain. The rounded regions thus exclude low gain, high phase-distortion regions.

Grant and Lee (2007) measured phase-error thresholds for pitch motion using two different gains applied to the inertial motion: 1 and 0.5. They found a significant effect of the motion gain on the phase-error thresholds. In fact, for a motion gain of 0.5, the phase-error thresholds were around 10 deg higher than for a motion gain of 1, suggesting that as the gain decreases, the allowed phase distortion increases. This would make the criteria regions neither rectangular nor rounded, but with a sloped upper boundary.(For an example of such a boundary, please see the low frequency level boundaries in Figure A.1) Presently, and until more data is collected using combined phase and amplitude mismatch, the rectangular regions are considered to be the simplest approximation.

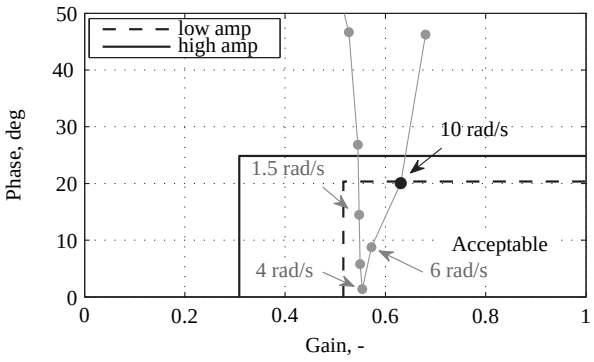
The different regions correspond to acceptable, or coherent, motion at different frequency ranges and amplitude levels. Generally speaking, for each simulation, and each type of vehicle being simulated, it is possible to identify a range of frequencies and amplitude levels that are crucial for the execution of the specific task at hand. Using different fidelity regions allows the tuning of the motion filter to have optimal performance at the frequencies and amplitudes of interest while providing insight into the effect of changes in the filter at other less crucial, but perhaps also important, regions.



(a) 1-2 rad/s



(b) 5-6 rad/s



(c) 10 rad/s

Figure 9.4: Representation of coherent motion regions for yaw based on phase-error thresholds and lower threshold gains for two amplitude levels: high (high amp) and low (low amp), and three frequency ranges, based on the method proposed by Advani and Hosman.

The usefulness of having different criteria for different frequencies becomes evident when comparing Figure 9.3 with Figure 9.4. In Figure 9.3 the designed motion filter and motion base performance seem to result in acceptable inertial motion only for frequencies around 10 rad/s. However, in Figure 9.4 it becomes possible to compare each frequency point with its corresponding criterion. In Figure 9.4 one can see that for high amplitude motion with frequencies around 1 to 2 rad/s, the designed filter still provides coherent motion. For the other frequency ranges, around 5 to 6 rad/s (Figure 9.4) and 10 rad/s (Figure 9.4), the same applies.

Figure 9.4 illustrates different regions for yaw, but the same can obviously be done for other degrees-of-freedom, and both rotational and translational channels can be tuned in this manner. As an example, Figure 9.5 shows criteria derived from the lower thresholds measured in the experiments described in Chapter 7, for sway motion. Here, no phase-error threshold data were collected in sway. For this reason, only gain criteria is presented.

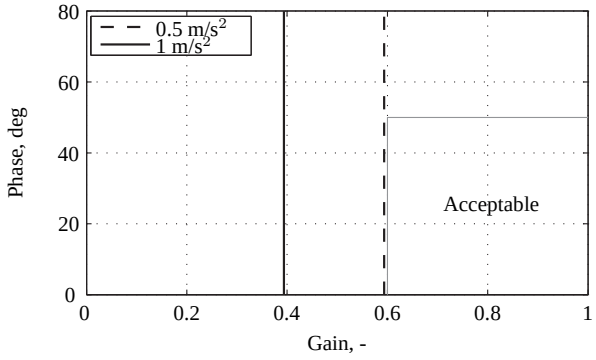
It is important to note that the maximum phase distortion indicated in the plots refers to the phase difference introduced by the combination of the motion filter and the dynamics of the motion platform. and the computational time delays. Typically, this maximum phase distortion includes phase lead introduced by the motion filters and high frequency phase lag introduced by the motion platform. Indirectly, it may or not include time delays due to simulator computation and communication times. These time delays are usually different for the visual and the motion systems.

Figure 9.6 shows a simplified schematic of the timing of different events during a simulation. The first event, designated by "vehicle" in the figure, represents the motion of the actual vehicle as it is calculated in the vehicle model. The events "motion" and "visual" represent the start of the platform motion and the motion represented through the visual displays, respectively. For simplification, the timing differences in the figure represent both time delays and phase differences. In this representation the "motion" is assumed to lead the "visual" but that is not necessarily always the case.

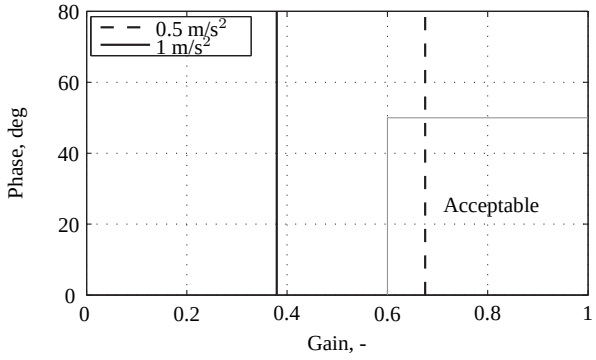
In the case of coherence zones, the phase difference was measured between inertial and visual stimuli and thus includes any timing differences between the visual and the inertial stimuli. This is represented by Δ_d in the figure. It does not, however, include any constraint on the overall time delay between computed vehicle motion and simulator motion.

In the Advani-Hosman criteria, the phase distortion criterion refers to the phase difference between vehicle motion and simulator motion, including the phase difference introduced by the motion filter and platform dynamics, and computational time delays. This is represented by Δ_m. Here, the time delay between vehicle and simulator motion is accounted for, but the time delay between visual and inertial motion is not.

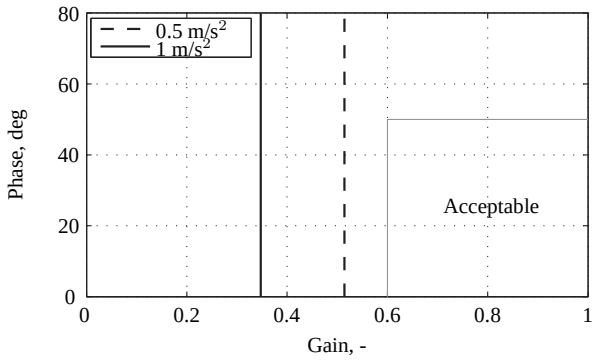
The Advani-Hosman criteria do not include any constraints on the differences between visual and inertial motion in the simulator. Nevertheless, that constraint is indirectly guaranteed by setting a maximum allowed time-delay for the visual system (Joint Aviation Author-



(a) 2 rad/s



(b) 3 rad/s



(c) 5 rad/s

Figure 9.5: Representation of coherent/acceptable motion regions for sway based on lower threshold gains for two amplitude levels and three frequencies. The acceptable region proposed by Advani and Hosman for the translational axes is also shown.

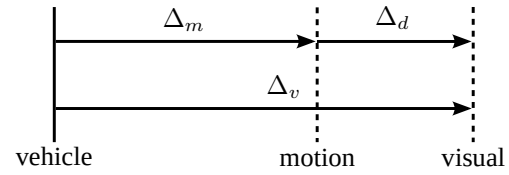


Figure 9.6: Schematic of time delay and phase distortion constraints.

ities, 2003; Royal Aeronautical Society, 2005; International Civil Aviation Organization, 2009), indicated by Δ_v in the figure.

For the coherence zone based criteria, the phase distortion and included time delays between inertial and visual motion are incorporated in the criteria, but there is no constraint regarding the delay between the vehicle motion and the visual and inertial stimuli in the simulator. Again, setting a maximum allowed visual system delay guarantees that both the visual and the inertial motion have an acceptable delay with respect to the vehicle motion.

For both criteria it holds that the overall timing of the simulation is only guaranteed when constraints are applied to two of the variables indicated in Figure 9.6. The third one can always be derived from the first two.

9.5 Upper thresholds

Up to this point only lower threshold values were used to derive criteria for motion cueing. However, measuring upper thresholds, apart from contributing to our general understanding of human motion perception mechanisms, can be useful for the design of new motion filtering techniques as well.

Generally, in a flight simulator it is desirable to provide the necessary motion feedback with the smallest possible motion space. The less motion space is used in one DOF, the more it can be used in other DOFs, for example. So, if two different motion gains provide the same perception of coherence between visual and inertial stimuli, the lower gain is the most advantageous. In such a scenario it may seem that upper thresholds have little relevance for motion cueing. However, when considering the repositioning motion of the simulator, for example, it makes sense to think about larger gains, so that the simulator returns as quickly as possible to its neutral position in order to have space available for the following maneuver.

One other application where it may be useful to know the highest gain that can be used for cueing a certain movement is during tilt coordination. During a tilt coordination maneuver, the simulator is placed at a non-zero angle with gravity, so that the specific force due to gravity is used to simulate a sustained linear acceleration. If the onset of linear acceleration is fast, then the rotational movement needed to place the cabin at the desired angle should

also be fast. However, to prevent subjects from perceiving that they are being rotated, the rotational motion has to be made slowly, so it remains below the perception threshold. If during a certain simulation scenario, besides an onset in specific force, there is also rotational vehicle motion in the same direction as the rotation needed for the tilt coordination maneuver, then the maximum rotational speed is no longer determined by the perception threshold, but by the upper threshold of the coherence zone. Since generally, the upper thresholds of coherence zones are above the one-to-one point, the rotational movement can then be performed at a higher speed.

This technique can be applied to improve the cueing of one of the more difficult problems in road vehicle simulation: curve driving. During a small radius turn, common in urban environments, the lateral specific force due to the circular trajectory of the car cannot be fully cued using only an onset in sway. Generally, it is also necessary to roll the simulator cabin to provide subjects with a sustained lateral cue during the curve. Entering and leaving the turns happens relatively fast, so it is difficult to provide a roll motion that remains below the perception threshold and still reaches a roll angle large enough to provide a realistic lateral specific force.

In this scenario, the slight actual roll motion that the car makes when entering a curve can be used to "mask" the addition of high angular velocity roll rotation in the simulator. The vehicle roll motion when entering the turn can be small, but it is visually perceptible. If roll motion is provided above the one-to-one point but below the upper threshold, the motion will be coherent with the visual roll cue, while simultaneously providing a larger roll angle in a shorter amount of time. This technique was used in the motion filter described in Chapter 2 for curve driving in the Desdemona simulator.

9.6 Coherence zone assessment method

To achieve simulator motion fidelity criteria that allow for the development of new simulation methods and at the same time serve the current simulator users, both human-centered and simulator-centered criteria should be combined.

From the coherence zone research it can be concluded that both the frequency and the amplitude of the visual stimulus, and hence of the vehicle motion, have an effect on the perception of coherence between inertial and visual cues.

From a simulator motion space point of view, there is always a trade-off between high motion gain and high phase distortion. For lower motion gains, also the phase distortion values should be lower to attain the same level of fidelity. This is visible in the rounded off regions proposed by Schroeder as opposed to the rectangular regions initially defined by Sinacori.

This trade-off was not studied during the coherence zone measurements, as only either gain or phase were measured and never both at the same time. Before coherence zones may

in fact be used as a fidelity criteria, combined measurements should be done. In Appendix A a method is proposed on how to measure coherence zones with both amplitude and phase mismatches and how to translate those measurements into motion fidelity criteria.

10

General Conclusions and Recommendations

10.1 Conclusions

The goal of the present study was to contribute to the improvement of inertial motion fidelity in simulation. From this goal, two more specific objectives were defined. The first was to understand the impact of different cueing solutions on subjects' perception, behavior and performance. The second was to extend our knowledge of combined perception of inertial and visual cues and to investigate the possibility of using this knowledge to derive guidelines for vehicle simulation.

The first part of this thesis presented the design and evaluation of a new motion cueing algorithm for curve driving in a large, centrifuge-based motion simulator.

The second part consisted of several experiments aimed at investigating both amplitude and phase coherence zones. Amplitude coherence zones were measured for yaw and sway and phase coherence zones were measured for yaw and pitch. The effects of stimulus amplitude and frequency were measured for both amplitude and phase coherence zones. Furthermore, for yaw amplitude coherence zones, the visual scene content and subjects' attention to the perception task were investigated. Three different simulators were compared using sway amplitude coherence zones.

In the third part of the thesis, the coherence zone data was used to derive motion fidelity criteria for vehicle simulation.

From the first part of this thesis it became clear that the impact of the quality of inertial motion stimuli on subjects perception, control behavior and performance is observable but difficult to quantify. There are no readily available methods, most criteria are based on a specific motion filter structure and the available metrics are both platform and task-dependent.

In the experiment described, as in most motion cueing design procedures, the design and choice of initial settings for the motion cueing algorithms were based on knowledge of human motion perception, the physical limitations of the simulator and the type of scenario and task being simulated. The optimization of those initial settings was largely based on human-based subjective judgments. The effect of each setting on motion perception, control behavior and performance could only be assessed through extensive experimentation with human subjects. This process implies a great time investment.

To expedite and improve process of design and tuning of motion filters, with respect to motion perception, more platform independent, human-based perception metrics and guidelines are needed. Coherence zones offer a systematic, human-based approach to the determination of motion boundaries that can guide the optimization process.

In the second part of this thesis coherence zones were measured as one way of studying the combined perception of visual and inertial cues. The collected data was summarized and discussed in Chapter 8. The advantage of using coherence zones in the context of vehicle

simulation is that it quantifies the level of perceived coherence between visual and inertial cues without necessarily having to determine the properties of the individual processes, that is, the perception of visual motion, and the perception of inertial motion.

In this respect, coherence zones represent a bridge between the large body of more fundamental research into human self-motion perception, and applied studies that are oriented towards motion cueing. The effects of stimulus amplitudes and frequencies on coherence zones provide a basic understanding of combined cue perception. Simultaneously, the coherence zone measurements can be directly converted into guidelines for vehicle simulation.

In the third part, coherence zones were used to define motion cueing criteria and to propose a motion cueing assessment method. The coherence zone assessment method and criteria provide three important additions to the already available criteria.

First, it presents not only criteria for desirable motion stimuli, but also offers a systematic, objective, human-perception-based method to measure the limits of the criteria for each desired degree-of-freedom and simulator configuration.

Second, the coherence zones' method and criteria introduce a third and fourth dimensions to the Sinacori plot: frequency and amplitude. Distinguishing between different frequencies results in criteria that make no assumptions regarding the motion filter structure. Moreover, with the possibility of constructing criteria for different amplitudes and frequencies, motion filters may be tuned optimally to motion profiles typical of specific tasks.

Third, by offering a measurement method and allowing different frequencies and amplitudes to be chosen, the coherence zones method presents simulator-based, task-specific criteria. This means that the ranges of desirable or acceptable motion may be defined independently for each simulator configuration and that these ranges can be used to find the best possible solution for each individual motion platform. In this respect, coherence zones as criteria are task and platform dependent. However, coherence zones as a metric, that is, as a measure of the perceived coherence of the inertial feedback provided, is platform and task independent. This allows one to compare motion cueing solutions across simulators, not on the basis of the inertial stimuli they provide, but on the perceived coherence between the visual and inertial stimuli, independently of each platforms' mechanical limitations.

10.2 Recommendations

10.2.1 Further research on amplitude coherence zones

This thesis extended the work on amplitude coherence zones to amplitudes and frequencies common to vehicle simulation. However, the motion profiles used were relatively simple: either acceleration steps or sinusoidal signals. The application of coherence zone measurements to pilot-in-the-loop simulation scenarios can be greatly improved by extending the type of profiles tested.

The effect of frequency has been studied by using single sinusoidal signals. However, in an actual simulation, the motion of the vehicle is likely to have a large combination of frequencies at different amplitudes. One option to extend our knowledge of coherence zones to different signal shapes would be to investigate how perception of coherence between visual and inertial stimuli works when the motion profile is a combination of two or more sines. Does perceived coherence at both frequencies individually guarantee that the combined motion profile is also perceived as coherent? Is it possible to predict the coherence zone of the combined, multiple frequency, motion profile from the coherence zones at the individual frequency signals? How does the relative amplitude of each sinusoid in the combined profile affect the perception of coherence?

10.2.2 Further research on phase coherence zones

The amount of data collected for phase coherence zones was considerably less than for amplitude coherence zones. Although different frequencies and amplitudes have been tested, there are not enough data points to fully understand the relationship between acceleration amplitude, velocity amplitude and frequency. For example, the effect of amplitude and frequency on phase coherence zones found by Grant and Lee (Grant and Lee, 2007) was not observed in the present work. It was hypothesized that their results might be explained by the fact that in one of their conditions very low amplitude stimuli were used. An amplitude level too close to the inertial and visual sensory thresholds, might hinder the perception of each stimulus individually and artificially increase the phase-error threshold. This effect has not been tested yet. It would be worthwhile to extend the amplitude and frequency ranges of the inertial and visual stimuli used in phase coherence zone measurements and observe whether or not, for extreme values, the phase-error threshold varies.

Moreover, as only one-sided phase-error thresholds were measured in the present work, it is recommended to investigate if the same conclusions can be drawn for phase-error thresholds when the inertial motion lags the visual motion. The choice of measuring only thresholds for leading inertial motion was based on the traditional motion filters used in flight simulation. Here, the use of high-pass filters introduces lead in the inertial motion with respect to the visual motion. However, when considering the characteristics of the motion platform in the tuning process, as suggested by Advani and Hosman (2006), this might change. The mechanical limits of the motion platform introduce extra lag in the inertial motion and, especially at high frequencies, may cause the inertial motion to actually lag the visual motion. In addition, phase-error thresholds for lagging inertial motion may also be used to fine-tune the tilt coordination channel of a motion cueing algorithm, where inertial linear accelerations are low-pass filtered.

10.2.3 Combined amplitude and phase coherence zones measurements

Coherence zones were measured in terms of amplitude or phase. During amplitude coherence zone measurements, the phase difference between the two stimuli was zero. The same way, for coherence zone measurements, the gain of the inertial motion with respect to the visual motion was one. As explained in Chapter 9 and further exemplified in Appendix A, for a better definition of the coherence zone criteria in a gain versus phase plot, it is also necessary to measure combined amplitude and phase coherence zones. This means that amplitude coherence zones should be measured when the two stimuli have a certain phase difference, and symmetrically, phase coherence zones should be measured for conditions where the inertial and visual motion have different amplitudes.

10.2.4 Coherence zones and degrees-of-freedom

Most of the amplitude and phase coherence zone measurements were performed in yaw. Phase coherence zones were also performed in pitch and amplitude coherence zones in sway. Although no differences were found between the pitch and yaw coherence zones, for example, in order to construct a complete set of criteria, it is important to gather data for all DOFs.

Moreover, it is useful to investigate how motion in one DOF affects the perception of coherence between inertial and visual stimuli in another DOF. For motion in multiple DOFs, is the separate contribution of each DOF perceived independently from the other(s) or is the motion perceived as a whole? Similar to what has also been proposed for multiple-frequency profiles, can perception coherence zones for multiple-DOF motion be predicted from perception coherence zones in each individual DOF? Besides amplitude and frequency do we need to add other dimensions such as the effects of motion in other DOFs and at other frequencies?

10.2.5 Other types of coherence zones

Besides amplitude and phase coherence zones, there are other measurements interesting for vehicle simulation. One example is the directional coherence zones (De Winkel et al., 2010a,b), where the direction of the inertial cue deviates from that of the visual cue. For example, for a surge visual cue (motion along the x axis), the inertial cue might be a few degrees to the left or right (motion along the x and y axes). These coherence zones might be interesting to apply in cueing solutions for simulators with rotating cabins mounted on top of linear rails or for the tuning of washout motion.

Other types of coherence zones that are interesting to investigate relate to the relationship between coherence zones and optimal gain. Studies that measured subjects' preferred inertial motion gain for different visual stimuli (Correia Grácio et al., 2010) have found that

there is not one, but a range of optimal gains. This range of values may be referred to as an optimal zone. Work has been done to compare perception coherence zones to optimal zones (Correia Grácio et al., 2013) and they have concluded that the optimal zone is contained within the coherence zone. It is likely that both the perception of coherence and the determination of an optimal gain are gradual processes that are not fully described by one threshold or one value. Further work on this topic should investigate whether coherence and optimal zones can be explained by the same perceptual mechanism, but measured at different levels of certainty (or acceptance).

10.2.6 Understanding the mechanisms behind the perception of coherence

Very recently an effort has been made to model perception coherence zones based on the uncertainty of the individual stimulus perception (Dos Santos Buinhas et al., 2013). The CZW was successfully modeled by the summation of the measured Just Noticeable Differences (JND) of the visual and of the inertial stimuli. Moreover, a theoretical model has also been developed to explain the "bending down" of the PMC for higher amplitudes. It would be very interesting to further validate the CZW model and to test the PMC model.

Although this line of experimentation may diverge from vehicle simulation scenarios, it is an indispensable contribution to our understanding of the basic sensing and perception mechanisms underlying the perception of combined stimuli. It is recommended that further research aimed at the development of motion simulation fidelity criteria and guidelines for motion cueing design is always accompanied, in parallel, by more fundamental research on human self-motion perception.

10.2.7 The link between perception and control behavior

As explained in the introduction of this thesis, there are several layers of fidelity and several approaches to study and improve simulator motion. Although the main focus of this thesis has been on perception, there has been some work done to bridge the gap between perception and control behavior studies.

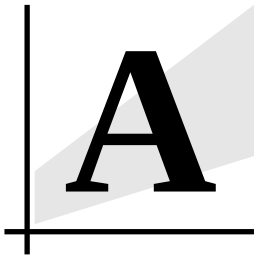
One of these studies investigated the possibility of including a sensory threshold model in a pilot control behavior model and identifying the value of the threshold during a manual tracking task (Valente Pais et al., 2012; Pool et al., 2012). This work used methods typical of control behavior studies and applied them to the measurement of human perception thresholds.

Beckers et al. (2012) compared perception coherence zones to behavior coherence zones. They measured phase coherence zones for different motion profiles from simple sinusoids to more complex multiple-frequency profiles. They then set up a manual control tracking task such that the inertial stimuli provided to subjects during the task resembled the

passively tested profiles. They performed the tracking task for different conditions of phase-error between the inertial and the visual stimuli. By identifying the pilot model parameters for all the conditions they attempted to establish a "behavior coherence zone". That means, a range of phase-error values for which subjects presented similar behavior to the condition with no phase-error. They concluded that subjects' control behavior changed gradually and that there was no steep thresholds for the behavior coherence zone. Nevertheless, they did show that it is possible for subjects to be outside their perception coherence zones and still control the system in a way similar to the one-to-one condition.

It would be very interesting to repeat such a study for amplitude coherence zones and different motion profiles. For that to happen, considerable effort has to be put into designing experimental conditions that fit both the passive and the active tasks. The improvement of the current identification techniques might allow more flexibility in the choice of motion profiles.

Experimental work combining perceptual and behavioral metrics is indispensable to the advancement of motion simulation. These type of studies will allow for a better comprehension of the role of perception in manual control and of the relative importance of perceptual fidelity and behavioral fidelity for the overall success of a simulation.



Perception Coherence Zones Assessment Method

A.1 Coherence zones method

When using coherence zone measurements to derive motion fidelity criteria it should be understood that the outcome is dependent on the specific characteristics of the motion simulator. As shown by the results from Chapter 4 and Chapter 7, the type of visual system and the performance of the motion base may have a large influence on the values found for the lower and upper thresholds. For that very reason, coherence zones may be used to develop cueing solutions that make the best use of each simulator individually.

This appendix describes an example of a possible, but definitely not the only, use of coherence zones as a motion cueing tool. It presents a step-by-step method on how to measure perception coherence zones and how to use those measurements to derive motion cueing criteria for typical high-pass motion filters.

It is assumed that, from a perception point of view, motion filters will only degrade the fidelity of the simulation if the amplitude attenuation and phase-distortion levels introduced are outside the coherence zone. Hence, designing motion filters that provide inertial motion within the coherence zone will lead to the best motion fidelity that a specific simulator can provide.

To use coherence zone measurements to design and tune motion filters three steps need to be performed: first, the measurement parameters, such as amplitude and frequency, have to be chosen; second, the measurements of the coherence zones for the chosen conditions have to be performed; and third, the results have to be converted into fidelity criteria. A brief description of these three steps and the desired results is presented below. The following sections will explain each of the three steps in more detail.

A.2 Selecting the measurement points

Coherence zones may be measured for different DOFs, different motion profiles with different frequencies and amplitudes and using different visual scenes. Since it is very impractical to do measurements for all possible occurring combinations of parameters, a selection should be made that may vary depending on the task to be performed by the subject in the simulator.

A.2.1 Frequency

For each specific task, the dominant and the “more often” occurring frequencies should be identified. Since the frequency of the stimulus has an effect on the coherence zones, more than one frequency level should be chosen. When the task at hand does not seem to generate motion at specific frequencies or the measurements are not aimed at a specific task, three frequency levels may be chosen: low, medium and high. These may be chosen to cover the full operational range during a specific task or to span the limits of the motion platform.

Steps	Description	Result
1. Selection	Based on the type of task/vehicle to be simulated, select the frequency and amplitude of the visual and inertial stimuli to be used.	Set of M×N conditions defined by M frequency levels and N amplitude levels.
2. Measurement	Define the measurement methods and the measurement points: phase and amplitude coherence zones, combined phase and amplitude coherence zones, lower and/or upper thresholds, etc. Perform the measurements.	Phase and amplitude coherence zone data measured at the selected conditions.
3. Criteria	Convert the measurement data into criteria that can then be used as a guideline for design and tuning of the motion cueing algorithm.	Motion fidelity criteria defined in terms of minimum motion gain and maximum phase-distortion for M frequency levels and N amplitude levels, resulting in M×N lines on a gain vs. phase plot.

Since the effect of frequency on the angular motion amplitude coherence zones is expected to be significant when comparing low ($<$ 1 rad/s) and high ($>$ 9 rad/s) frequencies to medium frequencies, collecting data in these three regions is likely to offer a good estimate of the coherence zones for all other intermediate frequencies. In other words, when for example collecting data at 0.6, 4 and 10 rad/s, the thresholds at 1 rad/s are likely to be somewhere between the values found at 0.6 and 4 rad/s and the threshold at 6 rad/s will be between the values found at 4 and 10 rad/s.

A.2.2 Amplitude

The selection of the stimulus amplitude is also important, especially for the current motion filtering solutions which do not discriminate between different amplitude inputs. Depending on the type of task, it could be expected that during a simulation the maximum amplitude of motion varies with time.

For example, when performing a heading capture task in an helicopter, the amplitude of motion is large for the first capture movement, and is very small when the pilot performs that last adjustments around the desired heading. For such a task, two amplitude levels can be chosen, one that matches the maximum amplitude during the capture movement and a second one that matches the average of the maximum amplitudes expected during the ad-

justment phase. This would result in two different criteria for the two different amplitudes. However, most motion filters in use today will not filter motion of different amplitudes differently. For that reason, the criteria corresponding to one or the other amplitude will have to be chosen.

One may then choose for the amplitude that is based on the most crucial part of the task. In the yaw capture task example, if the goal of the simulation is to study the importance of inertial feedback on pilot control behavior, then one might say that the first capture motion is made based primarily on the pilot's knowledge of the helicopter dynamics and the desired heading and that for the small adjustments inertial feedback is actually more relevant. In this case, one might choose for the amplitude level expected during the adjustments phase, as here the fidelity of the inertial motion provided is crucial for the goal of the study.

On the other hand, imagining that the goal of the simulation is to teach naive subjects how to perform a yaw capture task in an helicopter, one might consider the immediate response of the vehicle to control inputs to be very important. In this case, the amplitude level expected during this first maneuver should be chosen as a measurement point.

A third solution, assuming that motion fidelity should be equally good during both phases, is to choose an amplitude level halfway between the amplitude expected in the two phases.

One last option, would be to combine different motion filters for different phases of the simulation. Such a solution severely complicates the control of the inertial motion system. The online switching between two motion filters, each more appropriate to a different phase of the simulation, is not a trivial problem and is out of the scope of this appendix.

It is important to note that choosing only one amplitude level for the measurements determines the type of compromise to be made in the motion filter design. Criteria derived from such measurements already reflect the compromise made between having high fidelity motion at one amplitude or the other. If more than one amplitude is chosen, then more than one amplitude related criteria can be derived. Then, if there is a compromise to be made, it can be done during the motion filter tuning phase, by adjusting motion filter parameters to meet one criterion or the other.

A.2.3 Reducing the number of measurement points

Assuming that two amplitude levels and three frequency levels have been chosen, six measurement conditions are obtained. If both amplitude and phase coherence zones are to be measured, each with lower and upper thresholds, there would be a total of 24 measurements to be made. Although 24 is a reasonable number of measurements, to allow for a faster data collection, this number can be reduced.

In traditional high-pass filters, motion gains are chosen such that the necessary motion cues are provided while saving as much motion space as possible. In that case, only the

lower limit of the amplitude coherence zone is of interest and measuring only the lower threshold should suffice. Measuring upper thresholds may be useful for the tuning of washout motion or tilt coordination maneuvers as explained in Section 9.5. Nevertheless, to simplify this explanation, upper threshold measurements will not be considered. Moreover, depending on the performance of the motion and visual systems, it could be assumed that the inertial motion will always lead the visual motion and hence, only one threshold for the phase coherence zone can be measured as well.

Furthermore, other considerations can be made that allow reducing the amount of measurements to be made. It has been shown that amplitude coherence zones are affected by both the amplitude and the frequency of the visual stimulus, hence for all the six measurements conditions amplitude coherence zones should be measured.

In contrast, for phase coherence zones during angular motion and for high enough amplitudes, no significant effects of amplitude and frequency were found. If the chosen amplitude levels are high enough and allow subjects to easily perceive both the inertial and the visual stimuli, then it should suffice to perform measurements at only one amplitude level. In general, subjects report that higher amplitude conditions facilitate the perception task at hand. For this reason, phase coherence zones should be measured for the higher amplitude level.

Regarding the frequency, despite the fact that no significant effect on phase coherence zones was found, most motion platforms have a frequency dependent performance. To be sure to accommodate possible timing differences between the inertial and the visual stimuli, phase coherence zones should be measured at the three different frequencies.

A.3 Measuring coherence zones

As discussed above, in the following, phase coherence zone measurements refer to one phase threshold (motion leads) and amplitude coherence zone measurements refer to one amplitude threshold (lower threshold). If rendered necessary, the exact same procedures can be followed to measure the other thresholds.

For amplitude coherence zones, the self-tuning method described in Chapters 3 and 4 should be used. As compared to a staircase method, the self-tuning method is not necessarily faster or more accurate but it engages subjects more actively, which reduces subjects boredom and fatigue.

For phase coherence zones, subjects are able to recognize when the two stimuli, visual and inertial, are not synchronized but are generally unable to say which signal is leading and which one is lagging. For this reason, the self-tuning method is not very effective. Instead, a staircase method as described in Chapter 4 should be used.

For each of the conditions described above, four measurements should be performed, named here A, B, C and D. A corresponds to the measurement of the lower threshold of

the amplitude coherence zone and B to the phase threshold. The C measurement can be obtained by using the value of the measured lower threshold (A), divided by the amplitude of the visual stimulus, as a motion gain applied to the inertial motion and then measuring another phase-threshold. The D measurement is obtained by using the first measured phase threshold (B) as a fixed mismatch between the inertial and visual stimuli and then measuring a second lower amplitude threshold.

As discussed above, the phase thresholds are not expected to change for different amplitudes and so the phase coherence zone, measurement B, can be measured at only one amplitude level. Measurements A and D refer to amplitude coherence zone measurements, which has been shown are influenced by both amplitude and frequency, and should therefore be performed for all experimental conditions. Measurement C, despite being a phase coherence zone measurement, depends on the value measured for A, so it should also be performed for all conditions. Table A.1 summarizes the type of measurements to be performed at each of the measurement conditions.

Table A.1: Measurement conditions and type of thresholds to be measured at each of the conditions. Lower amplitude thresholds are indicated in black and phase-error thresholds, in gray.

	Frequencies		
Amplitudes	Low	Medium	High
Low	**A**,**C**,**D**	**A**,**C**,**D**	**A**,**C**,**D**
High	**A**,**B**,**C**,**D**	**A**,**B**,**C**,**D**	**A**,**B**,**C**,**D**

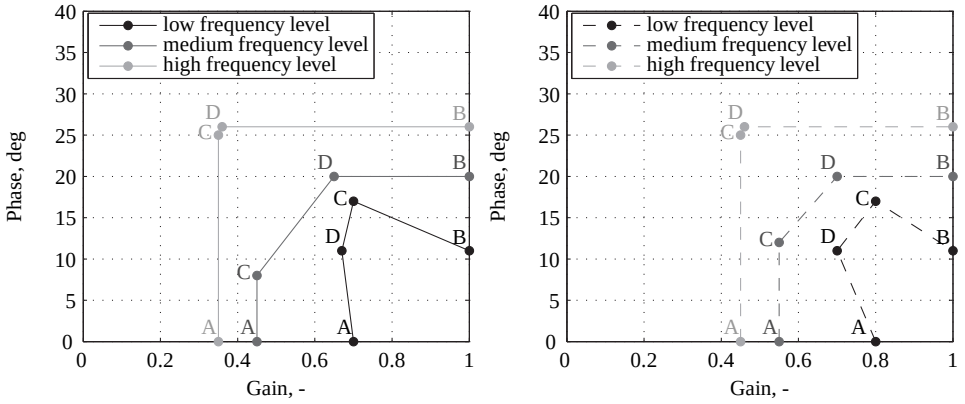
A.4 Converting coherence zones into criteria

The measured data can now be converted into different criteria for all the amplitude and frequency levels tested. Each of the tested conditions will result in one boundary line, each defined by four points, corresponding to the four measurements made, A to D.

To obtain those boundary lines, first, the measured lower thresholds should be converted into gains by dividing the measured value by the corresponding visual stimulus amplitude. The phase-error thresholds can be used directly. These gains and phase-errors can then be presented in a motion-gain vs. phase-distortion plot.

Figure A.1 shows hypothetical threshold values, translated into points along criteria boundary lines, for six experimental conditions.

As it was discussed in the previous section, measurement B was only performed at one amplitude level. To complete the set of criteria, this value is used twice, for both amplitude levels.



(a) High amplitude level. **(b)** Low amplitude level.

Figure A.1: Hypothetical amplitude lower thresholds and phase-error thresholds for three different stimulus frequencies and two amplitude levels.

Obviously, the more points are measured, the better defined each line will be. Nevertheless, four points are sufficient to conclude whether the coherent motion region is rectangular, rounded or with sloped upper and left boundaries.

Having defined the criteria bounds, different cueing solutions can now be tested offline by comparing the gain and phase distortion introduced by the motion filter in the inertial stimulus with the desired motion gains and phase errors.

The execution of the steps described above lead to motion fidelity criteria for high-pass motion filters. The step after that, and not described here, is to tune motion filters based on this criteria and then compare the found solutions with other motion filter settings obtained through expert judgments. The comparison of these solutions with respect to perceptual metrics would fully validate the coherence zone method and its results as perception-based motion cueing criteria. Doing so, would allow to close the loop presented in the thesis introduction (see Figure 1.5), from cueing, to perception, and then back to cueing.

References

Advani, S. K. and Hosman, R. J. A. W. (2006). Towards Standardizing High-Fidelity Cost-Effective Motion Cueing in Flight Simulation. In *Proceedings of the Royal Aeronautical Society Conference on Cutting Costs in Flight Simulation: Balancing Quality and Capability, London, UK, November 7-8*.

Advani, S. K., Potter, M., and Fernie, G. (2010a). CEAL - Flight Simulation Technology Applied to Rehabilitation Research. In *Proceedings of the AIAA Modeling and Simulation Technologies Conference and Exhibit, Toronto, ON, Canada, August 2-5*, AIAA 2010-8100.

Advani, S. K., Schroeder, J. A., and Burks, B. (2010b). What Really Can Be Done in Simulation to Improve Upset Training? In *Proceedings of the AIAA Modeling and Simulation Technologies Conference and Exhibit, Toronto, ON, Canada, August 2-5*, AIAA 2010-7791.

Angeles, J. (2003). *Fundamentals of Robotic Mechanical Systems: Theory, Methods, and Algorithms*. Springer-Verlag, 175 Fifth Avenue New York, NY 10010, USA, second edition. ISBN:0-387-95368-X.

Atencio, Jr., A. (1993). Fidelity Assessment of a UH-60A Simulation on the NASA Ames Vertical Motion Simulator. Technical Memorandum NASA 104016, NASA, Ames Research Center, Moffet Field, CA, USA.

Beckers, N. W. M., Pool, D. M., Valente Pais, A. R., Van Paassen, M. M., and Mulder, M. (2012). Perception and Behavioral Phase Coherence Zones in Passive and Active Control Tasks in Yaw. In *Proceedings of the AIAA Modeling and Simulation Technologies Conference and Exhibit, Minneapolis, MN, USA, August 13-16*, AIAA 2012-4794.

Benson, A. J., Hutt, E. C. B., and 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., and 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.

Berkouwer, W. R., Stroosma, O., Van Paassen, M. M., Mulder, M., and Mulder, J. A. (2005). Measuring the Performance of the SIMONA Research Simulator's Motion System. In *Proceedings of the AIAA Modeling and Simulation Technologies Conference and Exhibit, San Francisco, CA, USA, August 15-18*, AIAA 2005-6504.

Berthoz, A. and Droulez, J. (1982). *Tutorials on Human Motion Perception*, chapter Linear Self Motion Perception, 157–199. Plenum Press, 233 Spring Street, New York, NY 10013, USA. ISBN: 0-306-41126-1.

Berthoz, A., Pavard, B., and 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., Collet, C., Espié, S., and Graf, W. (2005). Objective Measurement of Simulator Sickness and the Role of Visual-Vestibular Conflict Situations. In *Proceedings of the Driving Simulation Conference North America, Orlando, FL, USA, November 30 - December 2*, 280–293.

Blaauw, G. J. (1982). Driving Experience and Task Demands in Simulator and Instrumented Car: A Validation Study. *Human Factors*, 24(4):473–486.

Bles, W., Hosman, R. J. A. W., and De Graaf, B. (2000). Desdemona – Advanced Disorientation Trainer and (Sustained-G) Flight Simulator. In *Proceedings of the AIAA Modeling and Simulation Technologies Conference and Exhibit, Denver, CO, USA, August 14-17*, AIAA 2000-4176.

Boer, E. R., Yamamura, T., Kuge, N., and Girshick, A. (2000). Experiencing the Same Road Twice: A Driver Centered Comparison Between Simulation and Reality. In *Proceedings of the Driving Simulation Conference, Paris, France, September 6-8*.

Borah, J., Young, L. R., and Curry, R. E. (1988). Optimal Estimator Model for Human Spatial Orientation. *Annals of The New York Academy of Sciences*, 545(Representation of Three-Dimensional Space in the Vestibular, Oculomotor, and Visual System):51–73.

Bos, J. E. and Bles, W. (2002). Theoretical Considerations on Canal-Otolith Interaction and an Observer Model. *Biological Cybernetics*, 86(3):191–207.

Bos, J. E., Groen, E. L., and Nooij, S. A. E. (2004). Further Thoughts on and Calculations by Spatial Orientation and Motion Sickness Modeling. Technical Report TM-04-I005, TNO Defense, Security and Safety, Soesterberg, The Netherlands.

Bos, J. E., Hosman, R. J. A. W., and Bles, W. (2002). Visual-Vestibular Interactions and Spatial (Dis)orientation in Flight and Flight Simulation. Technical Report TM-02-C009, TNO Defense, Security and Safety, Soesterberg, The Netherlands.

Bos, J. E., MacKinnon, S. N., and Patterson, A. (2005). Motion Sickness Symptoms in a Ship Motion Simulator: Effects of Inside, Outside, and No View. *Aviation, Space, and Environmental Medicine*, 76(12):1111–1118.

Van Boxtel, J. J. A., Van Ee, R., and Erkelens, C. J. (2006). A Single System Explains Human Speed Perception. *Journal of Cognitive Neuroscience*, 18(11):1808–1819.

Brandt, Th., Dichgans, J., and Koenig, E. (1973). Differential Effects of Central Versus Peripheral Vision on Egocentric and Exocentric Motion Perception. *Experimental Brain Research*, 16(5):476–491.

Brünger-Koch, M., Briest, S., and Vollrath, M. (2006). Virtual Driving With Different Motion Characteristics - Braking Manoeuvre Analysis and Validation. In *Proceedings of the Driving Simulation Conference, Paris, France, October 4-6*, 69–78.

Burns, D. and Osfield, R. (2004). Open Scene Graph A: Introduction, B: Examples and Applications. In *IEEE Virtual Reality*, 265. IEEE Computer Society.

Casali, J. G. and Wierwille, W. W. (1980). The Effects of Various Design Alternatives on Moving-Base Driving Simulator Discomfort. *Human Factors*, 26(6):741–756.

Chapron, T. and Colinot, J. (2007). The new PSA Peugeot-Citroën Advanced Driving Simulator: Overall Design and Motion Cue Algorithm. In *Proceedings of the Driving Simulation Conference North America, Iowa City, IA, USA, September 12-14*.

Cheung, B. S. K., Howard, I. P., Nedzelski, J. M., and Landolt, J. P. (1989). Circularvection About Earth-horizontal Axes in Bilateral Labyrinthine-defective Subjects. *Acta Otolaryngologica*, 108(5-6):336–344.

Chung, W. W. and Schroeder, J. A. (1997). Visual and Roll-Lateral Motion Cueing Synchronization Requirements for Motion-Based Flight Simulations. In *American Helicopter Society 53rd Annual Forum, Virginia Beach, VA, USA, April 29 - May 1*, 994–1006.

Chung, W. W. Y., Sweet, B. T., and Lewis, E. (2003). Visual Cueing Effects Investigation for a Hover Task. In *Proceedings of the AIAA Modeling and Simulation Technologies Conference and Exhibit, Austin, TX, USA, August 11-14*, AIAA 2003-5524.

Clark, C. C. and Hardy, J. D. (1959). Preparing Man for Space Flight. *Astronautics*, 88(IV):18–21.

Correia Grácio, B. J., Van Paassen, M. M., Mulder, M., and Wentink, M. (2010). Tuning of the Lateral Specific Force Gain Based on Human Motion Perception in the Desdemona Simulator. In *Proceedings of the AIAA Modeling and Simulation Technologies Conference and Exhibit, Toronto, ON, Canada, August 2-5*, AIAA 2010-8094.

Correia Grácio, B. J., Valente Pais, A. R., Van Paassen, M. M., Mulder, M., Kelly, L. C., and Houck, J. A. (2013). Optimal and Coherence Zone Comparison Within and Between Flight Simulators. *Journal of Aircraft*, 50(2):493–507.

Dagdelen, M., Reymond, G., Kemeny, A., Bordier, M., and Maïki, N. (2004). MPC based motion cueing algorithm: Development and application to the ULTIMATE driving simulator. In *Proceedings of the Driving Simulation Conference, Paris, France, September 8-10*, 221–233.

Damveld, H. J. (2009). *A Cybernetic Approach to Assess the Longitudinal Handling Qualities of Aeroelastic Aircraft*. Ph.D. thesis, Delft University of Technology, Faculty of Aerospace Engineering.

Dearing, M. G., Schroeder, J. A., Sweet, B. T., and Kaiser, M. K. (2001). Effects of Visual Texture, Grids, and Platform Motion on Unpowered Helicopter Landings. In *Proceedings of the AIAA Modeling and Simulation Technologies Conference and Exhibit, Montreal, Quebec, Canada, August 6-9*, AIAA 2001-4251.

Van Egmond, A. A. J., Groen, J. J., and Jongkees, L. B. W. (1949). The Mechanics of the Semicircular Canal. *Journal of Physiology*, 110(1-2):1–17.

Ellerbroek, J., Stroosma, O., Mulder, M., and Van Paassen, M. M. (2008). Role Identification of Yaw and Sway Motion in Helicopter Yaw Control Tasks. *Journal of Aircraft*, 45(4):1275–1289.

Elsner, W. (1971). Power Laws for the Perception of Rotation and the Oculogyral Illusion. *Perception & Psychophysics*, 9(5):418–420.

Feenstra, P. J., Wentink, M., Roza, Z. C., and Bles, W. (2007). Desdemona, an Alternative Moving Base Design for Driving Simulation. In *Proceedings of the Driving Simulation Conference North America, Iowa City, September 12-14*.

Fernandez, C. and Goldberg, J. M. (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.

Field, A. (2005). *Discovering Statistics Using SPSS*. SAGE Publications, 1 Oliver's Yard, 55 City Road, London EC1Y 1SP, UK, second edition. ISBN 978-0-7619-4452-2.

Fortmüller, T. and Meywerk, M. (2005). The Influence of Yaw Movements on the Rating of the Subjective Impression of Driving. In *Proceedings of the Driving Simulation Conference North America, Orlando, FL, USA, November 30 - December 2*, 362–373.

Fortmüller, T., Tomaske, W., and Meywerk, M. (2008). The Influence of Sway Accelerations on the Perception of Yaw Movements. In *Proceedings of the Driving Simulation Conference Europe, Monaco, January 31 - February 1*, 161–170. INRETS-Renault.

Frank, L. H., Casali, J. G., and Wierwille, W. W. (1988). Effects of Visual Display and Motion System Delays on Operator Performance and Uneasiness in a Driving Simulator. *Human Factors*, 30(2):201–217.

Godthelp, H. (1986). Vehicle Control During Curve Driving. *Human Factors*, 28:211–221.

Godthelp, H., Milgram, P., and Blaauw, G. J. (1984). The Development of a Time-Related Measure to Describe Driving Strategy. *Human Factors*, 26:257–268.

Goldstein, E. B. (2002). *Sensation and Perception*. Wadsworth, 511 Foresst Lodge Road, Pacific Grove, CA 93950, USA, sixth edition. ISBN 0-534-53964-5.

Gough, V. E. (1956-1957). Remarks on Automobile, Stability, Control, and Tyre Performance. In *Proceedings of the Automobile Division*, 392–394. The Institution of Mechanical Engineers, Incorporating the Institution of Automobile Engineers, The Institution of Mechanical Engineers, Westminster, London, UK.

Grant, P. R., Artz, B., Greenberg, J., and Cathey, L. (2001). Motion Characteristics of the VIRTTEX Motion System. In *Proceedings of the 1st Human-Centered Transportation Simulation Conference, Iowa City, IA, USA, November 4-7*, 1–11. The University of Iowa.

Grant, P. R. and Haycock, B. (2006). The Effect of Jerk and Acceleration on the Perception of Motion Strength. In *Proceedings of the AIAA Modeling and Simulation Technologies Conference and Exhibit, Keystone, CO, USA, August 21-24*, AIAA 2006-6253.

Grant, P. R. and Lee, P. T. S. (2007). Motion-Visual Phase-Error Detection in a Flight Simulator. *Journal of Aircraft*, 44(3):927–935.

Grant, P. R., Papelis, Y., Schwarz, C., and Clark, A. (2004). Enhancements to the NADS Motion Drive Algorithm for Low-Speed Urban Driving. In *Driving Simulation Conference Europe, Paris, France, September 8-10*, 67–77.

Grant, P. R., Yam, B., Hosman, R. J. A. W., and Schroeder, J. A. (2005). The Effect of Simulator Motion on Pilot's Control Behavior for Helicopter Yaw Control Tasks. In

Proceedings of the AIAA Modeling and Simulation Technologies Conference and Exhibit, San Francisco, CA, USA, August 15-18, AIAA 2005-6304.

Grant, P. R., Yam, B., Hosman, R. J. A. W., and Schroeder, J. A. (2006). Effect of Simulator Motion on Pilot Behavior and Perception. *Journal of Aircraft*, 43(6):1914–1924.

Gray, W. R., III (2009). Handling Qualities Evaluation at the USAF Test Pilot School. In *AIAA Atmospheric Flight Mechanics Conference, Chicago, IL, USA, August 10-13*, AIAA 2009-6317.

Graybiel, A., Clark, B., and Zarriello, J. J. (1960). Observations of Human Subjects Living in a "Slow Rotation Room" for Periods of Two Days. *Archives of Neurology*, 3(1):55–73.

Greenberg, J., Artz, B., and Cathey, L. (2003). The Effect of Lateral Motion Cues During Simulated Driving. In *Proceedings of the Driving Simulation Conference North America, Dearborn, MI, USA, October 8-10*.

Greig, G. L. (1988). Masking of Motion Cues by Random Motion: Comparison of Human Performance with a Signal Detection Model. Technical Report 313, UTIAS.

Van de Grind, W. A. (1988). The Possible Structure and Role of Neuronal Smart Mechanisms in Vision. *Cognitive Systems*, 2(2):163–180.

Groen, E. L. and Bles, W. (2004). How to Use Body Tilt for the Simulation of Linear Self Motion. *Journal of Vestibular Research*, 14(5):375–385.

Groen, E. L., Hosman, R. J. A. W., Bos, J. E., and Dominicus, J. W. (2004). Motion Perception Modelling in Flight Simulation. In *Proceedings of the RAeS Conference, "Flight Simulation 1929-2029: A Centennial Perspective", London, UK, May 26-27*, 13.1–13.9.

Groen, E. L., Howard, I. P., and Cheung, B. S. K. (1999). Influence of Body Roll on Visually Induced Sensations of Self-Tilt and Rotation. *Perception*, 28(3):287–297.

Groen, E. L., Smaili, M. H., and Hosman, R. J. A. W. (2005). Simulated Decrab Maneuver: Evaluation with a Pilot Perception Model. In *Proceedings of the AIAA Modeling and Simulation Technologies Conference and Exhibit, San Francisco, CA, USA, August 15-18*, AIAA 2005-6108.

Groen, E. L., Smaili, M. H., and Hosman, R. J. A. W. (2007). Perception Model Analysis of Flight Simulator Motion for a Decrab Maneuver. *Journal of Aircraft*, 44(2):427–435.

Groen, E. L., Valenti Clari, M. S. V., and Hosman, R. J. A. W. (2001). Evaluation of Perceived Motion During a Simulated Takeoff Run. *Journal of Aircraft*, 38(4).

Guedry, F. E., Graybiel, A., and Collins, W. E. (1962). Reduction of Nystagmus and Disorientation in Human Subjects. Joint Report NASA-CR-55756, United States Naval School of Aviation Medicine and NASA, (Unspecified center).

Gundry, A. J. (1977). Thresholds to Roll Motion in a Flight Simulator. *Journal of Aircraft*, 14(7):624–631.

Gurnee, H. (1931). The Effect of a Visual Stimulus upon the Perception of Bodily Motion. *The American Journal of Psychology*, 43(1):26–48.

Heerspink, H. M., Berkouwer, W. R., Stroosma, O., Van Paassen, M. M., Mulder, M., and Mulder, J. A. (2005). Evaluation of Vestibular Thresholds for Motion Detection in the SIMONA Research Simulator. In *Proceedings of the AIAA Modeling and Simulation Technologies Conference and Exhibit, San Francisco, CA, USA, August 15-18*, AIAA 2005-6502.

Held, R., Dichgans, J., and Bauer, J. (1975). Characteristics of Moving Visual Scenes Influencing Spatial Orientation. *Vision Research*, 15(3):357–365.

Hoffman, J. D., Lee, J. D., Brown, T. L., and McGehee, D. V. (2002). Comparison of Driver Braking Responses in a High-Fidelity Simulator and on a Test Track. *Journal of the Transportation Research Board*, 1803(1):59–65.

Hood, D. C. and Finkelstein, M. A. (1968). *Handbook of Perception and Human Performance*, volume I: Sensory Processes and Perception, chapter 5: Sensitivity to Light. A Wiley-Interscience Publication, 111 River Street, Hoboken, NJ 07030-5774, USA. ISBN 0-471-88544-4.

Horberry, T., Anderson, J., Regan, M. A., Triggs, T. J., and Brown, J. (2006). Driver Distraction: The Effects of Concurrent In-vehicle Tasks, Road Environment Complexity and Age on Driving Performance. *Accident Analysis and Prevention*, 38(1):185–191.

Hosman, R. J. A. W. (1996). *Pilot's Perception and Control of Aircraft Motions*. Ph.D. thesis, Delft University of Technology, Faculty of Aerospace Engineering.

Hosman, R. J. A. W. and Van der Steen, F. A. M. (1993). False Cue Detection Thresholds in Flight Simulation. In *Proceedings of the AIAA Flight Simulation Technology Conference, Monterey, CA, USA, August 9-11*, AIAA 1993-3578-CP, 193–201.

Hosman, R. J. A. W. and Van der Vaart, J. C. (1978). Vestibular Models and Thresholds of Motion Perception. Results of Tests in a Flight Simulator. Internal Report LR-265, Delft University of Technology, Faculty of Aerospace Engineering, Delft, The Netherlands.

Hosman, R. J. A. W. and Van der Vaart, J. C. (1980). Thresholds of Motion Perception and Parameters of Vestibular Models Obtained from Tests in a Motion Simulator. Effects of

Vestibular and Visual Motion Perception on Task Performance. Technical Report M-372, Delft University of Technology, Delft, The Netherlands.

Hosman, R. J. A. W. and Van der Vaart, J. C. (1981). Effects of Vestibular and Visual Motion Perception on Task Performance. *Acta Psychologica*, 48(1-3):271 – 287.

Howard, I. P. (1968). *Handbook of Perception and Human Performance*, volume I: Sensory Processes and Perception, chapter 11: The Vestibular System. A Wiley-Interscience Publication, 111 River Street, Hoboken, NJ 07030-5774, USA. ISBN 0-471-88544-4.

Huang, J. and Young, L. R. (1981). Sensation of Rotation About a Vertical Axis with a Fixed Visual Field in Different Illuminations and in the Dark. *Experimental Brain Research*, 41(2):172–183.

International Civil Aviation Organization (2009). Manual of Criteria for the Qualification of Flight Simulation Training Devices. Volume 1 – Airplanes. ICAO Doc 9625, International Civil Aviation Organization (ICAO). Third edition.

Jamson, A. H., Horrobin, A. J., and Auckland, R. A. (2007a). Driving: Whatever Happened to the LADS? Design, Development and Preliminary Validation of the New University of Leeds Driving Simulator. In *Proceedings of the Driving Simulation Conference North America, Iowa City, IA, USA, September 12-14*.

Jamson, A. H., Whiffin, P. G., and Burchill, P. M. (2007b). Driver Response to Controllable Failures of Fixed and Variable Gain Steering. *International Journal of Vehicle Design*, 45(3):361–378.

Jex, H. R., Magdaleno, R. E., and Junker, A. M. (1978). Roll Tracking Effects of G-Vector Tilt and Various Types of Motion Washout. In *Proceedings of the Fourteenth Annual Conference on Manual Control, Los Angeles and Moffet Field, CA, USA, April 25-27*, 463–502.

Joint Aviation Authorities (2003). Aeroplane Flight Simulators. JAR-STD 1A, Joint Aviation Authorities (JAA).

Kamphuis, H. H. (1994). *Visual-Vestibular Interaction and the Perception of Self-motion*. Master of Science Thesis, Delft University of Technology. Unpublished.

Kawakita, T., Kuno, S., Miyake, Y., and Watanabe, S. (2000). Body Sway Induced by Depth Linear Vection in Reference to Central and Peripheral Visual Field. *The Japanese Journal of Physiology*, 50(3):315–321.

Kemeny, A. and Panerai, F. (2003). Evaluating Perception in Driving Simulation Experiments. *TRENDS in Cognitive Sciences*, 7(1):31–37.

Kennedy, R. S., Lane, N. E., Berbaum, K. S., and Lilienthal, M. G. (1993). Simulator Sickness Questionnaire: An Enhanced Method for Quantifying Simulator Sickness. *International Journal of Aviation Psychology*, 3(3):203–220.

Kennedy, R. S., Lilienthal, M. G., Berbaum, K. S., Baltzley, D. R., and McCauley, M. E. (1989). Simulator Sickness in U.S. Navy Flight Simulators. *Aviation, Space, and Environmental Medicine*, 60(1):10–16.

Levitt, H. (1971). Transformed Up-Down Methods in Psychoacoustics. *Journal of the Acoustical Society of America*, 49(2):467–477.

Loose, R. and Probst, Th. (2001). Velocity not Acceleration of Self-Motion Mediates Vestibular - Visual Interaction. *Perception*, 30(4):511–518.

Mallery, R. M., Uchanski, R. M., and Hullar, T. E. (2010). Human Discrimination of Rotational Velocities. *Experimental Brain Research*, 204(1):11 – 20.

McLean, J. R. and Hoffmann, E. R. (1975). Steering Reversals as a Measure of Driver Performance and Steering Task Difficulty. *Human Factors*, 17(3):248–256.

McRuer, D. T., Graham, D., Krendel, E. S., and Reisener, W. (1965). Human Pilot Dynamics in Compensatory Systems. Theory, Models and Experiments With Controlled Element and Forcing Function Variations. Technical Report AFFDL-TR-65-15, Air Force Flight Dynamics Laboratory, Wright-Patterson Air Force Base, OH, USA.

Melcher, G. A. and Henn, V. (1981). The Latency of Circular Vection During Different Accelerations of the Optokinetic Stimulus. *Perception and Psychophysics*, 30(6):552–556.

Mercedes-Benz Driving Simulator Helps Fine Tune 2012 B-Class (2011). Retrieved on November 5, 2012, from http://www.emercedesbenz.com/autos/mercedes-benz/b-class/mercedes-benz-driving-simulator-helps-fine-tune-2012- b-class/.

Mergner, T. and Becker, W. (1990). *Perception and Control of Self Motion*, chapter Perception of Horizontal Self-rotation: Multisensory and Cognitive Aspects. Lawrence Erlbaum Associates, Inc., 365 Broadway, Hillsdale, New Jersey 07642, NY, USA. ISBN 0-8058-0909-0.

Mesland, B. S. (1998). *About Horizontal Self-Motion Perception....* Ph.D. thesis, Universiteit van Utrecht.

Mesland, B. S. and Wertheim, A. H. (1995). Visual and Nonvisual Contribuitions to Perceived Ego-motion Studied with a New Psychophysical Method. *Journal of Vestibular Research*, 5(4):227–288.

Miller, Jr., G. K. and Riley, D. R. (1977). Visual/Motion Cue Mismatch in a Coordinated Roll Maneuver. Technical Note NASA TN D-8364, NASA, Langley Research Center, Hampton, VA, USA.

Mudd, S. (1968). Assessment of the Fidelity of Dynamic Flight Simulators. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 10(4):351–358.

NADS-1 (2010). Retrieved on November 5, 2012, from http://www.nads-sc.uiowa.edu/sim_nads1.php.

Nakamura, S. (2001). The Perception of Self-Motion Induced by Central and Peripheral Visual Stimuli Moving in Opposite Directions. *Japanese Psychological Research*, 43(3):113–120.

Naseri, A. and Grant, P. R. (2011). Difference Thresholds: Measurement and Modeling. In *Proceedings of the AIAA Modeling and Simulation Technologies Conference, Portland, OR, USA, August 8-11*, AIAA 2011-6245.

Naseri, A., Grant, P. R., and Dufort, P. (2008). Modeling the Perception of Acceleration and Jerk using Signal Detection Theory. In *Proceedings of the AIAA Modeling and Simulation Technologies Conference and Exhibit, Honolulu, HI, USA, August 18-21*, AIAA 2008-6846.

Nieuwenhuizen, F. M. (2012). *Changes in Pilot Control Behaviour Across Stewart Platform Motion Systems*. Ph.D. thesis, Delft University of Technology, Faculty of Aerospace Engineering.

Nilsson, L. (1993). Behavioural Research in an Advanced Driving Simulator – Experiences of the VTI System. In *Proceedings of the Human Factors and Ergonomics Society 37th Annual Meeting, Seattle, WA, USA, October 11-15*, 612–616.

Van Paassen, M. M. and Stroosma, O. (2000). DUECA - Data-Driven Activation in Distributed Real-Time Computation. In *Proceedings of the AIAA Modeling and Simulation Technologies Conference and Exhibit, Denver, CO, USA, August 14-17*, AIAA 2000-4503.

Panerai, F., Droulez, J., Kelada, J., Kemeny, A., Balligand, E., and Favre, B. (2001). Speed and Safety Distance Control in Truck Driving: Comparison of Simulation and Real-World Environment. In *Proceedings of the Driving Simulation Conference, Sophia Antipolis, Nice, France, September 5-7*.

Pavard, B. and Berthoz, A. (1977). Linear Acceleration Modifies the Perceived Velocity of a Moving Visual Scene. *Perception*, 6(5):529–540.

Pool, D. M. (2012). *Objective Evaluation of Flight Simulator Motion Cueing Fidelity Through a Cybernetic Approach*. Ph.D. thesis, Delft University of Technology, Faculty of Aerospace Engineering.

Pool, D. M., Damveld, H. J., Van Paassen, M. M., and Mulder, M. (2011a). Tuning Models of Pilot Tracking Behavior for a Specific Simulator Motion Cueing Setting. In *Proceedings of the AIAA Modeling and Simulation Technologies Conference, Portland, OR, USA, August 8-11*, AIAA 2011-6322.

Pool, D. M., Valente Pais, A. R., De Vroome, A. M., Van Paassen, M. M., and Mulder, M. (2012). Identification of Nonlinear Motion Perception Dynamics Using Time-Domain Pilot Modeling. *Journal of Guidance, Control, and Dynamics*, 35(3):749–763.

Pool, D. M., Zaal, P. M. T., Damveld, H. J., Van Paassen, M. M., Van der Vaart, J. C., and Mulder, M. (2011b). Modeling Wide-Frequency-Range Pilot Equalization for Control of Aircraft Pitch Dynamics. *Journal of Guidance, Control, and Dynamics*, 34(5):1529–1542.

Pool, D. M., Zaal, P. M. T., Van Paassen, M. M., and Mulder, M. (2010). Effects of Heave Washout Settings in Aircraft Pitch Disturbance Rejection. *Journal of Guidance, Control, and Dynamics*, 33(1):29–41.

Probst, Th., Straube, A., and Bles, W. (1985). Differential Effects of Ambivalent Visual-Vestibular-Somatosensory Stimulation on the Perception of Self-Motion. *Behavioural Brain Research*.

Reid, L. D. and Nahon, M. A. (1985). Flight Simulation Motion-Base Drive Algorithms. Part 1: Developing and Testing the Equations. Technical Report 296, UTIAS, Toronto, Canada.

Reid, L. D. and Nahon, M. A. (1986a). Flight Simulation Motion-base Drive Algorithms. Part 2: Selecting the system parameters. Technical Report 307, UTIAS, Toronto, Canada.

Reid, L. D. and Nahon, M. A. (1986b). Flight Simulation Motion-base Drive Algorithms. Part 3: Pilot evaluations. Technical Report 319, UTIAS, Toronto, Canada.

Repa, B. S., Leucht, P. M., and Wierwille, W. W. (1982). The Effect of Simulator Motion on Driver Performance. *SAE Technical Paper 820307*.

Reymond, G., Droulez, J., and Kemeny, A. (2002). Visuovestibular Perception of Self-motion Modeled as a Dynamic Optimization Process. *Biological Cybernetics*, 87:301–314.

Reymond, G., Kemeny, A., Droulez, J., and Berthoz, A. (2001). Role of Lateral Acceleration in Curve Driving: Driver Model and Experiments on Real Vehicle and a Driving Simulator. *Human Factors*, 43(3):483–495.

Riecke, B. E. (2010). *Virtual Reality*, chapter 8: Compelling Self-Motion Through Virtual Environments Without Actual Self-Motion Using Self-Motion Illusions ("Vection") to Improve VR User Experience. InTech. ISBN: 978-953-307-518-1.

Ritchie, M. L., McCoy, W. K., and Welde, W. L. (1968). A Study of the Relation Between Forward Velocity and Lateral Acceleration in Curves During Normal Driving. *Human Factors*, 10(3):255–258.

Roark, M. and Junker, A. (1978). The Effects of Closed Loop Tracking on a Subjective Tilt Threshold in the Roll Axis. In *Proceedings of the 14th Annual Conference on Manual Control, Los Angeles and Moffet Field, CA, USA, April 25-27*, 443 – 450.

Rodchenko, V., Boris, S. Y., and White, A. D. (2000). In-Flight Estimation of Pilot's Acceleration Sensitivity Thresholds. In *Proceedings of the AIAA Modeling and Simulation Technologies Conference, Denver, CO, USA, August 14-17*, AIAA 2000-4292.

Royal Aeronautical Society (2005). Aeroplane Flight Simulator Evaluation Handbook – International Standards for the Qualification of Aeroplane Flight Simulators. Technical report, Royal Aeronautical Society (RAeS), London, UK. Third edition.

Samji, A. and Reid, L. D. (1992). The Detection of Low-Amplitude Yawing Motion Transients in a Flight Simulator. *IEEE Transactions on Systems, Man, and Cybernetics*, 22(2):300–306.

Dos Santos Buinhas, L., Correia Grácio, B. J., Valente Pais, A. R., Mulder, M., and Van Paassen, M. M. (2013). Modeling Coherence Zones in Flight Simulation During Yaw Motion. In *Proceedings of the AIAA Modeling and Simulation Technologies Conference and Exhibit, Boston, MA, USA, August 19-22*, submitted.

Schroeder, J. A. (1996). Evaluation of Simulation Motion Fidelity Criteria in the Vertical and Directional Axes. *Journal of the American Helicopter Society*, 41(2):44–57.

Schroeder, J. A. (1999). Helicopter Flight Simulation Motion Platform Requirements. Technical Report NASA/TP-1999-208766, NASA, Moffet Field, CA, USA.

Schwarz, C., Gates, T., and Papelis, Y. (2003). Motion Characteristics of the National Advanced Driving Simulator. In *Proceedings of the Driving Simulation Conference North America, Dearborn, MI, USA, October 8-10*.

Shirachi, D. K. and Shirley, R. S. (1977). The Effect of a Visual/Motion Display Mismatch in a Single Axis Compensatory Tracking Task. Contractor Report NAS2-7806 NASA CR-2921, NASA, Ames Research Center, Moffet Field, CA, USA.

Shirley, R. S. and Young, L. R. (1968). Motion Cues in Man-Vehicle Control – Effects of Roll-Motion Cues on Human Operator's Behavior in Compensatory Systems with Disturbance Inputs. *IEEE Transactions on Man-Machine Systems*, 9(4):121–128.

Siegler, I., Reymond, G., Kemeny, A., and Berthoz, A. (2001). Sensorimotor Integration in Driving Simulator: Contribuitions of Motion Cueing in Elementary Driving Tasks. In *Driving Simulation Conference, Sophia Antipolis, Nice, France, September 5-7*.

Sinacori, J. B. (1977). The Determination of Some Requirements for a Helicopter Research Simulation Facility. Contractor Report NAS2-9421 NASA-CR-152066, NASA, Ames Research Center, Moffet Field, CA, USA.

Soyka, F., Robuffo Giordano, P., Barnett-Cowan, M., and Bülthoff, H. H. (2012). Modeling Direction Discrimination Thresholds for Yaw Rotations Around an Earth-Vertical Axis for Arbitrary Motion Profiles. *Experimental Brain Research*, 220(1):89–99.

Soyka, F., Robuffo Giordano, P., Beykirch, K., and Bülthoff, H. H. (2011). Predicting Direction Detection Thresholds for Arbitrary Translational Acceleration profiles in the Horizontal Plane. *Experimental Brain Research*, 209(1):95–107.

Soyka, F., Teufel, H. J., Beykirch, K. A., Giordano, P. R., Butler, J. S., Nieuwenhuizen, F. M., and Bülthoff, H. H. (2009). Does Jerk Have to Be Considered in Linear Motion Simulation? In *Proceedings of the AIAA Modeling and Simulation Technologies Conference, Chicago, IL, USA, August 10-13*, AIAA 2009-6245.

Stapleford, R. L., Peters, R. A., and Alex, F. R. (1969). Experiments and a Model for Pilot Dynamics with Visual and Motion Inputs. Technical Report NASA CR-1325, Systems Technology, Inc., Hawthorne, CA, USA.

Van der Steen, F. A. M. (1998). *Self-Motion Perception*. Ph.D. thesis, Delft University of Technology, Faculty of Aerospace Engineering.

Stewart, J. D. (1971). Human Perception of Angular Acceleration and Implications in Motion Simulation. *Journal of Aircraft*, 8(4):248–253.

StRoadDesign (2008). Retrieved on July 10, 2008, from http://www.stsoftware.nl/StRoadDesign.html.

Stroosma, O., Van Paassen, M. M., and Mulder, M. (2003). Using the Simona Research Simulator for Human-Machine Interaction Research. In *Proceedings of the AIAA Modeling and Simulation Technologies Conference and Exhibit, Austin, TX, USA, August 11-14*, AIAA 2003-5525.

Stroosma, O., Van Paassen, M. M., Mulder, M., and Postema, F. N. (2007). Measuring Time Delays in Simulator Displays. In *Proceedings of the AIAA Modeling and Simulation Technologies Conference and Exhibit, Hilton Head, SC, USA, August 20-23*, AIAA 2007-6562.

Sweet, B. T. and Kaiser, M. K. (2005). Modeling of Perception and Control Attitude with Perspective Displays. In *Proceedings of the AIAA Modeling and Simulation Technologies Conference and Exhibit, San Francisco, CA, USA, August 15-18*, AIAA 2005-5891.

Telban, R. J. and Cardullo, F. M. (2001). An Integrated Model of Human Motion Perception with Visual-Vestibular Interaction. In *Proceedings of the AIAA Modeling and Simulation Technologies Conference and Exhibit, Montreal, Canada, August 6-9*, AIAA 2001-4249.

Toyota Develops World-class Driving Simulator. Real-as-possible Environment to Aid Development of Active Safety Technology (2007). Retrieved on September 15, 2008, from http://www.toyota.co.jp/en/news/07/1126_1.html.

Van der Vaart, J. C. (1992). *Modelling of Perception and Action in Compensatory Manual Control Tasks*. Ph.D. thesis, Delft University of Technology, Faculty of Aerospace Engineering.

Valente Pais, A. R., Mulder, M., Van Paassen, M. M., Wentink, M., and Groen, E. L. (2006). Modeling Human Perceptual Thresholds in Self-Motion Perception. In *Proceedings of the AIAA Modeling and Simulation Technologies Conference and Exhibit, Keystone, CO, USA, August 21-24*, AIAA 2006-6626.

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

Valente Pais, A. R., Van Paassen, M. M., Mulder, M., and Wentink, M. (2010b). Perception of Combined Visual and Inertial Low-Frequency Yaw Motion. In *Proceedings of the AIAA Modeling and Simulation Technologies Conference and Exhibit, Toronto, ON, Canada, August 2-5*, AIAA 2010-8093.

Valente Pais, A. R., Pool, D. M., De Vroome, A. M., Van Paassen, M. M., and Mulder, M. (2012). Pitch Motion Perception Thresholds During Passive and Active Tasks. *Journal of Guidance, Control, and Dynamics*, 35(3):904–918.

VTI's simulator facilities, Sim IV (2012). Retrieved on November 5, 2012, from http://www.vti.se/en/research-areas/vehicle-technology/vtis-driving-simulators/.

Wachtel, J. (1995). Brief History of Driving Simulators. *TR News*, (179):26–27, 45.

Walsh, E. G. (1961). Role of the Vestibular Apparatus in the Perception of Motion on a Parallel Swing. *Journal of Physiology*, 155(3):506–513.

Wentink, M., Bles, W., Hosman, R. J. A. W., and Mayrhofer, M. (2005). Design & Evaluation of Spherical Washout Algorithm for Desdemona Simulator. In *Proceedings of the AIAA Modeling and Simulation Technologies Conference and Exhibit, San Francisco, CA, USA, August 15-18*, AIAA 2005-6501.

Wentink, M., Correia Grácio, B. J., and Bles, W. (2009). Frequency Dependence of Allowable Differences in Visual and Vestibular Motion Cues in a Simulator. In *Proceedings of the AIAA Modeling and Simulation Technologies Conference and Exhibit, Chicago, IL, USA, August 10-13*, AIAA 2009-6248.

Wertheim, A. H. (1994). Motion Perception During Self-Motion: The Direct Versus Inferential Controversy Revisited. *Behavioral and Brain Sciences*, 17(2):293–355.

Wertheim, A. H. and Bles, W. (1984). A Re-Evaluation of Cancellation Theory: Visual, Vestibular and Oculomotor Contribuitions to Perceived Object Motion. Technical Report IZF 1984-8, Institute for Perception, National Defence Research Organization TNO, Soesterberg, The Netherlands.

Wertheim, A. H., Ooms, J., de Regt, G. P., and Wientjes, C. J. E. (1992). Incidence and Severeness of Seasickness: Validation of a Rating Scale. Technical Report IZF 1992 A-41, TNO Institute for Perception, Soesterberg, The Netherlands.

Westheimer, G. (1968). *Handbook of Perception and Human Performance*, volume I: Sensory Processes and Perception, chapter 4: The Eye as an Optical Instrument. A Wiley-Interscience Publication, 111 River Street, Hoboken, NJ 07030-5774, USA. ISBN 0-471-88544-4.

White, A. D. and Rodchenko, V. V. (1999). Motion Fidelity Criteria Based on Human Perception and Performance. In *Proceedings of the AIAA Modeling and Simulation Technologies Conference and Exhibit, Portland, OR, USA, August 9-11*, AIAA 1999-4330, 485–493.

De Winkel, K. N., Correia Grácio, B. J., Groen, E. L., and Werkhoven, P. (2010a). Visual-Inertial Coherence Zone in the Perception of Heading. In *Proceedings of the AIAA Modeling and Simulation Technologies Conference and Exhibit, Toronto, ON, Canada, August 2-5*, AIAA 2010-7916.

De Winkel, K. N., Weesie, J., Werkhoven, P., and Groen, E. L. (2010b). Integration of Visual and Inertial Cues in Perceived Heading of Self-Motion. *Journal of Vision*, 10(12):1–10.

Van Winsum, W. and Godthelp, H. (1996). Speed Choice and Steering Behavior in Curve Drving. *Human Factors*, 38(3):434–441.

Wong, S. and Frost, B. (1981). The Effect of Visual-Vestibular Conflict on the Latency of Steady-State Visually Induced Subjective Rotation. *Perception & Psychophysics*, 30:228–236.

Young, L. R., Dichgans, J., Murphy, R., and Brandt, T. (1973). Interaction of Optokinetic and Vestibular Stimuli in Motion Perception. *Acta Oto-Laryngologica*, 76:24–31.

Zaal, P. M. T. (2011). *Pilot Control Behavior Discrepancies Between Real and Simulated Flight Caused by Limited Motion Stimuli*. Ph.D. thesis, Delft University of Technology, Faculty of Aerospace Engineering.

Zaal, P. M. T., Pool, D. M., Chu, Q. P., Van Paassen, M. M., Mulder, M., and Mulder, J. A. (2009). Modeling Human Multimodal Perception and Control Using Genetic Maximum Likelihood Estimation. *Journal of Guidance, Control, and Dynamics*, 32(4):1089–1099.

Zacharias, G. L. and Young, L. R. (1981). Influence of Combined Visual and Vestibular Cues on Human Perception and Control of Horizontal Rotation. *Experimental Brain Research*, 41:157–171.

Zaichik, L. E., Rodchenko, V., Rufov, I. V., Yashin, Y. P., and White, A. D. (1999). Acceleration Perception. In *Proceedings of the AIAA Modeling and Simulation Technologies Conference and Exhibit, Portland, OR, USA, August 9-11*, AIAA 1999-4334, 512–520.

Zeppenfeldt, P. (1991). *Bepaling van Drempelwaarden voor het Waarnemen van Verschillen Tussen Visuele en Vestibulaire Stimulatie Tijdens Eigenbeweging (in Dutch)*. Master of Science Thesis, Faculty of Aerospace Engineering, Delft University of Technology. Unpublished.

Samenvatting

Perceptiecoherentiezones in Voertuigsimulatie

Ana Rita Valente Pais

Bewegingssimulatoren worden wereldwijd gebruikt voor piloot- en rijopleidingen. Ze bieden een relatief goedkoop, veilig, toegankelijk en milieuvriendelijk alternatief voor training in het echte voertuig. Simulatoren zijn alleen effectief indien de aangeboden stimuli de personen in staat stellen om verschillende situaties te herkennen en er adequaat op te reageren. Het bestuderen van en toezicht houden op de waarheidsgetrouwheid van de aangeboden stimuli is daarom van essentieel belang om het huidige veiligheidsniveau te waarborgen, en om de toekomstige veiligheid van het weg- en luchtverkeer te vergroten.

Een van de belangrijkste aspecten van de waarheidsgetrouwheid betreft het bewegingssysteem van de simulator. De beperkte bewegingsvrijheid van simulatoren belemmert de één-op-één weergave van de beweging van het voertuig. Bewegingsfilters worden gebruikt om de simulator binnen de mechanische grenzen te houden terwijl inertiële bewegingsstimuli worden aangeboden. Deze filters, ookwel bewegingsalgorithmes genaamd, veroorzaken een afname van de amplitude en een vervorming van de fase van de inertiële bewegingsterugkoppeling vergeleken met de visuele beweging gepresenteerd via het buitenzichtsysteem. Hierdoor ontstaat een discrepantie tussen de inertiële en de visuele stimuli. Indien deze discrepantie, of verschil, groot genoeg is om door de persoon in de simulator te worden waargenomen, kan het zijn dat de waarheidsgetrouwheid van de simulator wordt aangetast.

Kennis van hoe deze verschillen tussen de visuele en de inertiële stimuli de mens in de simulator beïnvloeden is nodig voor de ontwikkeling van nieuwe bewegingsalgorithmes,

simulatorontwerpen en de optimalisatie van de bewegingsfilters in bestaande bewegingsplatformen. Bovendien zal een beter begrip van de menselijke zelfbewegingprocessen, met name van de mechanismen die bij de perceptie van gecombineerde visuele en inertiële beweging een rol spelen, leiden tot betere en completere criteria om de kwaliteit van de waarheidsgetrouwheid van de inertiële bewegingssimulatie te beoordelen.

Het belangrijkste doel van dit proefschrift is een bijdrage te leveren aan de verbetering van de waarheidsgetrouwheid van de inertiële beweging in voertuigsimulaties. Vanuit dit doel zijn drie specifiekere doelstellingen gedefinieerd.

Doelstelling 1

> De eerste doelstelling is om het ontwerp- en afregelproces van bewegingsalgorithmes te bestuderen, en om de gevolgen op de perceptie, stuurgedrag en prestaties van personen te evalueren ten gevolge van verschillende aangeboden bewegingsmogelijkheden, en dus van verschillende inertiële stimulaties.

Doelstelling 2

> De tweede doelstelling is om in detail te focussen op een specifiek onderdeel van de waarheidsgetrouwheid van simulaties: de perceptie van gecombineerde visuele en inertiële stimuli. Hier is het doel om onze kennis van de visueel-inertiële bewegingswaarneming in voertuigsimulatiescenario's uit te breiden.

Doelstelling 3

> Het derde doel is om te onderzoeken of de kennis van gecombineerde perceptie van inertiële en visuele stimuli gebruikt kan worden om er richtlijnen voor perceptiegebaseerde bewegingsaanbieding van af te leiden die niet afhankelijk zijn van een bepaalde vorm van het bewegingsalgorithme of bewegingsplatform.

Om de eerste doelstelling te onderzoeken is er een nieuw bewegingsalgorithme ontworpen en geëvalueerd ten behoeve van het rijden van bochtige wegen in een stedelijke omgeving, in een grote, op een centrifuge gebaseerde bewegingssimulator. Drie bewegingsmogelijkheden werden getest: een klassiek filter zoals gebruikelijk voor Stewartplatformen, een algoritme met alleen wegtrillingen, en een nieuwe oplossing.

Om het effect van inertiële beweging op de prestatie en stuurgedrag van personen te evalueren zouden zeer specifieke prestatiedoelen moeten worden gesteld en het type stuurtaak zou zeer goed moeten worden gedefinieerd, waardoor er weinig ruimte voor personen zou overblijven om hun eigen stuurstrategie te kiezen. Om een taak te simuleren die zo veel mogelijk leek op echt stadrijden werden er geen beperkingen opgelegd aan de prestatie en

stuurgedrag. Als gevolg daarvan was het erg moeilijk om de drie bewegingsoplossingen te vergelijken op basis van prestatie en stuurgedrag.

Een analyse van de perceptie op basis van de antwoorden van personen op vragenlijsten toonde aan dat bepaalde inertiële bewegingen belangrijk waren voor ofwel de aanvaarding, ofwel de afwijzing van de betreffende bewegingsoplossing, maar het was niet mogelijk om te meten hoe verschillende stimuli bijdroegen aan de algemene acceptatie.

Het proces van het ontwerpen, afregelen en evalueren van een bewegingsalgorithme voor een onconventionele simulator, waarvoor er weinig richtlijnen en geteste oplossingen beschikbaar zijn, toonde aan dat er meer platformonafhankelijke, op de mens gebaseerde perceptiemetrieken en -richtlijnen nodig zijn.

De aanpak van de tweede doelstelling bestond uit het meten en evalueren van de perceptie van gecombineerde visuele en inertiële stimuli met behulp van het concept van de coherentiezone. Aangezien het gebruik van bewegingsfilters zowel een afname van de amplitude als een vervorming van de fase van de inertiële stimulus in vergelijking tot de visuele stimulus tot gevolg heeft, zijn er twee type coherentiezones gedefinieerd: de amplitudecohereentiezone en de fasecoherentiezone.

Een amplitudecoherentiezone geeft het bereik van de amplitude van inertiële bewegingen weer die, hoewel ze niet perfect overeenkomen met de visuele beweging, nog steeds door de proefpersoon als coherent worden waargenomen. Op dezelfde wijze geeft een fasecoherentiezone de waarden van het faseverschil tussen de inertiële en visuele stimuli aan die nog als coherent worden beschouwd.

Amplitudecoherentiezones kunnen worden gedefinieerd door een boven- en ondergrens, of op alternatieve wijze, door een gemiddelde coherentie (Engels: point of mean coherence, PMC) en de breedte van de zone (Engels: coherence zone width, CZW). De bovengrens geeft de grootste inertiële bewegingsamplitude aan die nog als coherent wordt waargenomen in combinatie met een bepaalde visuele stimulusamplitude. De ondergrens is de laagste inertiële beweging die nog als coherent wordt waargenomen. De PMC is de waarde van de inertiële amplitude die halverwege de boven- en ondergrens ligt, en de CZW is de afstand tussen boven- en ondergrens.

Voor fasecoherentiezones werd slechts één grens gemeten. Omdat algemene hoogdoorlaatfilters een fasevoorsprong van de inertiële beweging ten opzichte van de visuele beweging introduceren zijn alleen grenswaarden voor vóórlopende inertiële beweging gemeten. Deze gemeten grens voor het faseverschil geeft de grootst mogelijke waarde van de fasevoorsprong die op de inertiële beweging kan worden toegepast en nog steeds als coherent wordt beschouwd. Laagdoorlaatfilters zijn niet beschouwd.

Diverse experimenten werden uitgevoerd die waren gericht op het onderzoeken van ofwel de amplitude- ofwel de fasecoherentiezones. Amplitudecoherentiezones werden gemeten voor gierrotaties en laterale translaties. Gieramplitudecoherentiezones werden gemeten

in twee simulatoren: de Simona simulator en de Desdemona simulator. Fasecoherentiezones werden gemeten voor gier- en stamprotaties. Ook werden de effecten gemeten van de stimulusamplitude en -frequentie op zowel de amplitude- en fasecoherentiezones.

Resultaten tonen aan dat gier- en laterale amplitudecoherentiezones werden beïnvloed door zowel de amplitude als de frequentie van de visuele stimulus. Over het algemeen werden de PMC en CZW groter als de amplitude van de visuele stimulus groter werd. Voor grotere amplitudes van de visuele stimulus lijkt het dat de bovengrens dichter bij, of zelfs lager dan, de één-op-één lijn kwam te liggen en dit werd gekenmerkt door het naar beneden afbuigen van de coherentiezones, waarbij de PMV waarden kleiner werden dan de één-op-één lijn.

Het effect van de frequentie was ook vergelijkbaar voor gier- en laterale amplitudecoherentiezones. Voor hogere frequenties werd zowel de boven- als de ondergrens kleiner. Het effect van de frequentie is gerelateerd aan de perceptie van "bewegingssterkte". Geconcludeerd werd dat de waargenomen bewegingssterkte het beste kan worden uitgelegd als een gewogen combinatie van waargenomen hoekversnelling en hoeksnelheid in het geval van gierbewegingen, en van waargenomen lineaire jerk en lineaire versnelling in het geval van laterale beweging.

Uit de vergelijking van fasecoherentiezones voor gier- en stampbewegingen volgde dat deze gelijkwaardig waren en dat ze niet werden beïnvloed door de amplitude en de frequentie van de stimuli. Proefpersonen gedroegen zich als faseverschildetectoren en niet als tijdsvertragingsverschildetectoren, en de gemeten grenswaarde voor het faseverschil bedroeg ongeveer 19 graden, waarbij de inertiële beweging voorliep op de visuele visuele beweging.

In de vervolgstap, waarbij werd geprobeerd realistischere scenario's te simuleren waarin proefpersonen vlieg- of rijtaken moesten uitvoeren, werd het effect van verminderde aandacht van proefpersonen voor de perceptie van de amplitudecoherentiezones onderzocht. Coherentiezones werden gemeten terwijl proefpersonen een handmatige stuurtaak moesten uitvoeren waarin ze het voertuig binnen bepaalde bepaalde grenzen moesten houden. De handmatige stuurtaak werd uitgevoerd aan de hand van visuele stimuli alléén, en in een andere vrijheidsgraad dan de coherentiezonemetingen.

De actieve taak had geen enkel effect op de waargenomen coherentie van de visuele en inertiële stimuli. Een mogelijke verklaring is dat de perceptuele taak en de actieve stuurtaak niet gelijktijdig door de proefpersonen werden uitgevoerd, maar na elkaar, dat wil zeggen, verschoven in de tijd. Mogelijk was er maar een korte periode nodig om een mening te vormen over de waargenomen coherentie, waardoor de rest van de tijd besteed kon worden om zich te concentreren op de actieve taak.

Deze resultaten ondersteunen de veronderstelling dat betrefende perceptuele waarheidsgetrouwheid simulatoronderzoeken kunnen worden uitgevoerd met passieve taken en dat deze resultaten mogen worden gegeneraliseerd en worden gebruikt voor actieve, piloot-in-

control situaties. Dit is mogelijk niet het geval voor de waarheidsgetrouwheid gebaseerd op prestatie of stuurgedrag.

Als laatste stap in de studie naar coherentiezones en als voorbereiding op de derde doelstelling werden laterale amplitudecohereentiezones bepaald in drie hexapod simulatoren: de Visual Motion Simulator (VMS), de Generic Flight Deck (GFD) en de Integration Flight Deck (IFD) simulator bij het NASA Langley Research Center in Hampton, Virginia, USA. Deze drie simulatoren hadden bewegings- en visuele systemen met verschillende eigenschappen.

De verschillen tussen de simulatorconfiguraties werden teruggevonden in de gemeten bovengrens van de coherentiezone. De ondergrens werd niet significant beïnvloed door de simulatorconfiguratie noch door de amplitude of frequentie van de visuele stimuli.

Ondanks de verschillen in de bovengrenswaarden werden in alle drie de simulatoren dezelfde trends waargenomen voor de effecten van frequentie en amplitude. Uit een vergelijking tussen giercoherentiezones gemeten in de Simona simulator en de Desdemona simulator volgde hetzelfde.

De bevinding dat de precieze waarden van de coherentiezone kunnen variëren tussen simulatoren, maar dezelfde algemene trends behouden wat de visuele stimulus amplitude en frequentie betreft, heeft tot twee conclusies geleid. In de eerste plaats is het, ondanks fundamentele verschillen tussen de simulatoren zoals de verschillen tussen de Simona en Desdemona simulatoren (gierbeweging) en de verschillen tussen de VMS, de GFD en de IFD (laterale beweging), nog acceptabel om coherentiezones te meten in verschillende apparaten en deze metingen te vergelijken.

In de tweede plaats is het niet alleen acceptabel, maar zelfs wenselijk om het op deze manier te doen. Met behulp van coherentiezones kunnen verschillende simulaties worden beoordeeld op basis van hoe goed de perceptie van de gecombineerde stimuli past binnen de coherentiezone. Dit houdt in dat simulatoren niet zouden moeten worden beoordeeld en vergeleken op basis van hun mechanische prestaties alléén, maar op basis van de gecombineerde weergave van visuele en inertiële stimuli. Coherentiezones bieden een metriek om de geschiktheid van deze gecombineerde presentatie van stimuli te kwantificeren.

Om de derde doelstelling te bereiken werden de gemeten coherentiezones gebruikt om criteria voor de bewegingswaarheidsgetrouwheid van af te leiden, en om een op coherentiezones gebaseerde methode voor te stellen om de bewegingswaarheidsgetrouwheid te beoordelen. De gemeten ondergrenzen zijn omgezet in minimumeisen voor de parameters in bewegingsfilters. De grenzen voor het faseverschil zijn gebruikt om criteria te verkrijgen voor de maximaal toegestane fasevervorming. Deze afgeleide criteria zijn vervolgens vergeleken met bestaande criteria in de literatuur.

De criteria gebaseerd op de coherentiezones zijn weergegeven in een aangepaste Sinacorigrafiek. In een Sinacorigrafiek worden criteria voor de waarheidsgetrouwheid van de

beweging weergegeven door de maximaal acceptabel geachte vervorming van de amplitude en fase van de beweging bij een frequentie van 1 rad/s. Aangezien de coherentiezones bij verschillende frequenties en amplitudes werden gemeten, was het mogelijk om de grafiek uit te breiden met verschillende frequenties en amplitudes.

Het gebruik van coherentiezones als een methode om de bewegingswaarheidsgetrouwheid te beoordelen berust op de veronderstelling dat de waargenomen coherentie tussen visuele en ineriële stimuli een indicatie is voor "goede" beweging. Net zoals bij andere beoordelingcriteria, zoals zijn voorgesteld door Sinacori, Schroeder, en Advani en Hosman, bieden coherentiezones slechts een maat voor de doeltreffendheid van de aangeboden inertiële beweging in relatie tot een gewenste ideale beweging. Deze gewenste beweging komt overeen met de gesimuleerde voertuigbeweging en is in dit proefschrift verondersteld gelijk te zijn aan de visuele beweging. Al bovenstaande beoordelingsmethoden vereisen een goede waarheidsgetrouwheid van het voertuigmodel en zijn dus, op zich, niet voldoende om een hoge waarheidsgetrouwheid van de inertiële beweging te garanderen. Desalnietemin bieden ze de broodnodige ondersteuning bij het ontwerp en het afregelen van bewegingsalgorithmes.

De voorgestelde beoordelingsmethode op basis van coherentiezones biedt drie belangrijke aanvullingen op de reeds beschikbare criteria. Ten eerste geeft de methode niet alleen criteria voor gewenste bewegingstimuli, maar biedt ook een systematische, objectieve, en op menselijke waarneming gebaseerde methode om de grenzen van de criteria te meten.

Ten tweede voegt de op coherentiezone gebaseerde methode een derde en vierde dimensie aan de Sinacorigrafiek toe: frequentie en amplitude. Hierdoor hangen de coherentiezonecriteria niet af van de vorm van een specifiek bewegingsfilter.

Ten derde kan de coherentiezonemethode, door het bieden van een meetmethode waarvan de verschillende frequenties en amplitudes kunnen worden gekozen, simulatorgebaseerde en taakspecifieke criteria bieden. Coherentiezones als metriek, dat wil zeggen als maat voor de waargenomen coherentie van de aangeboden inertiële feedback, is echter platform- en taakonafhankelijk.

Acknowledgments

Doing a PhD involves as much scientific research as it does personal development. When searching for the boundaries of human knowledge, your own boundaries are tested and you learn a great deal about yourself. It is then not surprising that in the end, I am thankful to those who helped me professionally but also in my personal life.

At the Control and Simulation division I am thankful to my promotor Professor Max Mulder and my co-promotor René van Paassen. They have provided invaluable feedback on my experimental work but also on the process, at times frustrating, of telling the story. René, thank you for the patience to listen to all my questions in all those five minute installments ("René, do you have five minutes?").

A big thank you to Mark Wentink from Desdemona B.V., for taking me under his wing and always trying to show me the bigger picture. He was there when I needed it.

At TNO in Soesterberg, I am thankful to all the people who took the time to share with me their knowledge, thoughts and ideas, among them Wim Bles, Jelte Bos, Suzanne Nooij, Ksander de Winkel and Eric Groen. A special thank you to Eric Groen, whose head full of ideas, great knowledge of human motion perception and enthusiasm for flight simulation, first got me interested in this field of research.

To all my colleagues at the Control and Simulation, past and present, thank you for all the camaraderie, all the Friday night drinks and dinners, all the awesome conference trips, all the bad jokes and dry humor, thank you for all the heated discussions on current events, history, politics, economics, religion, philosophy, you name it! It is truly a great privilege to have such "tertulias" during coffee breaks.

For being an awesome colleague, a great friend and the best assistant coach I have ever had, a big thank you to Jan Comans.

I am also thankful to my colleague and countryman Bruno Correia Grácio. We have worked together on several experiments and he has shown to be a true team player. His unshaken good spirits have always kept the morale high in the more stressful days.

My office mate, Daan Pool, has been a true friend throughout these years. I am grateful not only for his friendship but also for his encouragement in the tougher days, for putting up with the Portuguese temperatures in the office, for sharing the pains and laughs of doing a PhD, for keeping me up to date on the old and new musical hits and most of all, for teaching me patience and kindness.

I am grateful to the master students I had the opportunity to work with during my PhD: Sebastian Malack, Peter Jonik, Aniek de Vroome, Niek Beckers and Luísa dos Santos Buinhas. It was a true pleasure to teach and learn from them.

In the course of my PhD I was lucky enough to be able to interact with many people related to my field of research. A few, in either a more direct or indirect way, have contributed to this thesis.

Ruud Hosman, from AMS Consulting, and Sunjoo Advani, from IDT, have provided me with many references and helped me understand the industry and aviation authorities point of view on simulation fidelity. Their input was a valuable contribution to Chapter 9.

Peter Grant, from UTIAS, has kindly provided me with experimental data from one of his own experiments for comparison with data from the experiment described in Chapter 5.

Professor Frank Cardullo from the Binghamton University, SUNY, has accompanied my path since I was a master student and has always been available for explanations, discussions, and the exchange of ideas.

Lon Kelly, from the NASA Langley Research Center, has greatly contributed to the work described in Chapter 7. It is due to his effort and perseverance that debugging C++ code through email did not end in disaster but instead, in a successful experiment in three of the NASA simulators. I am thankful for all his time preparing and running the experiments and for going out of his way to make me feel so welcome at Langley during my visit. I am also thankful to Jake Houck and Victoria Chung for supporting this cooperation project, and everyone involved in the implementation and running of the experiments.

At the Control and Simulation division, at TNO and Desdemona, and at the NASA Langley Research Center, I am thankful to the many volunteers who have so kindly offered their time and their patience to participate in my, at times not so exciting, experiments.

Outside the work environment, I was fortunate to count with many friends that have given me strength in the not so good times and rejoiced with me in the good ones.

I am grateful to my good old friends André, Bruno, Cristina, Joana e Rui, who are far away but always feel close by; and my friends from the Phoenix and Licor Café teams, who have became my basketball family.

DAS, my basketball club, has been a constant presence during my PhD years and it was the one place where I could briefly forget about work. I am thankful to the many people there who have helped me grow as a player, as a coach and as a person. Among them, my team mates and coaches of the past years, who have given me the chance to be part of something bigger than me; Bert van Blitterswijk, who gave me my first Dutch bike and made me feel welcome to the club; and my girls, the girls I coached for the past years. They taught me an incredible lot about teaching, motivation, discipline, group dynamics, basketball, patience and teenagers, not necessarily in this order.

DAS, mijn basketbalclub, is een aangename constante geweest tijdens mijn PhD-jaren en het was de enige plaats waar ik even niet aan mijn werk hoefde te denken. Ik ben de vele mensen bij DAS dankbaar die mij als speler, coach en persoon hebben helpen groeien. Zoals mijn teamgenoten en coaches van de afgelopen jaren, die mij de kans hebben gegeven om deel uit te maken van iets groters dan ik; Bert van Blitterswijk, die mij mijn eerste Nederlandse fiets gaf en die er erg aan heeft bijgedragen dat ik mij thuisvoelde bij de club; en mijn meisjes, de meisjes die ik de afgelopen jaren heb gecoached. Ze hebben mij ongelooflijk veel geleerd over lesgeven, motivatie, discipline, groepsdynamica, basketbal, geduld en tieners, en niet per se in deze volgorde.

I am also grateful for the *liefde* and *gezelligheid* of my Dutch family, and especially to Ineke, Herman, Karen, Robbert, Enora, Lauren, Mieke, Tineke and Kees. They have welcomed me in their homes, spoke English while patiently waiting for me to learn Dutch, taught me about Dutch culture and bravely accepted my Portuguese cooking, and most of all, as only a family can do, have provided me with unconditional practical and moral support.

Ik ben ook dankbaar voor de liefde en gezelligheid van mijn Nederlandse familie, vooral Ineke, Herman, Karen, Robbert, Enora, Lauren, Mieke, Tineke en Kees. Zij hebben mij verwelkomd in hun huizen, ze hebben Engels tegen mij gesproken terwijl ze met veel geduld wachtten tot ik Nederlands leerde, ze hebben mij de Nederlandse cultuur bijgebracht en ze waren zelfs dapper genoeg om mijn Portugese kookkunsten te proberen. Bovenal, zoals alleen een gezin kan doen, hebben ze mij onvoorwaardelijke praktische en morele steun gegeven.

My family in Portugal, my parents Inês and António, my brother Hugo and his wife Katy, and my aunt Irene have followed from close by the ups and downs of the PhD years. I am grateful for their love and their support. I am grateful to my aunt Irene for being like a second mother to me. She has taught me by example, to always stand up for what I believe in. I am grateful to my father for always providing, sometimes resorting to some creativity, any means necessary to perform in school. By doing so he showed me that in school, but also in life, one should never search for an excuse to be less than we can possibly be. I am grateful to my mother for being an inspiration to me. Thinking of how she stubbornly

decided to attend and complete six years of highschool, while holding a full time job and raising two small children, has forever raised the standard by which I measure my personal success.

A minha família em Portugal, os meus pais Inês e António, o meu irmão Hugo e a sua mulher Katy, e a minha tia Irene têm seguido de perto os altos e baixos dos meus anos de doutoramento. Estou grata pelo amor e apoio que me têm dado. Estou grata à minha tia Irene por ser como uma segunda mãe para mim. Dando o exemplo, ela ensinou-me a defender sempre aquilo em que se acredita. Estou grata ao meu pai por ter sempre proporcionado, às vezes recorrendo a um pouco de criatividade, todos os meios necessários para eu ter um bom desempenho escolar. Ao fazê-lo, ele mostrou-me que na escola, mas também na vida, nunca devemos procurar desculpas para sermos menos do que podemos. Estou grata à minha mãe por ser uma inspiração para mim. Pensando em como ela teimosamente decidiu frequentar e terminar os seis anos de liceu, enquanto trabalhava a tempo inteiro e criava dois filhos pequenos, elevou para sempre o padrão pelo qual eu meço o meu sucesso pessoal.

Last, I am grateful and forever indebted to the person who has helped me the most through this endeavor, Herman. I am very happy you are in my life.

Curriculum Vitae

Ana Rita Valente Pais was born on February 19, 1982 in Vila Franca de Xira, Portugal. From 1994 to 1997 she attended the Escola Secundária de Alves Redol in Vila Franca de Xira and from 1997 to 2000 the Escola Secundária Quinta do Marquês in Oeiras, Portugal.

In 2005 she obtained her Aerospace Engineering degree at Instituto Superior Técnico, Universidade Técnica de Lisboa, in Lisbon, Portugal. In the academic year of 2004/2005, under the European exchange program Erasmus, she followed courses at the Faculty of Aerospace Engineering of the Delft University of Technology. The next year, she enrolled as a master student at the Control and Simulation Division of that Faculty. During that year she did an internship at TNO, Soesterberg, The Netherlands. Here, she worked on extending a human self-motion perception model developed at that institute with models of motion perception thresholds. During her M.Sc. thesis, as a continuation of the internship work, she investigated how motion perception thresholds are affected by the presence of other motion cues in a simulator. In 2006 she obtained her M.Sc. degree (with Honours).

Also in 2006, she started working as a researcher at the Control and Simulation Division of the Faculty of Aerospace Engineering at the Delft University of Technology, to develop a urban car driving motion cueing algorithm for the Desdemona simulator, at TNO, in Soesterberg. In 2007 she was granted a Toptalent grant from The Netherlands Organisation for Scientific Research (NWO) to start a Ph.D. on motion perception in vehicle simulation. In 2011 she was awarded a Zonta International Amelia Earhart Fellowship that allowed her to visit the NASA Langley Research Center with whom she was cooperating in a motion simulation study.

Publications

Dos Santos Buinhas, L., Correia Grácio, B. J., Valente Pais, A. R., Mulder, M., and Van Paassen, M. M. (2013). Modeling Coherence Zones in Flight Simulation During Yaw Motion. In *Proceedings of the AIAA Modeling and Simulation Technologies Conference and Exhibit, Boston, MA, USA, August 19-22*, accepted for publication.

Correia Grácio, B. J., Valente Pais, A. R., Van Paassen, M. M., Mulder, M., Kelly, L. C., and Houck, J. A. (2013). Optimal and Coherence Zone Comparison Within and Between Flight Simulators. *Journal of Aircraft*, 50(2):493–507.

Beckers, N. W. M., Pool, D. M., Valente Pais, A. R., Van Paassen, M. M., and Mulder, M. (2012). Perception and Behavioral Phase Coherence Zones in Passive and Active Control Tasks in Yaw. In *Proceedings of the AIAA Modeling and Simulation Technologies Conference and Exhibit, Minneapolis, MN, USA, August 13-16*, AIAA 2012-4794.

Valente Pais, A. R., Pool, D. M., De Vroome, A. M., Van Paassen, M. M., and Mulder, M. (2012). Pitch Motion Perception Thresholds During Passive and Active Tasks. *Journal of Guidance, Control, and Dynamics*, 35(3):904–918.

Pool, D. M., Valente Pais, A. R., De Vroome, A. M., Van Paassen, M. M., and Mulder, M. (2012). Identification of Nonlinear Motion Perception Dynamics Using Time-Domain Pilot Modeling. *Journal of Guidance, Control, and Dynamics*, 35(3):749–763.

Valente Pais, A. R., Van Paassen, M. M., Mulder, M., and Wentink, M. (2011). Effect of Performing a Boundary-Avoidance Tracking Task on the Perception of Coherence Between Visual and Inertial Cues. In *Proceedings of the AIAA Modeling and Simulation Technologies Conference and Exhibit, Portland, OR, USA, August 8-11*, AIAA 2011-6324.

Correia Grácio, B. J., Wentink, M., Valente Pais, A. R., Van Paassen, M. M., and Mulder, M. (2011b). Driver Behavior Comparison Between Static and Dynamic Simulation for Advanced Driving Maneuvers. *Presence: Teleoperators and Virtual Environments*, 20(2):143–161.

Correia Grácio, B. J., Valente Pais, A. R., Van Paassen, M. M., Mulder, M., Kelly, L. C., and Houck, J. A. (2011a). Relationship Between Optimal Gain and Coherence Zone in Flight Simulation. In *Proceedings of the AIAA Modeling and Simulation Technologies Conference and Exhibit, Portland, OR, USA, August 8-11*, AIAA 2011-6556.

Jonik, P. M., Valente Pais, A. R., Van Paassen, M. M., and Mulder, M. (2011). Phase Coherence Zones in Flight Simulation. In *Proceedings of the AIAA Modeling and Simulation Technologies Conference and Exhibit, Portland, OR, USA, August 8-11*, AIAA 2011-6555.

Valente Pais, A. R., Van Paassen, M. M., Mulder, M., and Wentink, M. (2010b). Perception of Combined Visual and Inertial Low-Frequency Yaw Motion. In *Proceedings of the AIAA Modeling and Simulation Technologies Conference and Exhibit, Toronto, ON, Canada, August 2-5*, AIAA 2010-8093.

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

Valente Pais, A. R., Wentink, M., Van Paassen, M. M., and Mulder, M. (2009b). Comparison of Three Motion Cueing Algorithms for Curve Driving in an Urban Environment. *Presence: Teleoperators & Virtual Environments*, 18(3):200–221.

Valente Pais, A. R., Van Paassen, M. M., Mulder, M., and Wentink, M. (2009a). Perception Coherence Zones in Flight Simulation. In *Proceedings of the AIAA Modeling and Simulation Technologies Conference and Exhibit, Chicago, IL, USA, August 10-13*, AIAA 2009-6242.

De Vroome, A. M., Valente Pais, A. R., Pool, D. M., Van Paassen, M. M., and Mulder, M. (2009). Identification of Motion Perception Thresholds in Active Control Tasks. In *Proceedings of the AIAA Modeling and Simulation Technologies Conference, Chicago, IL, USA, August 10-13*, AIAA 2009-6243.

Wentink, M., Valente Pais, A. R., Mayrhofer, M., Feenstra, P., and Bles, W. (2008). First Curve Driving Experiments in the Desdemona Simulator. In *Driving Simulation Conference Europe, Monaco, January 31-February 1*, 135–146.

Valente Pais, A. R., Wentink, M., Mulder, M., and Van Paassen, M. M. (2007). A Study on Cueing Strategies for Curve Driving in Desdemona. In *Proceedings of the AIAA Modeling and Simulation Technologies Conference and Exhibit, Hilton Head, SC, USA, August 20-23*, AIAA 2007-6473.

Groen, E. L., Wentink, M., Valente Pais, A. R., Mulder, M., and Van Paassen, M. M. (2006). Motion Perception Thresholds in Flight Simulation. In *Proceedings of the AIAA Modeling and Simulation Technologies Conference and Exhibit, Keystone, CO, USA, August 21-24*, AIAA 2006-6254.

Valente Pais, A. R., Mulder, M., Van Paassen, M. M., Wentink, M., and Groen, E. L. (2006). Modeling Human Perceptual Thresholds in Self-Motion Perception. In *Proceedings of the AIAA Modeling and Simulation Technologies Conference and Exhibit, Keystone, CO, USA, August 21-24*, AIAA 2006-6626.

Valente Pais, A. R. and Mulder, M. (2006). The Secrets of Simulation: the Art of Deceiving the Senses. Leonardo Times, Periodiek van de Vliegtuigbouwkundige Studievereniging "Leonardo da Vinci", 10e jaargang, nummer 3, pages 6-8.

Bles, W., Wentink, M., Groen, E. L., and Valente Pais, A. R. (2006). Perceptual Consequences of Simulating Turning in Motion Based Simulators. In *Proceedings of the Driving Simulation Conference Europe, Guyancourt, France, September 15*.