MESTRADO
MULTIMÉDIA - ESPECIALIZAÇÃO EM MÚSICA INTERATIVA E DESIGN DE SOM

AUDITORY INDUCED VECTION: EXPLORING ANGULAR ACCELERATION OF SOUND SOURCES

Ivo Neves Santos

M

2019

FACULDADES PARTICIPANTES:

FACULDADE DE ENGENHARIA
FACULDADE DE BELAS ARTES
FACULDADE DE CIÊNCIAS
FACULDADE DE ECONOMIA
FACULDADE DE LETRAS

FACULDADE DE ENGENHARIA DA UNIVERSIDADE DO PORTO

Auditory Induced Vection: Exploring Angular Acceleration Of Sound Sources

Ivo Neves Santos

MASTER THESIS

Mestrado em Multimédia - Música Interativa e Design de Som

Supervisor: Prof.ª Inês Salselas, PhD

Co-Supervisor: Prof. Eduardo Magalhães

July 30, 2019

Auditory Induced Vection: Exploring Angular Acceleration Of Sound Sources

Ivo Neves Santos

Mestrado em Multimédia - Música Interativa e Design de Som

July 30, 2019

Resumo

Vection é a terminologia usada para designar a sensação de movimento num observador estático. O exemplo mais evidente da sensação de *Vection* ocorre quando um observador se encontra sentado em um de dois veículos parados lado a lado e, quando o outro veículo inicia a marcha, existe a sensação que é o veículo onde o observador se encontra que se move, criando desta forma, uma ilusão de movimento. Nesta pesquisa, propômo-nos a analisar três fatores derivados da aceleração angular de fontes sonoras e os seus efeitos na indução de *Vection* auditivo, nomeadamente: aceleração rápida, aceleração lenta e factores elevados de aceleração. Recorrendo à manipulação digital de duas fontes sonoras com características mecânicas, foram elaborados quatro estímulos auditivos, apresentados individualmente a participantes vendados e sentados no centro de um sistema multicanal ambisonics. Os resultados da análise de variância indicam que a aceleração angular rápida poderá ser um factor de potenciação na sensação de *Vection* auditivo. Para trabalho futuro propomos testar as diferenças entre reprodução binaural com headphones e sistemas multicanal, assim como a importância da sincronização entre estímulos auditivos e tácteis, na indução de *Vection* auditivo.

Abstract

Vection is the terminology used to designate the motion sensation of a static observer. The most obvious example of *Vection* occurs when an observer is seated in one of two stationary vehicles standing side by side and, when the other vehicle begins to march, there is the sensation that is the vehicle where the observer is that moves, creating an illusion of movement. In this research, we propose to analyze three factors derived from the angular acceleration of sound sources and its effects on auditory induced *Vection*, respectively: fast acceleration, slow acceleration and high acceleration factors. Using the digital manipulation of two sound sources with engine characteristics, four auditory stimuli were elaborated and presented individually to blindfolded participants seated at the center of a multi-channel ambisonics system. The results of the analysis of variance indicate that rapid angular acceleration may be an enhancer factor on the auditory *Vection* sensation. For future work we propose to study the differences between binaural reproduction through headphones and multichannel systems, as well as the importance of auditory and tactile synchronization, on auditory induced *Vection*.

Acknowledgements

I want to express my deepest appreciation to my supervisor Ines Salselas and co-supervisor Eduardo Magalhães, to whom I am extremely grateful for their generous guidance, patience and understanding, both prior and throughout this research.

Many thanks to my parents, girlfriend and friends for all the support and patient demonstrated on the roughest times and, finally, thanks to all academic staff and post-graduate students from the Faculty of Engineering of University of Porto that participated in either experiments and who made this possible.

Ivo Neves Santos

Contents

List of Figures

LIST OF FIGURES

Chapter 1

Introduction

1.1 Theoretical Framework

Vection effect is the terminology for embodied illusions or self-motion illusions and it can be perceived when there is a sensation of self movement in a stationary position.

The high quality and realism, both for visual and sound inputs, aims for even better resolution, providing the perceiver sensory information with a very close approximation to the real world. With the emancipation of multi and cross modal platforms, in many occasions, illusory self motion can be a desirable plus for industries in the professional training such as simulators, or even in the entertaining media industry (Väljamäe et al., 2008a).

Previous studies about *Vection* demonstrated that auditory cues establish an important role on self-motion illusions, specially when combined to the visual *stimuli*. The auditory induced self-motion illusion, when compared to visually induced *Vection*, is much weaker and less compelling and only occurs in about 25-60% of the participants (Riecke et al., 2005).

Research from Valjamae et al. (2005) has shown that different scenarios with moving sound sources and the use of auditory landmarks improve auditory *Vection*. Our research takes in consideration the use of sounds with engine characteristics, so they could be easily recognized and associated to motion. We hypothesized that the sonic characteristics of this type of sound sources may establish a contextual motion connection on the perceivers.

Most of the studies uses binaural reproduction through headphones, which has been shown a great effectiveness for auditory induced *Vection* (Valjamae et al., 2005). Despite this effectiveness, we also aim to study if it is possible to induce *Vection* with multichannel reproduction, which is not a conventional experimental setup of auditory induced *Vection*. The experimental setup used in our research is similar to the one used in previous research by Sakamoto et al. (2004), who also used an 8 speaker array reproduction system on *Vection* study.

The Ambisonics (Fellgett, 1975; Gerzon, 1975; Frank et al., 2015) is appropriate for virtual 3D environments and 360 degrees spatial audio. The ambisonics codifies a sound-field according its directional properties and it allows new approaches not only to music composition and mixing, but also in sound design and academic studies, due to the versatility of the format and its ability to create realist soundscapes that can have a huge influence on our perception.

Gathering the considerations above referred, our research aims to explore the angular acceleration of sound sources on auditory induced *Vection*. The angular acceleration corresponds to the variation of the angular velocity in time and it is represented in degree per second squared.

With a similar experimental setup found in Sakamoto et al. (2004) research, two experiments were conducted. First, a pilot study was conducted to observe which movement characteristics suggests more movement to the participants, regarding the sound sources. Despite the different variables, we've constructed our stimuli taking in consideration Valjamae et al. (2005) methodologies, who've designed different stimuli with different motion characteristics and presented to blindfolded participants, to enhance the auditory perception. The pilot study also aimed for possible flaws that might be encountered during the experimental process.

Experiment 1 explores the angular acceleration of sound sources on auditory induced *Vection*. Using two distinguished sound sources representing motorcycle engine and boat engine, respectively, four stimuli - Placebo (control test), slow acceleration, fast acceleration and high factors of acceleration - were designed based on the same methodologies referred above. In this experiment, we've also attached a vibrotactile device to the respective chair, serving as "anchor" (Valjamae et al., 2005) for the blindfolded participants.

1.2 Motivation

The technological evolution led to the exploration of new paradigms of our perception. The use of multiple inputs help to modulate perception and, for immersive environments such as Virtual Reality, this cross-modal perception is fundamental to bring new sensations for multimedia content, such as video-games or conventional cinema. The high quality and realism, both for visual and sound inputs, aims for even better resolution, providing the perceiver sensory information with a very close approximation to real world.

Vection effect can induce a high-level realism, by the ability of creating motion sensations on the perceiver in static environments and, in many occasions, illusory self motion can be a desirable plus for industries in the professional training such as simulators or entertaining media industry (Väljamäe et al., 2008a).

We are motivated to bring new possible approaches for immersive environments and multimedia content. Also, we aim to explore the effects of angular acceleration on self motion illusions recurring to sound design techniques that might be implemented on conventional setups, helping this way, to bring new sensations to multimedia and immersive content.

1.3 Dissertation Layout

Besides the Introduction, this dissertation contains four more chapters.

The chapter 2 presents the related content and methodologies found in previous researches. In the beginning, concepts related to how humans encode the surrounding auditory information are introduced. In addition, notions about the *vestibular system* are also presented, given the importance that the *vestibular system* has in detecting biological motion and acceleration. The *Vection* section presents the main considerations and methodologies about visual and auditory *Vection* found in previous studies. Besides, it is also presented some studies about cross-modal perception. To close this chapter, we present a set of software that we've considered suitable for movement induction and for the study of *vection* proposed in this dissertation. Chapter 3 presents all methodologies and technical details used in the experimental setup and stimuli design. This chapter is divided into the *Pilot* and *Experiment 1* sections, respectively. On each respective section, the methods and considerations used during the conduction of the experiments are presented, as well as the respective results and discussion. Chapter 4 presents a general discussion from both experiments conducted in this research, *Pilot* and *Experiment 1*, respectively. Chapter 5 presents the main conclusions of the further studies and considerations about future research related to auditory induced *Vection*.

Chapter 2

Literature Review

In modern days, people get a valuable contribution of audiovisual contents provided in various formats, most of them being consumed statically such as cinema or video-games. The high quality and realism, both for visual and sound inputs, aims for even better resolution, providing the perceiver sensory information with a very close approximation with real world.

Vection effect can induce a high-level realism, by the ability of creating motion sensations on the perceiver in static environments. In many occasions, illusory self motion can be a desirable plus for industries in the professional training such as simulators, or even in the entertaining media industry (Väljamäe et al., 2008a).

This chapter presents a review of the main studies and considerations required for the development of our research, gathering perception concepts related to how the brain works on the interpretation of various stimuli inputs, focusing on the auditory inputs.

2.1 Sound and Movement

Sound is a fundamental part of humans perception. It provides multiple information both for perceiving and understanding the surrounding world, as well as the direction, intensity and respective approximation distance of the sound source, so it can be defined as the process of sound localization based on intensity and time arrival differences of sound.

2.1.1 Ear Morphology

Hearing is a very complex biological process that assures the translation of the real world into meaningful sound events and occurs in human ear, being defined as a "conscious appreciation of vibration, perceived as sound". (Alberti, 2001, page 1).

When the elastic medium vibrates, this vibrations enters the external ear by the external acoustic Meatus, making the Timpani membrane vibrate. Timpani translates the non amplified vibrations, which flows the signal to Incus and Malleus, very small bones that provide the necessary amplification that Cochlea, the spiral organ of the internal ear that has the function of gathering frequency and amplitude information, needs to translate the signal.

Cochlea filter, rather than being just a Fourier Transform translator, splits the acoustic information into band-limited channels (De Cheveigné, 2001). It has thousands of tiny hair cells (Alberti, 2001) that analyze the frequency spectrum of the incoming signal (from 20Hz to 20kHz), allowing the human brain to separate it into reliable data that contains all the necessary information to decode correctly the given content, and attribute it to mental maps intimately related to the individual background living experience and knowledge.

2.1.2 Spatial Hearing (*HRTF/ITD/ILD*)

There are some psycho-acoustic and psycho-physic characteristics that enrich our spatial listening perception.

Hearing process occurs within a space where sound events happen and, the way that brain encodes information can be measured by a mathematical function denominated as Head Related Transfer Function (*HRTF)*, which "is a system of coordinates shifts in conjunction with movements of subject's head" (Blauert, 1997, page 14).

As Geronazzo (2018) describes, HRTFs are into the frequency domain, being related to the spatio-temporal acoustic properties resulting from the interaction between the listener's head shape (ears included), torso and the surrounding soundfield.

This interaction between body/space, in a digital system, can be reproduced by gathering data from the subject's head and it has been showed that personalized *HRTFs* provide a greater user acoustic data, essential for realistic immersive scenarios (Geronazzo, 2018).

HRTFs are based on Interaural Time Differences (*ITD*) and Interaural Level Differences (*ILD*), which provide crucial cues for azimuthal sound localization. *ITD* refers to the time difference of the sound between the two ears, providing directional localization and being more effective in the low frequency domain, up to 700 Hz approximately. *ILD*, respectively, works better for frequencies above 700Hz and refers to the level difference. It provides the approximated distance due to the sound pressure between the left and right ear, resulting in some natural frequency filtering in the farthest ear caused by the head's shape, and a boost on nearest ear (Palomäki et al., 2005; Xie, 2013).

2.1.3 Auditory Scene Analysis

In the book Spatial Hearing, from (Blauert, 1997), characteristics of the input signals are suggested and correlated with the spatial direction of the auditory event. Everyday, the human ear has the ability of processing multiple auditory events coming from various spatial directions, with different intensity, pitch, timbre, reverberation or delay. This capacity of the brain to process, distinguish, analyze and categorize sound events independently in real-time is called Auditory Scene Analysis (Bregman, 1994) and it has brought multiple topics of investigation along the years, related to the brain processes in the auditory system and how much impact the Auditory Scene Analysis has when perceiving and decoding a sound.

People create their own categories to understand the various stimulus perceived on everyday life, in order to contextualize the information given by a determined environment. These categorizations are designated as mental maps and refers to the association of a sound to the action that actually created the sound (Jones et al., 2010). For example, in many cultures, when people hear a siren, brain forms a mental image of a police car or an ambulance. These are mental attributions that, constructed over a global state of events that occured in the perceiver's life and are based on the previous knowledge acquired, help to separate a set of events into groups of relevant information based on Gestalt grouping principles (Bregman, 1994).

According to Bregman (1994), there are two grouping processing: sequential grouping and simultaneous grouping.

The first grouping processing, sequential grouping, is the one responsible for connecting "sense over time" (Bregman, 1994), which allows the user to understand the big picture of all data and contextualize it, so brain can reduce the amount of incoming information and apply it the analysis over the given context. Simultaneous grouping is the processing of selecting which sound components correspond to the same acoustic source. Given that, grouping is an essential process to transform a vast amount of sensory information into a contextual representation, helping people to understand the surrounding environment.

Summarizing, our brain, in conjunction with the auditory system, has the ability to collect all signals and order them so it can be easier to access the correct information of a given sound source or event. Auditory Scene Analysis is the process that allows us, as humans, to be selective on the hearing process on complex environments which contains, simultaneously, multiple sonic information.

2.1.4 Vestibular System

The sense of motion provoked by sound is mediated by the *vestibular system*, which assumes primal functions related to humans sensory system and its stimulation is a prominent source of various illusions which the human interprets as a subjective orientation (Meiry, 1965; Angelaki and Cullen, 2008), due to its link with the biological motion.

Although the profound understanding of the *vestibular system* is not the main focus of this study, it might be useful for further result analysis and so due to its connection with our motion and acceleration.

In the auditory human physiology, the inner ear is constituted by a group of organs responsible for two distinct functions, related respectively to the auditory translation mediated by the cochlea and a non-auditory group of organs named as *vestibular system* (Meiry, 1965). These organs assume multiple functions related to perception as individuals and are respectively the *vestibular nerve*, the *otolith* organs and the *semicircular canals* (Khan and Chang, 2013; Day and Fitzpatrick, 2005).

These organs have a huge importance in daily routine. When people walk, get up or sit, dance and bob the head while listening to music, these biological movements associated to the human being are mediated by these *vestibular sensors* (Meiry, 1965) which are linked with the omnipresent gravity force (Stoffregen and Riccio, 1988). Precisely, the *otolith* and the *semicircular organs* are responsible for the sensation of different types of acceleration, being the two *otolith sensors* responsible for linear acceleration and the *semicircular organs*, respectively, are responsible for rotational movements (Day and Fitzpatrick, 2005; Meiry, 1965; Angelaki and Cullen, 2008). Also, the information contained from the *semicircular canals* and the *otoliths* is sent to the central nervous system providing the perception of motion sensation (Meiry, 1965).

Like a modern cell phone such as Iphones or Androids, which have multiple sensors that allows the cell phone to detect the rotation of the screen, the vestibular system has the same role but in a biological context. It provides the sensory information necessary for the interpretation of the acceleration, rotation and translation of the head movement in the space and, even where there is not movement at all, the *vestibular system* provides information about gravity, letting terrestrial and aquatic animals perceive the gravitational direction Day and Fitzpatrick (2005).

Given this, as Day and Fitzpatrick (2005) states on the first page of their study about the *vestibular system*, "perception of self and non-self motion, spatial orientation, navigation, voluntary movement, oculomotor control, and autonomic control, comes from their unique and complete description of head motion and orientation in three dimensions". Perception of self-motion is based mainly on the integration of visual and *vestibular* motion information (Brandt et al., 1975). In one hand, visual information is fundamental to perceive constant velocity self-motion. On the other hand, *vestibular system* has the functionality of only detecting body acceleration Brandt et al. (1975); Nakamura and Shimojo (1999).

2.2 Vection Effect

Vection is the terminology for embodied illusions or self-motion illusions. It is the conjunction between Motion and Vector and it can be perceived when there is a sensation of self movement in a stationary position.

The main example given for *Vection* sensation is when someone is sitting in a stationary train and another train starts moving alongside, inducing in the perceiver the illusion of movement Seno and Fukuda (2012). In terms of immersiveness, *Vection* effect can induce a high-level realism, by the ability of creating motion sensations on the perceiver in static environments such as video-games or cinema. In many occasions, illusory self motion can be a desirable plus for industries in the professional training such as simulators, or even in the entertaining media industry Väljamäe et al. (2008a).

For induced *Vection*, we can attribute different types of this effect: *Circular Vection* (angular motion) (Howard and Howard, 1994; Post, 1988) and *Linear Vection* (background/foreground motion) (Trutoiu et al., 2009). *Circular Vection* can be described as an illusion induced by a rotation stimuli in the perceiver and can be easily reproduced in a laboratory setting. *Linear Vection* is related with the linear motion of human navigational system (Trutoiu et al., 2009).

This *Vection* literature review will be divided into visual *Vection* and the importance of auditory cues on self motion illusions, showing different approaches, methodologies and results of the thematic in question.

2.2.1 Visual Vection

Although the visual *Vection* is not the main focus of our research, we may find relations between concepts and methodologies found in previous researches about visual *Vection*, that might be suitable on auditory *Vection* studies.

This main example of *Vection* was studied by applying the embodied self-motions on the train example of *Vection* (Seno and Fukuda, 2012). This study consisted on the recreation of a static in-train foreground texture and a moving background train texture as well as a gratin with same motion with the same foreground/background stimuli. By asking the observer to press a button when they perceived *Vection* effect, they were able to measure *Vection* starting point, latency and respective duration. Results showed that *Vection* latency was shorter in the train illusion stimuli than the grating stimuli, and the duration was greater in the train illusion than in the grating one. An important factor is that using motion stimuli of the same size, speed and depth, *Vection* latency and duration can be altered by changing the stimuli meanings (Bregman, 1994) .

Research from Nakamura (2006) demonstrated that a stationary object behind a moving pattern inhibits *Vection*. However, a static stimulus in the front of a moving background facilitates the sensation of self-motion. Also, Howard and Howard (1994) discovered that at low velocities, small stationary point increased *Vection* magnitude.

In order to test this, Howard and Howard (1994) used a vertical cylinder of translucent plastic. This cylinder has the purpose of hosting the perceiver in the experiment and, inside of it, the visual field was filled with a random array of white spots. The cylinder capsule of the observer was rotated from left to right, from the subject's point of view and results have shown that the maximum latency of *Vection* was 60s (1 minute) due to stimulus duration (60s). Moreover, studies confirmed that when there aren't stationary objects in view, *Vection* latency is longer and *Vection* magnitude is smaller than when stationary objects are in view (Howard and Howard, 1994).

Previous study investigated the importance of the foreground in *Vection* effect induction, by applying the same stimuli attributes as the background stimuli (Nakamura and Shimojo, 1999). Each stimuli considers foreground pattern, background pattern and a fixation cross.

Foreground/background patterns were random dots, one moving from left to right at a determined speed and the other one remained stationary. Results showed that a faster movement of the stimuli induced stronger *Vection*. In the foreground-motion condition as perceived a very weak *Vection*, indicated by shorter durations and lower strength estimates, even in the fast-motion condition.Foreground slower motion did not report any self motion. However, on the background motion condition, strong *Vection* was perceived on a slower background motion.

Background movement is one of the big research topics related to *Vection*. Many studies of background motion related to *Vection* led to the object and background hypothesis (Seno et al., 2009).

Vection is also strengthen by a slower background motion (Nakamura and Shimojo, 1999). The visual field in *Vection* is a fundamental part for it to happen. It needs a large area of visual stimuli moving uniformly, giving the perceiver a sensation of self-motion in the opposite direction of stimuli (Nakamura and Shimojo, 1999).

Previous studies on *Vection* were collected by Seno et al. (2009) in order to theorize a standard object/background for self-motion illusions stimuli. They further demonstrated that motion stimuli that had a property of an object could not induce *Vection* efficiently. The farther away the perceived motion stimuli are, the stronger the *Vection* that is induced. By testing older theories and methodologies, Seno et al. (2009) demonstrated that perceptual background dominantly induced on *Vection*. Background area always induced on *Vection* and the amount of attention to the background movement paid by the observer's, dominantly induced on *Vection* (Seno et al., 2009).

2.2.2 Visual and Auditory Interaction

Research from Riecke et al. (2005) investigated the influence of auditory cues on visual *circular vection*. This study aims for the understanding of the importance of auditory cues on the sensation of self-motion and consisted on the projection of rotational images. Three conditions were considered for the stimuli design: no sound, mono sound and spatialized sound reproduced. The visual stimuli consisted in the presentation an image representing an external environment and the auditory cues were presented through headphones, using generic HRTF's. All of the participants executed, individually, 48 trials corresponding to the combination of the three stimuli design, considered by the researchers. Results from the participants rating indicated that the addition of spatialized sound suggested more convincingness. Also, auditory cues enables a better resolution on the perception of the virtual space. Moreover, adding spatialized auditory cues on virtual environments indicates a rise on the amount of self-motion perceived (Riecke et al., 2005).

Recently, Keshavarz et al. (2014) combined also studied the effects of the combination of auditory and visual stimuli and its contribute to self-motion illusions. The presented stimuli to the participants consisted in visual or auditory cues, individually, or a combination of both. The visual stimuli was present though six projectors and for the auditory stimuli it was used an array of 7 speakers and a sub-woofer. The on *Vection* onset and strength were measured through verbal responses given by the participants during the experimental process. The results of this research corroborates previous studies on *Vection* (Riecke et al., 2005) and indicates that the self-motion sensation was enhanced when both visual and auditory stimuli were presented.

2.2.3 Auditory cues

Although sound may not be as effective when comparing with visual stimuli, it has been proven that it is possible to induce auditory self-motion illusions.

Previous research from Kapralos et al. (2004) made a huge contribution on the understanding of the auditory cues in self-motion perception. With the combination with different scenarios implying, and quote, “(i) physical motion only, (ii) moving audio-cues only, (iii) decreasing intensity cues, and (iv) physical motion coupled with moving audio-cues” (Kapralos et al., 2004, page 1) could demonstrate that the addition of auditory cues can determine the self-motion perception. By creating a system with eight speakers located in an anechoic room, Sakamoto et al. (2004) proved that self-motion was induced by linearly moving sound images. Four trials were applied with different sonic characteristics, such as (i) movement front to back, (ii) movement back to front, (iii) movements from left to right and (iv) movements from right to left. Results have shown that self-motion magnitude induced by images moving from back to front was greater than images moving from front to back, suggesting that the auditory system is more sensitive to back to front motions (Sakamoto et al., 2004).

2.2.4 Ecological validity of auditory cues

Realistic rendered environments suggests an increase of the sensation of presence both for single and multiple sound sources (Larsson et al., 2004). The characteristics of sound sources may play an important role on self-motion illusions given the necessity of attributing meaning of the surrounding inputs. Previous studies consider that sounds with characteristics resembling perceiver’s motion can reveal the embodiment on the real or virtual environment (Väljamäe et al., 2008a).

Researchers, who have a study about sound as an enhancer of linear self-motion on virtual reality environments, first hypothesized that self-motion sounds such as footsteps or engine sounds, represent a specific type of acoustic body-centered feedback in virtual environments (Väljamäe et al., 2008a). Assenting on the mental maps (Bregman, 1994) previously referred, they found a stronger sensation of self-motion when moving sound fields that contained ecological sounds instead of noises or pure tones (Väljamäe et al., 2008a).

Twelve ecological sounds with variations in the auditory scene content and spatio-temporal moving sounds were presented with three initial positions on the scene: distant, closed and mixed. “Distant” simulates the situation where the listener is approaching the landmarks. “Closed” is related with the transitions from one landmark to another and “mixed” is moving one landmark toward another one. With a 0-100 scale, researchers Väljamäe et al. (2008a) were able to rate the three measures applied. Self-motion intensity showed a significant effect of the engine sound. 52-83 % of the participants perceived self-motion in different type of stimuli and 23 to 50 % perceived self-motion by the rotation of the acoustic field. Hence, auditory scenes containing distant approaching sounds resulted in higher on *Vection* reports than the close and receding ones.

The experiment conducted by Valjamae et al. (2005) also considers the importance of the ecological consistency on self-motion perception. Results suggest that sounds with tonal characteristics are more appropriate to represent distinguished sound objects and, even with noises instead of sounds, people tended to attribute a specific context to the perceived stimuli.

2.2.5 Vection modulation factors

For self-motion illusions, research has been conducted to study the influence of external variables that might influence *Vection* perception. The use of multiple sensory inputs have different implications on human's perception and, researches have been conducted in order to understand their influence on the perception of self-motion illusions.

In 2008, research from Riecke et al. (2008) demonstrated that the usage of vibrations can enhance the *Circular Vection* sensation. The experimental setup consisted in a static hammock chair and a vibration platform. The stimuli consisted on the synchronous rotational movement of two sound sources at 60^{o}/s. The results indicated that the addition of vibrations enhanced self-motion on the participants who've reported to feel self-motion sensation.

Later, in 2013, Farkhatdinov et al. (2013) research corroborates Riecke et al. (2008) previous study, pointing more evidence for the importance of the vibrations on *Vection* effect. For the haptic stimuli were considered three different types vibrations induced by *sine wave, Pink Noise and Chirp signal*. This way, they could not only observe the influence of vibrations on *Vection* but they also could examine which characteristics of vibrations elicits more self-motion sensation. Results indicate that adding vibrations stimuli to the feed augmented *Vection* perception. Also, results suggest that *Vection* sensation is modulated by the frequency of the stimuli. Moreover, Farkhatdinov et al. (2013) indicated that humans might be more sensitive to determined type of vibrations. The real life example given is related to vehicles, which vibrations may vary, pointing for further studies related to the modulation of the haptic stimuli.

In the same year, Seno (2013) demonstrated that music modulates the strength of *Vection*. In order to test the effects of music in *Vection*, Seno (2013) applied five music conditions. In two experiments was used slower tempo music and two with fast tempo. Finally, the no-music condition. With fourteen naive volunteers under these conditions, results showed that average latency's were shortest and duration was longer in two fast music conditions. Slow in the slower music condition and slowest in no-music condition. The same thing happens with the average magnitudes, that were largest in the two fast music tempo conditions. Seno (2013) demonstrated that presence of background fast tempo music had facilitated the *Vection* effect.

2.3 Cross-modal Perception

On everyday life we have the ability to perceive the surrounding environment by the encoding of multiple sensory inputs, such as vision and hearing. For example, someone waiting on a red traffic light to cross the street, while the cars are passing through, that person can see, hear and feel the cars passing by, through the various sensory inputs provided by that environment. Although we have the mental capacity to distinguish the different inputs individually, they are processed as a single coherent event.

Cross-modal perception occurs when a sensory input is modulated by another distinguished sensory input. One clear example of cross-modal perception can be attributed to synesthesia, which corresponds to the visual perception modulation by the auditory inputs, also known as color-hearing (Marks, 1975). Research from Marks (1975) demonstrated that, for synesthetic people, the brightness of the images varies with the density of the inducing sounds. Moreover, the size of the images were modulated by the volume of the auditory inputs.

As described on previous studies (Vroomen and Gelder, 2000), synchronized sounds along a visual stimuli of light flashes with different spatial locations tend to be perceived close together, modulating the spatial attention. Another interesting example of cross-modal perception is the *ventriloquism*. The sound comes from the ventriloquist who doesn't move its mouth to produce speech. However, due to the mouth synchronization of the external element, the puppet, there is a perceptual grouping driven by the visual lip sync between sound and mouth, despite the different spatial localization of both events.

For the study of cross-modal integration of auditory and visual motion signals, Meyer and Wuerger (2001) hypothesized if simultaneous auditory motion alters the visual motion detection. For the visual stimuli it was used 500 dots moving randomly. Presented through two speakers, for the design of the auditory stimuli it was used modulated white noise. Results demonstrated that visual motion bias induced by the auditory motion stimuli was consistent with the direction of the auditory motion.

2.4 Movement induction and technology

Nowadays, immersive environments requires a high level of visual and sonic detail, in order to provide a realistic virtual worlds for the users, improving immersion and presence. Along with technological evolution, new tools and software for the digital manipulation of sound starts to urge, resulting in new approaches and methodologies to researches that might need technology with such characteristics.

On this section, we present the main considerations related to technological software which enable the manipulation of sound sources in a 3D space as we found suitable to our research, regarding the focus on the auditory system and vestibular stimulation.

2.4.1 Ambisonics

Ambisonics sound technique represents a hierarchical reproduction system, able to target a number of varying loudspeaker arrays. The Ambisonics (Fellgett, 1975; Gerzon, 1975; Frank et al., 2015) is very appropriate for virtual 3D environments and 360^{o} spatial audio due to its format, which codifies a sound-field according its directional properties and, in opposition to a conventional multichannel sound system, the ambisonics format instead of having each channel associated to a speaker, each speaker represents physical properties of the acoustic field (Arteaga, 2015).

These characteristics of the ambisonics format allow the easy manipulation of the movement of sound sources in a 3D space. Artistically, it allows the exploitation of new approaches not only to music composition and mixing, but also in sound design and academic studies due to the versatility of the format and its ability to create realist soundscapes that can be have a huge influence on our perception.

For VR applications, companies like Facebook, Google and companies which own digital audio workstation software like Ableton Live from Cycling' 74 or Reaper from Cuckoo, Ambisonics have become the state of the art when comes to 360º spatial audio, due to its easy way to decode the sound-field to different reproduction formats like Dolby Atmos, 5.1 systems or binaural decoding, allowing for conventional headphones reproduction.

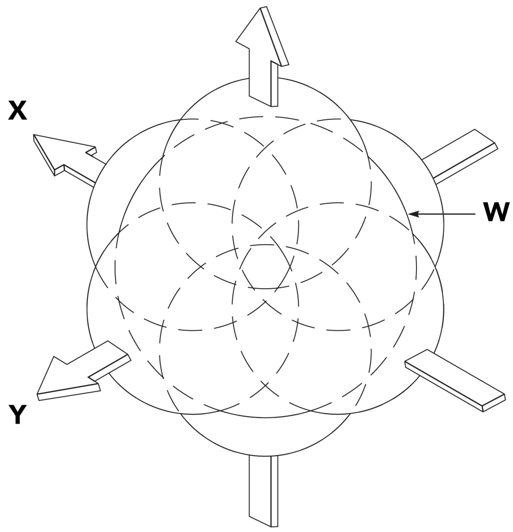

Figure 2.1: Ambisonics Axis

The figure 2.1 above represents the ambisonics soundfield. For complete description, see Soundfield website[1]

2.4.2 Reaper and Ambisonics Toolkit (ATK)

Reaper is an open-source fully customizing Digital Audio Workstation from Cuckoo, which allows to record, edit and process audio.

Among other plug-ins, when it comes to Ambisonics format, the user has the possibility to install the software Ambisonics Toolkit for Reaper. The ATK Cuckoos allows the user to send multiple sound sources to a sphere that represents the 360º environment, by encoding the signals through its encoders. ATK's transformation plug ins follows the processing chain, being a very effective tool to manipulate the sound sources throughout the space. For more detailed information, check the Ambisonics Toolkit website[2]

2.4.3 Sound Particles

"Sound Particles is something completely different from any other professional 3D audio software that exists today." Sou

[1] https://www.soundfield.com
[2] http://www.ambisonictoolkit.net/documentation/reaper/

Sound Particles is a standalone application which has multiple psycho-acoustic effects and physics incorporated by the algorithm, being an extraordinary tool to manipulate spatial audio and create custom sound design.

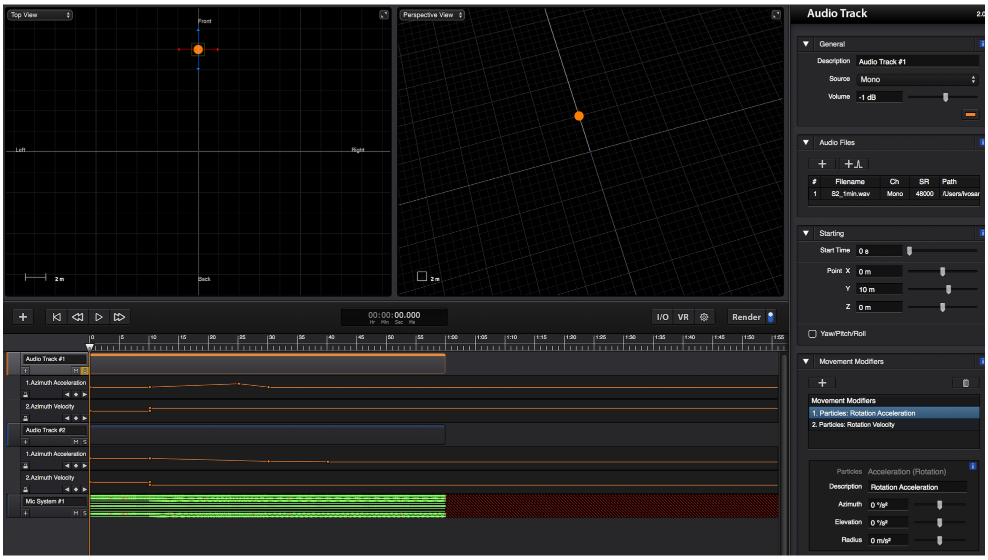

Figure 2.2: Sound Particles Interface

The *Sound Particles* interface allows the manipulation of multiple sound sources in a 3D space. To each sound source, there are multiple parameters that can be defined, such as the placement of the sound source on the axis and respective starting point. Also, it is possible to apply modifiers to each sound source, enabling the user to define velocity and acceleration, as well as random delays or equalization. The automation mode enables precise control of these parameters, according to the user preference. The technical details and full description can be accessed on *Sound Particles* website[3].

2.5 Final Considerations

There are some conclusions that connects motion and acceleration perception: the vestibular system. *Vection* occurs due to the stimulation of the vestibular system, which is responsible for motion detection (Todd and Lee, 2015). The sense of motion provoked by sound is mediated by the *vestibular system* which assumes primal functions related to humans sensory system and its stimulation is a prominent source of various illusions which the human interprets as a subjective orientation Meiry (1965), due to its link with the biological motion

The *vestibular system* gives the brain the information needed to interpret the head's rotation, translation and acceleration movements (Day and Fitzpatrick, 2005). Located in human's inner ear, is the one responsible for movement detection and movement interpretation. *Vection* occurs due to a *vestibular* stimulation, which has the ability to trick our brain and perception.

[3] https://www.soundparticles.com/

Although the auditory *Vection* tends to be weaker when compared to visual *Vection*, auditory self-motion sensation occurs in 20 to 75 % of blindfolded listeners, depending on various stimuli factors (Väljamäe, 2009).

Previous findings about *Vection* demonstrated that auditory cues establish an important role on self-motion illusions. Research from (Valjamae et al., 2005; Väljamäe et al., 2009) has shown that different scenarios with moving sound sources and the use of auditory landmarks improve auditory *Vection*. Studies about the influence of auditory cues on self-motion (Sakamoto et al., 2004) have shown that self-motion magnitude induced by images moving from back to front was greater than images moving from front to back, suggesting that the auditory system is more sensitive to back to front motions. Moreover, previous research points evidence for the importance of realistic rendered environments, which indicates to increase *Vection* perception (Larsson et al., 2004).

Most of the studies uses binaural reproduction through headphones, which has been shown a great effectiveness for auditory induced *Vection* (Valjamae et al., 2005). Despite this effectiveness, we aim to study if it is possible to induce *Vection* with multichannel reproduction systems, which is not a conventional experimental setup of auditory induced *Vection*. The experimental setup considered for our research is similar to the one used in previous research by Sakamoto et al. (2004), who also recurred to an 8 speaker array reproduction system on *Vection* study.

For measurement, we found relevant studies from Sakamoto et al. (2004); Valjamae et al. (2005), which used numeric scales to rate the self-motion convingingness. In this research, due to the lack of technological resources required to measure objectively subjective sensations (Keshavarz et al., 2015; Palmisano et al., 2015) such as *Vection*, we found the numeric scale the most appropriated method to measure self-motion convincingness. In addition, we took in consideration same methods from previous studies (Valjamae et al., 2005; Riecke et al., 2008), which considered necessary to blindfold the participants, so they could maximize the auditory system.

To test our hypothesis, different stimuli will be designed and presented to the participants through a listening experiment, to test all the previous studies assumptions. For the auditory stimuli design we consider appropriated the usage of the software Sound Particles.

2.6 Research questions and objectives

Given the information collected from the literature review, is it possible to induce auditory *Vection* using angular acceleration of sound sources?

Moreover, are sound sources resembling engines effective on *Vection* sense of presence? Previous studies point evidence for the relevance of the sound sources characteristics on auditory induced *Vection*. Although the study of presence is not the focus of our research, we found appropriate to consider the relevance of the ecological validity of the virtual environment.

The objective of our research aims to bring new approaches for immersive environments and multimedia content. In one hand, we aim to explore sound design techniques based on spatialized audio. In another hand, we also aim to find possibilities of conventional setups of auditory induced *Vection*.

Chapter 3

Methods

Preambul

Vection is the terminology for embodied illusions or self-motion illusions.

In this chapter we present the considerations and methodologies used for the conduction of our research, which aims to test the effects of angular acceleration of sound sources on auditory induced *vection*.

The methodology chapter is divided into two experiments. In the Pilot we've manipulated four motion characteristics of the sound sources that might have more impact on movement perception, regarding the sound sources.

To do so, we've created four different auditory stimuli, each one with a specific type of movement of the sound sources. The stimuli were separated, respectively, by *direction, acceleration, cadence (rhythm)* and *Doppler effect*, and it was intended to observe which one is more prominent in inducing movement sensation.

This study served as an experimental lab, from which we could point some flaws when designing and conducting the experiment, so we could improve on the methods used on further research.

In the Experiment 1, we further explore the results obtained in the first one, particularly, the effects of angular acceleration, focusing on auditory induced *Vection*.

For the experimental design of this research, we've subjectively chosen two sound sources with mechanical characteristics - motorcycle and boat engines - and we've designed three acceleration stimuli and one *Placebo*, for control test.

The angular acceleration stimuli we're divided in *slow acceleration*,*fast acceleration* and *high factors of acceleration*.

On both experiments, *Pilot* and Experiment 1, respectively, we've used a similar setup of a circular 8 speaker array disposed circularly around the blindfolded participant. In the *Experiment 1* we've attached a vibrotactile device to the chair, serving as an *anchor* sound (Valjamae et al., 2005).

3.1 Pilot Study

3.1.1 Summary

For this first step of the experiment our approach was to investigate which types of movement of the sound sources elicits more movement perception.

This has the objective to extract information from the participants about which kind of soundscapes suggests more movement, so we can have a solid structure and accuracy when constructing the final experiment.

For this study it was intended to create different 3D sound scenarios, each one with a specific type of movement of the sound sources. The scenarios were separated respectively by direction, acceleration, cadence (rhythm) and Doppler effect, being reproduced one at a time.

This is a fundamental step to understand which types of movement elicits more the sense of motion on the respective listeners, and based on that, the final experiment gathers considerations from the obtained feedback.

By applying a mean and standard deviation formulas, we were able to know which scenarios had more impact on movement perception.

The virtually created 3D ambisonic audio soundscapes were made with software such as Ambisonics Toolkit (ATK) for Reaper, Facebook Spatial Workstation and Sound Particles.

For the creation of the soundscapes, first, we've chosen two distinguishable sound sources. At this point, we only wanted to have the feedback about the related movement and so, we opted to use thirty seconds of two wavetables, C1 (32.70 Hz) and G1 (49.00 Hz), respectively, placed on a 120 *BPM* timeline. These sound sources were created on the software Ableton Live 10, which provides adequate presets of wavetable synth.

3.1.2 Working hypothesis

The experimental design of this Pilot took into consideration multiple motion characteristics that may be applied digitally to the sound sources. According to data collected from the literature review, studies from Valjamae et al. (2005) applied different motions characteristics to the sound sources such as "distant, closed and mixed".

We considered four main factors of motion that might have bigger impact on motion perception - *direction* of the sound sources, *acceleration* of the sound sources, the *cadence* of the sound sources and *doppler effect* - and it was intended to observe not only which one is more prominent in inducing movement sensation, but also possible flaws in the experimental design.

Our investigation aims to respond to the given research question:

Which movement characteristics of the sound sources elicits more movement sensation?

From these motion related concepts, we've designed four experimental stimuli in which we've considered the characteristics referred above, respectively presented to blindfolded participants, sitting in the center of a circular system reproduction, with an 8 speaker array.

3.1.3 Stimuli

For this pilot, we aim to search what kind movement characteristics transmits more sense of motion, regarding the sound sources. Given this, we have divided the soundscapes into four movement characteristics categories: direction (i), acceleration (ii), cadence (iii) and doppler (iv).

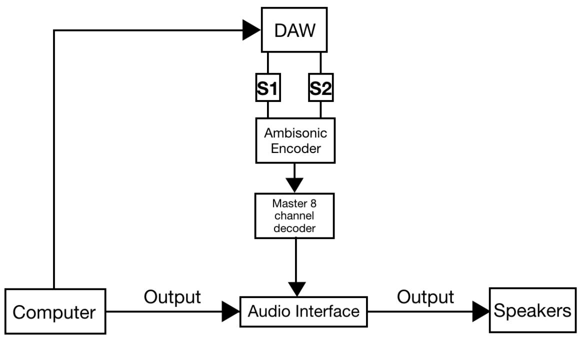

Figure 3.1: Pilot - Signal Flux

All of the stimuli are composed by two sound sources. The selected sound sources were chosen subjectively, and both of them were used in the experimental design of the stimuli. For the design of the stimuli we've chosen two wavetable from Ableton Live presets - C1 (32.70 Hz) and G1 (49.00 Hz) - placed on a 120 BPM timeline.

Each stimuli was presented with a random order to the participants and all of them were presented by four introductory beeps, which have the function to get the listener's attention, before the reproduction of the actual auditory stimuli. These beeps are followed by a couple seconds of silence, to give the listener a brief time to focus the attention on the incoming stimuli.

3.1.3.1 Direction

This stimuli scenario refers to the direction of the sound sources. In this stimuli scenario we used angular motion from back to front and vice versa. One sound source does back to front motions and the other one does the opposite. First, the sound sources are presented individually and then the stimuli evolves to mixed type, by having both of sound sources moving at the same time.

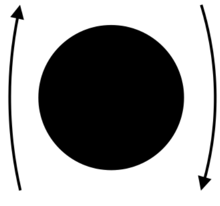

Figure 3.2: Direction

3.1.3.2 Angular Acceleration

In the acceleration stimuli we applied random accelerations to the sound sources. The variation between the two sound sources goes from slow to fast accelerations and it was created by varying randomly the saw cycle automation associated to the audio clip of the respective sound sources.

Figure 3.3: Angular acceleration

3.1.3.3 Cadence

For the cadence stimuli we've applied random sound transitions in order to create rhythms with the sound sources movements. The sound sources were manipulated to create different random patterns based on shorter and longer audio clips with motion associated.

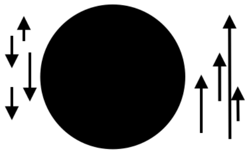

Figure 3.4: Cadence

3.1.3.4 Doppler

The Doppler stimuli was created with software Sound Particles 2.0 and consisted in moving the sound sources back to front and front to back at 50m/s and 100m/s.

Figure 3.5: Doppler

3.1.4 Measures

The results of the Pilot were measured by an evaluation of the overall rating of each stimuli scenario.

As demonstrated in previous studies (Kapralos et al., 2004), dark rooms and blindfolds are commonly associated with high values of perceptual enhancement. On our research, in order to enhance the auditory system, all the participants were instructed to use a blindfold and to be focused when listening to each stimuli. At the end of the scenario reproduction, a small light is turned on and a pause is made for them to take the blindfold and answer to a small form that classifies the movement in a numeric scale from 1 to 10, being 1 no movement at all.

3.1.5 Participants and Procedure

At this point, 16 participants (12 male and 4 female) with a mean age of 26 (SD. 4.8), agreed to participate in the study and none of them reported health issues related to the auditory system. Participants gave their written informed consent prior to the experiment.

The subjects of study were previously instructed about the active process of listening to four 3D audio scenarios reproduced through a 2nd order ambisonics array (point source*), using 8 speakers (4 Adam Audio SX3 and 4 Makie HR624MKII) positioned on a circular distribution with an angle of 45º in between previously calibrated.

The system was being amplified by the Metric Halo 2882 hardware and all stimuli were triggered with REAPER DAW[1].

The participants were sitting in an immobilized chair placed in the center of the 8 speaker array (1.64m radius and 1.10m height) and all the stimuli were presented with a SPL measure of approximately 70dB.

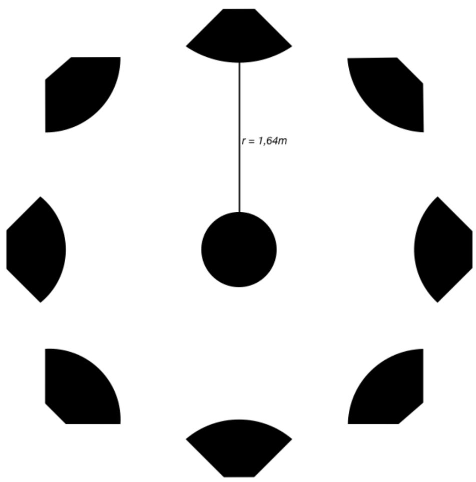

Figure 3.6: Speaker array

The room was dark and people were asked if they could see anything. All participants claimed to be fully blind before the reproduction of the stimuli.

[1]Digital Audio Workstation

3.1.6 Results

From the answers collected by a group of 16 participants, composed respectively by 12 male and 4 female, aged between 22 and 42 and with a median age of 26 (SD.4.8), we were able to gather the needed feedback for the conduction of the experimental design of the final experience.

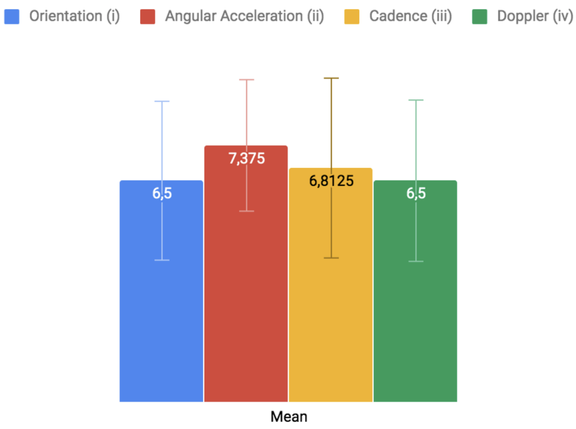

Figure 3.7: Pilot - Mean and Standard Deviation

The most rated stimuli was the one attributed to condition (ii), *angular acceleration*, with a mean of 7.38 and a standard deviation of 1.63, respectively.

	Direction (i)	Angular Acceleration (ii)	Cadence (iii)	Doppler (iv)
Mean	6.5	**7.38**	6.81	6.5
Standard Deviation	1.97	**1.63**	2.23	2

Figure 3.8: Pilot: Mean and Standard Deviation

The values extracted from the four stimuli comparison presented to participants indicate that there wasn't a significant difference between the (i), (iii) and (iv) conditions.

Both (i) (M = 6.5, STD = 1.97) and (iv) (M = 6.5, STD = 2) stimuli presented the same mean and *STD* values. Condition (iii) values, cadence, presented a slight difference between between (i) and (iv) condition.

3.1.7 Discussion

The approach to the experimental design of the stimuli was to attribute different motion characteristics to the sound sources of the different stimuli, to test which sound sources movement had bigger impact on the perception of motion. All participants were naive with respect to the experimental setup and to the purpose of the experiment.

This Pilot study served as an experimental lab, from where we could extract some mistakes that occurred during the conduction of this experiment. This way, we could improve about the methodologies used in the Experiment 1 (3.2 of chapter 3).

The first flaw found in this study starts with the attributed duration of the stimuli. In this pilot, different durations of exposure were attributed to the four different stimuli presented to the participants, which implies a variable - time of exposure - that was not premeditated to measure and did not correspond to our research question.

When designing the stimuli, the technical approach of the conducted pilot was not the most appropriated due to the use of different ambisonic decoders and the uncontrolled audio edition, which led to the introduction of different variables, such as the duration of the audio clips and the duration of the motion automation attributed to each clip that compose the respective stimuli. Instead, we should have established specific stimuli durations and coherent audio edition, so we could have more accurate information to measure.

Apart from the *doppler* (iv) condition, which was constructed with Sound particles 2.0, the accelerations and velocities could not be extracted due to the nature of the tools used, which demonstrated to be inappropriate for accurate control of the values.

The gender distribution of the participants was not even, with 12 male and 4 female on the collected sample for the Pilot. Also, participants reported some confusion when rating the provided stimuli and they tended to compare each one the stimuli. This happened due to the lack of clarity of the instructions given to the participants, who were not being told that it wasn't supposed for them to establish a comparison between stimuli.

From this mistakes we've learned fundamental steps for more accurate measurements and variables control. Also, from the feedback given by the participants, we could rethink the approach about how the instructions should be conducted in further experiments.

However, despite the flaws found in this pilot, the combination of the extreme rotation of both sound sources in the (ii) condition, angular acceleration, was the characteristic that suggested more movement to the participants and served as an indicator for the Experiment 1. Moreover, we found interest in studying the effects of the angular acceleration of the sound sources on auditory induced *vection*.

For future research we consider the study of the variable - *time of exposure* - on auditory induced *Vection*, as we found to be a promissory factor for studies that might use 2ND or higher ambisonics orders, with a circular speaker setup.

3.2 Experiment 1 - Angular Acceleration Effects on Auditory Induced *Vection*

3.2.1 Summary

From collected data of the previous Pilot, results indicated that people tended to consider the movement characteristic of sound found at stimuli (ii), angular acceleration, the one which elicited more movement for the perceivers, regarding the sound sources.

In this study, we approach similar methods to test angular *Vection* with the angular acceleration of sound sources, reproduced with an 8 speaker circular array and a simulation chair, with a vibrotactile device serving as an "anchor".

3.2.2 Working hypothesis

The question research that we've committed to study is “Can angular acceleration induce auditory *Vection*?”.

Most of the studies are made with binaural synthesis, which has been shown a great effectiveness for auditory induced *Vection* (Valjamae et al., 2005). Despite this effectiveness, we also aim to study if it is possible to induce *Vection* with multichannel reproduction, which is not a conventional experimental setup of auditory induced *Vection*.

For the experimental design of this research, we've subjectively chosen two sound sources with mechanical characteristics and we've designed three acceleration stimuli and one *Placebo*, for control test and it can be listened following the link[2].

The sound sources were manipulated with the software Sound Particles 2.0, which provides an intuitive user interface for 3D audio sound design.

3.2.3 Stimuli

From the obtained results on the previous experiment, we've tried to study the angular effects of two mechanical sound sources.

In this experiment, all the stimuli were related to acceleration and we've explored three different approaches to acceleration: *Placebo*, *Slow Acceleration*, *Fast Acceleration* and *High Factors of Acceleration*.

All of the stimuli are composed by the same sound sources. The first sound source represents a motorcycle and the other one represents a boat engine.

Both of the sound sources were edited with Ableton live, to attribute the duration and the desired audio portion of the collected sound sources used for the stimuli.

Each stimuli was presented with a random order to the participants and all of them were initiated by four introductory beeps, which have the simple function to get the listener's attention. Before the reproduction of the actual auditory stimuli the beeps are followed by a couple seconds of silence, to give the listener a brief time to focus the attention on the incoming stimuli.

All stimuli are based on sound sources moving on opposite direction and all of them start with five seconds of null velocity and acceleration. After 5seconds, the sound sources initiate their angular trajectory. On stimuli (ii) and stimuli (iii), both sound sources have the same radius (10m) and are placed in opposite positions from each other. The acceleration was set to a maximum of 20º/s2 and a constant velocity of 10m/s was attributed to both sound sources. Stimuli (iv) explores higher rates of acceleration and both sound sources were set to a constant velocity of 25m/s and a maximum acceleration of 150º/s2.

[2]https://soundcloud.com/user-605419593

The stimuli were created with the software Sound Particles 2.0, which not only provides accurate control of the angular velocity and acceleration of multiple sound sources, but also allows the user to export to ambisonics in AmbiX or FuMa.

In this particular study we've exported to FuMa due to the nature of the ATK decoder, which uses FuMa normalization.

3.2.3.1 Stimuli i - *Placebo*

This stimuli was intended to be the control stimuli scenario or *Placebo*. Contrary to the other acceleration stimuli, it is important to underline that the engine characteristics of the selected sound sources transmits motion by itself, even if they are statically placed in the ambisonics sphere.

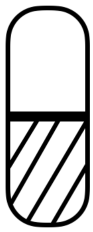

Figure 3.9: Placebo

3.2.3.2 Stimuli ii - Slow Acceleration

This stimuli corresponds to larger time intervals that the sound sources take to reach the maximum acceleration.

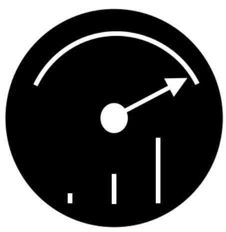

Figure 3.10: Slow Acceleration

3.2.3.3 Stimuli iii - Fast Acceleration

This stimuli corresponds to shorter time intervals that the sound sources take to reach the maximum acceleration.

Figure 3.11: Fast Acceleration

3.2.3.4 Stimuli iv - High Acceleration Factors

On the opposite, these sound sources were set to a higher velocity and much higher accelerations (max. 150º/s2) than the other acceleration stimuli.

Figure 3.12: High Acceleration

3.2.4 Measures

For the Experiment 1, we took in consideration the elaboration of a pre-experiment questionnaire. Through that evaluation, we were able to gather demographic data, auditory health issues, or if the participants experienced simulation environments before.

Taking in consideration that this specific type of sensation deals with subjective parameters that are quite complex and expensive to measure objectively, we've opted to ask the participants to verbally classify their self-motion convincingness.

To classify the convincingness of self-motion, we've opted for a short numerical scale from 0 to 7 and the participants were always blindfolded throughout the entire experiment. Study from Valjamae et al. (2005) used larger numerical scales for the verbal classification of *Vection* convincingness, from 0 to 100. The use of a joystick (Riecke et al., 2005) to measure the time-stamps and offsteps could be a plausible approach, however, due to the lack of technological resources, we could not measure neither the time, neither the offset points where auditory induced *Vection* occurred.

At the end of each stimuli, the answers and observations needed for the qualitative measurement were recorded by an audio recorder application from an Android cellphone.

The classification of the auditory induced *Vection*, previously instructed, was evaluated by two verbal instructions - "Classify your self-motion sensation.", "Was there a specific direction felt on the self-motion?". This way we could analyze the rate of the subjective self-movement sensation and the correspondent direction.

3.2.5 Participants and Procedure

In this experiment, from the answers collected by a group of 20 naive people (13 males and 7 females), aged between 23 and 48 with a median age of 29 (SD. 8.32) agreed to participate in the study. The participants gave their written informed consent prior to the experiment, in which informs the participants about the terms, nature and characteristics of the presented study.

Then, the subjects of study were previously instructed about the active process of listening to four 3D audio scenarios reproduced through a 2nd order ambisonics array (point source*), using 8 speakers (4 adam audio SX3 and 4 Makie HR624MKII) positioned on a circular distribution with an angle of 45º in between previously calibrated. The system was being amplified by the Metric Halo 2882 hardware and all stimuli were triggered with REAPER DAW[3].

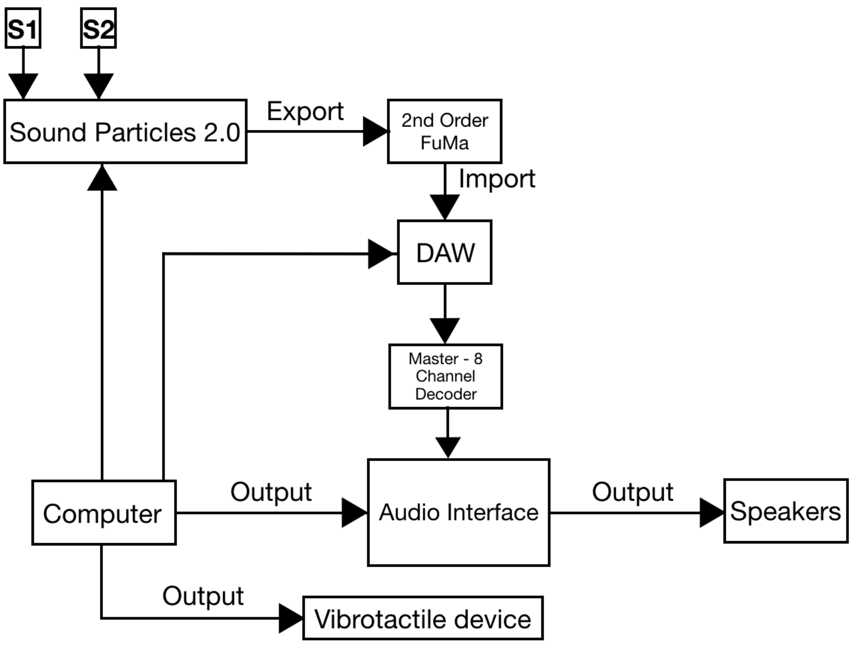

Figure 3.13: Experiment 1 - Signal Flux

The participants were sitting in an immobilized chair placed in the center of the 8 speaker array (1.64m radius and 1.10m height) and all the stimuli were presented with a SPL measure of approximately to 70dB.

The chair contained a vibrotactile device attached to the arc where people were instructed to position their feet.

[3]Digital Audio Workstation

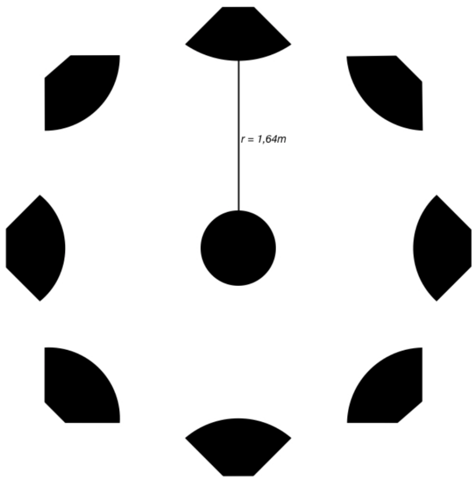

Figure 3.14: 8 channel speaker array

Being an experimental study that uses only the auditory system, participants were instructed for the use of a blindfold, to focus on the auditory perception of the stimuli. To improve convincingness, the participants were also verbally instructed for the simulation environment in which the experiment occurred.

With binaural reproduction it is possible to position virtually a sound source in the center or near the "head", designated as anchor sound (Valjamae et al., 2005) (Riecke et al., 2005). However, the same doesn't apply to multichannel reproduction. This way, the vibrotactile device was found to be an useful solution to establish a "fixation cross" (Nakamura and Shimojo, 1999) between the static subject and the moving sound sources.

The cutoff frequency of the vibrotactile device was set to 160Hz. Taking in consideration that to much vibrations may distract the subjects, it's threshold was previously regulated subjectively, giving primacy to the auditory stimuli and minimize possible distracting factors that these vibrations may cause.

The participants were previously instructed to be focused when listening to each stimuli and to keep their feet on the arc of the chair. They were also instructed about the related questions they needed to answer verbally in order to classify the subjective perceptual sensation of *Vection*.

All participants were presented randomly with the same group of four stimuli twice with a permutation plan for repeated measures. In this case, we've used partial counterbalance method due to the difference between the number of scenarios (4 x 4 = 16) and the number of participants (20), which was resolved by random distribution of the orders between the participants. Consequently, 4 random people had to listen to a group of repeated order of stimuli, also attributed randomly.

After each stimuli, participants were asked to rate verbally their self-motion sensation from zero to seven and specify the direction, in case they perceived any self-motion oscillation. The verbal answers were recorded in the audio format in order to collect the qualitative statements from the participants.

3.2.6 Results

3.2.6.1 Questionnaire

Results from the pre-questionnaire answers of 20 participants, 85% of participants experienced simulation environments before and the remaining 15% reported to never have had any contact or experience with these environments.

	Male	Female
Number	13	7
Percentage (%)	65%	35%

Figure 3.15: Experiment 1 - Participants genre

3.2.6.2 Convincingness of self motion

To measure the convincingness of self-motion from the 20 subjects, a one-way ANOVA was conducted to compare four conditions of angular accelerations of mechanical sounds sources on auditory induced *Vection*. We found this method appropriated to determine if there is a difference between the three types of accelerations, when compared to the *Placebo* (i).

The mean score indicated that the *fast acceleration* condition (M = 4.85, SD = 1.53) was significantly greater on the slow condition (M = 4.38, SD = 2.20), followed by the *high factor of acceleration* condition (M = 4.35, SD = 1.75).

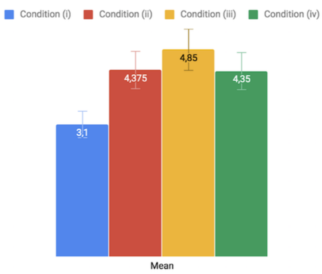

Figure 3.16: Experiment 1 - Mean and Standard Deviation

For the *Placebo* condition we expected lower or null values, when compared to stimuli with sound sources in motion. The data collected from the counting of the *Placebo* and (iii) condition presented below, displays a representative variation between the placebo and this conditions.

On the left side of the figure 3.17 presented below, it is possible to observe that most of participants considered the stimuli (i), *Placebo*, the one that suggested less or none self-motion on the subjects of study. On the opposite, it is possible to verify that the stimuli (iii), *fast acceleration*, was the one that suggested more self-motion on the participants.

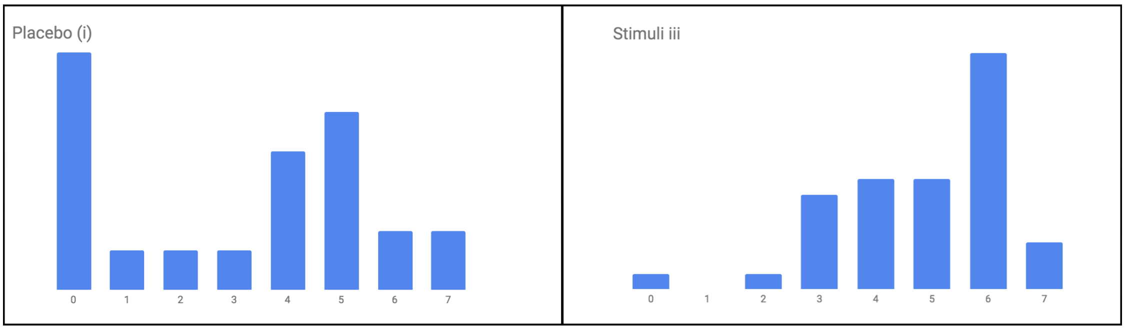

Figure 3.17: Rating - Placebo (i) vs. Stimuli iii

3.2.6.3 ANOVA - First four stimuli

From the collected ratings from the first four conditions presented before repetition, the results from the one-way ANOVA indicated that there was an amount of reported auditory induced *vection* remembered at the $p<.05$ level for the condition (iii) [$F(1, 38) = 9.011$, $p = 0.0047$].

Placebo (i) vs. Condition iii						
Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	38,025	1	38,025	9,011	0,0047	4,098
Within Groups	160,35	38	4,219			
Total	198,375	39				

Figure 3.18: Placebo (i) vs. Stimuli iii ANOVA

At the $p<.05$ level remembered, the results from both conditions (ii) [$F(1, 38) = 3.09$, $p = 0.08659$] and (iv) [$F(1, 38) = 1.27$, $p = 0.26683$], respectively, presented values of o above 0.05, rejecting the null hypothesis.

3.2.6.4 ANOVA - Repetition of the same stimuli

From the collected ratings of the four stimuli repetition presented to the participants, the results from the one-way ANOVA demonstrates that there was a significant difference amount of reported auditory induced *vection* convincingness remembered at the $p<.05$ level between the conditions (iii), *fast acceleration* - [$F(1, 38) = 5.58$, $p = 0.0234$] - and condition (iv), *high acceleration factors* - [$F(1, 38) = 7.00$, $p = 0,0117$] -, respectively.

For the four stimuli repetition presented to the participants, the null hypothesis is rejected by the results obtained from condition ii [F(1, 38) = 2.73 , p = 0.1065] which presented values above the level of the remembered $p < 0.05$.

3.2.6.5 ANOVA - General

From the collected ratings of the total amount of 8 stimuli presented, the results from the one-way ANOVA demonstrates that there was a significant amount of reported auditory induced *vection* convincingness remembered at the $p < .05$ level for condition iii [F(1, 78) = 14.624, p = 0.0003].

Placebo (i) vs. Condition iii						
Source of Variation	*SS*	*F*	*MS*	*df*	*P-value*	*F crit*
Between Groups	61,25	14,624	61,25	1	0,0003	3,963
Within Groups	326,7		4,188	78		
Total	387,95			79		

Figure 3.19: Global ANOVA

3.3 Discussion

The one-way ANOVA for all stimuli presentation indicate that there was a significant amount of reported auditory induced *vection* at the $p < .05$ level for (iii) condition - fast acceleration - [F(1,78) = 14.6, p = 0.0003] which suggests that the short time intervals from 2s - 5s, from lower (>0m/s2) to maximum angular acceleration values of 20m/s2, along with a constant velocity of 10 degrees/second, seems to induce auditory *vection*.

Besides the results obtained through the one way analysis of variance of the different groups of presented stimuli, we also complemented the data along with two verbal questions asked during the experimental process. At the end of the experiment, some participants demonstrated their will to provide observations about the related study and their individual experience of self-motion. From the feedback given by the participants we were able to collect qualitative information for data support, which will be discussed more in depth below.

Chapter 4

General Discussion

In this research we hypothesized that auditory induced *Vection* can be influenced by angular acceleration. To test this hypothesis, two experimental studies were conducted - Pilot and Experiment 1. Taking into account the literature review, this chapter presents not only the inferences and considerations of our results, but also the limitations of the presented study.

4.1 Pilot

The Pilot aims the research of which kind movement characteristics transmits more sense of motion, regarding the sound sources. Four stimuli were created and presented individually to each one of the 16 blindfolded participants (12 male and 4 female) with a mean age of 26 (SD. 4.8), who were sitting in the center of a 2ND order ambisonics, reproduced through an 8 speaker array.

Results from the condition (ii) (M = 4.85, SD = 1.53), angular acceleration, indicated that most of the participants found this condition the one which suggested more movement, regarding the sound sources. The results obtained on the condition (ii) may have a relation with the nature of the experimental setup and the circular disposition of the speaker array.

Time of Exposure

Taking into consideration the duration of the different stimuli, the condition (ii), angular acceleration, was the stimuli presented with a larger duration (60s), when compared to the other conditions. These values indicate that the duration of the stimuli might have some influence on the perception of angular movement.

4.2 Experiment 1

In Experiment 1 we've created three variables of acceleration and a placebo, to serve as the control test.

The experimental setup of the Experiment 1 was similar to the setup used in the previous study, Pilot, apart from the introduction of a vibrotactile device attached to the arc of the simulation chair.

Binaural reproduction as demonstrated to be effective on auditory induced *Vection*. However, we've proposed to investigate the effects of angular acceleration on auditory induced *Vection* with an 8 channel speaker array reproduction system and a chair with a vibrotactile device attached, serving as an "anchor" (Valjamae et al., 2005).

Self-motion Convincingness

We've compared the statistical results from the one-way ANOVA observed between three groups of results. These one way variance analysis corresponds to the ratings of convincingness given by the first four stimuli presentation, the repetition of the same four stimuli and of all the stimuli, respectively.

For the first four stimuli presented, the results of the one-way ANOVA indicated that there was a significant amount of reported auditory induced *Vection* convincingness remembered at the $p<.05$ level for the condition iii [$F(1, 38) = 9.01$, $p = 0.00472$]. For the repetition of the stimuli, demonstrates that there was an amount of reported auditory induced *Vection* convincingness remembered at the $p<.05$ level between the conditions iii [$F(1, 38) = 5.58$, $p = 0.02344$] and iv [$F(1, 38) = 7.00$, $p = 0{,}01174$].

When compared to the first four stimuli, the values from the repetition of the condition iii, fast acceleration, presented lower values of p. We must consider the influence of the priming effect on the expectancy decrease, which considers that the exposure to one stimulus influences a response to a subsequent stimulus, without a conscious guidance or intention (Weingarten et al., 2016).

The auditory induced self-motion illusion, when compared to visually induced vection, is much weaker and less compelling and only occurs in about 25-60% of the participants (Riecke et al., 2005). Our results suggest that about 45% reported to feel self-motion.

Cognition

Our results suggest that cognitive factors may have a role on auditory induced *Vection*, as some participants reported to imagine "someone cutting grass with a machine" or "cars passing by", we found plausible to consider possible implications related to the capacity of "suspending disbelief".

"I know I am steady but I feel like I am moving" is another report from a different participant, which suggests about the importance of cognition on self-motion perception, as (Riecke et al., 2006) research with a different experimental setup suggests.

By eliminating the visual stimuli with the blindfold, the sensibility of other senses are boosted, implying other type of mental and physical representations that might not happen in the same way as with the visual input. Research from Guastavino et al. (2005) provided evidence that the same acoustic phenomenon can give rise to two different cognitive representations, which integrate properties of mental representations into physical descriptions of the stimuli. Moreover, results from previous study of Valjamae et al. (2005) indicated that, even using noises instead of concrete sounds, people tended to attribute a specific context to the perceived stimuli. Our results suggest the same, despite the distinguished characteristics of the sound sources and the respective experimental setup.

Angular Acceleration

The one-way variance analysis results from the entire Experiment 1, demonstrated that angular acceleration is a factor that deserves further attention to auditory induced self-motion.

The one-way ANOVA for all stimuli presentation indicate that there was a significant amount of reported auditory induced *Vection* at the $p<.05$ level for one condition - fast acceleration - $[F(1, 78) = 14.6, p = 0.0003]$.

The usage of short time intervals from 2s - 5s, from lower (>0m/s2) to maximum angular acceleration values of 20m/s2, along with a constant velocity of 10 degrees/second, seems to enhance auditory induced *Vection*.

According to Larsson et al. (2004), velocity influences *Vection* for multiple sound sources, being that faster velocity simulations (60^o/s) suggests to induce more *Vection*.

Our study points to another direction, given the attributed constant velocity of 10^o/s. Regardless, it must be underlined that there was a big difference about the conditions and design of this experimental study.

Results indicate that auditory *Vection* occurred in condition (iii), fast acceleration. However, we could not be sure about the factors that had more or less influence on self-motion perception, due to the chair with the vibrotactile device.

Vibrotactile Device

Research from Riecke et al. (2009) demonstrated that adding vibrations simulating the rotation of a chair could indeed induce *Vection*.

Although we could not measure that variable, statements such as “the whole sensation of being in a vehicle is very strong, which makes you expect movement, along with motion of the sound field.” and “This really gives you the sensation that you’re driving, and so, it really feels like you’re moving”, not only might indicate a relationship between the nature of the sound sources and a synchronized vibrotactile device, but also expectations.

Since the vibrotactile device was being feeded by the signal of the stimuli, the correspondent modulation of the increase and decrease of acceleration suggests an evidence about the influence of vibrations on auditory induced *Vection*.

For higher values presented on placebo (i), as we hypothesized, we should consider the importance of the nature of the sound sources, which resembles engines. Research of Väljamäe et al. (2008b), about auditory induced *Vection* with ecological sounds, demonstrated that a stationary sound resembling engine noise had a positive effect on self-motion and presence ratings.

Also, this results accentuate even more the importance of further research related to the vibrotactile modulation on auditory vection.

Auditory Ecological Coherence

On our research, some participants reported to feel a counter-balance of riding a vehicle, and considered the whole experience "very realistic". Reports lead us to hypothesize that the conjunction between the nature of the sound sources and the vibrotactile device seems to establish an evidence for self-motion convincingness.

This follows along with the importance of auditory scene coherency and ecological validity found in previous studies (Väljamäe et al., 2009; Larsson et al., 2004). For example, a soundscape that contains a rapidly rotating ocean sound provides more ecological incoherence than a soundscape that contains a rotating motorcycle. Statements from the participants reinforces previous findings about the importance of auditory landmarks and ecological sounds on auditory induced *Vection*.

Direction of self-motion

In this research we could not measure objectively the direction of self-motion. However, from the collected statements provided by the participants during the experiment, as we hypothesized, the direction of self-motion that was mostly reported was related to head rotation.

Our results also suggest a tendency to report to the left - "I felt I was moving to the left and then I returned to my central position." and some of them reported back-to-front and front-to-back motion.

Although there isn't a solid foundation, that we're aware and acknowledged of, this effect might be related to the opposite direction of the sound sources movement and to the correspondent ratios of angular acceleration, when the sound sources passes through each other.

Ambisonics reproduction system

The Ambisonics (Fellgett, 1975; Gerzon, 1975; Frank et al., 2015) is very appropriate for virtual 3D environments and 360º spatial audio. Artistically, it allows the exploration of new approaches not only to music composition and mixing, but also in sound design and academic studies due to the versatility of the format and its ability to create soundscapes with ecological validity, which has been demonstrated to have a huge influence on our perception (Valjamae et al., 2005; Larsson et al., 2004).

Although most studies related to vection uses binaural reproduction (Valjamae et al., 2005; Larsson et al., 2004; Riecke et al., 2008, 2005) through heaphones, according to Sakamoto et al. (2004) previous study, results suggested that it is possible to induce vection by moving images linearly with an array of 8 speakers disposed circularly around the participants. Our results follow along Sakamoto et al. (2004) results but for angular movements, suggesting an evidence for the viability of multichannel reproduction on *Vection* studies, despite the resolution of the system.

Chapter 5

Conclusions and Future Work

5.1 Conclusion

In this research we hypothesized that auditory induced *Vection* can be influenced by angular acceleration.

To test this hypothesis, two researches were conducted - Pilot and Experiment 1.

Despite of the non validation of the Pilot, results from the angular acceleration stimuli indicated that most of the participants found this condition the one which suggested more movement, regarding the sound sources. This suggested a relation to the nature of the experimental setup and the circular disposition of the speaker array.

Also, angular acceleration (ii) was the stimuli with the larger duration (60s), when compared to the other conditions. These values indicate that the duration of the stimuli might have some influence on the perception of movement.

In Experiment 1 we've created three variables of acceleration and a placebo, to serve as the control test.

The one-way variance analysis results from the entire experiment 2, demonstrated that angular acceleration is a factor that deserves further attention to auditory induced self-motion.

Although we could not be precise about what caused self motion, the usage of short time intervals from 2s - 5s, from lower (>0m/s2) to maximum angular acceleration values of 20m/s2, along with a constant velocity of 10 degrees/second, seems to enhance auditory induced *Vection*.

Research from Riecke et al. (2009) demonstrated that adding vibrations simulating the rotation of a chair could indeed induce *Vection*. Although we could not measure that variable, statements tend to indicate a relationship between the nature of the sound sources and the synchronized modulation of the stimuli, provided by the vibrotactile device.

The auditory scene coherency and ecological validity of the sound sources points to be a determinant factor when dealing to auditory *Vection*, as studies from Valjamae et al. (2005) suggest.

Regarding the quantitative and qualitative analysis, results from of the suggests that about 45% reported to feel self-motion.

5.2 Future Work

Nowadays, immersive environments requires a high level of visual and sonic detail, in order to provide a realistic virtual worlds for the users, improving immersion and presence.

The Ambisonics (Fellgett, 1975; Gerzon, 1975; Frank et al., 2015) is very appropriate for virtual 3D environments and 360º spatial audio. Artistically, it allows the exploitation of new approaches not only to music composition and mixing, but also in sound design and academic studies due to the versatility of the format.

For future work, we intend to compare binaural reproduction through headphones with conventional reproduction systems and their efficiency on auditory induced *Vection*. Also, the use of different speaker placement for adapted stimuli seems to be an interest approach that we may took in consideration for future researches.

Moreover, we consider a necessity to determine offset points of the vibrotactile modulations and its impact auditory induced *Vection*, so we could bring new technical sound design approaches for immersive audio environments.

There are several lines of research arising from this work which should be pursued, so we can have a better understanding of how does our body and brain reacts and perceives the self-motion illusions induction.

References

SoundParticles overview. https://www.soundparticles.com/products/soundparticles/overview/. Accessed: 2019-02-08. Cited on page 14.

P. W. Alberti. The anatomy and physiology of the ear and hearing. *Occupational exposure to noise: Evaluation, prevention, and control*, pages 53–62, 2001. Cited on pages 5 and 6.

D. E. Angelaki and K. E. Cullen. Vestibular system: the many facets of a multimodal sense. *Annu. Rev. Neurosci.*, 31:125–150, 2008. Cited on pages 7 and 8.

D. Arteaga. Introduction to ambisonics. 2015. Cited on page 13.

J. Blauert. *Spatial hearing: the psychophysics of human sound localization*. MIT press, 1997. Cited on page 6.

T. Brandt, E. R. Wist, and J. Dichgans. Foreground and background in dynamic spatial orientation. *Perception & Psychophysics*, 17(5):497–503, 1975. Cited on page 8.

A. Bregman. *Auditory Scene Analysis: The Perceptual Organization of Sound*. A Bradford book. MIT Press, 1994. ISBN 9780262521956. URL https://books.google.pt/books?id=jI8muSpAC5AC. Cited on pages 6, 7, 9, and 11.

Cuckoos. ATK documentation. http://www.ambisonictoolkit.net/documentation/reaper/. Accessed: 2019-02-09. Cited on page 14.

B. L. Day and R. C. Fitzpatrick. The vestibular system. *Current biology*, 15(15):R583–R586, 2005. Cited on pages 7, 8, and 15.

A. De Cheveigné. The auditory system as a separation machine. *Physiological and Psychophysical Bases of Auditory Function, DJ Breebart, AJM Houtsma, A. Kohlrausch, VF Prijs, and R. Schoonhoven, eds.(Maastricht, The Netherlands: Shaker Publishing)*, pages 453–460, 2001. Cited on page 6.

I. Farkhatdinov, N. Ouarti, and V. Hayward. Vibrotactile inputs to the feet can modulate vection. In *2013 World Haptics Conference (WHC)*, pages 677–681. IEEE, 2013. Cited on page 12.

P. Fellgett. Ambisonics. part one: General system description. *Studio Sound*, 17(8):20–22, 1975. Cited on pages 1, 13, 36, and 38.

M. Frank, F. Zotter, and A. Sontacchi. Producing 3d audio in ambisonics. In *Audio Engineering Society Conference: 57th International Conference: The Future of Audio Entertainment Technology–Cinema, Television and the Internet*. Audio Engineering Society, 2015. Cited on pages 1, 13, 36, and 38.

M. Geronazzo. User acoustics with head-related transfer functions. *Encyclopedia of Computer Graphics and Games. Springer, Cham*, pages 1–5, 2018. Cited on page 6.

M. A. Gerzon. Ambisonics. part two: studio techniques. *studio sound*, 17(8):24–26, 1975. Cited on pages 1, 13, 36, and 38.

C. Guastavino, B. F. Katz, J.-D. Polack, D. J. Levitin, and D. Dubois. Ecological validity of soundscape reproduction. *Acta Acustica united with Acustica*, 91(2):333–341, 2005. Cited on page 34.

I. P. Howard and A. Howard. Vection: the contributions of absolute and relative visual motion. *Perception*, 23(7):745–751, 1994. Cited on pages 8 and 9.

M. R. Jones, R. R. Fay, and A. N. Popper. *Music perception*, volume 36. Springer Science & Business Media, 2010. Cited on page 7.

B. Kapralos, D. Zikovitz, M. R. Jenkin, and L. R. Harris. Auditory cues in the perception of self motion. In *Audio Engineering Society Convention 116*. Audio Engineering Society, 2004. Cited on pages 11 and 21.

B. Keshavarz, L. J. Hettinger, D. Vena, and J. L. Campos. Combined effects of auditory and visual cues on the perception of vection. *Experimental brain research*, 232(3):827–836, 2014. Cited on page 10.

B. Keshavarz, J. L. Campos, and S. Berti. Vection lies in the brain of the beholder: Eeg parameters as an objective measurement of vection. *Frontiers in psychology*, 6:1581, 2015. Cited on page 16.

S. Khan and R. Chang. Anatomy of the vestibular system: a review. *NeuroRehabilitation*, 32(3): 437–443, 2013. Cited on page 7.

P. Larsson, D. Västfjäll, and M. Kleiner. Perception of self-motion and presence in auditory virtual environments. In *Proceedings of seventh annual workshop presence*, pages 252–258, 2004. Cited on pages 11, 16, 35, and 36.

L. E. Marks. On colored-hearing synesthesia: cross-modal translations of sensory dimensions. *Psychological bulletin*, 82(3):303, 1975. Cited on page 13.

J. L. Meiry. *The vestibular system and human dynamic space orientation.* PhD thesis, Massachusetts Institute of Technology, 1965. Cited on pages 7, 8, and 15.

G. F. Meyer and S. Wuerger. Cross-modal integration of auditory and visual motion signals. *Neuroreport*, 12(11):2557–2560, 2001. Cited on page 13.

S. Nakamura. Effects of depth, eccentricity and size of additional static stimulus on visually induced self-motion perception. *Vision Research*, 46(15):2344–2353, 2006. Cited on page 9.

S. Nakamura and S. Shimojo. Critical role of foreground stimuli in perceiving visually induced self-motion (vection). *Perception*, 28(7):893–902, 1999. Cited on pages 8, 9, 10, and 28.

S. Palmisano, R. S. Allison, M. M. Schira, and R. J. Barry. Future challenges for vection research: definitions, functional significance, measures, and neural bases. *Frontiers in psychology*, 6:193, 2015. Cited on page 16.

K. J. Palomäki, H. Tiitinen, V. Mäkinen, P. J. May, and P. Alku. Spatial processing in human auditory cortex: the effects of 3d, itd, and ild stimulation techniques. *Cognitive brain research*, 24(3):364–379, 2005. Cited on page 6.

R. B. Post. Circular vection is independent of stimulus eccentricity. *Perception*, 17(6):737–744, 1988. Cited on page 8.

B. E. Riecke, J. Schulte-Pelkum, F. Caniard, and H. H. Bülthoff. Influence of auditory cues on the visually-induced self-motion illusion (circular vection) in virtual reality. In *Proceedings of Eigth Annual Workshop Presence*, 2005. Cited on pages 1, 10, 26, 28, 34, and 36.

B. E. Riecke, J. Schulte-Pelkum, M. N. Avraamides, M. V. D. Heyde, and H. H. Bülthoff. Cognitive factors can influence self-motion perception (vection) in virtual reality. *ACM Transactions on Applied Perception (TAP)*, 3(3):194–216, 2006. Cited on page 34.

B. E. Riecke, D. Feuereissen, and J. J. Rieser. Auditory self-motion illusions (circular vection) can be facilitated by vibrations and the potential for actual motion. In *Proceedings of the 5th Symposium on Applied Perception in Graphics and Visualization*, pages 147–154. ACM, 2008. Cited on pages 12, 16, and 36.

B. E. Riecke, D. Feuereissen, and J. J. Rieser. Auditory self-motion simulation is facilitated by haptic and vibrational cues suggesting the possibility of actual motion. *ACM Transactions on Applied Perception (TAP)*, 6(3):20, 2009. Cited on pages 35 and 37.

S. Sakamoto, Y. Osada, Y. Suzuki, and J. Gyoba. The effects of linearly moving sound images on self-motion perception. *Acoustical Science and Technology*, 25(1):100–102, 2004. Cited on pages 1, 2, 11, 16, and 36.

T. Seno. Music modulates the strength of vection. *Psychology*, 4(07):566, 2013. Cited on page 12.

T. Seno and H. Fukuda. Stimulus meanings alter illusory self-motion (vection)—experimental examination of the train illusion. *Seeing and perceiving*, 25(6):631–645, 2012. Cited on pages 8 and 9.

T. Seno, H. Ito, and S. Sunaga. The object and background hypothesis for vection. *Vision research*, 49(24):2973–2982, 2009. Cited on page 10.

T. A. Stoffregen and G. E. Riccio. An ecological theory of orientation and the vestibular system. *Psychological review*, 95(1):3, 1988. Cited on page 8.

N. P. Todd and C. S. Lee. The sensory-motor theory of rhythm and beat induction 20 years on: a new synthesis and future perspectives. *Frontiers in human neuroscience*, 9:444, 2015. Cited on page 15.

L. C. Trutoiu, B. J. Mohler, J. Schulte-Pelkum, and H. H. Bülthoff. Circular, linear, and curvilinear vection in a large-screen virtual environment with floor projection. *Computers & Graphics*, 33 (1):47–58, 2009. Cited on pages 8 and 9.

A. Väljamäe. Auditorily-induced illusory self-motion: A review. *Brain research reviews*, 61(2): 240–255, 2009. Cited on page 16.

A. Valjamae, P. Larsson, D. Vastfjall, and M. Kleiner. Travelling without moving: Auditory scene cues for translational self-motion. Georgia Institute of Technology, 2005. Cited on pages 1, 2, 11, 16, 17, 18, 24, 26, 28, 34, 36, and 37.

A. Väljamäe, P. Larsson, D. Västfjäll, and M. Kleiner. Sound representing self-motion in virtual environments enhances linear vection. *Presence: Teleoperators and Virtual Environments*, 17 (1):43–56, 2008a. Cited on pages 1, 2, 5, 8, and 11.

A. Väljamäe, A. Tajadura-Jiménez, P. Larsson, D. Västfjäll, and M. Kleiner. Handheld experiences: Using audio to enhance the illusion of self-motion. *IEEE MultiMedia*, 15(4):68–75, 2008b. Cited on page 35.

A. Väljamäe, P. Larsson, D. Västfjäll, and M. Kleiner. Auditory landmarks enhance circular vection in multimodal virtual reality. *Journal of the Audio Engineering Society*, 57(3):111–120, 2009. Cited on pages 16 and 36.

J. Vroomen and B. d. Gelder. Sound enhances visual perception: cross-modal effects of auditory organization on vision. *Journal of experimental psychology: Human perception and performance*, 26(5):1583, 2000. Cited on page 13.

E. Weingarten, Q. Chen, M. McAdams, J. Yi, J. Hepler, and D. Albarracín. From primed concepts to action: A meta-analysis of the behavioral effects of incidentally presented words. *Psychological Bulletin*, 142(5):472, 2016. Cited on page 34.

B. Xie. *Head-related transfer function and virtual auditory display*. J. Ross Publishing, 2013. Cited on page 6.