The influence of different audio representations on linear visually induced self-motion

Henrik Rampendahl, Stefan M. Grünvogel, Arnulph Fuhrmann

TH Köln - University of Applied Sciences
henrik.rampendahl@smail.th-koeln.de
stefan.gruenvogel@th-koeln.de
arnulph.fuhrmann@th-koeln.de

Abstract: We investigate the influence of four different audio representations on visually induced self-motion (vection). Our study followed the hypothesis, that the feeling of visually induced vection can be increased by audio sources while lowering negative feelings such as visually induced motion sickness. The test participants wore a head-mounted display (HMD) with headphones and were moved forward linearly in four different virtual scenes to create visually induced vection. One of the four scenes had no auditive stimulus, the other three had mono, stereo or spatial auditive stimuli. The last two were set up for the directional localization of the same sounds from the same places, while the mono source has been homogeneous wind noise with raised and lowered frequency according to the virtual speed. The scenes with stereo and spatial audio did not lead to the expected significant increase of linear vection in contrast to the reference scene without visual stimuli and also not among each other. However, the mono wind scene significantly increased vection to the visual reference scene.

Keywords: Self-Motion, Vection, Presence, Virtual Reality, Spatial Audio

1 Introduction

Vection is the feeling of motion although one's own body does not move itself and remains in place. Not only in the context of virtual environments, this feeling is often accompanied by symptoms of motion sickness [SHA$^+$90]. Thus creators of virtual environments often simply address this problem by looking for and applying techniques to reduce vection, making their product more user friendly. Vection, however, is an impression that coexists with other impressions from the real world, making it a desirable feature for certain virtual reality applications to support presence and immersion.

In order to support the developers of virtual environments, it is important to gain knowledge about the control of the vection, e.g. to increase or decrease the perceived size of the vection. In this article, we examine the extent to which acoustic stimuli influence the feeling of vection. In our investigations, we focus specifically on the effect of different spatial presentations of the acoustic stimuli. We compare monophonic, stereo and spatial sound over headphones, as well as the complete absence of external sound sources.

After a review of related work in the following section, we describe the setup of the experiment in Section 3. This is followed by a short discussion on our questionnaires in Section 4. The test procedure is explained in Section 5, followed by the evaluation of the results in Section 6 and the conclusion in Section 7.

2 Related Work

Vection is commonly referred to as "visual illusion of self-motion in a stationary observer" [PASB15] and has been studied for over a century. One common example is a person, sitting in a stationary train while another train is passing by. This person can have a visual illusion of self-motion by visually witnessing the other train passing by. During the years, research has expanded and many definitions of vection occurred. Palmisano et al. tried to collect and overview this definitions [PASB15]. They divided the definition of vection in four fields, all referring to combination(s) of stimulation and experience type(s). Stimulation types can be vestibular, biomechanical, haptic, auditory or visual, where the experience type is illusory and/or real. The four fields start from the traditional restrictive "visual illusion of self-motion in a stationary observer" to an all including "conscious subjective experience of self-motion (real or illusory)". Our study would fall under this definition in the field of auditory and/or visual stimuli with illusion as experience type.

Since vection is always a subjective experience, it is difficult to detect and measure it reliable. There is a need for objective indicators [PASB15] which are used to validate the results of subjective measurements based on self-reporting.

Some of the objective measurement methods for detecting vection, like measuring eye movements [RSM+14, TKG03] or brain activities [RSM+14, TKG03], require special technical equipment and expertise that were not available for our experiments.

Body sway, the movement of the body in several directions while standing at the same position, has been reported to correlate with the feeling of vection [KKMW00]. On the other hand, body sway may also be induced by visual stimuli without induced vection [AP12, BLSV79]. To measure body sway, different methods are available ranging from using a force platfrom [EH85], marker-based or marker-free motion capturing [GSCD06] up to voxel based approaches [WSAK10]. In our experiments we opted for marker-based motion detection due to the availability of a suitable device.

A comprehensive explanation of how and to what extent audio sources have an influence on many aspects around VR applications, such as vection, presence, immersion, and virtual motion, is given by Riecke et al. [RVSP09]. They investigated the influence of mono and spatial sound cues on circular vection and state that appropriate audio cues are important for high perceptual and behavioral realism in VR applications, especially when it is expected from real world experiences. Audio can be used to simulate the temporal and spatial properties of motion [Väl05]. It is considered as important for self-motion simulators and even spatial sound with low quality and non-individualized head related transfer function (HRTF)s is assumed capable of producing a self-motion experience [Väl05]. In contrast

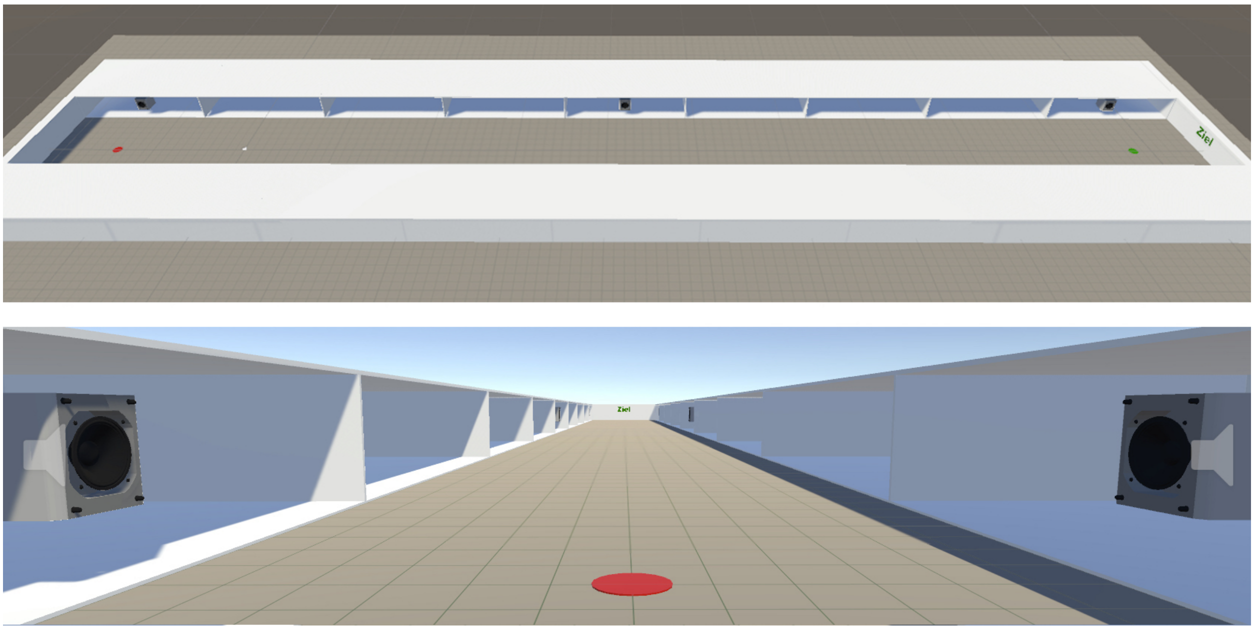


Figure 1: Top: a bird view of the VE with black speakers, red start and green finish point. — Bottom: A screenshot from a third-person perspective behind the starting point.

to circular vection [RVSP09], we compare the influence of different audio representations such as mono, stereo and spatial on linear visually induced self-motion and presence.

3 Experimental Setup

The study took place in a motion capture studio in a building of our university. For visual and auditory presentation, an Oculus Rift head-mounted display (HMD) with integrated headphones was used. The test participants where standing throughout the experiment.

It was intended to measure the body sway of the subjects as an indicator of the strength of the vection. For this purpose, the position of the subjects' body parts was measured with an optical motion capture system (OptiTrak with 24 cameras). We will not go into further details here, as we have not obtained relevant results with this approach.

In order to investigate the influence of four different audio representations on visually induced self-motion, a virtual environment (VE) has been created in the game engine Unity. Each subject was moved through this VE four times, with the same motion, consisting of 2s acceleration at the beginning, 9s constant velocity and 2s deceleration at the end. We created four different scenes in Unity with this setup. Each scene consists of the same VE and motion-sequence, but with different audio representations, namely *mono*, *stereo*, *spatial* and *vision* only, i.e. a scene with no sound at all.

The subjects are moved linearly forward through the scene, starting on a red circular area and ending on a green one (Figure 1). Behind the green circular area the word "goal" is written on a wall. To allow visual distance assessment, all environmental objects are of the same size. The ground is a green grid with dark gray squares, each 1x1 meters. Nine open boxes that appear like garages are lined up left and right. The passage between the boxes is 12m wide. Each box is 14m wide, 8m deep and 3m high. Speaker models representing sound sources are placed in the middle of the first, center and last box on each side. The motion distance from the red to the green circular area is 112m. Acceleration and deceleration are $5\frac{m}{s^2}$, the constant velocity is $10\frac{m}{s}$. These motion values are designed to produce a noticeable but not too strong impression of visually induced self-motion in the VE. The virtual sky was nearly cloudless with the sun coming from behind.

In the *visual scene*, there is no sound presented to subjects at all. It serves as reference for visually induced vection of each subject. The main requirement to all the different representations of audio is that they relate to the movement depicted. Subjects should experience motion as a plausible combination of visual and auditive stimuli. For this purpose, six different audio loops where applied to the six speaker models in the stereo and spatial sound scenes. Each audio loop represents a different instrument. Over the first left speaker a 3s bongo loop was played, over the right a 9s e-guitar lick loop. The center left speaker got a 6 seconds piano loop and the right a 5 seconds flute loop. The left speaker in the last box played a 4 seconds tuba loop and the right a 4s drum loop. For equality of loudness, all sound are leveled to around 89dB. Reverb zones are bypassed in both, the stereo and spatial scene and effects like Doppler where off. A logarithmic curve was customized for the volume roll-off. This made sounds between the speaker of the first and center box audible or between the speaker of the center and last box. However, sound from the speaker of the first box was not audible at the speaker of the last box and vise versa.

For the *stereo scene*, a spatial blend parameter was set to 3D for each audio source of a speaker. This completely suppressed the 2D effect, which plays all sound sources in the scene at the same volume, regardless of orientation and distance. The 3D setting made it possible to localize sound sources in the VE, so that the virtual movement could also be perceived auditory. Each loudspeaker appeared as a mono sound source, whose sound was attenuated depending upon alignment and distance of the head and could be heard over the headphones as stereo sound.

For the *spatial scene*, we used headphone-based sound field synthesis, where every sound direction is modeled with a HRTF. A sound library from Google called Resonance Audio was used for the implementation. Simply put, it converts all sound sources in the scene to ambisonic objects and then applies a preset of HRTFs to them.

In comparison to the 3D setting of the spatial blend parameter we use for stereo in Unity, the experimenters' impressions was, that spatial audio improved the localizing of sound sources, especially if one is in front and the other behind. However, the auditive perception of sound origins with spatial audio did not seem as good as expected, compared to the stereo scene. This can have different reasons. Localization may be negatively affected by different headphone models [SFK08] or non-individualized HRTFs [YC13, Yao17].

For the *mono scene*, all audio sources in the scene could simply have been set to mono, as it was done for the sound of a single fountain in a circular motion by Riecke et al. [RVSP09].

But to hear all sounds at the same volume and regardless of the distance is perceived as very irritating. One solution could be to reduce the volume of each sound source with its distance to the observer. However, the monophonic representation of several visible sound sources in a VE lacks audible directional information, which can be considered necessary for the assignment of a sound to its sound source in the VE and for a plausible auditory impression of the sound field. Thus, different monophonic sound of multiple objects in a VE seems very unnatural. In order to meet the aforementioned requirement for a plausible combination of visual and acoustic stimuli we selected a 15s loop with the sound of wind flowing past a window for the mono representation of motion. The loop was pitched from 1 to 2.8 times its frequency according to the virtual movement. This was intended to remind the test persons of headwind. The combination of monophonic headwinds without visual stimulus can be regarded as very plausible, but to the disadvantage of a uniform auditory scene structure and thus the direct comparability of spatial and monophonic representations of the same sound sources in space.

4 Questionnaires

A demographic questionnaire was made before the actual test runs. It includes, beside others, questions about subjects gender, age, height, weight, sportiness, highest educational attainment, job, impairment to the visual or auditive system or physical impairments. Also previous knowledge in VR, how much they use the internet or a computer in their daily lives or how often they play video games.

For the four runs of the visual, mono, stereo and spatial scene, a unified questionnaire was created, which the subjects completed after each run. It included questions about perceived self-motion, presence, auditory aspects and simulator sickness. Some questions for presence were taken from *A Presence Questionnaire* of Witmer and Singer [WS98]. The simulator sickness questions were a selection of the *Simulator Sickness Questionnaire* [KLBL93].

5 Procedure

Each subject first completed the demographic questionnaire. Afterwards they had to put on the rigid bodies using the hook-and-loop tape and gloves. The attachment of the rigid bodies was carried out carefully by visual judgment and in the same way for each subject. Then there was a brief introduction with a sample scene of the VE. The subject saw the red starting point on the ground and the goal text at the end of the passage. Bird sounds were presented in 3D stereo from the virtual speaker to the left. The subject could turn his/her head to acknowledge the different levels on each headphone. The subject was told that there are four of these scenes and after each scene the same questionnaire is filled out. They were instructed that, at the start of each scene, they had to stand relaxed, let their arms hang down, and look toward the goal text at the end of the passage.

In a pilot study we found that almost all subjects had their strongest impression of self-motion on the first run. Therefore, in the final experiment each subject was presented with a different order of the visual, mono, stereo and spatial scene.

6 Results

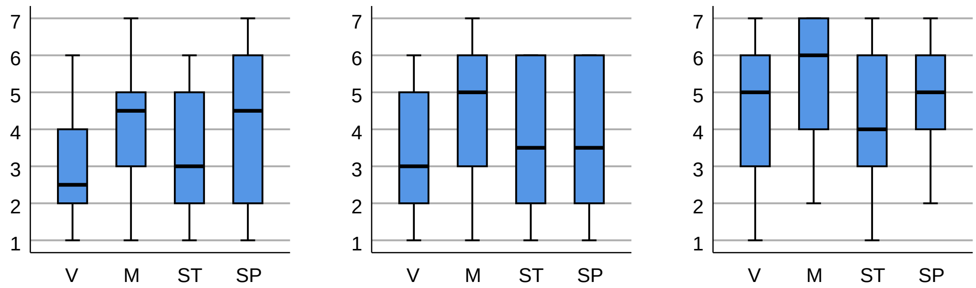
The study was conducted with 22 subjects, including 16 male and 6 female. The results of the demographic questionnaire before the actual runs are as follows. The average age was 27. The proportion of students was 63.6%, those of employees and self-employed each 13.6%. No one has stated to have physical impairments. Five people reported a debility of sight, four of them explicitly a nearsightedness. Four others said they had hearing problems, three of them tinnitus and one 70% hearing on the right ear, which is assumed to have no adverse effect on the results. Half of the subjects already had experience with virtual reality (VR) applications, most of them with the Oculus Rift, but also HTC-Vive, Google Cardboard and others. Most of the subjects use the internet and computers a few ours a day. Only 4 subjects said they play digital games every day. The rest play very rarely and 3 don't play digital games at all.

The results of the questionnaire completed after each run are presented in the next three subsections. Because most of the data did not follow a normal distribution (after a *Shapiro-Wilk* test) and are ordinal, we conducted for the statistical analysis in each case first a *Friedman* two-way analysis of variance by rank with related samples. In the case we found significant differences between the scenes, we carried out a *Dunn-Bonferroni* post-hoc test ($\alpha = 0.05$) and calculated the effect size after Cohen.

6.1 Perception of Self-Motion

The results on the question about the perceived self-motion in the acceleration phase at the beginning, in the middle of the simulation with constant velocity and at the deceleration phase to the end of each round of the experiment are shown in Figure 2. The average perceived self-motions at the beginning and middle of the simulations are not normally distributed according a *Shapiro-Wilk* test, except at the end phase.

For the beginning phase the Friedman analysis shows significant differences $\chi^2(3) = 10,411, p = 0.015$ between the simulations but a following *Dunn-Bonferroni* post-hoc test ($\alpha = 0.05$) does not show any significant differences. For the middle phase the Friedman analysis shows significant differences $\chi^2(3) = 12.205, p = 0.007$ between the simulations. The post-hoc test shows significant differences between the visual and mono sound scene ($p = 0.036$) with low effect size $r = 0.23$. For the end phase the Friedman analysis shows significant differences $\chi^2(3) = 9.089, p = 0.028$ between the simulations but the following *Dunn-Bonferroni* post-hoc does not show any significant differences.



a) Start. **b)** Middle. **c)** End.

Figure 2: Perceived self-motion during the start (acceleration), middle (constant velocity) and end (deceleration) of the simulation in our four different sound scenes: V = visual, M = mono, ST = stereo, SP = spatial sound scene.

6.2 Auditory Aspects

Figure 3 a) presents the results on the question, on how strongly what they heard influenced the overall impression (range 1 = “not at” all up to 7 = “very convincing”). The Friedman analysis shows significant differences $\chi^2(3) = 23.939, p < 0.001$ between the scenes. The post-hoc test shows significant differences between the visual-spatial ($p < 0.004$, $r = 0.29$) resp. between visual-mono sound scene ($p < 0.001$ $r = 0.37$).

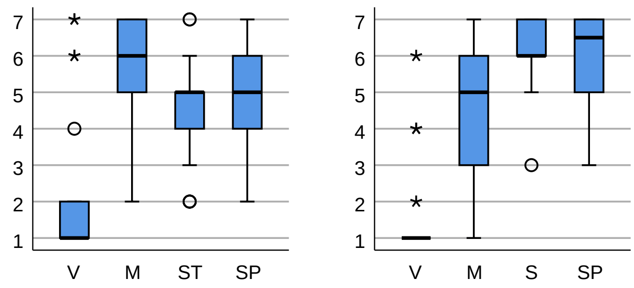
Figure 3 b) presents the results on the question, on how well they could locate what they heard (range 1 = “not at all” to 7 = “very well”). The Friedman analysis shows significant differences $\chi^2(3) = 50.921, p < 0.001$ between the scenes. The post-hoc test shows significant differences between the pairs visual-mono ($p = 0.012$), visual-stereo ($p < 0.001$) and visual-spatial ($p < 0.001$), which is not surprising, because in the visual sound scene no sound was played at all. Furthermore, the post-hoc test also shows significant differences between the pairs mono-stereo ($p = 0.043$) with the low effect size $r = 0.22$.

If we conduct the same statistic procedures with only the mono, stereo and spatial sound scene, we also get significant difference $\chi^2(3) = 17.710$, $p < 0.001$ between the scenes. The post-hoc tests shows significant differences between mono-spatial ($p = 0.01$, $r = 0.189$) and between mono-stereo ($p = 0.005$, $r = 0.21$) but with low effect sizes.

6.3 Presence Aspects

Figure 4 a) presents the results on the question, on how strong the feeling was to be at the virtual place (range 1 = “not at all” up to 7 = “equal to normal feeling”). The Friedman analysis shows that there is a significant difference $\chi^2(3) = 18.124$, $p < 0.001$ between the sound scenes. The post-hoc test shows significant differences between the pairs visual-stereo ($p = 0.030$, $r = 0.23$), visual-mono ($p = 0.012$, $r = 0.26$) and visual-spatial ($p = 0.08$, $r = 0.27$).

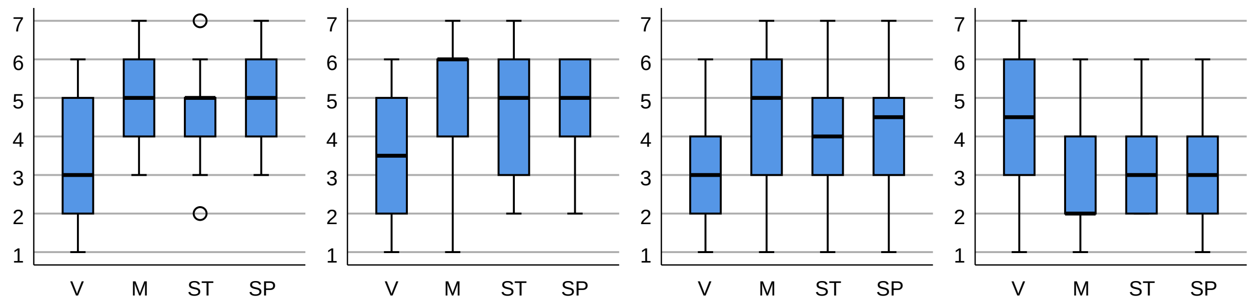
Figure 4 b) presents the results on the question, on how convincing the feeling was to move in the virtual word (range 1 = “not convincing” up to 7 = “very convincing”). The



a) Influence on overall impression.

b) Localization.

Figure 3: Results regarding questions on auditory aspects in different sound scenes: V = visual, M = mono, ST = stereo, SP = spatial sound scene.



a) Being there.

b) Moving in VE.

c) VE reminds reality.

d) Being still in lab.

Figure 4: Results regarding presence aspects in different sound scenes: V = visual, M = mono, ST = stereo, SP = spatial sound scene.

Friedman analysis shows that there is a significant difference $\chi^2(3) = 14.536$, $p = 0.002$ between the sound scenes. The post-hoc tests shows significant differences between the pair visual-mono ($p = 0.04$, $r = 0.28$).

Figure 4 c) presents the results on the question, on how strong the memories about the virtual environment resemble memories about a real environment (range 1 = "not at all" to 7 = "very strong"). The Friedman analysis shows that there is a significant difference $\chi^2(3) = 18.925$, $p < 0.001$ between the sound scenes. The post-hoc tests shows significant differences between the pair visual-mono ($p = 0.002$, $r = 0.295$).

Figure 4 d) presents the results on the question, on how strong the impression was to be in a real laboratory environment (range 1 = "no longer available" up to 7 = "very strong"). The Friedman analysis shows that there is a significant difference $\chi^2(3) = 12,633$, $p = 0.006$ between the sound scenes. The post-hoc tests shows significant differences between the pair visual-mono ($p = 0.014$, $r = 0.25$).

6.4 Discussion

In the following sections the results of the questionnaire conducted after each run is discussed, including the perception of self-motion, auditory influences and presence. In addition, the results of the center of gravity (CoG) measurements are evaluated and discussed. In some of the diagrams (3 and 4) outliers can be seen, which would usually be removed. However, because the study consisted of a smaller number of only 22 subjects, they were not removed for completeness.

The great majority of subjects answered all simulator sickness questions with “not at all”. No subject reported strong effects. We hence conclude that simulator sickness did not affect our study negatively.

6.4.1 Perception of Self-Motion

The results of the questionnaire about the intensity of perceived self-motion during acceleration, constant velocity or deceleration are shown Figure 2. It can be seen, that the most intense self-motion seems to have been experienced at the deceleration. This agrees with the verbal feedback of subjects and could also be observed during experiment. It is also visible in the calculated CoG from the motion capturing (MoCap) data, which is discussed in more detail below.

We expected that the perceived self-motion will always be higher when audio cues are present (mono, stereo or spatial audio). But our experiments did not confirm this. At the start (acceleration) and at the end (deceleration) we did not find any significant differences between the four conditions. We only found a significant difference in the middle (constant velocity) between visual and mono, with mono being rated better. This was also reported by the subjects. They stated that mono sounded like engines, an airplane or a car. Some said the noise gave them a better relation to the speed. Others stated that with only directional sounds from the environment something is missing, that is usually there and makes the impression more plausible. Also interestingly, many subjects stated that they could not hear a difference between stereo and spatial sound, which could explain that we did not find any significant differences between these two conditions with respect to perception of self-motion.

6.4.2 Auditory Aspects

The results of the questionnaire about auditive aspects are shown in Figure 3. In the question about the influence of audio on the overall impression (Figure 3 a), we could find significant differences between visual and mono, as well as between visual and spatial. This was expected, as the visual scene does not present audio at all.

We also expected the localization of sound sources to work better with spatial audio than with stereo and not at all in the mono and visual scene. However, the post-hoc test shows no significant difference between stereo and spatial, whereas in visual and mono it surprisingly does. Mono was rated quite high for localization, although the wind noise can neither be localized nor its direction determined. This could indicate doubts or ambiguities of the subject with regard to the question. Uncertainty could be found by the comprehension questions of several subjects regarding localization, whereby it was made clear that the question was related to the position determination of audio sources in the VE. Other possible uncertainties were found which may also due to misinterpretation of the questions, for example, during the visual scene in which people heard nothing through the headphones. Nevertheless they rated at least some influence (Figure 3 a) and three, that they could localize the sound (Figure 3 b).

6.4.3 Presence Aspects

The results of the questionnaire about presence aspects are shown in Figure 4. The expectation, that audio supports the sense of presence, is confirmed by the feeling of being in the VE (Figure 4 a), which is rated significantly higher in all audio scenes than in the visual only scene. This in turn is consistent with the feeling of being still in the lab, where the visual only scene has the highest rate (Figure 4 d). The expectation, however, that spatial audio will influence presence more than stereo, could not be confirmed. This is consistent with reports from many subjects that they could not detect any audible difference between stereo and spatial scenes and also to the examiners' experience, that the spatial representation was not as much better than stereo, as it was expected, which was discussed in section 3. In the following conversations, most subjects stated that they experienced a greater influence through the mono scene than through the stereo or spatial scene, both for the movement and the reality of the scene. Although there is only a significant difference in the post-hoc tests for visual and mono, the rating (median) of the mono scene for moving in the VE (Figure 4 b)) and reality (Figure 4 c)) are consistent with the subjects' reports and indicate slightly better experiences than for stereo or spatial audio.

7 Conclusion

Our study followed the hypothesis, that the feeling of visually induced vection can be improved by audio sources while lowering negative feelings such as visually induced motion sickness (VIMS). For the three motion stages of acceleration, constant velocity and deceleration, we could find only a significant difference between the visual and mono scene for constant velocity. This is in contrast to the results of [RVSP09] for a circular motion with mono sound. The reason might be a more plausible coherence of monophonic sound and movement in space for our simulated airflow than for a fixed sound source in space. Also, we could not observe the expected improvement of perceived self-motion of the spatial over the stereo scene, which both also did not differ significantly from the visual reference scene. Negative effects such as VIMS did not occur, which is attributed to the short test runs.

The presence, however, was enhanced by all audio representations for the feeling of being in the VE, which was expected. But to our surprise, only the mono scene increased the feeling of movement and reality of the scene significantly over the visual reference scene. The unexpectedly good results for mono at constant speed and for presence might also be due to the plausibility of the generated airflow and not necessarily to the type of audio presentation itself. Subjects reported mono sounded like engines, an airplane or a car, and that the noise gave them a better relation to the speed or that with only directional sound, something was missing what is usually there.

The usage of body segment tracking data to assess body sway as comparable measure for perceived self-motion was difficult, due to tracking errors and diverging curves. A further possibility of evaluation, which was not carried out here, could be to examine the differences between the individual runs per test person to compare them or, if desired, to join them together later. This could be a possibility to exclude constant test person specific characteristics such as reaction time from the overall body sway evaluation.

While stereo and spatial showed no significant differences in vection, presence or localization of the sound sources, mono produced unexpected good results. This could stimulate further research on a more simple and plausible support of presence and perceived self-motion, which, according to our results, are more promising than the use of spatial over stereo representation. However, it cannot be ruled out that longer test runs may lead to different results.

References

[AP12] April Ash and Stephen Palmisano. Vection during conflicting multisensory information about the axis, magnitude, and direction of self-motion. *Perception*, 41(3):253–267, 2012.

[BLSV79] Alain Berthoz, Michel Lacour, John F. Soechting, and Pierre P. Vidal. The role of vision in the control of posture during linear motion. In *Progress in Brain Research*, volume 50, pages 197–209. Elsevier, 1979.

[EH85] Pertti Era and Eino Heikkinen. Postural sway during standing and unexpected disturbance of balance in random samples of men of different ages. *Journal of Gerontology*, 40(3):287–295, 1985.

[GSCD06] Michela Goffredo, Maurizio Schmid, Silvia Conforto, and Tommaso D'Alessio. A markerless sub-pixel motion estimation technique to reconstruct kinematics and estimate the centre of mass in posturography. *Medical engineering & physics*, 28(7):719–726, 2006.

[KKMW00] Tetsuya Kawakita, Shinsuke Kuno, Yozo Miyake, and Satoru Watanabe. Body sway induced by depth linear vection in reference to central and peripheral visual field. *The Japanese journal of physiology*, 50(3):315–321, 2000.

[KLBL93] Robert S. Kennedy, Norman E. Lane, Kevin S. Berbaum, and Michael G. Lilienthal. Simulator sickness questionnaire: An enhanced method for quantifying simulator sickness. *The International Journal of Aviation Psychology*, 3(3):203–220, 1993.

[PASB15] Stephen Palmisano, Robert S. Allison, Mark M. Schira, and Robert J. Barry. Future challenges for vection research: definitions, functional significance, measures, and neural bases. *Frontiers in Psychology*, 6, Feb 2015.

[RSM+14] Barry Robert, Palmisano Stephen, Schira Mark, De Blasio Frances, Karamacoska Diana, and MacDonald Brett. EEG Markers of Visually Experienced Self-motion (Vection). *Frontiers in Human Neuroscience*, 8, 2014.

[RVSP09] Bernhard E. Riecke, Aleksander Väljamäe, and Jörg Schulte-Pelkum. Moving sounds enhance the visually-induced self-motion illusion (circular vection) in virtual reality. *ACM Transactions on Applied Perception*, 6(2):1–27, Feb 2009.

[SFK08] David Schonstein, Laurent Ferré, and Brian F. Katz. Comparison of headphones and equalization for virtual auditory source localization. *The Journal of the Acoustical Society of America*, 123(5):3724–3724, 2008.

[SHA+90] Robert M. Stern, Senqi Hu, Richard B. Anderson, Herschel W. Leibowitz, and Kenneth L. Koch. The effects of fixation and restricted visual field on vection-induced motion sickness. *Aviation, space, and environmental medicine*, 61(8):712–715, 1990.

[TKG03] Kai V. Thilo, Andreas Kleinschmidt, and Michael A. Gresty. Perception of self-motion from peripheral optokinetic stimulation suppresses visual evoked responses to central stimuli. *Journal of Neurophysiology*, 90(2):723–730, 2003.

[Väl05] Aleksander Väljamäe. *Self-motion and presence in the perceptual optimization of a multisensory virtual reality environment*. PhD thesis, Chalmers tekniska högskola, 2005.

[WS98] Bob G. Witmer and Michael J. Singer. Measuring presence in virtual environments: A presence questionnaire. *Presence: Teleoperators and Virtual Environments*, 7(3):225–240, 1998.

[WSAK10] Fang Wang, Marjorie Skubic, Carmen Abbott, and James M. Keller. Body sway measurement for fall risk assessment using inexpensive webcams. In *2010 Annual International Conference of the IEEE Engineering in Medicine and Biology*, pages 2225–2229. IEEE, 2010.

[Yao17] Shu-Nung Yao. Headphone-based immersive audio for virtual reality headsets. *IEEE Transactions on Consumer Electronics*, 63(3):300–308, August 2017.

[YC13] Shu-Nung Yao and Li Jen Chen. HRTF adjustments with audio quality assessments. *Archives of Acoustics*, 38(1):55 – 62, 2013.