Influence of Auditory Cues on the visually-induced Self-Motion Illusion (Circular Vection) in Virtual Reality

Bernhard E. Riecke, Jörg Schulte-Pelkum, Franck Caniard, & Heinrich H.Bülthoff
Max Planck Institute for Biological Cybernetics, Tübingen, Germany
{bernhard.riecke@tuebingen.mpg.de, joerg.sp@tuebingen.mpg.de, franck.caniard@tuebingen.mpg.de, heinrich.buelthoff@tuebingen.mpg.de}

Abstract

This study investigated whether the visually induced self-motion illusion ("circular vection") can be enhanced by adding a matching auditory cue (the sound of a fountain that is also visible in the visual stimulus). Twenty observers viewed rotating photorealistic pictures of a market place projected onto a curved projection screen (FOV: 54˚x45˚). Three conditions were randomized in a repeated measures within-subject design: No sound, mono sound, and spatialized sound using a generic head-related transfer function (HRTF). Adding mono sound increased convincingness ratings marginally, but did not affect any of the other measures of vection or presence. Spatializing the fountain sound, however, improved vection (convincingness and vection buildup time) and presence ratings significantly. Note that facilitation was found even though the visual stimulus was of high quality and realism, and known to be a powerful vection-inducing stimulus. Thus, HRTF-based auralization using headphones can be employed to improve visual VR simulations both in terms of self-motion perception and overall presence.

Keywords---**Vection, self-motion perception, spatial orientation, virtual reality, motion simulation, human factors, psychophysics, multi-modal cue integration, auditory cues, HRTF**.

1. Introduction

This paper addresses the visually induced self-motion illusion known as vection, and investigates whether additional matching auditory cues might be able to facilitate the illusion – if this were the case, it would have important implications for both our understanding of multi-modal self-motion perception and optimizing virtual reality applications that include simulated movements of the observer. Most people know the phenomenon of vection from real-world experience: When sitting in a train waiting to depart from the train station and watching a train on the neighboring track pulling out of the station, one can have the strong impression of moving oneself, even though it was in fact the train on the adjacent track that just started to move. A similar effect can be observed when sitting in the car waiting for the traffic light to turn green and when a close-by large truck slowly starts to move.

Such self-motion illusions can be reliably elicited in more controlled laboratory settings. Typically, vection has been investigated by seating participants in the center of a rotating optokinetic drum that is painted with simple geometrical patterns like black and white vertical stripes. When stationary observers are exposed to such a moving visual stimulus, they will at first correctly perceive motion of the visual stimulus (object motion). After a few seconds, however, this perception typically shifts toward oneself being moved and the moving visual stimulus slowing down and finally becoming earth-stationary. This self-motion illusion is referred to as circular vection, and the illusion has been studied extensively for more than a century [10, 21]. Excellent reviews on the phenomenon of vection are provided by [6, 15, 43]. More recently, the vection literature has also been revisited in the context of virtual reality (VR) and ego-motion simulation applications [13, 30]. So why is the phenomenon of illusory self-motion interesting in the context of VR?

Being able to move about one's environment and change one's viewpoint is a fundamental behavior of humans and most animals. Hence, being able to simulate convincing self-motions is a key necessity for interactive VR applications. There are a number different approaches to simulating ego-motion in VR, including motion platforms, free walking using head-mounted displays (HMDs), locomotion interfaces such as treadmills, or simply just presenting visual information about the self-motion. Each of these approaches offers distinct disadvantages: The drawback of using motion platforms is that they require a considerable technical and financial effort, and even then performance in VR is not necessarily comparable to corresponding real-world tasks like driving or flight simulations [1, 3, 23]. An often used alternative is to allow users to freely walk around while wearing a position-tracked head-mounted display. For most tasks, however, this requires a rather large walking area in which the observer's position is precisely tracked. This is, however, often infeasible or simply too costly. Using locomotion interfaces like treadmills or bicycles to allow for proprioceptive cues from physically walking or cycling etc. is often believed to be an optimal solution – there are, however, many open design and implementation issues that need to be carefully evaluated to come up with an optimal (and affordable) solution for a given task, especially if self-rotations are involved [14]. There has been only little research on the perception of ego-motion (vection) using treadmills, and informal observations suggest that participants hardly ever report compelling sensations of self-motion that is comparable to vection as experienced in optokinetic drums, even in the most advanced linear treadports. Durgin and Pelah state, for example, that “during treadmill locomotion, there is rarely any illusion that one is actually moving forward” [8]. Finally, when only visual information about the self-motion is provided, users hardly ever have a convincing sensation of self-motion, especially for the relatively small field of views that are common for off-the-shelf VR display devices.

In sum, despite tremendous progress in VR simulation technology, self-motion simulation in VR still poses a major challenge, and self-motion simulation is typically not as effective and convincing as corresponding real-world motions. This can lead to a number of problems including disorientation, reduced or misadapted task performance, general discomfort, and motion sickness (see, e.g., the discussion in [5, 29, 30]).

Nonetheless, it is known that moving visual stimuli *can* in certain situations be sufficient for triggering a compelling sensation of (illusory) self-motion, as is illustrated by the train illusion described above. This motivated us to investigate how far we can get without moving the observer at all, and how using VR technology might allow to optimize self-motion perception compared to the traditionally used optokinetic drums displaying abstract black and white patterns (instead of a natural scene as in the train illusion example).

Recent studies demonstrated that vection can indeed be reliably induced and investigated using VR setups that used video-projection setups [20, 13, 31, 32]. Lowther and Ware [20], Palmisano [25], and Riecke et al. [28] showed, for example, that the ability of VR to provide stereoscopic cues and to display naturalistic scenes instead of more abstract geometrical patterns can enhance vection reliably. Multi-modal contributions to vection have, however, received only little attention in the past. A noteworthy exception is the study by Wong and Frost [44], which showed that circular vection can be facilitated if participants receive an initial physical rotation (“jerk”) that accompanies the visual motion onset. One could imagine that the physical motion – even though it did not match the visual motion exactly – nevertheless provided a qualitatively correct motion signal, which might have reduced the visuo-vestibular cue conflict and thus facilitated vection. More recently, Schulte-Pelkum et al. [35] and Riecke et al. [31] showed that simply adding vibrations to the participant’s seat and floor plate during the visual motion can also enhance the self-motion sensation of the otherwise stationary participants. Post-experimental interviews revealed that the vibration were often associated with an actual motion of the VR setup (which never happened), thus making the simulation more believable.

Even though the auditory modality plays a rather important role in everyday life when moving about, there has been surprisingly little research on the relation between auditory cues and induced self-motion sensations.

This is all the more striking as auditorily induced circular vection and nystagmus have been reported as early as 1923 [7] and later been replicated several times [12, 17, 22]. Lackner demonstrated, for example, that an array of speakers simulating a rotating sound field can indeed induce vection in blindfolded participants [17]. Only recently has auditory vection received more interest, and a small number of studies were able to induce auditory vection in at least some of the participants, both for rotational and translational motions [16, 18, 32, 33, 38, 41, 39, 40]. While most researchers used artificial sounds (e.g., pink noise) [16, 17, 33], Larsson et al. [18] and Riecke et al. [32] hypothesized that the nature or interpretation of the sound source might also be able to affect auditory vection. In line with their hypothesis, they were able to demonstrate that sound sources that are typically associated with stationary objects (so-called “acoustic landmarks” like church bells) are more effective in triggering auditory circular vection than artificial sounds like pink noise or sounds that normally stem from moving objects (e.g., footsteps). These results strongly suggest the existence of higher cognitive or top-down contributions to vection, as the interpretation or meaning associated with a sound source affected the illusion. These results challenge the prevailing opinion that vection is mainly a bottom-up driven process. A more in-depth discussion of top-down and higher level influences on auditory as well as visual vection can be found in [32]. A similar benefit for using “acoustic landmarks” has recently been shown for translational vection [39]. Even non-spatialized sound was found to enhance vection if it resembled the sound of a vehicle engine [40].

Other factors that have been shown to facilitate auditory vection include the realism of the acoustic simulation and the number of sound sources [18, 32]. So far, though, there has been hardly any research on cross-modal contributions to auditory vection, and we are only aware of a study by Väljamäe et al. that showed that vibrations can enhance auditory vection [39], in line with experiments by Schulte-Pelkum et al. that showed a similar benefit of vibrations for visually-induced vection [35]. A comparable enhancement of auditory vection was observed when infrasound was added to the rotating sound sources (15Hz) [39].

Compared to visually induced vection, which is quite compelling and can even be indistinguishable from real motion [2], the auditory induced self-motion illusion is much weaker and less compelling. Furthermore, auditory vection occurs only in about 25-60% of the participants. Hence, even though auditory vection *can* occur, auditory cues alone are clearly insufficient to reliably induce a compelling self-motion sensation that could be used in applications. Therefore, the current study investigated whether *additional* spatial auditory cues can be utilized to *enhance* visually induced self-motion. Even though there is a large body of literature on visual vection, audio-visual interactions for vection have hardly if at all been investigated before.

Figure 1: ***Top:*** **360° roundshot photograph of the Tübingen market place, which was wrapped onto a cylinder to provide an undistorted view of the scene for the simulated viewpoint centered in the cylinder.** ***Bottom:*** **Participants were seated at a distance of about 1.8m from a curved projection screen (left) displaying a view of the market place (right).**

Instead of using the classic black-and-white striped patterns as vection-inducing visual stimulus – which is not really suitable for VR applications – we opted here for using a naturalistic visual stimulus that has previously been shown to be quite powerful in inducing visual vection [28].

2 Hypotheses

Two main hypotheses on how adding auditory cues could potentially facilitate visual vection were investigated in the current study:

Hypothesis 1: Influence of adding non-spatialized auditory cues: First, one might imagine that there is a rather unspecific facilitation of vection by the auditory cues increasing the overall believability of the simulation and the resulting presence and involvement in the simulated scene, independent of the spatial content of the auditory cues. To address this issue, we compare a no-sound condition with a simple mono rendering of an auditory landmark in the scene (the sound of the fountain on the market place scene that was used as the visual stimulus).

Hypothesis 2: Influence of adding spatialized acoustic landmarks: Second, the spatial content of the auditory simulation could directly enhance vection by providing additional information about the spatial location of an acoustic landmark and hence the current orientation of the observer. This hypothesis was tested by comparing the above-mentioned mono-condition with a proper spatialized acoustic rendering of the correct location of the landmark using a generic head-related transfer function (HRTF). Furthermore, the simulation might appear more realistic in the spatialized condition, as the acoustic landmark should appear properly externalized and spatialized. This might also increase overall believability and presence in the simulated scene [11, 24, 38].

3 Methods

Twenty naive participants (eight male) took part in this experiment and were paid at standard rates[1]. All

[1] A subset of the experimental conditions with a smaller number of participants has previously been presented in an overview talk at the IEEE VR 2005 conference in Bonn [31].

participants had normal or corrected-to-normal vision and were able to locate the spatialized sound source without any problems.

3.1 Stimuli and Apparatus

Participants were comfortably seated at a distance of 1.8m from a curved projection screen (2m curvature radius) on which the rotating visual stimulus was displayed (see Fig. 1, bottom). The visual stimulus consisted of a photorealistic view of the Tübingen market place that was generated by wrapping a 360˚ roundshot (4096 × 1024 pixel) around a virtual cylinder (see Fig. 1, top). The simulated field of view (FOV) was set to 54˚×45˚and matched the physical FOV under which the projection screen was seen by the participants. Black curtains covered the side and top of the cabin surrounding the projection screen in order to increase immersion and block vision of the outside room. A force-feedback joystick (Microsoft force feedback 2) was mounted in front of the participants to collect the vection responses. Visual circular vection was induced by rotating the stimulus around the earth-vertical axis with alternating turning direction (left/right). Auditory cues were displayed using active noise-canceling headphones (Sennheiser HMEC 300) that participants wore throughout the experiment. Active noise cancellation was applied throughout the experiment to eliminate auditory cues from the surrounding room that could have interfered with the experiment.

In the spatialized auditory condition, a generic HRTF and a Lake DSP system (Huron engine) with multiscape rendering were used. Note that in the spatialized auditory condition, the fountain sound was always audible (as we have omni-directional hearing), even when the visual counterpart was outside of the current field of view. Participants perceived the spatialized fountain sound properly externalized and associated it readily with the visual counterpart as intended. None of the participants commented on any mismatch between the spatialized auditory cues and visual counterpart. In the mono sound condition, the sound was perceived “inside the head” as is to be expected for mono sound, and we are not aware that any participant experienced any ventriloquism effect in the sense that the moving visual stimulus might have created the illusion of a rotating sound.

3.2 Procedure and experimental design

Each participants performed 48 trials, consisting of a factorial combination of 3 auditory conditions (no sound, mono sound, HRTF-spatialized sound; these conditions were randomized within each session) × 2 turning directions (left/right; alternating) × 2 sessions × 4 repetitions of each condition. Participants were instructed to indicate the **onset of vection** by deflecting the joystick in the direction of perceived self-motion as soon as it was sensed. The amount of deflection indicated the **vection intensity**, and the time between vection onset and maximum vection (joystick deflection) reached indicated the **vection buildup time**. After each trial, participants indicated the **convincingness** of the perceived self-motion on a 0-100% rating scale (in steps of 10%) using a lever next to the joystick.

Participants started each trial by pressing a dedicated button on the joystick, which caused the static image to start rotating clockwise or counterclockwise (alternating, in order to reduce motion after-effects) around the earth-vertical axis with constant acceleration for 3s, followed by a constant velocity (30˚/s) phase. The maximum duration of constant velocity rotation was 46s, after which the stimulus decelerated at a constant rate for 3s. Stimulus motion stopped automatically once maximum joystick deflection (vection intensity) was sustained for 10s (otherwise it continued for 46s) to reduce the potential occurrence of motion sickness. Participants were asked to initiate each trial themselves to ensure that they could prepare for the next trial and paid attention to the stimulus[2].

Between trials, there was a pause of about 15 seconds to reduce potential motion aftereffects. In order to familiarize participants with the setup, a practice block containing 4 trials preceded the main experimental blocks. Furthermore, because none of the participants had experienced vection in the laboratory before, they were exposed, prior to beginning the practice block, to a vection stimulus for about 2 minutes or until they reported a strong sense of self-motion.

Overall between-subject differences in vection responses were removed using the following normalization procedure: Each data point per participant was divided by the ratio between the mean performance of that participant across all conditions and the mean of all participants across all conditions. In addition to the vection measures, **spatial presence** was assessed after the experiment using the Igroup Presence Questionnaire (IPQ) [34].

Participants were always instructed to watch the stimuli in a natural and relaxed manner, just as if looking out of the window of a moving vehicle. Furthermore, they were told to neither stare through the screen nor to fixate on any position on the screen (in order not to suppress the optokinetic reflex). Instead, they were instructed to concentrate on the central part of the projection screen.

4 Results

The vection data for the three sound conditions are summarized in Figure 2. The results of paired t-tests are indicated in the top inset of each plot. Adding mono sound increased the convincingness ratings slightly but insignificantly by about 10%. All other vection measures showed no difference between the no sound and mono sound condition.

2 This procedure is not uncommon in psychophysical studies and implies that they might have been able to anticipate vection. We are, however, not aware of any study showing that this anticipation has any detrimental effect on the resulting data. If anything, we would rather expect that it might reduce the within-subject variability or random noise, as participants could start the next trial when they were ready for it and focusing on the stimulus to be presented.

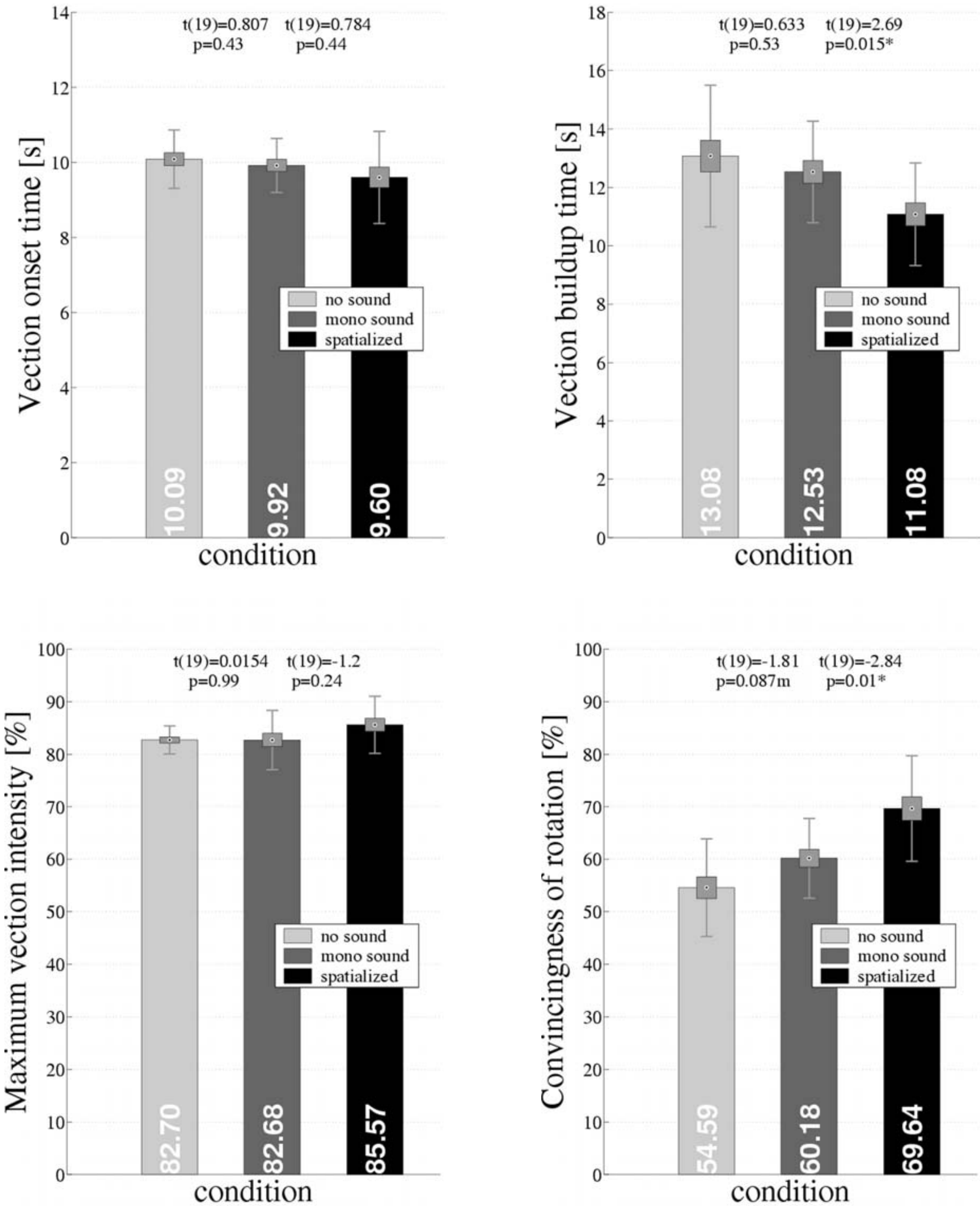

Figure 2: Mean of the four vection measures, averaged over the 20 participants. Boxes indicate one standard error of the mean, whiskers depict one standard deviation. The results of pairwise comparisons between the three sound conditions using paired t-tests are indicated in the top inset of each plot. An asterisk '*' indicates that the two conditions differ significantly from each other on a 5% level, an 'm' indicates that the difference is only marginally significant ($p < 0.1$). Note the small but consistent vection-facilitating effect of the proper spatialized auditory rendering of the fountain sound (right bars) as compared to simple mono display (middle bars). There were no significant differences between using mono sound and no sound at all.

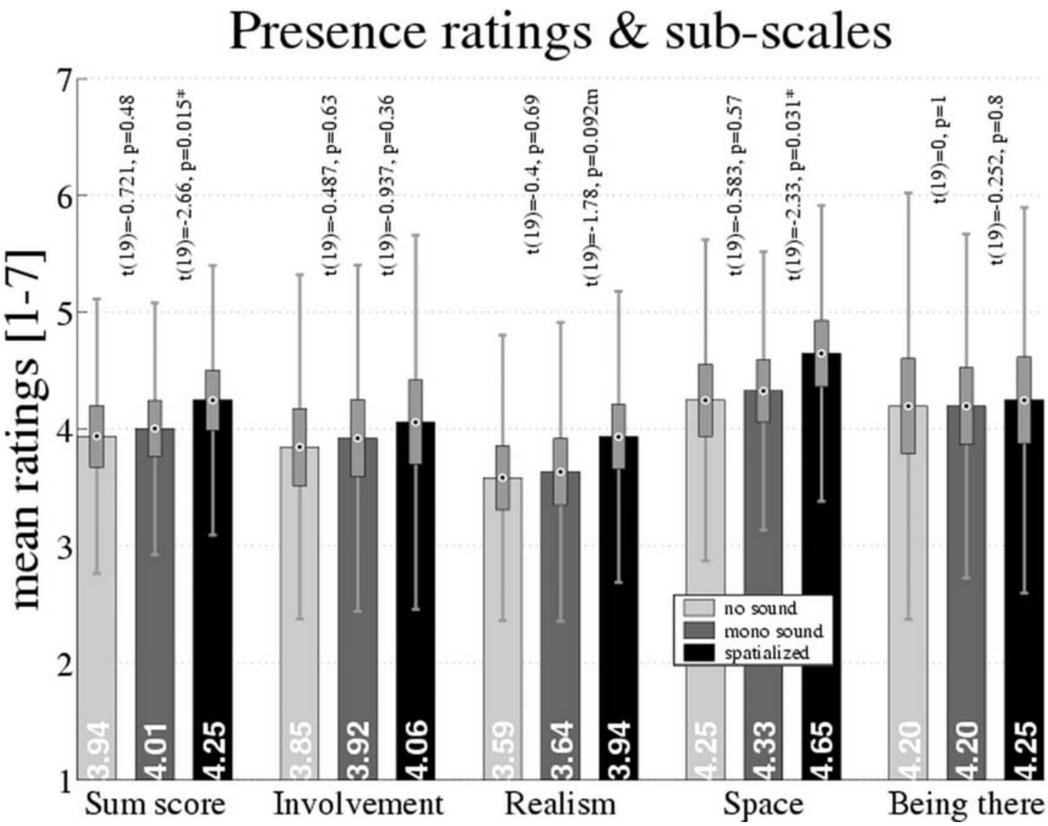

Figure 3: **Presence ratings for the three sound conditions. The sum score over all 14 items of the Igroup Presence Questionnaire (left three bars) were split up according to the four original sub-scales described by Schubert el al. [34]: "Involvement", "realism", "space", and "being there". Even though the effect size was quite small (<6%), the presence ratings were consistently higher for the spatialized sound condition.**

Comparing the mono condition with the spatialized sound condition demonstrates, however, a small but consistent vection-facilitating effect of the sound spatialization. The strongest effect was observed for the convincingness ratings (16% increase) and the vection buildup time (12% decrease). The other vection measures show only small and insignificant effects, albeit in the correct direction.

A similarly small, but consistent advantage for the spatialized sound can be observed for the presence ratings, which are summarized in Figure 3. This effect reached significance for the presence sum score and the "space" sub-scale. In addition, the "realism" sub-scale showed a marginally significant effect. The other presence sub-scales did not show any significant effects.

5 Discussion

Even though adding mono sound increased (insignificantly) the convincingness of the motion simulation by about 10%, neither the presence ratings nor any of the other vection measures were affected. That is, merely adding an audio cue that is associated with the fountain on the market place but not spatially aligned with it did not increase vection or presence significantly. This argues against an unspecific benefit of just adding audio cues. Only when the sound source was actually perceived to originate from the same location as its visual counterpart did we observe a significant facilitation of both vection and presence, which argues for a *specific* facilitation due to the spatialization of the sound source. This indicates that cross-modal consistency is indeed an important factor in improving VR simulations. This is all the more relevant as most existing VR simulations have rather poor audio quality, especially in terms of localizability of the sound sources (and externalization if headphone-based auralization is used).

As this study demonstrated, adding HRTF-based auralization using headphones can reliably be used to improve self-motion perception as well as presence in VR, even when the visual rendering is already of high quality and realism. This has many practical advantages, especially for applications where speaker arrays are unsuitable or where external noise must be excluded.

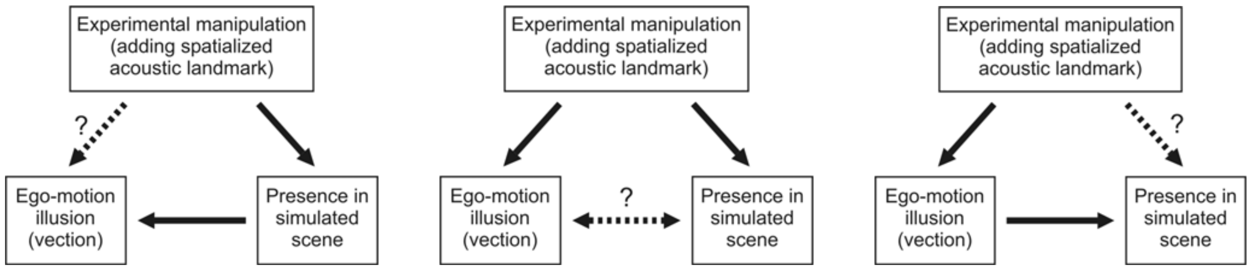

Figure 4: Schematic illustration of potential causal relations between adding the acoustic landmarks and the resulting facilitation of both vection and presence, as described in the text.

From the current data, it is, however, unclear whether there might also be a *causal* relationship or mediation between presence and vection, as is illustrated in Figure 4. On the one hand, it is conceivable that the observed increase in self-motion sensation might be mediated by the increase in presence (cf. Fig. 4, left).

A study by Riecke et al. [28] on visually induced circular vection suggests that an increase in presence might indeed be able to enhance vection: As an attempt to indirectly manipulate spatial presence without altering the physical stimulus properties too much, a photorealistic view onto a natural scene (just like in the current experiment) was compared to several globally inconsistent visual stimuli that were generated by scrambling image parts in a random manner. Thus, the stimulus could no longer be perceived as a globally consistent three-dimensional scene, which was expected to decrease spatial presence. The data showed both a decrease in presence and in vection for the globally inconsistent, scrambled stimuli. The authors suggest that higher-level factors like spatial presence in the simulated scene, global scene consistency, and/or consistent pictorial depth cues might have mediated the change in self-motion perception.

On the other hand, it is also feasible that an increase in the self-motion sensation might in some situations also be able to enhance overall presence and involvement (cf. Fig. 4, right), as suggested by Riecke et al. [27] and discussed in more detail in [30]. This seems sensible, as actual self-motions in the real world are typically accompanied by a corresponding sensation of self-motion. Hence, if self-motions simulated in VR are unable to evoke a natural percept of self-motion, the overall believability of the VR simulation and presence in the virtual environment in particular might also be affected.

In the long run, a deeper understanding of any potential causal relations between presence and the effectiveness of a simulation for a given task or goal (here: self-motion perception) would be rather helpful for optimizing VR simulations from a perceptual point of view. Further, carefully designed experiments are, however, required to tackle these issues.

In the debriefing after the experiment, participants rated the motion simulation as much more convincing when the spatialized sound was included. Nevertheless, the effect size of adding spatialized sound was rather small, both in terms of vection and rated presence. We propose two potential reasons here. First, it might reflect a ceiling effect, as the visually induced vection was already quite strong and showed relatively low onset latencies without the auditory cues. Second, auditory cues are known to be far less powerful in inducing vection than visual cues, which might explain the small effect size. Hence, we would expect a larger benefit of adding spatialized auditory cues if the auditory and visual vection inducing potential were equated in terms of their effect strength. On the one hand, the vection-inducing potential of the auditory cues could probably be increased by using more sound sources and rendering acoustic reflections and later reverberations in the simulated scene properly [18]. On the other hand, one could try to reduce the vection-inducing potential of the visual cues to a level comparable to the auditory cues by degrading the visual stimulus or by reducing the visual field of view. According to the latter, we would predict that the benefit of adding spatialized sound to VR simulations should be highest for low-cost simulators with poor image quality and/or a small field of view. Further experiments are currently being performed to investigate these hypotheses.

Apart from a specific vection-enhancing effect, adding spatialized auditory cues to VR simulations can have a number of further advantages, as is discussed in more detail in [19, 42, 37]: Adding auditory cues is known to increase presence in the simulated world, especially if spatialized auditory cues are used that are perceived as properly externalized and can be well localized, for example by using individualized HRTFs [11, 24, 38]. This is in agreement with the observed presence-facilitating effect of spatialized auditory cues in the current study. Furthermore, auditory cues provide the advantage of extending the perceivable virtual space beyond the limits of the visual field of view of the setup. This makes auditory cues perfectly suited for warning signals or for guiding attention. The omni-directional characteristics of human hearing enables us to get also a decent impression of the size and layout of a (real or simulated) scene without the need to turn our head and face the direction or object of interest

[26]. In general, whenever the corresponding situation in the real world would be accompanied with specific sounds, one would probably expect to hear those sounds in VR, too. This is of particular importance for achieving high perceptual realism in specific applications like driving and flight simulations, where adding appropriate engine sounds or environmental sounds is of crucial importance. One of the most frequent usages of audition is probably due to its clear potential to elicit emotional responses, a fact that is well-known and frequently employed by, for example, the movie industry. Last but not least, including auditory cues can also be particularly important for people who's preferred modality or cognitive style is auditory (and not visual or kinesthetic).

Hence, adding spatialized auditory cues to (predominately visual) VR simulations and ego-motion simulations in particular can have a number of advantages including an increase in the perceived self-motion. Relatively little research has been performed in this area, and additional studies are required to investigate these issues further. It is conceivable, however, that the requirements for visual rendering quality could be relaxed when appropriate simulation of the auditory modality (and potential other modalities) is provided [9]. As high quality auditory rendering can be achieved at relatively low cost, adding spatialized auditory cues might allow us in the future to increase simulation effectiveness while reducing the overall simulation effort, especially when the attention guiding potential of auditory cues is employed. Using a selective rendering approach, guiding attention has, for example, been shown to reduce computational costs of the visual rendering considerably [4, 36]. This is promising for the usage of auditory cues for optimizing VR simulations both on a computational and perceptual level.

Acknowledgements

This research was funded by the European Community (IST-2001-39223, FET Proactive Initiative, project "POEMS" (Perceptually Oriented Ego-Motion Simulation, see www.poems-project.info) and the Max Planck Society.

References

[1] E. R. Boer, A. R. Girshik, T. Yamamura, and N Kuge. Experiencing the same road twice: A driver-centred comparison between simulation and reality. In *Proceedings of the Driving Simulation Conference 2000*, Paris, France, 2000.

[2] T. Brandt, J. Dichgans, and R. Held. Optokinesis affects body posture and subjective visual vertical. *Pflugers Archiv-European Journal of Physiology*, 339:97–97, 1973.

[3] J. Burki-Cohen, T. H. Go, W. Y. Chung, J. Schroeder, S. Jacobs, and T. Longridge. Simulator fidelity requirements for airline pilot training and evaluation continued: An update on motion requirements research. In *Proceedings of the 12th International Symposium on Aviation Psychology, April 14-17*, pages 182–189, Dayton (OH), USA, 2003.

[4] K. Cater, A. Chalmers, and G. Ward. Detail to attention: Exploiting visual tasks for selective rendering. In *EGRW '03: Proceedings of the 14th Eurographics Workshop on Rendering*, pages 270–280, 2003.

[5] S. S. Chance, F. Gaunet, A. C. Beall, and J. M. Loomis. Locomotion mode affects the updating of objects encountered during travel: The contribution of vestibular and proprioceptive inputs to path integration. *Presence - Teleoperators and Virtual Environments*, 7(2):168–178, April 1998.

[6] J. Dichgans and T. Brandt. Visual-vestibular interaction: Effects on self-motion perception and postural control. In R. Held, H. W. Leibowitz, and H.-L. Teuber, editors, *Perception*, volume VIII of *Handbook of Sensory Physiology*, pages 756–804. Springer, 1978.

[7] R. Dodge. Thresholds of rotation. *Journal of Experimental Psychology.*, 6:107–137, 1923.

[8] F. H. Durgin and A. Pelah. Self-motion perception during locomotor recalibration: More than meets the eye. *Journal of Experimental Psychology: Human Perception and Performance*, in press.

[9] N. I. Durlach and A. S. Mavor, editors. *Virtual reality: Scientific and technological challenges*. National Academy Press, 1995.

[10] M. H. Fischer and A. E. Kornmüller. Optokinetisch ausgelöste Bewegungswahrnehmung und optokinetischer Nystagmus [Optokinetically induced motion perception and optokinetic nystagmus]. *Journal für Psychologie und Neurologie*, pages 273–308, 1930.

[11] C. Hendrix and W. Barfield. The sense of presence within auditory virtual environments. *Presence – Teleoperators and Virtual Environments*, 5(3):290–301, 1996.

[12] P. E. Hennebert. Audiokinetic nystagmus. *Journal of Auditory Research*, 1(1):84–87, 1960.

[13] L. J. Hettinger. Illusory self-motion in virtual environments. In Kay M. Stanney, editor, *Handbook of Virtual Environments*, chapter 23, pages 471–492. Lawrence Erlbaum, 2002.

[14] J. M. Hollerbach. Locomotion interfaces. In Kay M. Stanney, editor, *Handbook of Virtual Environments*, chapter 11, pages 239–254. Lawrence Erlbaum, 2002.

[15] I. P. Howard. The perception of posture, self motion, and the visual vertical. In K. R. Boff, L. Kaufman, and J. P. Thomas, editors, *Sensory processes and perception*, volume 1 of *Handbook of human perception and performance*, pages 18.1–18.62. Wiley, New York, 1986.

[16] B. Kapralos, D. Zikovitz, M. Jenkin, and L.R. Harris. Auditory cues in the perception of self-motion. In *Proceedings of the 116th AES convention, Berlin 2005*, Berlin, Germany, 2004.

[17] J. R. Lackner. Induction of illusory self-rotation and nystagmus by a rotating sound-field. *Aviation Space and Environmental Medicine*, 48(2):129–131, 1977.

[18] P. Larsson, D. Västfjäll, and M. Kleiner. Perception of self-motion and presence in auditory virtual environments. In *Proceedings of Seventh Annual Workshop Presence 2004*, pages 252–258, 2004. Available: www.kyb.mpg.de/publication.html?publ=2953.

[19] P. Larsson, A. Väljamäe, D. Västfjäll, and M. Kleiner. Auditory-induced presence in mediated environments and related technology. In *Handbook of Presence*. Lawrence Erlbaum, 2005. (submitted).

[20] K. Lowther and C. Ware. Vection with large screen 3d imagery. In *ACM CHI '96*, pages 233–234, 1996.

[21] E. Mach. *Grundlinien der Lehre von den Bewegungsempfindungen*. Leipzig, Germany: Engelmann, 1875.

[22] A. M. Marmekarelse and W. Bles. Circular vection and human posture ii: Does the auditory-system play a role. *Agressologie*, 18(6):329–333, 1977.

[23] M. Mulder, M. M. van Paassen, and E. R. Boer. Exploring the roles of information in the control of vehicular locomotion - from kinematics and dynamics to cybernetics. *Presence - Teleoperators and Virtual Environments*, 13:535–548, 2004.

[24] K. Ozawa, Y. Chujo, Y. Suzuki, and T. Sone. Psychological factors involved in auditory presence. *Acoustical Science and Technology*, 24:42–44, 2004.

[25] S. Palmisano. Consistent stereoscopic information increases the perceived speed of vection in depth. *Perception*, 31(4):463–480, 2002.

[26] J. Pope and A. Chalmers. Multi-sensory rendering: Combining graphics and acoustics. In *Proceedings of the 7th International Conference in Central Europe on Computer Graphics*, pages 233–242, Czech Republic, 1999.

[27] B. E. Riecke, J. Schulte-Pelkum, M. N. Avraamides, and H. H. Bülthoff. Enhancing the visually induced self-motion illusion (vection) under natural viewing conditions in virtual reality. In *Proceedings of Seventh Annual Workshop Presence 2004*, pages 125–132, 2004. Available: www.kyb.mpg.de/publication.html?publ=2864.

[28] B. E. Riecke, J. Schulte-Pelkum, M. N. Avraamides, M. von der Heyde, and H. H. Bülthoff. Influence of scene consistency and spatial presence on circular vection in virtual reality. In *ACM SIGGRAPH Symposium on Applied Perception in Graphics and Visualization* (APGV), 2005. (accepted).

[29] B. E. Riecke, J. Schulte-Pelkum, and H. H. Bülthoff. Perceiving simulated ego-motions in virtual reality - comparing large screen displays with HMDs. In *SPIE - Invited paper on VALVE: Vision, Action, and Locomotion in Virtual (and Real) Environments*, San Jose, CA, USA, 2005. Available: www.kyb.mpg.de/publication.html?publ=3233.

[30] B. E. Riecke, J. Schulte-Pelkum, and Franck Caniard. Using the perceptually oriented approach to optimize spatial presence & ego-motion simulation. In *Handbook of Presence*. Lawrence Erlbaum, 2005. (submitted).

[31] B. E. Riecke, J. Schulte-Pelkum, Franck Caniard, and H. H. Bülthoff. Towards lean and elegant self-motion simulation in virtual reality. In *Proceedings of IEEE VR2005*, pages 131–138, Bonn, Germany, 2005. www.vr2005.org.

[32] B. E. Riecke, Daniel Västfjäll, Pontus Larsson, and J. Schulte-Pelkum. Top-down and multi-modal influences on self-motion perception in virtual reality. In *HCI international 2005 (accepted)*, Las Vegas, NV, USA, 2005. www.hci-international.org.

[33] S. Sakamoto, Y. Osada, Y. Suzuki, and J. Gyoba. The effects of linearly moving sound images on self-motion perception. *Acoustical Science and Technology*, 25:100–102, 2004.

[34] T. Schubert, F. Friedmann, and H. Regenbrecht. The experience of presence: Factor analytic insights. *Presence - Teleoperators and Virtual Environments*, 10(3):266–281, June 2001.

[35] J. Schulte-Pelkum, B. E. Riecke, and H. H. Bülthoff. Vibrational cues enhance believability of ego-motion simulation. In *International Multisensory Research Forum (IMRF)*, 2004. Available: www.kyb.mpg.de/publication.html?publ=2766.

[36] V. Sundstedt, K. Debattista, and A. Chalmers. Selective rendering using task-importance maps. In *ACM SIGGRAPH Symposium on Applied Perception in Graphics and Visualization (APGV)* (p. 175), 2004.

[37] A. Väljamäe, A. Kohlrausch, S. van de Par, D. Västfjäll, P. Larsson, and Kleiner M. Audio-visual interaction and synergy effects: implications for cross-modal optimization of virtual and mixed reality applications. In *Handbook of Presence*. Lawrence Erlbaum, 2005. (submitted).

[38] A. Väljamäe, P. Larsson, D. Västfjäll, and Kleiner M. Auditory presence, individualized head- related transfer functions, and illusory ego-motion in virtual environments. In *Proceedings of Seventh AnnualWorkshop Presence 2004*, pages 141–147, Valencia, Spain, 2004. Available: www.kyb.mpg.de/publication.html?publ=2954.

[39] A. Väljamäe, P. Larsson, D. Västfjäll, and Kleiner M. Effects of vibratory stimulation on auditory induced self-motion. In *Proceedings of IMRF 2005*, Rovereto, Italy, 2005. Poster presented at IMRF 2005; Available: www.kyb.mpg.de/publication.html?publ=3379.

[40] A. Väljamäe, P. Larsson, D. Västfjäll, and Kleiner M. Sonic self-avatar and self-motion in virtual environments. In *Proceedings of the 8th Annual Workshop of Presence, London 2005*, London, England, 2005. (submitted).

[41] A. Väljamäe, P. Larsson, D. Västfjäll, and Kleiner M. Travelling without moving: Auditory scene cues for translational self-motion. In *Proceedings of ICAD 2005*, Limerick, Ireland, 2005.

[42] A. Väljamäe, D. Västfjäll, P. Larsson, and Kleiner M. Perceived sound in mediated environments. In *Handbook of Presence*. Lawrence Erlbaum, 2005. (submitted).

[43] R. Warren and A. H. Wertheim, editors. *Perception & Control of Self-Motion*. Erlbaum, New Jersey, London, 1990.

[44] S. C. P. Wong and B. J. Frost. The effect of visual-vestibular conflict on the latency of steady-state visually induced subjective rotation. *Perception & Psychophysics*, 30(3):228–236, 1981.