Towards Lean and Elegant Self-Motion Simulation in Virtual Reality

Bernhard E. Riecke* Jörg Schulte-Pelkum† Franck Caniard‡ Heinrich H. Bülthoff§

Max Planck Institute for Biological Cybernetics, Tübingen, Germany

ABSTRACT

Despite recent technological advances, convincing self-motion simulation in Virtual Reality (VR) is difficult to achieve, and users often suffer from motion sickness and/or disorientation in the simulated world. Instead of trying to simulate self-motions with physical realism (as is often done for, e.g., driving or flight simulators), we propose in this paper a perceptually oriented approach towards self-motion simulation. Following this paradigm, we performed a series of psychophysical experiments to determine essential visual, auditory, and vestibular/tactile parameters for an effective and perceptually convincing self-motion simulation. These studies are a first step towards our overall goal of achieving lean and elegant self-motion simulation in Virtual Reality (VR) without physically moving the observer. In a series of psychophysical experiments about the self-motion illusion (circular vection), we found that (i) vection as well as presence in the simulated environment is increased by a consistent, naturalistic visual scene when compared to a sliced, inconsistent version of the identical scene, (ii) barely noticeable marks on the projection screen can increase vection as well as presence in an unobtrusive manner, (iii) physical vibrations of the observer's seat can enhance the vection illusion, and (iv) spatialized 3D audio cues embedded in the simulated environment increase the sensation of self-motion and presence. We conclude that providing consistent cues about self-motion to multiple sensory modalities can enhance vection, even if physical motion cues are absent. These results yield important implications for the design of lean and elegant self-motion simulators.

CR Categories: H.1.2 [Models and Principles]: User/Machine Systems—Human factors, Human information processing H.5.1 [Information Interfaces and Presentation, (e.g. HCI]: Multimedia Information Systems—Artificial, augmented, and virtual realities H.5.2 [Information Interfaces and Presentation, (e.g. HCI]: User Interfaces—Input devices and strategies, Interaction styles J.4 [Social and Behavioral Sciences]: Psychology

Keywords: Vection, self-motion illusion, spatial orientation, virtual reality, motion simulation, human factors, psychophysics, multi-modal cue integration, vibrations, auditory cues

1 INTRODUCTION

Over the last decades, Virtual Reality (VR) technology has been used increasingly in research as well as industrial and consumer market applications. One area that has received more and more interest is self-motion simulation, often with a specific focus such as driving, flying, or architecture walkthroughs. Low-end self-motion simulators consist mainly of a monitor and sometimes simple stereo sound renderings. These motion simulations, however, generally lack convincingness and users get lost in the virtual environment rather easily. Typical high-end self-motion simulation setups generally include a more immersive visualization device (e.g., projection screen or head-mounted display) that is used in combination with a motion platform that can physically move the observer in up to six degrees of freedom. Such motion platforms are, however, typically rather large, expensive, and require a high degree of technical expertise and support. Despite the extensive effort associated with motion simulators, users are still easily disoriented in the simulated environment and often confronted with side-effects like motion sickness.

Hence, something essential seems to be missing in those traditional approaches. In this paper, we will highlight some recent results that followed a different approach: Instead of focusing on displaying motion cues with physical realism (as is often being done for, e.g., high-end driving or flight simulators), we were concerned with the *perceptual* convincingness and effectiveness of a motion simulation. That is, by performing psychophysical experiments that investigate the individual contributions and interactions of different cues and sensory modalities, we hope to eventually be able to devise an alternative approach to self-motion simulation. Ultimately, we aim to achieve lean and elegant, and above all, effective and convincing self-motion simulation, ideally without even moving the observer physically. This goal is being tackled within an ongoing EU project on Perceptually Oriented Ego-Motion Simulation ("POEMS", IST-2001-39223, see www.poems-project.info). The main idea of the POEMS project is to provide users only with the essential cues about self-motion, relying heavily on multi-modal presentation. To achieve this, we need to define the perceptually relevant rendering parameters for effective self-motion simulation as well as enhanced spatial presence[1] in the simulated environment. Our long-term aim is to enable developers and designers of VR systems to optimize their simulations both technically and perceptually, thus allowing for lean, elegant, and low-cost VR simulations with a high sense of spatial presence and effective self-motion simulation. This is a prerequisite for the usability of VR for, e.g., training purposes.

In this paper, we will provide an overview of some of the first findings of the POEMS project. Based on a recent theoretical framework for spatial orientation processes [16], we followed the proposition that convincing self-motion perception as well as sufficient spatial presence are essential, necessary prerequisites for robust and effortless spatial orientation in VR. Accordingly, these first experiments were mainly concerned with the subjective convincingness and effectiveness of illusory self-motions as well as on spatial presence in VR. Most people know the phenomenon of illusory self-motion ("vection") from real-world experience: When sitting in a stationary train, one can have the convincing impression that one has begun moving even though it was in fact the train on the adjacent track that just started to move. Here, we report a selection of four psychophysical experiments that investigated relevant stimulus parameters for the visual, tactile, and auditory modality that influence the perception of illusory self-motion (vection) in a VR setup.

*e-mail: bernhard.riecke@tuebingen.mpg.de
†e-mail: joerg.schulte-pelkum@tuebingen.mpg.de
‡e-mail: franck.caniard@tuebingen.mpg.de
§e-mail: heinrich.buelthoff@tuebingen.mpg.de

[1]We follow here Witmer & Singer's definition of presence as "the subjective experience of being there in one place or environment, even when one is physically situated in another" [21].

IEEE Virtual Reality 2005
March 12-16, Bonn, Germany
0-7803-8929-8/05/$20 ©2005 IEEE

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 displaying either a consistent view of the market place (middle) or a sliced version of the same scene (right).

These vection experiments are the basis for later experiments, that will investigate if the previously found self-motion simulation parameters also allow for robust, natural, and effortless spatial orientation in the simulated world - an issue that is still largely unsolved. As rotations seem to be harder to simulate and to spatially update than translations (e.g., [9]), only upright rotations were investigated in the current study as a first step. Further research is planned to investigate how these results compare to translations and combined rotations and translations.

When stationary observers are exposed to a large moving visual stimulus, they can experience a compelling illusion of self-motion: At first, they will "correctly" perceive the motion of the visual stimulus (object motion). After a few seconds, however, this perception typically shifts towards oneself being moved and the moving visual stimulus appearing earth-stationary. This self-motion illusion is referred to as vection, and has been studied extensively for more than a century. A good overview is provided in [2, 6]. Vection has typically been studied using a rotating optokinetic drum that is painted with simple geometrical patterns like black and white vertical stripes. Typically, the onset latency until the observer perceives vection is measured, as well as the duration and speed of illusory self-motion. The prevailing explanation for visually induced illusory self-motion perception (vection) is that the illusion arises from bottom-up perceptual processes. Therefore, past research has focused primarily on examining how low-level physical parameters of the visual stimulus (contrast, spatial frequency distribution, velocity etc.) affect vection.

In this paper, we would like to focus instead on top-down influences and parameters that are more closely related to VR applications, issues that have to a large degree been neglected previously. Questions addressed include the relation between vection and presence in the simulated environment, and the effect of providing consistent multi-sensory cues about self-motion. For the sake of simplicity and space limitations, we would like to focus on just one key issue from each of the four experiments, even though the experiments include more conditions.

The first experiment is concerned with the influence of scene consistency and the notion of "presence" in the simulated scene on the visually induced self-motion illusion (section 2). The second experiment investigated whether subtle modifications of the projection screen used to present the moving stimulus can enhance the self-motion illusion (section 3). This experiment has already been described in more detail in [15]. The third and forth experiment examined the influence of multi-modal stimulus presentation using additional vibrations (experiment 3, section 4) and 3D auditory cues (experiment 4, section 5).

2 Experiment 1: Influence of top-down factors (scene consistency and presence) on vection

To address the influence of top-down (i.e., cognitive) process on visually-induced ego-motion perception, global scene consistency[2], and thereby spatial presence in the simulated scene, was manipulated in the first experiment by presenting either a photorealistic image of the Tübingen market place or a sliced (inconsistent) version of the same stimulus (see Fig. 1). Global scene consistency was expected to increase the believability of visual stimuli (top-down effect), as it allows for locomotion and spatial presence in the simulated scene. Conversely, scene slicing reduces the global scene consistency in the sense that individual objects are still easily recognizable, but their spatial arrangement is nonsensical and inconsistent with the real world, which is in turn expected to decrease spatial presence. Note, however, that the physical stimulus properties (bottom-up factors) are hardly altered in the sliced condition. That is, any consistent effect of scene slicing on vection should consequently be attributed to top-down effects, and might be mediated by spatial presence.

[2]Global "scene consistency" refers here the coherence of a scene layout that is consistent with our natural environment, where, e.g., houses are not floating in mid-air, and a market place consists of houses not jumbled-up or upside-down, but arranged meaningfully around an open place.

2.1 Methods

Twelve naive participants took part in this experiment and were paid at standard rates. All participants had stereo vision and normal or corrected-to-normal vision. For all of the experiments reported in this paper, participants were comfortably seated at a distance of 1.8m from a curved projection screen (1400×1050 pixel resolution) on which the rotating visual stimulus was displayed (see Fig. 1, bottom). The visual stimulus consisted either of a photorealistic view of the Tübingen market place that was generated by wrapping a 360° roundshot of 4096×1024 pixel around a virtual cylinder, or of a sliced version thereof that was generated by slicing the image horizontally and reassembling it in an inconsistent manner (see Fig. 1). The simulated FOV matched the physical FOV of $54^\circ \times 40.5^\circ$. Circular vection was induced by rotating the stimulus around the earth-vertical axis. Visibility of the surrounding room was prevented using black curtains and switching off the room light. To mask spatial auditory cues, the sound of several layers of flowing water was played through active noise-canceling headphones that participants wore throughout the experiment. Responses were collected using a joystick that was mounted in front of the participants at a comfortable distance.

Participants started each trial by pressing a button on the joystick, at which time the static image smoothly started rotating clockwise or counterclockwise (alternating across trials) around the earth-vertical axis. After a 3s acceleration phase, the stimulus kept rotating with a constant velocity (40°/s) for up to 60s, followed by a 6s deceleration phase. Participants were instructed to indicate the onset of their perceived self-motion by pulling the joystick in the direction of perceived self-motion.

Four dependent measures were used to quantify vection: The time interval between the onset of stimulus rotation and the first deflection of the joystick indicated the **vection onset time** and was the primary dependent measure. Participants were also asked to deflect the joystick more the stronger the perceived self-motion intensity became. This continuous recording allowed us to measure the maximum **vection intensity** (joystick deflection) as well as the time between vection onset and maximum vection ("**vection buildup time**") reported by the participant in each trial. Thus, vection facilitation could be quantified as a reduction of vection onset time and the time between vection onset and maximum vection, and/or an increase in reported vection intensity. The rotation of the stimulus stopped automatically once maximum joystick deflection (vection intensity) was sustained for 10s (otherwise it lasted for up to 60s) to reduce the potential occurrence of motion sickness. Finally, at the end of each trial participants were asked to provide a "**convincingness rating**" of perceived self-motion by moving a lever next to the joystick to select one of the 11 possible values of a 0-100% rating scale. The value of 0% corresponded to no perceived motion at all (i.e., perception of a rotating visual stimulus and a stationary self) and that of 100% to a very convincing sense of vection (i.e., perception of a earth-stationary stimulus and a rotating self).

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 a standard presence questionnaire by Schubert, Friedmann, & Regenbrecht [18].

A repeated-measures, within-subject design was used with randomized presentation order of the intact and sliced stimulus, and two repetitions per condition. Furthermore, the turning direction (left/right) was alternated to reduce adaptation and motion-aftereffects. Participants were always instructed to watch the stimuli in a natural and relaxed manner. 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.

2.2 Results

The behavioral data for the four dependent variables are summarized in Figure 2. As predicted, global scene consistency improved vection consistently: For the unsliced stimulus, both vection onset time and the time between vection onset an maximum vection ("vection buildup time") were reduced. Furthermore, rated vection intensity and convincingness of the self-motion were significantly increased. A comparable benefit for the consistent scene was found for the presence ratings (Fig. 3). The overall presence sum score as well as three of the four subscales showed a significant increase in reported presence for the intact (consistent) stimulus, as compared to the sliced (inconsistent) stimulus.

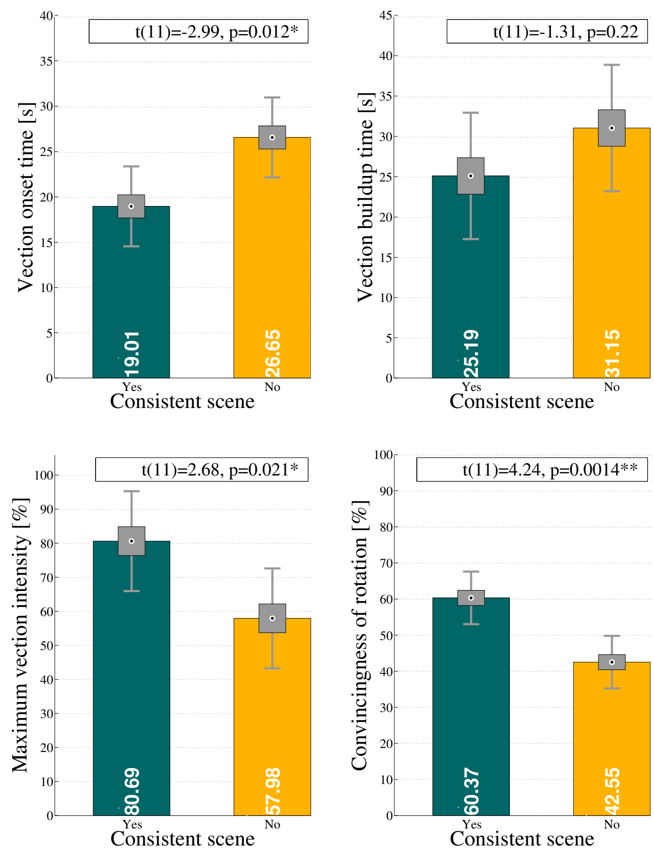

Figure 2: Mean performance for experiment 1, averaged over the 12 participants. Boxes and whiskers depict one standard error of the mean and one standard deviation, respectively. The results of paired, two-tailed t-test are indicated in the top inset. Note the consistent vection-facilitating effect of global scene consistency (left bars).

2.3 Discussion

Since the scene slicing left the physical stimulus properties and image statistics mostly unchanged, we believe that the considerable vection-degrading effect of the slicing cannot be fully explained by low-level bottom-up influences. Instead, we posit that top-down factors mediated the effect. Such factors might include presence in

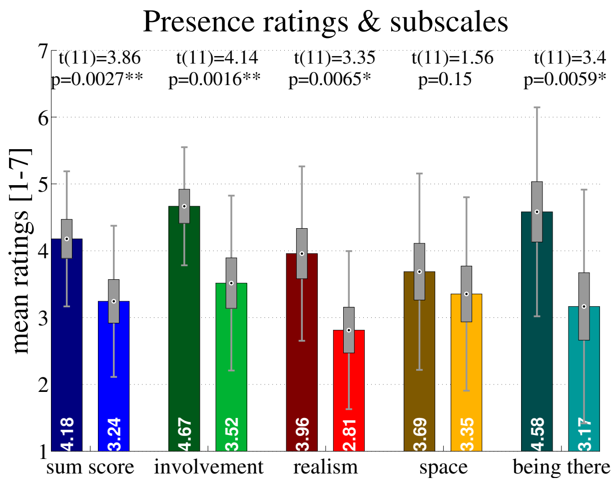

Figure 3: Mean presence ratings and subscales for the globally consistent scene (darker bars) and inconsistent (sliced) scene (lighter bars). The mean sum score over all 14 items (left pair of bars) is split up into four subscales: involvement/attention, realism, space, and being there. The statistical comparison of the two conditions using paired, two-tailed t-test are displayed above each pair of bars. Note the overall higher presence ratings for the globally consistent scene.

the simulated scene, globally consistent depth cues, and/or the implied possibility of traveling the scene. According to the "presence hypothesis" put forth by Prothero and colleagues [14], "the sense of presence in the environment reflects the degree to which that environment influences the selected rest frame." As the intact, naturalistic market scene yielded higher presence, one could hypothesize that it might accordingly more readily be accepted as a stable, primary rest frame with respect to which motions are being judged. This, in turn, might be able to explain the vection-enhancing effect of presenting a consistent, naturalistic scene.

Even though we cannot pinpoint the exact nature of these top-down influences, and there are always alternative explanations conceivable, this experiment supports the hypothesis that top-down or cognitive influences do play a considerable role in self-motion perception and should in the future be taken more into consideration both in fundamental research (where they have mostly been neglected) and in motion simulation applications.

3 Experiment 2: Subtle facilitation of vection through minor modifications of the projection screen

Vection is known to be facilitated by both static fixation points [1, 4] and foreground stimuli that are perceived to be stationary in front of a moving background stimulus [7, 13], whereas stationary objects behind the moving objects impair vection [8]. Fixating on a static foreground stimulus while moving, however, is rather unnatural and cumbersome. When driving a real or simulated vehicle, it might even be dangerous to pay attention to, for example, some dirt on the windshield (acting as a fixation point) instead of the path you want to follow. Here, we will present preliminary observations that subtle scratches in the periphery of the projection screen can have similar vection-facilitating effects, even under relaxed viewing conditions and without fixating or suppressing the optokinetic reflex. This could be exploited in driving or flight simulators by including for example subtle spots or dirt on the (real or simulated) windshield.

3.1 Methods

Methods are comparable to the first experiment, apart from the differences described below. Twenty-two naive participants were randomly assigned to one of two groups: Twelve participants were presented with the unmarked screen (just as in experiment 1), and ten participants were presented with the marked screen. For the latter condition, a different projection screen of identical size was used that contained subtle scratches in the periphery of the projection screen (see Fig. 4). Marks were located at the upper left part of the screen, and they were unobtrusive to the extent that, in fact, only one of the participants was able to report having noticed them in a post-experimental interview. Apart from these marks, size, material, and reflection properties of the two screens were identical. Hence, any difference in results between the two screens should be attributed to the minor scratches on the screen. Due to the between-subject experimental design, between-subject (unpaired) two-tailed t-tests were used (cf. Fig. 5 and 6).

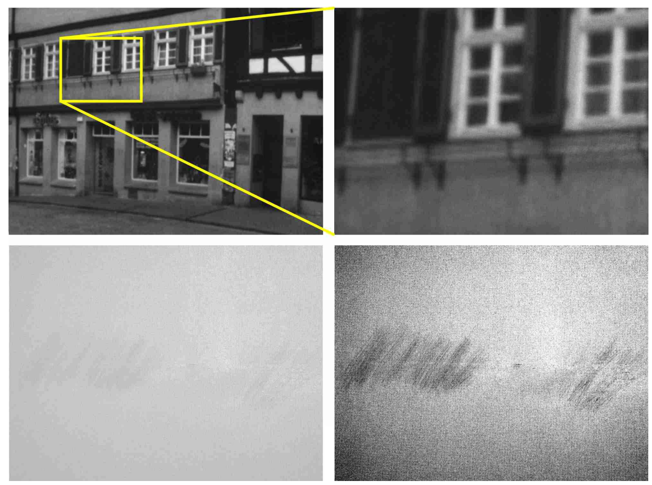

Figure 4: **Top left:** View of the projection screen displaying the market scene. The marks are located at the upper-left part of the screen, as illustrated by the close-ups to the right and below. **Bottom:** Close-up of the same region as above (right), but illuminated with plain white light to illustrate the marks. **Left:** The original photograph demonstrating the unobtrusive nature of the marks (diagonal scratches). **Right:** Contrast-enhanced version of the same image to illustrate the marks.

3.2 Results

Figure 5 reveals a considerable and highly significant vection-enhancing influence of the marks on the screen for all four dependent measures. Vection onset time and vection buildup time were reduced by more than a factor of two, and vection intensity and convincingness ratings were close to ceiling level. Furthermore, the marks on the screen considerably increased presence ratings as well (Fig. 6). This was found for both the overall presence sum score and all four subscales.

3.3 Discussion

This experiment demonstrated that simply adding marks in the periphery of the projection screen can reliably induce convincing self-motion illusions with quick vection onset in a non-obtrusive way, without any explicit fixation and under natural, relaxed viewing conditions. The magnitude of the effect is rather striking and comparable to results obtained by a fixation point in traditional studies

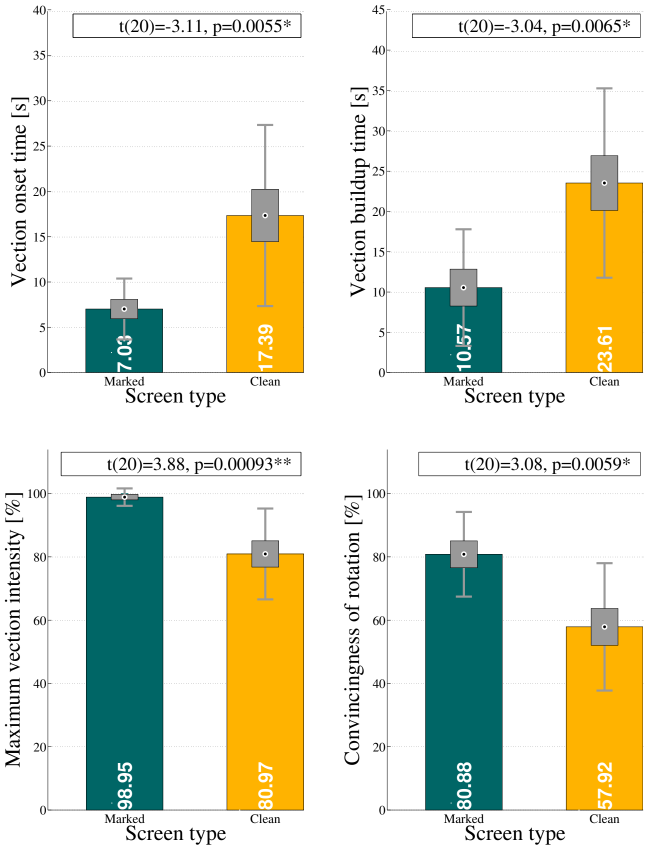

Figure 5: Mean performance for experiment 2, averaged over the ten and twelve participants. Note the considerable vection-facilitating effect of the additional marks on the screen (left bars) for all four dependent variables.

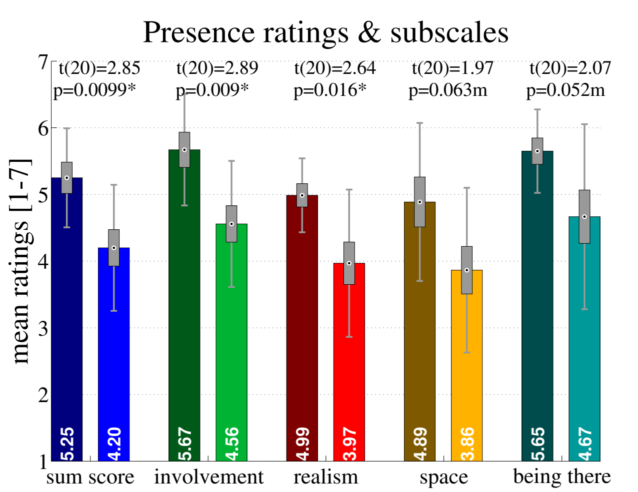

Figure 6: Mean presence ratings and sub-scales for the marked screen (darker bars) and unmarked screen (lighter bars). Note the consistently higher presence ratings for the marked screen.

using full-field stimulation in an optokinetic drum: Becker et al. [1] found for example a reduction of vection onset latencies from 30s without fixation to 10s with fixation at a rotational velocity of 30°/s. Further experiments with more carefully controlled parameters are needed, however, to corroborate the preliminary observations of the present study.

A related study by Lowther and Ware [12] demonstrated a consistent vection-facilitating effect when a 5×5 grid was overlayed on a large flat projection screen that was used to present the moving stimuli. Note that the grid in that study was clearly visible to the participants and extended over the whole screen. In the current study, however, the experimental manipulation was much more subtle. In fact, only one of the participants in the post-experimental interview was able to report the imperfections of the projection screen. Nevertheless, the hardly visible marks facilitated vection consistently in all dependent variables, and the effect size was even stronger than for Lowther and Ware's clearly visible grid that extended over the whole screen [12].

The additional presence-enhancing effect of the scratches on the screen was rather unexpected. It is not clear why scratches on the screen should *directly* increase presence or involvement in the simulation. In fact, one might expect a presence decline due to the degradation of the simulation fidelity due to the scratches. Nevertheless, the additional marks on the screen did significantly increase presence and even involvement. We posit that this effect might be attributed to the dynamical component of the visual stimulus, in the sense that the increase in the self-motion illusion might have caused or mediated the increase in presence and involvement.

In many motion simulation applications, convincing self-motion is desired without restricting user's eye or head movements unnaturally. As this experiment suggests, adding for example dirt or stains on the (real or simulated) windshield of a motion simulator should reliably increase the convincingness and strength of the self-motion illusion without imposing unnatural constraints on users' behavior. The effect should be even stronger when stereoscopic depth cues separate the windshield from the background [12, 7, 13]. The underlying factors are, however, not fully understood yet. We can only speculate that the additional marks on the screen might have provided some kind of stable reference frame with respect to which the moving stimulus is being judged.

4 Experiment 3: Multi-sensory cue integration - Can vibrations enhance the visually induced self-motion illusion?

The above two studies demonstrated that the visually induced perception of illusory self-motion (vection) can be reliably enhanced by both top-down factors and visual display parameters. Experiments 3 and 4 investigated whether non-visual modalities like vibrations and audition, respectively, can also contribute to the illusion. Vibrations might enhance vection by two different mechanisms: On the one hand, they might affect the weighting between the visual and vestibular/proprioceptive modality, since adding vibrational noise renders the vestibular/proprioceptive system less reliable [3]. This, in turn, should decrease the visuo-vestibular cue conflict and thereby facilitate the visually-induced self-motion illusion. On the other hand, vibrations might make the visually simulated motion appear more realistic or convincing - as most physical motions are not perfectly smooth but accompanied by some kind of vibration. This experiment was not designed to distinguish between these two different mechanisms, but to investigate whether there is any affect at all of simply adding vibrations to the participant's seat and foot plate. Even though vibrations are commonly used in many motion simulation applications such as fun-rides in theme parks, we were not aware of any published research on this issue.

4.1 Methods

A repeated-measures, within-subject design with 24 participants was used in the current study. Turning direction was alternated, and the order of the trials with and without additional vibrations was randomized. In half of the trials, low-frequency broad-band vibrations were applied to the participants' seat and floor plate during the visual motion phase using special force transducers (shakers). Vibration frequencies ranged from 15-90Hz, and vibration amplitude was set to a clearly perceivable level. Maximum visual stimulus velocity was set to 30°/s, acceleration/deceleration time was 12s, and each condition was repeated four times. Apart from this, the methods and experimental procedures were comparable to the first experiment.

4.2 Results and discussion

Figure 7 shows a small but significant contribution of the added vibrations: Vection onset times and vection buildup times were significantly reduced and convincingness ratings significantly increased. Hence, we conclude that adding vibrations did facilitate vection, even though just simple (on/off) vibrations were used that had no relation whatsoever to movement velocity or direction. The observed effect indicates multi-sensory interactions in the illusory perception of self-motion, which has important implications for both the theory of self-motion perception and for motion simulator design.

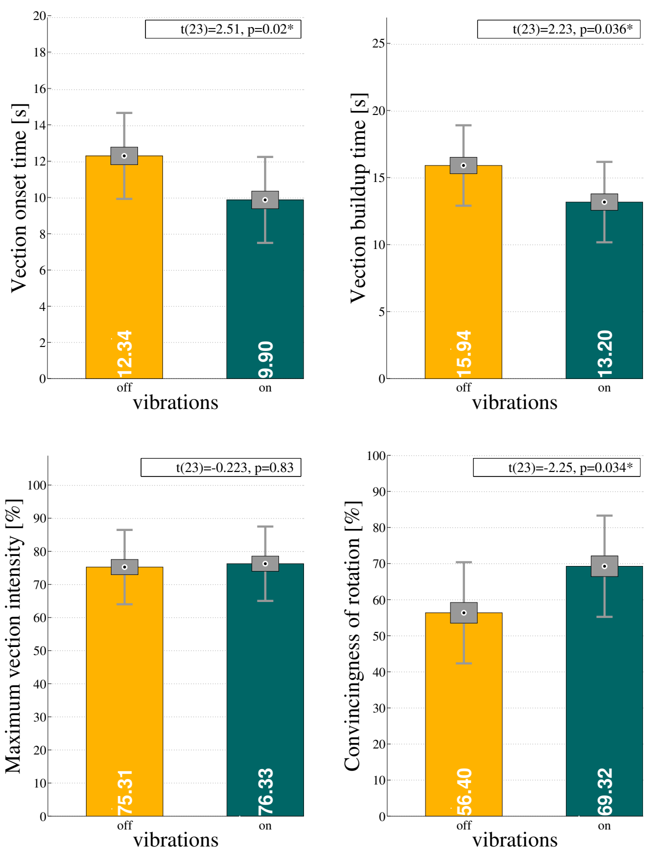

Figure 7: Mean performance for experiment 3, averaged over the 24 participants. Note that vibrations significantly reduced vection onset and buildup times and increased convincingness ratings, even though the effect size was rather small.

Post-experimental interviews revealed that even the simple vibrations used rendered the motion more realistic. Most participants reported that they had imagined themselves being on the simulated market scene, and that the vibrations supported the interpretation of moving in some vehicle or turntable in that scene. Without the vibrations, this interpretation was not supported. This suggests that cognitive (top-down) factors might have contributed to the vection-enhancing effect of vibrations.

Interestingly, three participants reported that the vibrations did not match the velocity profile of the visual motion. They would have expected vibration strength and/or frequency to increase with visual stimulus velocity. For these three participants, vibrations actually decreased vection, and they reported that the conflict made the self-motion illusion less realistic and convincing. Taken together, this suggests that more refined and meaningful vibrations that better match the motion and motion metaphor (e.g., riding a car) might enhance vection even more. This will be investigated in future experiments, and could have important implications for all motion simulation where a convincing self-motion sensation and good spatial orientation are required (like, e.g., architecture walk-throughs).

5 Experiment 4: Multi-sensory cue integration - Can auditory cues enhance vection?

It is known that some people can get the illusion of self-motion just from a rotating sound field while being blindfolded. This was demonstrated as early as 1977 by Lackner [10] who used an array of speakers. Recently, this auditory self-motion illusion has been replicated by the POEMS project with headphone-based 3D sound rendering using a generic head-related transfer function (HRTF) [11]. In the latter study, a realistically rendered 3D sound environment was found to increase auditory presence and vection. Furthermore, multiple sound sources induced significantly more self-motion responses than a single sound source. The type of sound source was also found to affect auditory vection: Sound sources that are normally stationary (like church bells or a fountain noise) were found to be more instrumental in inducing auditory vection than sound sources that are normally moving (like footsteps or the sound of a driving bus). However, compared to the visually induced self-motion illusion, which is quite strong and convincing, the auditorily induced illusion is much less compelling, and only about 25-50% of the participants perceive any illusory self-motion at all. Hence, auditory cues alone are clearly insufficient for reliably enabling a convincing self-motion simulation in VR.

In this final experiment, we tested whether *additional* spatial auditory cues can nevertheless be utilized to enhance the visually-induced self-motion illusion. To test this, we compared a simple mono sound rendering of a landmark in the scene (the sound of the fountain on the market place scene) with a proper 3D acoustic rendering of the correct location of the fountain using a generic head-related transfer function (HRTF) and a Lake DSP system (Huron engine). In the 3D acoustic rendering ("spatialized") condition, the spatial content of the auditory simulation might enhance vection by providing consistent additional information about the spatial location of acoustic landmark. Note that the fountain was always audible (as we have omni-directional hearing), even when the visual counterpart was outside of the current field of view. 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, independent of the spatial content of the auditory cues.

5.1 Methods

A repeated-measures, within-subject design with 19 participants was used in the current study, with alternating left/right turns and randomized trials with mono or 3D spatialized auditory cues (50% each). Maximum stimulus velocity was set to 30°/s, and each condition was repeated six times. Apart from this, the experimental procedures were comparable to the first experiment.

5.2 Results and discussion

Figure 8 demonstrates a small but consistent vection-facilitating effect of the 3D sound spatialization. The strongest effect was observed for the convincingness ratings, the other dependent variables show only a small effect size, albeit in the correct direction. Presence ratings showed a similarly small, but consistent advantage for the spatialized sound (Fig. 9). This effect reached significance for the overall presence sum score.

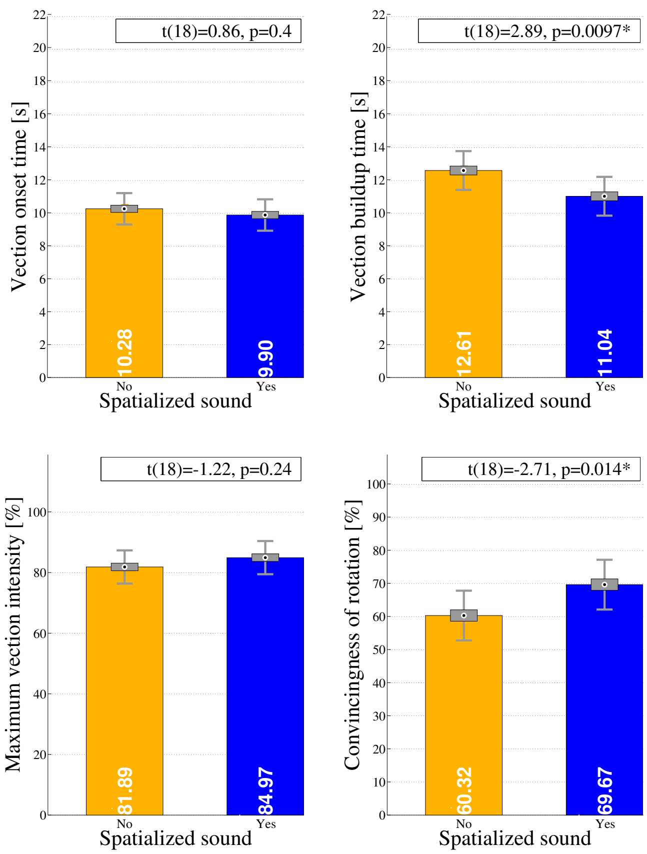

Figure 8: Mean performance for experiment 4, averaged over the 19 participants. Note the small but consistent vection-facilitating effect of the proper 3D auditory rendering of the fountain sound (right bars) as compared to simple mono display (left bars).

We conclude that adding 3D sound to self-motion simulations increases overall presence, but also affects the perception of self-motion itself - even though this effect might, in fact, be mediated be the increased presence. As the study by Larsson et al. [11] suggests, the benefit from additional 3D auditory cues should be stronger if multiple sound sources are used and acoustic reflections in the 3D environment are rendered properly.

Even though participants in a post-experimental interview rated the simulation as much more convincing with the added 3D sound, the effect on the self-motion illusion was rather small. This might reflect a ceiling effect, as visually induced vection was already quite strong without the 3D sound. We might expect a stronger influence of auditory cues if the auditory and visual vection-inducing potential were equated in terms of effect strength. This could, for example, be achieved by degrading the visual stimulus or reducing the visual field of view. Thus, low-cost motion simulators without a large

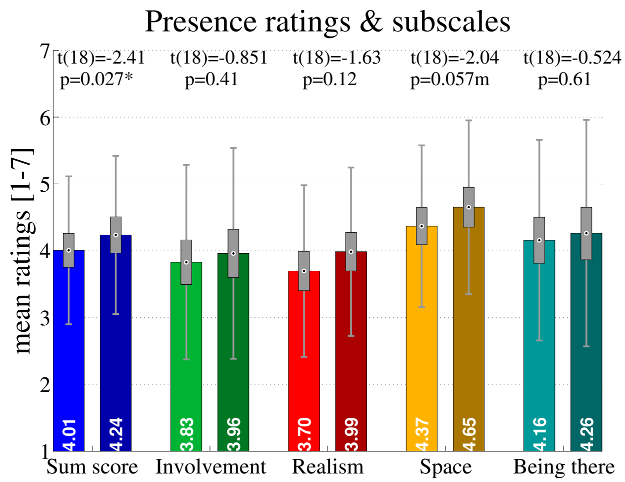

Figure 9: Mean presence ratings and sub-scales for the mono presentation (lighter bars) and 3D sound rendering (darker bars). Even though the effect size was small, the presence ratings were consistently higher for the 3D sound rendering.

field of view in particular might benefit from the addition of spatialized sound. Adding 3D sound to an application is furthermore known to increase the overall presence in the simulation [5] and has the advantage of extending the perceivable virtual space beyond the limits of the visual field of view of the setup. As this study and the study by Larsson et al. [11] demonstrated, a headphone-based auralization can reliably be used to improve self-motion perception and presence in VR. This has many practical advantages, especially for applications where speaker arrays are unsuitable or where external noise must be excluded.

6 Summary and Conclusions

Compared to the traditional approach in circular vection studies that uses rotating optokinetic drums, the present study demonstrates that vection can indeed be reliably induced and studied using a virtual reality setup, even when only a relatively small field of view is used ($54° \times 40.5°$). The experiments presented here yielded mean vection onset times around 10s, which compares well with traditional vection research using optokinetic drums (where onset latencies typically range from 2s to 10s for circular vection [2]). Note, however, that in these classic studies, full-field stimulation of the entire visual field was provided to the participants, compared to the relatively limited field of view of $54° \times 40.5°$ in our study. This highlights the power of using naturalistic scenes in VR for obtaining vection, compared to more classical approaches such as presenting abstract geometric patterns in optokinetic drums. In line with this argument, Steen & Brockhoff [20] achieved saturated vection after only 3s using a high-fidelity flight simulator with a FOV of $142° \times 110°$.

Several factors were found to increase both the self-motion illusion and presence: Experiment 1 demonstrated the benefits of displaying a consistent, naturalistic scene. This suggests that there are at least some relevant cognitive (top-down) contributions to self-motion perception, a line of inquiry that has traditionally been largely neglected. Especially the interpretation and meaning associated with particular stimuli seem to consistently affect self-motion perception and presence and should receive more attention. Experiment 2 provides some initial evidence that subtle modifications in the visualization setup can have a tremendous effect on the simulation, even when they are not explicitly noticed by most of the participants. The minor marks on the screen improved the simulation not only in terms of effectiveness of the self-motion simulation, but also in terms of overall presence, involvement, and realism. This highlights the importance of a careful design and experimental evaluation of VR setups with respect to their respective design goals. The last two experiments were concerned with multi-modal interactions, and showed a reliable, albeit small, benefit for adding consistent multi-sensory information, both in terms of vibrations (experiment 3) and 3D sound (experiment 4). It should be noted that multi-sensory effects on the vection illusion have hardly been studied, as Hettinger [6] points out in his comprehensive overview on the self-motion illusion in VR (p.487). Our study provides thus some initial investigations into this field which is highly relevant for VR-based motion simulations.

Even though the experimental manipulations allowed for mean vection onset times below 10s, most participants did not sense any self-movement until almost one revolution of the virtual scene. For VR applications like navigating virtual environments, however, convincing self-motion simulation over much smaller turning ranges is required, demanding vection onset times close to zero. We are currently investigating two approaches to tackle this challenge of immediate self-motion perception in VR without physically moving the observer much: On the one hand, we will refine the simulation to yield a better cross-modal consistency and believability for visual, auditory, and vibrational cues. Ultimately, we aim at embedding such a multi-modal simulation in a consistent motion metaphor, according to the desired application. In a flight or driving simulator, this could, for example, be done by including subtle stationary foreground cues like dirt or stains on a windshield, and carefully designing vibrations and spatialized sound to match the visual motion as well as the current state of the engine. Furthermore, an acoustic simulation of the 3D environment including reflections and reverberations might be beneficial [11]. On the other hand, we are experimenting with small physical jerks to specify motion onsets, similar to the much larger physical motions that were successfully employed by Wong & Frost [22] to reduce vection onset times.

In this paper, we presented a few initial results of the POEMS project suggesting that a perceptually oriented approach to self-motion simulation that includes carefully designed psychophysical experiments can indeed be quite valuable both in terms of designing and improving self-motion simulations and understanding the underlying relevant factors. We are, of course, aware of the potential problems involved in using introspective and questionnaire-based measures (see, e.g., [19]). To address this issue, the results from this study will be compared with behavioral measures of spatial updating and spatial orientation performance in future experiments. This will allow us to determine not only whether a self-motion simulation is subjectively convincing, but also whether it enables robust and effective spatial orientation in the virtual world. This issue is still largely unsolved, but recent results are quite promising [17].

By combining behavioral measures with the already established vection and presence measures, we hope to gradually be able to come closer to our ultimate goal: To devise a truly lean and elegant self-motion simulation paradigm through a better understanding of the underlying factors and multi-modal interactions. Thus, by further pursuing the perceptually oriented approach, we hope to eventually be able to reduce overall simulation effort and costs by focusing on the essential aspects both from a bottom-up and top-down perspective.

Acknowledgments: This research was funded by the EU grant POEMS-IST-2001-39223 (see www.poems-project.info) and the Max Planck Society.

References

[1] W. Becker, S. Raab, and R. Jürgens. Circular vection during voluntary suppression of optokinetic reflex. *Experimental Brain Research*, 144(4):554–557, June 2002.

[2] 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.

[3] M. O. Ernst and M. S. Banks. Humans integrate visual and haptic information in a statistically optimal fashion. *Nature*, 415:429–433, Jan. 2002.

[4] H. Fushiki, S. Takata, and Y. Watanabe. Influence of fixation on circular vection. *Journal of Vestibular Research-Equilibrium & Orientation*, 10(3):151–155, 2000.

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

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

[7] I. P. Howard and T. Heckmann. Circular vection as a function of the relative sizes, distances, and positions of 2 competing visual-displays. *Perception*, 18(5):657–665, 1989.

[8] I. P. Howard and A. Howard. Vection - the contributions of absolute and relative visual- motion. *Perception*, 23(7):745–751, 1994.

[9] R. L. Klatzky, J. M. Loomis, A. C. Beall, S. S. Chance, and R. G. Golledge. Spatial updating of self-position and orientation during real, imagined, and virtual locomotion. *Psychol. Sci.*, 9(4):293–298, 1998.

[10] 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.

[11] P. Larsson, D. Västfjäll, and M. Kleiner. Perception of self-motion and presence in auditory virtual environments. In *Presence*, (www.presence2004.org), 2004.

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

[13] S. Nakamura and S. Shimojo. Critical role of foreground stimuli in perceiving visually induced self-motion (vection). *Perception*, 28(7):893–902, 1999.

[14] J. D. Prothero. *The Role of Rest Frames in Vection, Presence and Motion Sickness*. PhD thesis, University of Washington, 1998. Available: www.hitl.washington.edu/publications/r-98-11/.

[15] 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 *Presence*, 2004. Available: www.kyb.mpg.de/publication.html?publ=2864.

[16] B. E. Riecke and M. von der Heyde. Qualitative modeling of spatial orientation processes using logical propositions: Interconnecting spatial presence, spatial updating, piloting, and spatial cognition. Technical Report 100, MPI for Biological Cybernetics, 2002. Avaliable: www.kyb.mpg.de/publication.html?publ=2021.

[17] B. E. Riecke, M. von der Heyde, and H. H. Bülthoff. Visual cues can be sufficient for triggering automatic, reflex-like spatial updating. *ACM Transactions on Applied Perception (TAP)*, 2005. (submitted).

[18] 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.

[19] M. Slater. How colorful was your day? Why questionnaires cannot assess presence in virtual environments. *Presence - Teleoperators and Virtual Environments*, 13(4):484–493, August 2004.

[20] F. A. M. van der Steen and P. T. M. Brockhoff. Induction and impairment of saturated yaw and surge vection. *Perception & Psychophysics*, 62(1):89–99, January 2000.

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

[22] 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.