Top-Down and Multi-Modal Influences on Self-Motion Perception in Virtual Reality

B. E. Riecke[1] *D. Västfjäll*[2] *P. Larsson*[3] *J. Schulte-Pelkum*[4]

Abstract

INTRODUCTION: Much of the work on self-motion perception and simulation has investigated the contribution of physical stimulus properties (so-called "bottom-up" factors). This paper provides an overview of recent experiments demonstrating that illusory self-motion perception can also benefit from "top-down" mechanisms, e.g. expectations, the interpretation and meaning associated with the stimulus, and the resulting spatial presence in the simulated environment. METHODS: Several VR setups were used as a means to independently control different sensory modalities, thus allowing for well-controlled and reproducible psychophysical experiments. Illusory self-motion perception (vection) was induced using rotating visual or binaural auditory stimuli, presented via a curved projection screen (FOV: 54x40.5°) or headphones, respectively. Additional vibrations, subsonic sound, or cognitive frameworks were applied in some trials. Vection was quantified in terms of onset time, intensity, and convincingness ratings. RESULTS & DISCUSSION: Auditory vection studies showed that sound sources participants associated with stationary "acoustic landmarks" (e.g., a fountain) can significantly increase the effectiveness of the self-motion illusion, as compared to sound sources that are typically associated to moving objects (like the sound of footsteps). A similar top-down effect was observed in a visual vection experiment: Showing a rotating naturalistic scene in VR improved vection considerably compared to scrambled versions of the same scene. Hence, the possibility to interpret the stimulus as a stationary reference frame seems to enhance the self-motion perception, which challenges the prevailing opinion that self-motion perception is primarily bottom-up driven. Even the mere knowledge that one might potentially be moved physically increased the convincingness of the self-motion illusion significantly, especially when additional vibrations supported the interpretation that one was really moving. CONCLUSIONS: Various top-down mechanisms were shown to increase the effectiveness of self-motion simulations in VR, even though they have received little attention in the literature up to now. Thus, we posit that a perceptually-oriented approach that combines both bottom-up and top-down factors will ultimately enable us to optimize self-motion simulations in terms of both effectiveness and costs.

1 Introduction

With virtual reality (VR) technology becoming more affordable and wide-spread during the last years, there are more and more applications that involve simulations of self-motion through computer-generated worlds. Most applications, however, are not yet able to convey a convincing and natural feeling of self-motion. This problem is may contribute to the phenomenon that users are often easily disoriented in virtual environments or cannot find their way around easily and without considerable cognitive effort (e.g., Riecke, 2003). So how can the effectiveness of self-motion simulations in VR be improved?

A number of studies have demonstrated that physically moving through the environment (e.g., walking while wearing an HMD or being moved on a motion platform) improves spatial orientation performance. Physical motions of the observer seem to be of particular importance when simulated rotations of the observer are involved (e.g., Chance, Gaunet, Beall, & Loomis, 1998; Klatzky, Loomis, Beall, Chance, & Golledge, 1998). For most applications, however, allowing for physical motion of the user is not practical due to constraints in space, technical equipment, or simply money. Hence, it would be advantageous to have a reliable means of producing the illusion of self-motion through simulated worlds *without* any physical motions.

---

[1] Max Planck Institute for Biological Cybernetics, Tübingen, Germany. E-mail: bernhard.riecke@tuebingen.mpg.de
[2] Chalmers University of Technology, Göteborg, Sweden. E-mail: daniel@ta.chalmers.se
[3] Chalmers University of Technology, Göteborg, Sweden. E-mail: pontus.larsson@ta.chalmers.se
[4] Max Planck Institute for Biological Cybernetics, Tübingen, Germany. E-mail: joerg.sp@tuebingen.mpg.de

In this manuscript, we provide an overview over some recent experiments that investigated novel approaches to quantify and increase the effectiveness and perceptual convincingness of self-motion simulations *without* moving the observer physically[5]. More specifically, we investigated how visual, auditory, vibrational, and subsonic cues in a VR simulator can contribute to the illusion of self-motion ("vection"). In particular, we were interested how cognitive, top-down factors like spatial presence and the interpretation or meaning associated with the presented stimuli affected the self-motion illusion.

Many people have experienced illusory self-motion in the real world, for example when sitting in a train waiting to depart from the train station. If the train on the adjacent track starts to pull out of the station, one can have the compelling illusion that one's own train just started to move, even though it is, in fact, still stationary. This illusion typically breaks down as soon as a conflict is noticed, for example, by looking out of the opposite window. Here, we propose that investigating such self-motion illusions could ultimately be used to improve the effectiveness of VR simulations and spatial orientation in VR in particular.

During the last century, there has been a large number of studies investigating the effect of various physical stimulus properties on visually induced illusory self-motion, especially for rotations around the earth-vertical axis ("circular vection"). See Dichgans & Brandt (1978), Hettinger (2002), and Warren & Wertheim (1990) for excellent reviews on this topic. The typical setup used in the literature to study visually induced circular vection consists of a rotating cylinder painted with a simple geometric pattern, such as black and white vertical stripes. Initially, the observers seated in the centre of the cylinder correctly perceive the visual stimulus as moving and themselves as stationary. After a so-called vection onset time of typically 5-30 seconds, a qualitative shift in the percept occurs, and the observers start perceiving *themselves* as rotating and the visual stimulus as slowing down and finally becoming earth-stationary. This illusion can be quite compelling, and observers typically have the impression of themselves physically moving, even when they are aware that this cannot be the case.

A number of physical stimulus properties (so-called bottom-up factors) have been identified that can enhance this self-motion illusion. First, a large visual field of view (FOV) covered by the moving stimulus will enhance the vection illusion (Hettinger 2002), even though it has been shown that FOVs as small as 7.5° in the central visual field can be sufficient for inducing vection if the moving stimulus is seen as being in the background behind the occluding edges (Andersen & Braunstein, 1985). Thus, the foreground-background relations of moving/stationary objects in the scenery are also important. Other relevant bottom-up or physical stimulus properties include the stimulus contrast and the number of vertical high-contrast edges, which is closely related to the spatial frequency spectrum of the stimulus. There is an abundance of studies in the literature investigating and demonstrating such bottom-up contribution of various physical stimulus parameters to self-motion perception and spatial orientation (e.g., Dichgans & Brandt, 1978; Hettinger, 2002; Tan, Gergle, Scupelli, & Pausch, 2003; Warren & Wertheim, 1990). It is conceivable, however, that self-motion perception can also be influenced by expectations as well as the interpretation or associated meaning of the stimuli (i.e., top-down effects). For example, the fusion of modality-specific representations of physical stimuli into a coherent (typically amodal) unified representation is most likely affected by non-sensory (top-down, cognitive) information. Thus, we expect that providing consistent, multi-modal stimulation in a VR setup should enhance the self-motion illusion. We are, however, only aware of two studies that explicitly addressed the influence of cognitive factors on vection: Lepecq, Giannopulu, & Baudonnniere (1995) demonstrated that seven year old children experience vection quicker when they are sitting on a potentially moving platform, compared to a stationary one. More recently, Palmisano & Chan (2004) showed that such cognitive priming can influence vection onset times for adults, too: Vection started earlier and lasted longer when they were seated on a potentially moving chair and were primed towards attending their self-motion sensation, compared to a condition where they were seated on a stationary chair and instructed to attend to object motion, not self-motion. Even though these two studies demonstrated a clear cognitive or top-down effect on illusory self-motion perception, we are not aware of any further publications that directly address the issue. One of the aims of this paper is thus to provide more experimental evidence that top-down effects can indeed play an important role in the perception of self-motion – an issue that might be of considerable importance for designing lean and elegant self-motion simulators.

We will start by reporting a study in detail that investigated circular vection induced purely by auditory cues, and how top-down factors and sub-threshold vibrations and subsonic cues can all contribute to this illusion (section 2). The remainder of this manuscript is concerned with providing an overview on some recent studies on visually induced circular vection and how vection can be influenced by top-down cues. We will briefly report experiments on

---

[5] This research has been conducted in the context of the POEMS project on Perceptually Oriented Ego-Motion Simulation, EU grant POEMS-IST-2001-39223 (see www.poems-project.info)

the influence of scene consistency and spatial presence on vection (section 3), followed by a study where the visually presented scene was simply put upside down to investigate the influence of ecological validity and naturalism of the simulation (which are both top-down factors; see section 4). We will conclude with a study that addressed the issue of participants' knowledge about whether they could potentially move or not (similar to studies by Lepecq et al. (1995) and Palmisano & Chan (2004)).

2 Top-down and multi-modal influences on auditory-induced circular vection

Illusory self-motion induced by sound has received little attention until recently. Furthermore, most of the studies in the field address linear self-motion (e.g., Kapralos, Zikovitz, Jenkin, & Harris, 2004; Sakamoto, Osada, Suzuki, & Gyoba, 2004). To the best of our knowledge, the only study which has addressed auditory-induced circular vection was performed by Lackner (1977). In this study, participants were exposed to rotating sound fields presented over loudspeakers or headphones. Lackner found that participants experienced self-rotation in both conditions, but that the loudspeaker technique was significantly more effective than headphones. Furthermore, when the contours of the experimental room were visible to the participant, auditory stimulation no longer elicited illusory self-rotation. This suggests that, although illusory circular self motion can be elicited by sound alone, this sensation can easily be dominated by visual input. Nevertheless, auditory cues can play an important role for self-motion simulations in VR. In particular, we believe that the self-motion illusion may be enhanced by the listeners' cognitive (top-down) interpretation of sound sources. This idea was inspired by the work of Gaver (1993a, 1993b), who proposed that in everyday listening we are primarily attending to the events that actually cause the sound, rather than the sound itself. Applied to the case of rotating sound fields, we posit that a listener can, when exposed to an everyday sound field, not only correctly and intuitively identify the sound sources contained therein, but also easily resolve whether the sound sources are likely to be an "acoustic landmark"[6] or not. If an acoustic landmark is rendered as moving with respect to the listener, e.g., by means of dynamic binaural spatialization, such a simulation should be more likely to induce self-motion percepts rather than source motion percepts. Conversely, if a sound source is identified as originating from a potentially moving object (e.g., a car), we would predict a higher probability for experiencing object motion rather than self-motion.

Furthermore, it is possible that auditory-induced illusory self-motion may be mediated by perceptual cues directly. It is for example known that strong, low-frequency sound stimuli can elicit responses from the vestibular organ (Blauert, 1999; Todd & Cody, 2000). Such acoustically induced vestibular sensations may in turn give rise to a feeling of physical acceleration, and may therefore be instrumental in amplifying the sensation of self-motion elicited by the rotating sound field. It has also been suggested that simple vibrotactile stimulation may provide as strong of a contribution to the self-motion illusion as physical motion cues (Bürki-Cohen, Soja, & Longbridge, 1998, Harris, Jenkin, & Zikovitz, 2000). It is, however, unclear whether this is due to the vibrations causing slight vestibular activation or to the fact that participants cognitively interpret the vibrations as being generated by the simulated vehicle, thus rendering the motion simulation more realistic and convincing. Nonetheless, it seems reasonable to assume that both acoustically induced vestibular sensations and vibrotactile stimulation, in addition to the previously discussed cognitive influences, should be considered when optimizing (acoustic) self-motion simulations.
In sum, we tested two main hypotheses for the study presented in this section:

1. Rotating auditory stimuli consisting of one or more sound sources will elicit a stronger, more compelling sensation of self-rotation if the sound sources can be identified as being still (i.e., acoustic landmarks) compared to if the sound sources are identified as being mobile or if they are unidentifiable (artificial sounds)[7].
2. The addition of diffuse (not easily recognizable or localizable) vibrations and subsonic sound may activate the vestibular system and in turn improve the sense of ego-motion.

2.1 Method

The auditory stimuli used to test the first hypothesis were all binaural simulations of a virtual listener surrounded by one or three virtual sound sources and rotating for 60s at a velocity of 20, 40 or 60°/s. The distance between the sources and the receiver was 5 meters. Non-individualized Head Related Transfer Functions (HRTFs) measured from a human participant were used in the creation of all stimuli. Three different source sounds in each category

---

[6] An acoustic landmark refers to a sound source that can be easily identified and can be used as a reference point, as it is expected to not move, like for example church bells.

[7] Parts of this research has previously been presented in Larsson et. al. (2004).

(Still, i.e., auditory landmarks (S), Moving (M) or Artificial (A)) were selected and assigned to the different virtual sound sources. These were: S1) Bus on idle, S2) Small fountain S3) Barking dog, M1) Footsteps, M2) Bicycle, M3) Driving bus, A1) Stationary pink noise, A2) Pink noise bursts, 250 ms + 250 ms of silence, and A3) Pink noise bursts of random length and temporal distribution. The parameters *type of sound source* (still, moving or artificial), *velocity* (20, 40, 60°/sec) and *number of sound sources* (1 or 3) were varied in a within-subjects design with 26 participants.

The second hypothesis was tested in a separate session with 14 participants. A within-subjects design was used where *sound only vs. sound + vibration, velocity (20, 40, 60 ^{o}/sec), number of sound sources (1 or 3)* and *type of sound source (still or moving)* were varied. In addition, for a selected number of four sounds, an additional experimental variable was tested: *sound only* vs. *sound + subsonic sound***.** The vibration and subsonic stimulus was an amplitude-modulated 15 Hz sinusoidal signal fed to either a shaker attached to a seating arrangement or subwoofers (see Figure 1).

Figure 1: **Left:** Experimental chair mounted on a turntable to suggest to participants that they could potentially be rotating physically (which never happened during the experiment). **Middle:** Participant fitted with headphones and blindfold (right). Four fake loudspeakers were mounted on the walls to suggest that the sound presented via headphones might in fact be originating from the stationary loudspeakers. One of the fake loudspeakers can be seen in the upper right corner of the picture. **Right:** Subwoofer array used for the subsonic stimulation.

The experiments were conducted in a semi-anechoic room. Auditory stimuli were presented with Beyerdynamic DT-990Pro circumaural headphones driven by a NAD Amplifier, model 3020. Participants were seated on an ordinary office chair placed on an electrically controllable platform to suggest that they might, in fact, be moved - although this never happened during any of the trials (see Figure 1). The top-down influence of setting up a mental framework to increase the believability of the stimulus is believed to contribute to illusory self-motions (Lepecq et al., 1995; Palmisano & Chan, 2004). This seating arrangement also prevented the participants from having any contact with footrests or the floor. The participants were blindfolded throughout the experiment. In the separate vibration/subsonic sound session, a shaker mounted on top of the movable platform was used to excite diffuse chair vibrations and a custom made subwoofer wall consisting of sixteen 15” loudspeaker drivers was used to produce subsonic sound. Participants in both sessions were instructed to report the onset of vection by indicating the direction (left/right) verbally. In addition, participants were asked at the end of each trial to rate the intensity and convincingness of perceived vection on a 0-100% scale.

2.2 Results and discussion

Overall, support was found for the hypothesis that still sound sources (“acoustic landmarks”) are more instrumental in inducing vection than both moving and artificial sound sources. The means for the intensity ratings for single sound sources are shown in Figure 2. An ANOVA for intensity ratings yielded a significant main effect of sound source (F(2,25) = 5.66, p<.001), with the still sound sources yielding significantly higher intensity ratings (M = 36.3) than the moving (M = 20.3) and artificial ones (M = 20.1). Neither the main effects of velocity, nor any of the interactions reached significance (F>1, p>.05).

Similar results were obtained for multiple sound sources: An ANOVA for intensity ratings showed a significant main effect of sound source (F(2,25) = 12.15, p<.001). Still sound sources resulted in significantly higher intensity ratings (M = 39.2) than the moving and artificial sound sources (M = 30.6 and M = 29.7, respectively). Furthermore, a significant main effect of velocity was observed (F(2,25) = 7.74, p<.001), with the faster rotation velocity (60°/s) yielding higher intensity ratings (M = 38.5) than both the 40°/s (M = 32.0) and 20°/s (M = 30.5) condition. Even though this effect was not predicted for auditory vection, it mimics recent results for visually-induced vection (Schulte-Pelkum, Riecke, von der Heyde, & Bülthoff, 2003). No other factors reached significance.

Furthermore, multiple sound sources induced significantly more vection responses than a single sound source. For instance, the frequency of “yes” responses on the binary vection measure was approximately 20% greater for multiple compared to single sound sources. Collapsed across conditions, McNemar tests showed that this increase was significant (p<.05) for all cases.

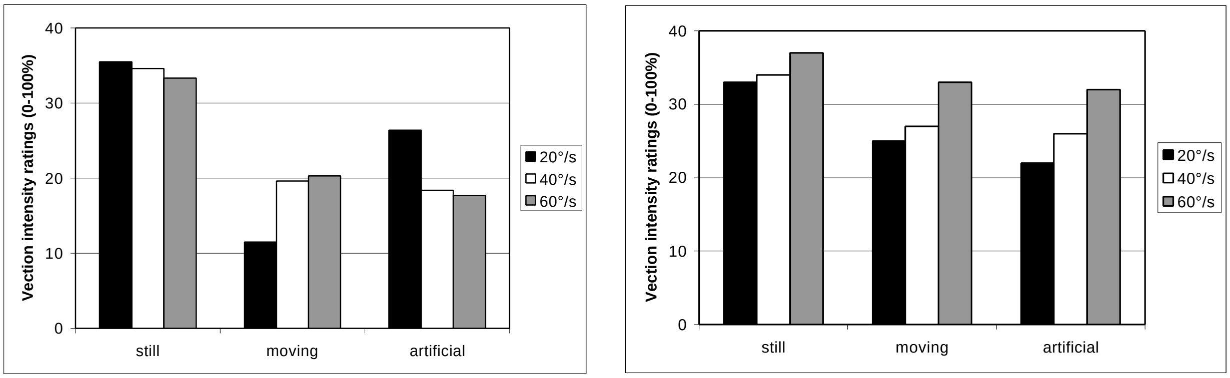

Figure 2: **Left:** Intensity ratings (0-100%) of auditory-induced vection for single sound sources. Larger values indicate higher perceived intensity of vection. **Right:** Intensity ratings for multiple sound sources. Note the higher intensity ratings for the still sounds sources (acoustic landmarks) as compared to both the moving and artificial sound sources.

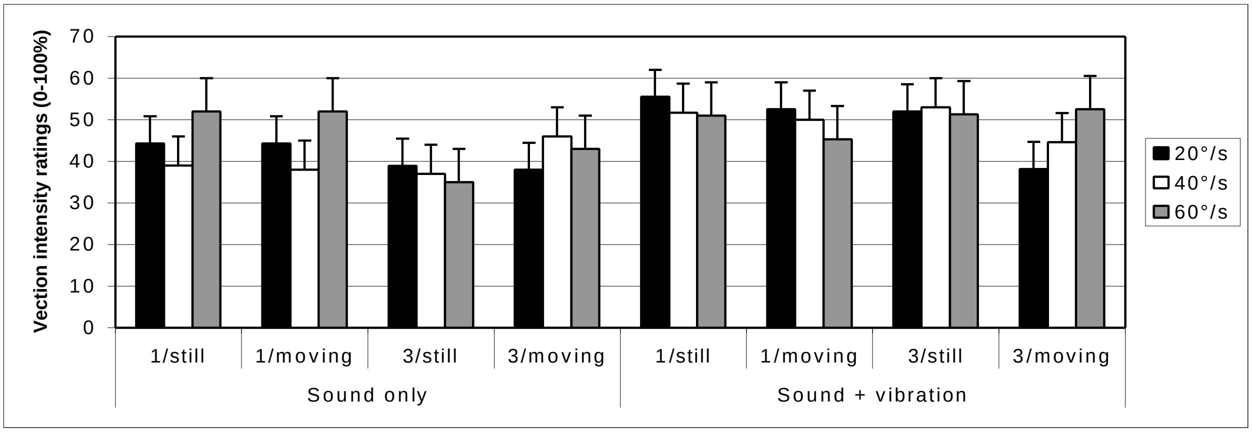

Figure 3: Mean intensity ratings for *sound only* vs. *sound+vibration* conditions. The graph shows the percentage of participants indicating auditory-induced vection for the different conditions. 1/still = one still sound source (acoustic landmark), 3/still = three still sound sources, 1/moving = one moving sound source, 3/moving = three moving sound sources. Note the vection-facilitating effect of additional vibrations and auditory landmarks (still sound sources).

Furthermore, additional vibrational and subsonic stimulation increased vection: Figure 3 shows vection intensity ratings across the *sound only* and *sound+vibration* experimental conditions. An ANOVA for vection intensity yielded two significant main effects. First, overall intensity for the *sound+vibration* condition was rated higher (M = 50.3) than for the *sound only* condition (M = 41.4), F(1, 13) = 12.5, p<.01. Second, as in the previous study , the still sound sources induced stronger ego-motion sensations (M = 47.9) than the moving sound sources (M = 40.5), F(1, 13) = 5.44, p<.05. No other factors reached significance.

Similarly, an ANOVA for the specific comparison *sound only* vs. *sound + subsonic* (Figure 4) showed a main effect of *sound+subsonic* (M = 54.2) vs. *sound only* (M = 40.7), F(1, 13) = 14.12, p<.01. No other factors were significant.

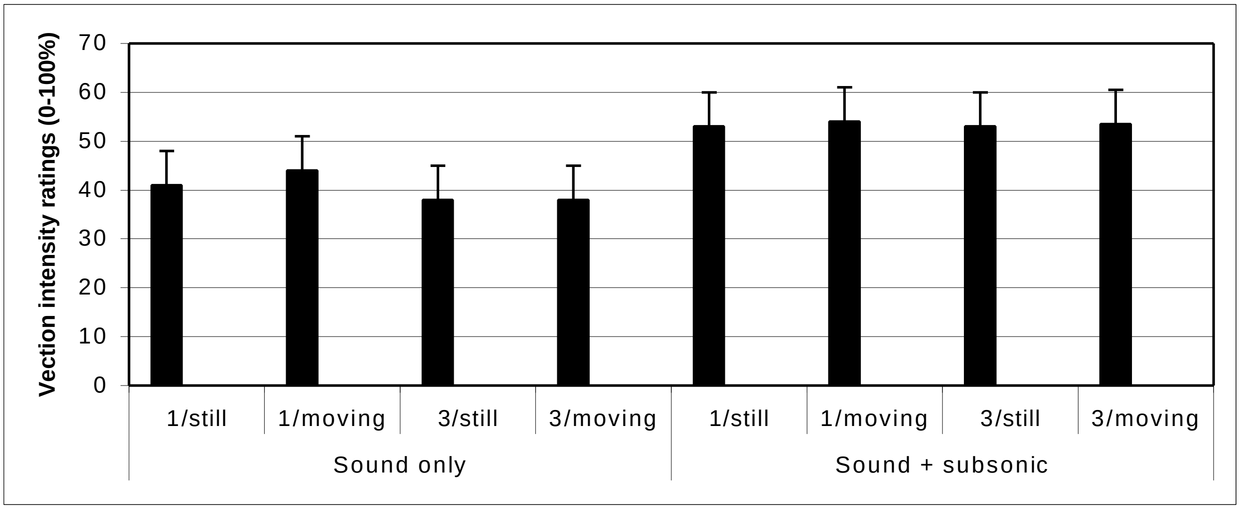

Figure 4: Mean intensity ratings for the *sound only* vs. *sound+subsonic* conditions for the 14 participants. The graph shows the percentage of participants indicating auditory-induced vection for the different conditions. 1/still = one still sound source, 3/still = three still sound sources, 1/moving = one moving sound source. Only the 20°/s velocity was used here. Note the increased vection intensity ratings for the additional subsonic stimulation.

In summary, the present results suggest a strong top-down effect on auditory-induced vection. The type of sound source seems to be a primary determinant of this illusion: Sounds that are usually associated with stationary sound sources (i.e., so-called acoustic landmarks) facilitated vection, compared to sounds that are typically associated with moving objects or sounds that have no clear association (artificial sounds).

The results also suggest that the type of sound source may play less of a role when the environment contains multiple sound sources, which might indicate a ceiling effect for auditory-induced vection. Strong support was found for the hypothesis that diffuse vibrations and subsonic sound increase auditory-induced vection, both in terms of binary vection responses and intensity/convincingness ratings. Across all conditions, we found an average increase of 10-15% for the auditory-induced vection responses when additional vestibular activation was applied through vibrations or subsonic stimulation. These results are promising for further work on how to increase the effectiveness of self-motion simulations (cf. Riecke, Schulte-Pelkum, Caniard, & Bülthoff, 2005).

In the remainder of this article, we will provide an overview on some recent experiments that investigated the influence of top-down and multi-modal contributions to visually-induced circular vection using a virtual reality projection system.

3 Influence of scene consistency on visually-induced circular vection

In a study on visually induced circular vection, the contribution of top-down influences was assessed by presenting different visual stimuli on a curved projection screen (see Figure 5). Vection was induced by rotating stimuli around the observers' earth-vertical axis. The stimulus consisted either of a photorealistic panoramic image of the Tübingen market place or three different scrambled versions thereof (see Figure 6). This scrambling effectively destroyed scene consistency (i.e., the interpretation of the stimulus as a "3D scene" that one could potentially move through, and thus changed the overall meaning associated with the visual stimulus, which are all top-down processes), while only slightly changing image statistics (bottom-up contributions).

Figure 5: Participant seated behind the curved projection screen, which displays a 54° × 40.5° undistorted view of the Tübingen market place (condition A).

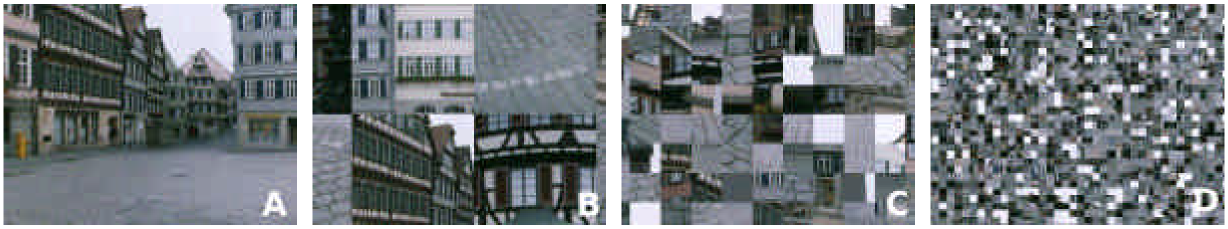

Figure 6: Stimuli used for the different experimental conditions. Condition A (left) shows a view onto the intact (globally consistent) market scene. Conditions B-D show views onto scrambled versions of the same stimulus, with increasing scrambling severity (2, 8, or 32 mosaics per 45° field of view for condition B, C, and D, respectively).

3.1 Hypotheses

The experimental design was based on the assumption that scene scrambling should reduce spatial presence (the feeling of "being there") in the virtual environment, as only the globally consistent stimulus can naturally be recognized or interpreted as a "scene", which might in turn allow for actions such as locomotion through the scene. Hence, we expected that spatial presence should be highest in the intact scene and decreasing for increased scrambling severity. We pursued three hypotheses here:

1. If global scene consistency of the visual stimulus and/or spatial presence in the simulated scene is critical for perceived self-motion, we would predict that the intact stimulus should yield stronger vection than any of the other stimuli, ideally in terms of all response measures (A>B=C=D).
2. If, however, object recognition and local consistency of the stimulus are important for perceiving self-motion, we would predict that the scrambling severity (i.e., number of mosaics per solid angle) should reduce vection. Especially the most severe scrambling level (see Figure 6, conditions C and D) should show the least vection, as individual objects like doors and windows are no longer recognizable, whereas they still are in condition B.
3. Apart from the top-down influences discussed above, the scene scrambling also affected physical stimulus properties or so-called bottom-up factors: The scrambled stimuli contained additional high-contrast vertical edges that are not present in the globally consistent stimulus (A). Such high-contrast edges are known to increase perceived stimulus speed (Diener, Wist, Dichgans & Brandt 1975) and vection (Dichgans & Brandt 1978). Hence, if these bottom-up factors dominate over top-down factors like scene consistency and object recognition, we would expect that the scrambled stimuli should actually result in stronger vection than the intact stimulus (A). Alternatively, the two effects might cancel each other out.

3.2 Results and Discussion

Scrambled stimuli yielded significantly longer vection onset times and lower ratings for the intensity and convincingness of the self-motion illusion than the intact market scene (all p's < .05). Furthermore, presence scores, assessed using the IPQ Igroup Presence Questionnaire by Schubert, Friedmann, & Regenbrecht (2001) were significantly reduced due to the scene scrambling. This result is in agreement with hypothesis 1, but disproves hypothesis

3. No differences were, however, found for the different scrambling levels (B, C, & D) for any of the dependent variables. That is, the recognizability of individual objects in the scene (like houses, doors, windows, etc.) did not seem to have a clear influence on the ego-motion illusion or presence, thus failing to support hypothesis 2.

Taken together, the data suggest that top-down contributions (i.e., the consistent reference frame provided by the intact market scene) dominated here over the low-level (bottom-up) factors (more contrast edges in the scrambled stimuli, which are known to facilitate vection, see hypothesis 3). We posit that stimuli depicting naturalistic scenes provide observers with a convincing reference frame for the simulated environment which enables them to feel more “spatially present”. This, in turn, might have facilitated the self-motion illusion. Thus, top-down mechanisms seem to affect visually induced vection as well, a phenomenon that was typically believed to be mainly bottom-up driven.

4 Influence of scene inversion on visually induced circular vection

A similar study on visually induced circular vection took a slightly different approach in investigating the influence of the naturalness and/or associated meaning of the visual stimulus. Using the same setup as the previous experiment and a comparable experimental paradigm, the naturalness and/or associated meaning of the visual stimulus was manipulated by either presenting the same, intact, visual stimulus of the Tübingen market place as in the previous experiment (condition A) or an upside-down version of the same stimulus. We hypothesized that this stimulus inversion should decrease vection measures if the naturalness or ecological validity of the stimulus plays a significant role. In addition, the feeling of “being there” (presence) in the simulated scene was expected to decrease with stimulus inversion. Note, however, that the stimulus inversion did not alter image statistics at all, and did not introduce any high contrast edges as did the scene scrambling in the previous experiment.

In accordance with this hypothesis, convincingness ratings were indeed significantly lower for the upside-down stimulus ($F(1,9)=13.2$, $p=.005$). Furthermore, presence ratings assessed using the IPQ (Igroup Presence Questionnaire) by Schubert et al. (2001) were slightly but significantly decreased for the upside-down stimulus ($t(9)=2.44$, $p<.05$). Neither vection onset time nor rated vection intensity, however, showed any clear difference between the two stimuli used. This suggests that the convincingness of the self-motion illusion is more directly affected by top-down contributions (such as presence in the simulated scene) than the onset and intensity of vection. Nevertheless, simply showing an upside-down version of the identical stimulus did significantly affect participants’ subjective ratings both in terms of convincingness and presence. This strongly suggests a top-down contribution to presence and the convincingness of self-motion illusions, as the physical stimulus properties (i.e., the factors responsible for bottom-up contributions) were not affected apart from a reversal in the up-direction.

5 Influence of cognitive bias on visually induced circular vection

Using a similar setup and experimental paradigm as in the previous two experiments, a further study investigated the influence of pre-knowledge and expectations on visually induced circular vection. The issue addressed in this experiment was the potential influence of cognitive bias and expectations on self-motion simulation in VR. That is, we investigated whether participants’ self-motion perception can be modified without changing any stimulus parameters whatsoever, just by changing their expectation and pre-knowledge whether physical motions might actually be possible or not. A study on linear vection by Lepecq et al. (1995) showed that seven year old children indeed perceived vection earlier when they were previously shown that the chair they were seated on can actually move physically, even though it never did during the actual experiment. 11 year old children, however, did not show any vection increase.

Here, we tested whether adults’ susceptibility to vection can be cognitively influenced by mounting the whole experimental setup on top of a motion platform and changing participants’ pre-conceptions whether this platform might actually move during the experiment or not. In two balanced conditions, participants were either (a) shown and convinced that the platform was not operating and completely switched off, or (b) shown before the experiments that the platform can actually move, and they were told that the platform might in fact be moving in some of the trials. In addition, vibrations were presented in half of the trials in both condition (a) and (b).

The cognitive manipulation did not influence vection responses recorded during the experiment. In post-experimental interviews, however, 2/3 of the participants reported that they had, in fact, moved physically on some trials of the platform on condition (b). In addition, more than 1/3 of all participants rated their certainty that the platform had physically moved as more than 50%. The reason that the cognitive manipulation did not affect vection onset times may be due to the fact that the highly realistic visual stimulus was already so effective that it might have reached a ceiling level. Moreover, the within-subject design might have been problematic since participants might have realized that there was no difference between the experimental conditions. Interestingly, the trials where physical self-motion was reported were mostly trials where the visual motion was accompanied by vibrations. That is, many participants reported having associated the vibrations to movement of the motion platform (which never happened in reality). This finding suggests that even simple vibrations can increase the believability of ego-motion simulations, especially if the cognitive framework allows for the interpretation that participants might, in fact, be moving physically.

6 Summary and Conclusions

AUDITORY CUES: In the auditory vection studies (section 2), participants were blindfolded and listened to rotating sound sources displayed via headphones. To assess the influence of top-down processes, acoustic landmarks that are associated with stationary objects (fountain, bus on idle etc.) were contrasted with mobile or artificial sound sources (driving bus, footsteps, pink noise etc.). These top-down processes showed a significant vection-facilitating effect in all dependent measures. A related (unpublished) study included trials comparing simulations where realistic wall reflections and reverberation were added to the various sound sources, with anechoic counterparts. The vection-facilitating effect was less pronounced for the anechoic simulation, even though it provides better sound source localizability. That is, the ecological probability or assumptions about the meaning of a stimulus (i.e., top-down processes) significantly affect self-motion perceptions.

VISUAL CUES: In a comparable visual vection study (section 3), the contribution of top-down influences was assessed by presenting either a photorealistic image of a market scene or various scrambled versions thereof. This scrambling effectively destroyed global scene consistency (i.e., the overall meaning and thus top-down processes) while hardly changing image statistics (bottom-up contribution). Scrambled stimuli yielded significantly longer vection onset times, lower perceived vection intensity, and lower convincingness ratings than the intact market scene. Furthermore, spatial presence scores were significantly reduced. Results suggest that high level information (consistent reference frame for the intact market scene) dominated over the low-level information (more contrast edges in the scrambled stimulus, which are known to facilitate vection). Even simply presenting the scene upside down affected the convincingness of the self-motion illusion significantly (section 4). We posit that stimuli depicting naturalistic scenes provide observers with a convincing reference frame for the simulated environment which enables them to feel "spatially present". This, in turn, facilitates the self-motion illusion. Thus, top-down mechanisms proved to affect vection, a phenomenon that was typically believed to be mainly bottom-up driven.

VIBRATIONS AND SUB-SONIC CUES: In the auditory vection study (section 2), even hardly noticeable vibrations improved vection intensity and convincingness significantly. We suppose that this influence might be due to bottom-up effects (e.g., adding noise to the vestibular system and thus changing vestibular thresholds and multi-modal weighting) as well as top-down effects (changing the stability assumption and the convincingness of the motion simulation). Similar vection-facilitating effects were found for additional subsonic stimuli, which significantly increased vection intensity and convincingness.

COGNITIVE INFLUENCE: The mere knowledge of whether one could potentially be physically moved (by seating participants on a mobile vs. static platform) affected both auditory and visually induced self-motion illusions. For visual stimuli, for example, more than 2/3 of the participants did believe that they physically rotated at least in some trials when they believed they *could* potentially be moved (even though that never happened).

In sum, we presented recent evidence that not only simple bottom-up processes, but also top-down processes and especially the interpretation and meaning associated with particular stimuli do affect self-motion perception consistently. Furthermore, the congruence and consistency of the motion metaphor and simulated scene seems to be relevant. Thus, we hoped to have raised the awareness for the importance of considering both bottom-up and top-down influences as two tightly intertwined aspects of the same stimulus. We posit that these two complementary views could help understanding the “big picture” and allow for a leaner and more elegant approach to ego-motion simulation (cf. Riecke et al., 2005). This might allow one to reduce overall simulation effort and costs by focusing on the essential aspects both from a bottom-up and top-down perspective.

Acknowledgments

This study was supported by the EU grant POEMS-IST-2001-39223 (see www.poems-project.info) and by the Max Planck Society.

References

Andersen, G.J. & Braunstein, M.L. (1985). Induced self-motion in central vision. *Journal of Experimental Psychology: Human Perception and Performance, 11,* 122-132.

Blauert, J. (1999). *Spatial Hearing: The Psychophysics of Human Sound Localization*. Cambridge, MA: MIT Press.

Bürki-Cohen, J., Soja, N. N., & Longbridge, T. (1998). Simulator platform motion: The need revisited. *The International Journal of Aviation Psychology, 8(3),* 293-317.

Chance, S. S., Gaunet, F., Beall, A. C., & Loomis, J. M. (1998). 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.

Dichgans, J. & Brandt, T. (1978). Visual-Vestibular Interaction: Effects on Self-Motion Perception and Postural Control. In R. Held, H. W. Leibowitz, & H.-L. Teuber (Eds.), *Perception*, Vol. VIII of *Handbook of Sensory Physiology* (pp. 756–804). Berlin Heidelberg: Springer.

Gaver, W. W. (1993a). How do we hear in the world? Explorations in ecological acoustics. *Ecological Psychology, 5,* 285-313.

Gaver, W. W. (1993b). What in the world do we hear? An ecological approach to auditory event perception. Explorations in ecological acoustics. *Ecological Psychology, 5,* 1-29.

Harris, L. R., Jenkin, M., & Zikovitz, D. C. (2000). Visual and non-visual cues in the perception of linear self motion. *Experimental Brain Research, 135,* 12-21.

Hettinger, L. J. (2002) Illusory self-motion in virtual environments. In K. M. Stanney (Ed), Handbook of Virtual Environments, (pp. 471-492). Lawrence Erlbaum,

Kapralos, B., Zikovitz, D., Jenkin, M., & Harris, L. R. (2004) Auditory cues in the perception of self-motion. *Presented at the 116th AES convention, Berlin, Germany*, (preprint 6078).

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

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

Larsson, P., Västfjäll, D., & Kleiner, M. (2004). Perception of self-motion and presence in auditory virtual environments. *Proceedings of the Seventh Annual International Workshop PRESENCE 2004,* (pp. 252-258).

Lepecq, J. C., Giannopulu, I., & Baudonniere, P. M. (1995). Cognitive effects on visually induced body motion in children. *Perception*, **24**(4), 435–449.

Palmisano, S., & Chan, A. Y. C. (2004). Jitter and size effects on vection are immune to experimental instructions and demands. *Perception*, (33), 987-1000.

Riecke, B. E. (2003). *How far can we get with just visual information? Path integration and spatial updating studies in Virtual Reality*. Ph.D. thesis, Eberhard-Karls-Universität Tübingen, Fakultät für Physik. Available: www.kyb.mpg.de/publication.html?publ=2088.

Riecke, B. E., Schulte-Pelkum, J., Caniard, F., & Bülthoff, H. H. (2005). Towards Lean and Elegant Self-Motion Simulation in Virtual Reality. *IEEE VR2005,* Bonn, Germany.

Sakamoto, S., Osada, Y., Suzuki, Y., & Gyoba, J. (2004). The effects of linearly moving sound images on self-motion perception. *Acoustical Science & Technology, 25(1), 100-102.*

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

Schulte-Pelkum, J., Riecke, B. E., von der Heyde, M., & Bülthoff, H. H. (2003). Circular vection is facilitated by a consistent photorealistic scene. *PRESENCE 2003.*

Tan, D. S., Gergle, D., Scupelli, P. G., & Pausch, R. (2003). With Similar Visual Angles, Larger Displays Improve Spatial Performance. *Proceedings of CHI 2002*, **5**(1), 195-202.

Todd, N. P. M., & Cody, F. W. (2000). Vestibular responses to loud dance music: A physiological basis of the “rock and roll threshold”? *Journal of the Acoustical Society of America, 107(1),* 496-500.

von der Heyde, M. & Riecke, B. E. (2002). Embedding presence-related terminology in a logical and functional model. In F. Gouveia (Ed.), *PRESENCE 2002*, pp. 37–52. Universidare Fernando Pessoa, Porto, Portugal.

Warren, R., & Wertheim, A. H. (1990), Perception & Control of Self-Motion. New Jersey, London: Erlbaum.

Wist, E. R., Diener, H. C., Dichgans, J., & Brandt, T. (1975). Perceived distance and perceived speed of self-motion - linear vs angular velocity. *Perception & Psychophysics*, **17**(6), 549–554.