Scene Consistency and Spatial Presence Increase the Sensation of Self-Motion in Virtual Reality

Bernhard E. Riecke*
Max Planck Institute for Biological Cybernetics, Tübingen, Germany

Jörg Schulte-Pelkum†
Max Planck Institute for Biological Cybernetics, Tübingen, Germany

Marios N. Avraamides‡
Department of Psychology, University of Cyprus

Markus von der Heyde§
SCC, Bauhaus-Universität Weimar, Germany

Heinrich H. Bülthoff¶
Max Planck Institute for Biological Cybernetics, Tübingen, Germany

Abstract

The illusion of self-motion induced by moving visual stimuli (“vection”) has typically been attributed to low-level, bottom-up perceptual processes. Therefore, past research has focused primarily on examining how physical parameters of the visual stimulus (contrast, number of vertical edges etc.) affect vection. Here, we investigated whether higher-level cognitive and top-down processes – namely global scene consistency and spatial presence – also contribute to the illusion. These factors were indirectly manipulated by presenting either a natural scene (the Tübingen market place) or various scrambled and thus globally inconsistent versions of the same stimulus. Due to the scene scrambling, the stimulus could no longer be perceived as a consistent 3D scene, which was expected to decrease spatial presence and thus impair vection. Twelve naive observers were asked to indicate the onset, intensity, and convincingness of circular vection induced by rotating visual stimuli presented on a curved projection screen (FOV: 54°x45°). Spatial presence was assessed using presence questionnaires. As predicted, scene scrambling impaired both vection and presence ratings for all dependent measures. Neither type nor severity of scrambling, however, showed any clear effect. The data suggest that higher-level information (the interpretation of the globally consistent stimulus as a 3D scene and stable reference frame) dominated over the low-level (bottom-up) information (more contrast edges in the scrambled stimuli, which are known to facilitate vection). Results suggest a direct relation between spatial presence and self-motion perception. We posit that stimuli depicting globally consistent, naturalistic scenes provide observers with a convincing spatial reference frame for the simulated environment which allows them to feel "spatially present" therein. We propose that this, in turn, increases the believability of the visual stimuli as a stable "scene" with respect to which visual motion is more likely to be judged as self-motion. We propose that not only low-level, bottom-up factors, but also higher-level factors such as the meaning of the stimulus are relevant for self-motion perception and should thus receive more attention. This work has important implications for both our understanding of self-motion perception and motion simulator design and applications.

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, ego-motion simulation, human factors, psychophysics, spatial presence, Virtual Reality, spatial orientation

1 Introduction

When exposed to a visual stimulus that depicts a translation or a rotation with constant velocity, most observers experience a compelling illusion of self-motion to the direction opposite to that of the moving stimulus [Dichgans and Brandt 1978; Hettinger 2002]. This illusory self-movement, known as vection [Fischer and Kornmüller 1930], occurs in natural settings when a stimulus occupying a large part of the visual field moves relative to a stationary observer. For example, when a train is leaving the railway station moving in the forward direction, a passenger looking out of the window of a motionless train in the adjacent track may suddenly feel a strong sensation of backward self-movement.

The fact that the illusion of self-motion is sometimes very compelling is demonstrated nicely in a study by Lepecq et al. [Lepecq et al. 1993]. In this study, participants aimed at the memorized locations of targets before and after a period of forward vection. Results revealed that when aiming at the targets after being exposed to vection, participants had modified their responses as if they had really moved forward somewhat (i.e., they overestimated the eccentricity of laterally-placed targets). In carefully designed studies, vection can even be indistinguishable from actual self-motion [Brandt et al. 1973].

When a large visual stimulus starts to move, there is a mismatch between the visual stimulus suggesting self-motion and the lack of any concurrent vestibularly sensed acceleration signal indicating the motion onset and increase in velocity. This visuo-vestibular conflict is thought to be responsible for the fact that it typically takes several seconds from the onset of visual motion before self-motion is perceived (reported values in the literature vary between 2s and 30s). The phenomenon of vection has been traditionally linked to the vestibular system’s limitation of only being able to register accelerations and decelerations and not velocity per se [Zacharias and Young 1981]. That is, as soon as a constant visual velocity is reached and no more accelerations are present, there is no longer any concurrent visuo-vestibular mismatch (apart from the lack of centrifugal and Coriolis forces for rotations). If there would have been an initial physical acceleration and a corresponding vestibular signal, that signal would decay over time, which might explain

---

*e-mail: bernhard.riecke@tuebingen.mpg.de
†e-mail: joerg.sp@tuebingen.mpg.de
‡e-mail: mariosav@ucy.ac.cy
§e-mail: markus.von.der.heyde@scc.uni-weimar.de
¶e-mail: heinrich.buelthoff@tuebingen.mpg.de

The definitive version will be published in Proceedings of APGV 2005,
http://doi.acm.org/10.1145/nnnnnn.nnnn

the possibility for experiencing compelling illusory self-motion if enough time has passed since the initial visual acceleration.

A number of studies have determined several factors relating to the stimulus and the experimental setting that can moderate the onset time, duration, and intensity of the vection sensation. Most of these factors have traditionally been bottom-up parameters, that is, physical stimulus properties. For example, it has been shown that the speed of illusory self-motion increases linearly with increasing stimulus movement velocity up to 120°/s [Brandt et al. 1973]. Furthermore, the solid angle (field of view) subtended by the moving visual stimulus shows a strong influence on vection, with full-field stimulation yielding the most compelling vection and the lowest vection onset time [Brandt et al. 1973; Dichgans and Brandt 1978]. The perceived depth structure of the stimulus has been shown to affect vection as well: When the visual stimulus is comprised of two parts (either superimposed or spatially separate), the part that is *perceived* as more distant primarily determines vection – even when it is, in fact, physically closer [Howard and Heckmann 1989; Ohmi et al. 1987]. However, later studies have provided evidence that the motion of the foreground can also affect vection. For example, when the foreground remains stationary relative to the background [Howard and Howard 1994] or moves slowly to the opposite direction [Nakamura and Shimojo 1999] vection is facilitated. Another factor known to influence vection is the pattern of eye-movement: When viewing a moving visual stimulus, the eyes normally follow the stimulus smoothly (optokinetic nystagmus). When a stationary fixation cross is presented on top of the moving stimulus, the resulting suppression of the optokinetic nystagmus reduces the onset latency of vection significantly, compared to a control condition where the eyes follow the motion of the stimulus in a natural way [Becker et al. 2002; Brandt et al. 1973; Fushiki et al. 2000].

The fact that the majority of studies on vection have focused on bottom-up parameters that affect vection means that relatively little work has been carried out to examine how higher-level and top-down processes like the semantic interpretation of the moving stimulus can affect vection. The possibility that psychological factors can affect the probability of sensing vection or at least modulate its onset latency and strength has been largely neglected by researchers. Nevertheless, we posit that such higher-level factors could have an important role in the perception of vection. For instance, it could be the case that vection is perceived because of our inherent assumption of a stable environment [Dichgans and Brandt 1978]. That is, while during the course of our daily lives we typically move around in the environment, it is only rarely the case that a large, distant portion of our surroundings moves relative to us. As a result, when this happens – in experimental settings or in some rare natural occasions – we are more inclined to attribute the movement to ourselves instead. Perhaps this is why the perceived background of a vection-inducing stimulus is typically the dominant determinant of the presence of vection and modulator of the strength of vection. In daily life, the more distant elements comprising the background of visual scenes are generally stationary and therefore any movement of the elements perceived as more distant is more likely to be interpreted as a result of self-motion [Nakamura and Shimojo 1999; Ohmi et al. 1987].

If indeed vection depends on the assumption of a stable environment, then one would expect that the sensation of vection should be enhanced if the presented visual stimulus (e.g., a virtual environment) is "accepted" as a real world-like stable reference frame. That is, we posit that vection in a simulated environment should be enhanced if participants feel immersed and spatially present in that environment and might thus more readily expect the virtual environment to be stable, just like the real world is expected to be stable. To the best of our knowledge, this hypothesis has not been examined before. We are only aware of a brief commentary paper which stressed the importance of an ecological context and a naturalistic optic array for studying self-motion perception [Wann and Rushton 1994]. Apart from that, past research on vection has traditionally used abstract stimuli like black and white striped patterns or random dot displays. The goal of the present study is to determine whether vection can be modulated by the nature of the stimulus depending on whether it depicts a natural scene that allows for presence or not. On the one hand, the existence of such higher-level contributions would be of considerable theoretical interest, as it challenges the prevailing opinion that the self-motion illusion is mediated solely by the physical stimulus parameters, irrespective of any higher cognitive contributions and the perceptual correlates of the stimuli. On the other hand, it would be important for increasing the effectiveness and convincingness of self-motion simulations: Physically moving the observer on a motion platform is rather costly, labor-intensive, and requires a large laboratory setup and safety measures. Thus, if higher-level and top-down mechanisms could help to improve the simulation from a perceptual level and in terms of effectiveness for the given task, this would be quite beneficial, especially because these factors can often be manipulated with relatively simple and cost-effective means.

1.1 Hypotheses

Participants in our experiment experienced circular vection induced by two types of stimuli: a photorealistic image of a natural scene and scrambled versions of it[1]. Various scrambled version of the stimulus were created by scrambling image parts either in a mosaic-like manner or by slicing the original image horizontally and randomly reassembling it (cf. Fig. 1). Questionnaires administered after the experiment assessed the extent of experienced presence for each experimental stimulus. The purpose of the scene scrambling was to decrease global scene consistency[2] while only slightly changing image statistics (bottom-up contributions, see also hypothesis 3 below). The experimental design was based on the assumption that global scene consistency should increase the believability of the visual stimulus (higher-level effect), as it allows for locomotion and spatial presence in the simulated scene. Conversely, scene scrambling should reduce believability and spatial presence in the virtual environment (i.e., the subjective experience of "being there" in one place or environment, even when one is physically situated in another [Witmer and Singer 1998]), as only the globally consistent stimulus can naturally be recognized or interpreted as a three-dimensional scene, which might in turn allow for actions such as locomotion through the scene. These are all highly cognitive or top-down processes. That is, spatial presence was expected to be highest in the globally consistent (unscrambled) scene and decreasing as scrambling severity increased. Three hypotheses were examined here:

1. **Global scene consistency & presence (higher-level factors):** Based on the stable environment hypothesis, we would predict that global scene consistency and presence should be important factors for vection, as they are expected to increase the believability and perceived depth of the visual stimulus. Hence, the globally consistent (unscrambled) stimulus depicting a natural scene should enhance vection, compared to any of the scrambled or sliced stimuli which are all globally inconsistent, ideally in terms of all response measures. If vection depended only on global scene consistency and presence, the

---

[1] A limited subset of the data has previously been briefly presented at conferences in overview presentations [Riecke et al. 2005b; Riecke et al. 2005c].

[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.

Figure 1: **Top:** 360° roundshot of the Tübingen Market Place. **Middle:** 54x45° view of the 4 stimuli used in one session: Original image and 2, 8, and 32 slices. **Bottom:** 54x45° view of the 4 stimuli used in the other session: Original image and 2x2, 8x8, and 32x32 mosaics per 45°x45° FOV.

various scrambled stimuli should not differ from each other in either the presence ratings or the vection measures. (i.e., A>B=C=D & a>b=c=d; see Figure 1).

2. **Object recognition (higher-level factor):** If, however, vection depends not (only) on presence and global scene consistency but (also) on whether the stimuli contained identifiable elements, then vection should decrease as the severity of scrambling (number of mosaics/slices per solid angle) is increased (i.e., B>C>D, b>c>d). The least vection would then be expected for the most severe scrambling levels (cond. C, D, c, d), where individual objects can hardly be recognized any more.

3. **Number of vertical high-contrast edges (bottom-up factor):** Apart from the higher-level influences discussed above, the mosaic-like scene scrambling also affected physical stimulus properties or so-called bottom-up factors: The mosaic-like scrambled stimuli contained additional vertical high-contrast edges – a bottom-up factor that is known to increase perceived stimulus speed [Distler 2003] and vection [Dichgans and Brandt 1978]. Hence, if these bottom-up factors dominate over higher-level factors like presence, scene consistency, and object recognition, the scrambled stimuli would be expected to increase vection, compared to the sliced stimuli which did not contain such additional vertical edges and thus had a horizontal spatial frequency spectrum quite similar to the globally consistent stimulus (i.e., B>b, C>c, D>d).

2 Methods

12 naive participants (5 female) participated in the study in exchange of monetary compensation. All participants had stereo vision and normal or corrected-to-normal vision.

2.1 Stimuli and Apparatus

Participants were comfortably seated at a distance of 1.8m from a curved projection screen on which the rotating visual stimuli were displayed (cf. Fig. 2). The image was projected using a JVC D-ILA DLA-SX21S video projector. The projection screen had a curvature radius of 2m, and the simulated FOV was set to 54° × 45° to match the physical FOV. Stimuli comprised of several variations of 360° roundshots of the Tübingen market place (4096 × 1024 pixel, see Fig. 1) wrapped around a virtual cylinder which rotated around the vertical axis of the participant producing thus the sensation of circular vection. The horizontally sliced stimuli were created by slicing the 360° roundshot image horizontally, randomly re-ordering the slices, and adding a random horizontal offset (between 0 and 360°) for each slice (see Fig. 1, a-d). Scrambled stimuli were created by subdividing the original image into individual mosaic-like patches which were subsequently reshuffled. The top and the side walls of the cabin were covered with black curtains to reduce visual cues and the physical reference frame of the room. Furthermore, spatial auditory cues were masked by the sound of several layers of flowing water that was played through active noise-canceling headphones that participants wore throughout the experiment. Responses were collected using a Microsoft force feedback 2 joystick that was mounted in front of the participants at a comfortable distance.

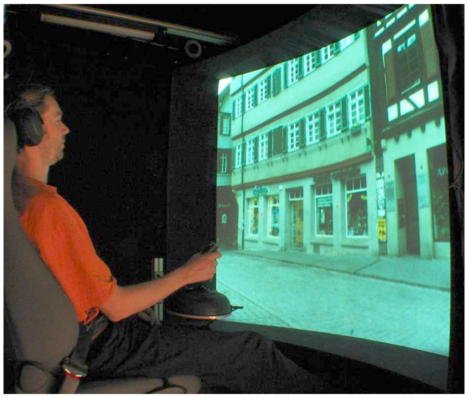

Figure 2: Participant seated in front of curved projection screen displaying a view of the Tübingen market place.

2.2 Procedure

Experimental trials were initiated by participants pressing a button on the joystick, upon which the static image started rotating clockwise or counterclockwise around the vertical axis with constant acceleration for 3s. Maximum rotational velocities were 20°/s and 40°/s. The assignment of trials to the two velocity levels and rotation directions was randomized within the experimental session. The maximum duration of constant velocity rotation was 60s, after which the stimulus decelerated at a constant rate for 6s and came to a stop. Participants were instructed to pull the joystick in the direction of their perceived self-motion as soon as it was sensed. 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 became; this allowed recording the time course of **vection intensity**. The rotation of the stimulus stopped automatically if maximum joystick deflection was sustained for 10s (otherwise it continued for 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 steps of a 0%-100% rating scale. The value of 0% corresponded to "no perceived motion at all" (i.e., perception of a rotating stimulus and a stationary self) and that of 100% to "very convincing sense of vection" (i.e., perception of a stationary stimulus and a rotating self).

During each of the two sessions, participants performed 2 blocks containing 16 trials each. In one session, the scrambled stimuli were presented, in the other session, the sliced stimuli were used. The presentation order of the two sessions was counterbalanced across participants. The experiment followed a 2 (session: mosaic, slices) × 4 (scrambling severity: intact, 2, 8, 32 mosaics/slices) × 2 (rotation velocity: 20, 40°/s) × 2 (turning direction) within-subject factorial design with two repetitions per condition. A pause of 15 seconds was inserted between trials to reduce potential motion aftereffects. In order to familiarize participants with the setup, a practice block containing 4 trials (one for each scrambling severity level) preceded the 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 until they reported a strong sense of ego-motion.

Participants were instructed to watch the stimuli "as relaxed and naturally" as possible. They were also told not to suppress the optokinetic reflex (OKR), and neither to stare through the screen nor to fixate on a static point on the screen, but to concentrate on the image in the central part of the projection screen. We did not use any fixation point, even though it is known that a fixation point reduces vection onset times [Becker et al. 2002; Fushiki et al. 2000]. The main reason was that from an applied perspective for ego-motion simulation, it is more relevant to investigate how one can induce vection under natural viewing conditions, i.e., without a fixation point. Furthermore, this also reduced the perceived flicker and ghost images due to the 60 Hz projection: Even the moderate rotation velocities of 20°/s and 40°/s can already produce strong flicker and ghost images if the eyes fixate one point and do not follow the image motion. For example, a single vertical line translating sideways is seen as multiple flickering lines as it moves across the fixation point.

3 Results

The data for vection onset time, convincingness, and vection intensity are summarized in Figure 3 for reference. Repeated-measures ANOVAs were performed for the three dependent variables using session, scrambling severity, and rotation velocity as factors. Furthermore, correlational analyses between vection measures and the presence questionnaire data were performed.

3.1 Vection onset time

The 3-way ANOVA for vection onset time revealed two significant main effects. First, participants were faster reporting the onset of vection when the stimuli rotated at 40°/s than at 20°/s, $F(1,10)=23.9$, $p<.001$. Second, vection onset times varied depending on scrambling severity, $F(3,30)=6.23$, $p<.01$. More specifically, participants indicated the onset of vection faster with the intact stimuli than any of the scrambled stimuli, all pairwise p's <.05. No differences for vection onset times were obtained among the 2, 8, and 32 slices/mosaics stimuli. Neither the effect of session nor any of the interactions were significant.

3.2 Vection intensity

As in the vection onset time analysis, the only significant statistics for vection intensity were the main effects for rotation velocity and scrambling intensity ($F(1,10)=42.0$, $p<.001$ and $F(3,30)=8.29$, $p<.001$, respectively). Participants indicated stronger vection for stimuli rotating at 40°/s than at 20°/s. Furthermore, vection was rated as stronger for the intact stimulus than any of the 2, 8, or 32 slices/mosaics, all pairwise p's<.05. Intensity ratings did not differ significantly among the 2, 8, and 32 slices/mosaics. Again neither the effect of session nor any interactions were significant.

3.3 Convincingness of vection

The analyses for the convincingness ratings revealed effects that paralleled those of the other two measures. Participants rated as more convincing the illusory self-movement produced by stimuli rotating at 40°/s than at 20°/s ($F(1,10)=23.7$, $p=.001$). Moreover, they rated vection as being more convincing for the globally consistent stimulus than any of the other stimuli ($F(3,30)=41.4$, $p<.001$; all pairwise p's<.001). There was also a significant difference between the 2 and the 8 slices/mosaics ($t(11)=-4.16$, $p<.01$). None of the remaining pairwise tests reached significance. Neither session nor any of the interactions were significant.

3.4 Questionnaires

After each session, participants completed the 14-item Igroup Presence Questionnaire (IPQ, [Schubert et al. 2001]) for each of the

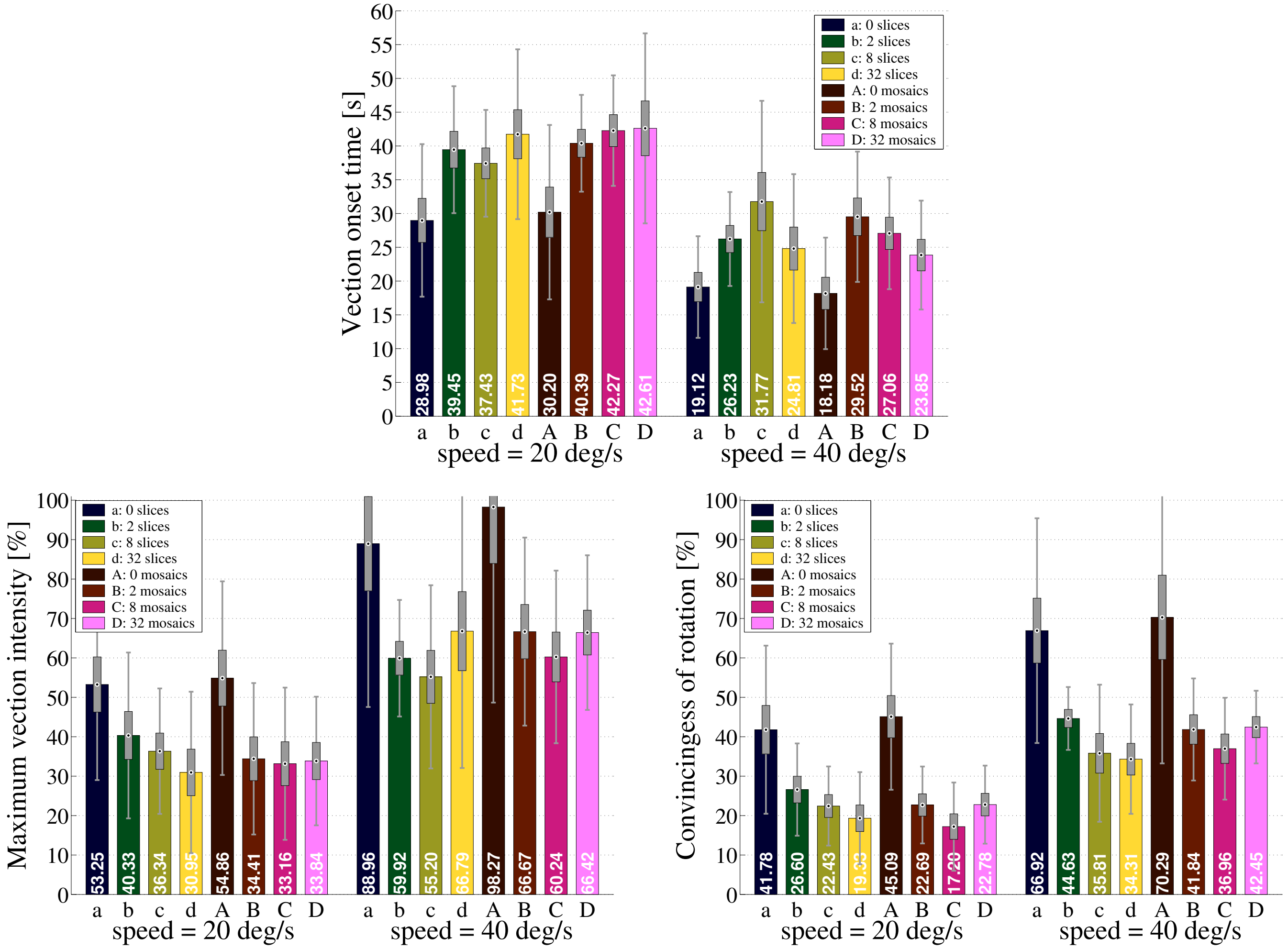

Figure 3: **Top:** Plotted are mean vection onset times for each of the 16 experimental conditions. The left and right group of eight bars represent the low and high velocity conditions (20°/s and 40°/s, respectively). Boxes and whiskers depict one standard error of the mean and one standard deviation, respectively. The eight visual stimulus conditions are explained in the figure legend at the top. For the sake of clarity in all three data plots, 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. Note that this procedure was only applied for the data plotting and not for any of the statistical tests. **Bottom left:** Perceived vection intensity, quantified as the maximum joystick deflection reached. **Bottom right:** Convincingness ratings.

four scenes that were presented in the experimental session. In total, we obtained eight sets of questionnaires from each participant. The questionnaire data are summarized in Figure 4. In our sample, the IPQ showed high reliability ($\alpha = .91$).

To examine the structure and constituent elements of the presence questionnaire, we analyzed similarities and correlations between the responses to the different questions of the IPQ using a factor analysis. The factor analysis revealed a 2-dimensional structure of presence: Factor 1 contained items about realism of the simulated scene and spatial presence (e.g., sense of acting in the virtual environment), while factor 2 contained items that addressed attentional aspects or involvement (e.g., awareness of real surroundings of the simulator vs. the simulated environment). Factor 1 and 2 correspond to the bottom right and middle plots, respectively, in Figure 4.

Furthermore, participants also completed a simulator sickness questionnaire (SSQ) before and after each session [Kennedy et al. 1992]. As expected, the simulator sickness ratings were somewhat higher after the experiment (0.336 ± 0.049SE vs. 0.173 ± 0.048 on a 0-3 point rating scale), but all participants felt comfortable finishing the experiment. An additional presence susceptibility questionnaire (unpublished), which is supposed to measure a person's general susceptibility to presence, did not show any clear results or correlations with any of the vection measures. Therefore, only the results form the IPQ presence questionnaire will be discussed in the following.

Mean presence scores obtained with the IPQ were computed for each level of the sliced or mosaic scenes (see Figure 4, top left plot). A repeated-measures ANOVA with session (slices vs. mosaic) and number of slices (unsliced, 2, 8, and 32) showed a significant effect only for the number of slices (F(3,18)=21.5, p=.001). A post-hoc analysis showed that only the presence ratings of the intact market scene differed significantly from all other levels (Bonferroni-corrected p=.003), but no significant differences between the 2, 8, and 32 slices were found (see Figure 4). That is, two slices were enough to impair presence significantly, and no further decrease in presence was observed for the 8 and 32 slices. Mean presence scores for each of the four original subscales of the IPQ (realism,

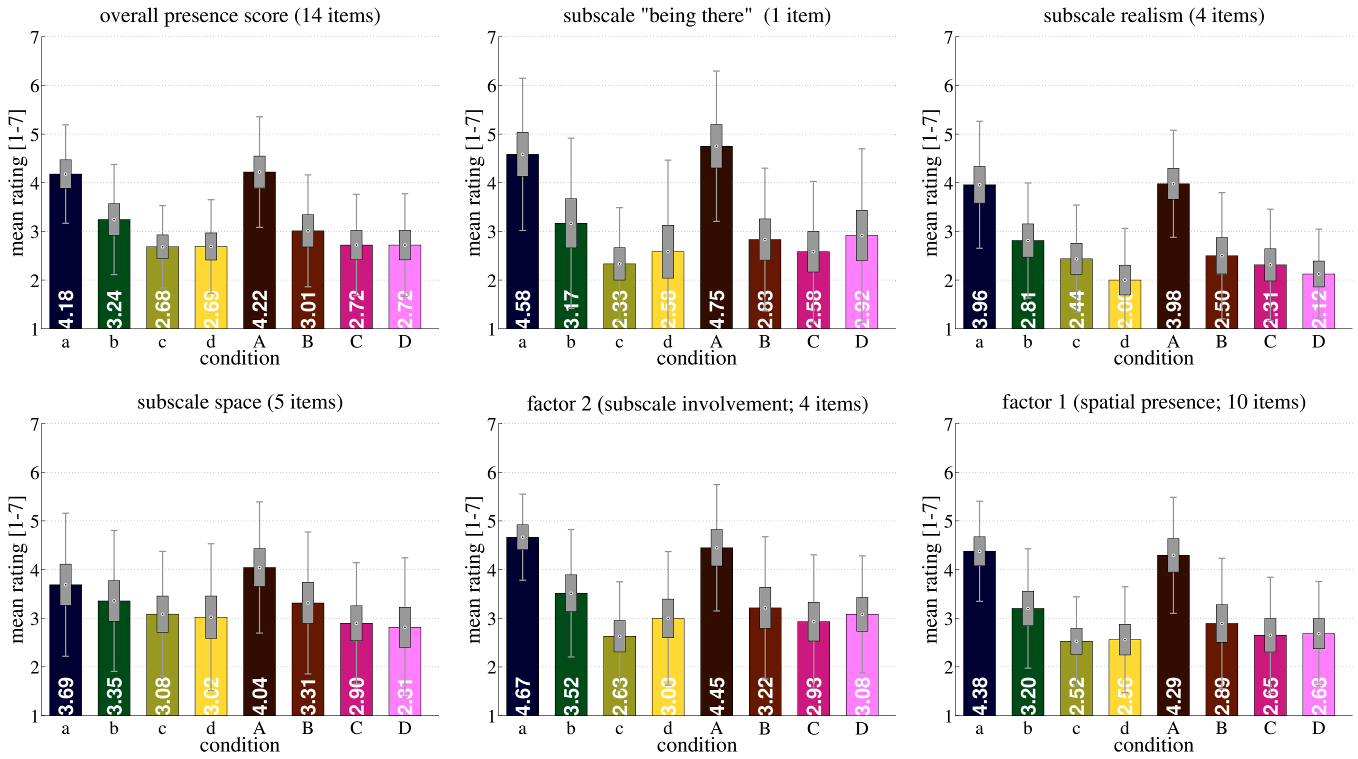

Figure 4: Presence ratings for the eight different visual stimuli. The top left plot shows the mean sum score over all 14 items of the Igroup Presence Questionnaire (IPQ). The four plots next to that show the data that were split up according to the four original subscales as described by Schubert el al.: "Being there", realism, space, and involvement. The remaining two plots on the bottom right show the mean presence ratings according to the results of the factor analysis. The involvement sub-scale coincides with factor 2 of the factor analysis, and the remaining 3 subscales (10 items) constitute factor 1, which can be seen as the spatial presence aspect. It can be seen that only the consistent scenes induced high spatial presence and high attentional involvement. Note the qualitatively similar pattern of results for all scales: Only the intact scenes (a, A) yielded high presence ratings, while all scene scrambling reduced presence consistently.

		overall presence score			factor 1 (spatial presence)			factor 2 (involvement)		
		onset time	convin-cingness	vection intensity	onset time	convin-cingness	vection intensity	onset time	convin-cingness	vection intensity
horizontal										
presence rating	r	-.289	.665*	.209	-.154	.731**	.146	-.427	.323	.247
unsliced	p	.36	.018	.515	.634	.007	.65	.166	.306	.44
presence rating	r	-.675*	.892**	.302	-.462	.853**	.218	-.868**	.631*	.611*
2 slices	p	.016	.001	.208	.131	.0001	.497	.0002	.028	.035
presence rating	r	-.790**	.625*	.313	-.571	.602*	.038	-.802**	.369	.653*
8 slices	p	.002	.03	.321	.051	.038	.908	.002	.238	.021
presence rating	r	-.319	.302	-.004	-.051	.262	-.286	-.62*	.202	.506
32 slices	p	.312	.34	.991	.875	.411	.368	.031	.529	.093
scrambled										
presence rating	r	.086	.776**	-.1	.246	.802**	-.197	-.292	.519	.140
unsliced	p	.791	.003	.757	.44	.002	.54	.357	.084	.664
presence rating	r	-.448	.596*	.364	-.24	.513	.222	-.687*	.467	.492
2 slices	p	.144	.041	.245	.453	.088	.487	.014	.126	.104
presence rating	r	-.728**	.753**	.385	-.564	.728**	.219	-.778**	.461	.600*
8 slices	p	.007	.005	.216	.056	.007	.493	.003	.132	.039
presence rating	r	-.086	.646*	-.057	.102	.518	-.241	-.445	.695*	.302
32 slices	p	.79	.02	.869	.752	.084	.451	.181	.012	.339

Table 1: Table of paired-samples correlations between vection measures (vection onset time, convincingness, and vection intensity) and presence ratings (sum scores (left) and factor 1 (middle) and factor 2 (right) of the factor analysis). N = 12. Note that convincingness ratings correlated highly with the factor 1 (spatial presence) but not with factor 2 (involvement). Conversely, vection onset time was negatively correlated with factor 2 (involvement), but not with factor 1 (spatial presence). Vection intensity was only moderately correlated with factor 2 (involvement), but not at all with factor 1 (spatial presence).

presence, space, and involvement) and also of the compound scales that were merged according to the factor analysis (factor 1 = "spatial presence" and factor 2 = "involvement") are shown in Figure 4.

In the next step, we investigated how presence in the simulated scene related to the different aspects of the self-motion illusion by preforming correlation analyses between the mean presence scores and the three measurands from the vection experiment (vection onset time, vection intensity, and convincingness). Table 1 shows the paired-samples correlations (r) and the corresponding p-values for the overall presence score and the two factor values from the factor analysis, which were interpreted as "spatial presence" (factor 1) and "attention/involvement" (factor 2).

For the overall presence sum score, we found significant correlations between presence ratings and convincingness ratings (7 out of 8 correlations reached significance). Vection onset time correlated negatively with presence (3 out of 8), which means that higher presence was associated with shorter onset times in these cases. No significant correlations were observed between presence ratings and vection intensity. The more detailed analysis revealed significant negative correlations of the involvement scale (factor 2) with vection onset time (4 out of 8), while the compound spatial presence scale (factor 1) correlated positively with the convincingness ratings (4 out of 8). Vection intensity correlated with the involvement scale (2 out of 8), but not with the compound spatial presence scale (0 out of 8). It should be pointed out that given the small sample size (N=12), these correlations are quite substantial.

Interestingly, there was an asymmetry in the correlation results: Factor 1 (spatial presence) correlated mainly with convincingness ratings, whereas factor 2 (involvement) correlated mainly with vection onset time (see Table 1).

4 Discussion & Conclusions

Previous studies have typically used abstract stimuli like black and white geometric patterns to induce vection. Here, we show that the illusion can be enhanced if a natural scene is used instead: The current experiment revealed that a visual stimulus depicting a natural, globally consistent scene can produce faster, stronger, and more convincing sensation of illusory self-motion than more abstract, sliced or scrambled versions of the same stimulus. A possible explanation for why this happens is that natural scenes are less likely to be interpreted as moving because of the assumption of a stable environment [Dichgans and Brandt 1978].

Results from the questionnaires show that the natural, globally consistent scene was also associated with higher presence ratings than any of the sliced or scrambled stimuli. This raises the possibility that presence and vection might be directly linked. It could be the case that vection was facilitated with the natural scene stimulus because participants felt more present in it. Compatible with this hypothesis are the results from the various scrambled stimuli: neither the presence ratings nor the vection onset time or intensity showed any consistent difference in the statistical tests. 2 slices/mosaics were sufficient to reduce presence and impair vection as compared to the natural scene.

In this study, the presence questionnaire showed a two-dimensional structure, namely spatial presence (factor 1) and attention/involvement (factor 2). Furthermore, we found a differential influence of these two factors: While spatial presence was closely related to the convincingness of the rotation illusion, involvement in the simulation was more closely related to the onset time of the illusion. This should be taken into consideration when attempting to improve VR simulations. Depending on task requirements, different aspects of presence might be relevant and should receive more attention or simulation effort.

Even though this study showed a clear correlation between vection and presence, further research is needed to determine if there is actually a *causal* relation between presence and vection, and whether presence might also be affected by the perception of self-motion, as suggested by a recent study using the same VR setup [Riecke et al. 2004]: In that study, vection onset times were unexpectedly decreased when minor scratches were added to the projection screen. These hardly noticed scratches also enhanced vection in terms of both intensity and convincingness ratings. We are not aware of any theoretical reason why these imperfections in the simulation setup should increase presence in the simulated environment. If anything, one might rather expect a decrease in presence. Nevertheless, these minor modifications increased presence ratings significantly, which suggests that the presence increase might have been mediated by the increase in vection.

The finding that both presence and vection measures were no better for the 2 slices/mosaics than the 8 and 32 slices/mosaics rules out the hypothesis that vection was facilitated with the natural scene because it contained identifiable objects (hypothesis 2). Many scene objects could be identified with the 2 slices/mosaic stimuli, whereas the 8 slices/mosaics and the 32 slices/mosaics in particular contained hardly any recognizable objects. Nonetheless, no consistent improvement of vection was observed for the 2 slices/mosaics (apart from a small benefit in terms of convincingness ratings). That is, in contrast to predictions from hypothesis 2, global scene consistency and not the perception of individual objects determined vection and presence.

The data are, however, in full agreement with hypothesis 1: Global scene consistency played the dominant role in facilitating vection, and any global inconsistency reduced vection as well as presence and involvement considerably. Even though it might seem plausible and to be expected that a naturalistic scene induces stronger vection than more abstract stimuli, we are not aware of any previous publications investigating this. Previous research has shown that adding vertical high contrast edges facilitates vection [Dichgans and Brandt 1978]. It has also been found that increasing contrast and spatial frequency of a moving stimulus leads to higher perceived velocity [Distler 2003]. In our study, we found that higher rotational velocities of the stimulus induce vection more easily than slower velocities. Therefore, one would predict that the mosaics should improve vection as compared to the horizontal slices (hypothesis 3). The results of this study showed, however, no such vection-facilitating effect of the additional vertical edges at all. Instead, adding the vertical high contrast edges actually *reduced* vection, compared to the intact stimulus. This suggests the data cannot be convincingly explained by low-level, bottom-up processes alone, and that the bottom-up contributions (more vertical contrast edges in the mosaic-like scrambled stimulus) were dominated by higher-level and top-down processes (consistent reference frame for the intact market scene). This is corroborated by the fact that the additional vertical contrast edges in the mosaic-like scrambled stimulus did not increase vection compared to the horizontally sliced stimulus (which did not have any more vertical contrast edges than the intact stimulus). Higher-level factors that might have contributed include global consistency of the scene and the resulting change in perceived depth, believability of the stimulus, presence and involvement in the simulated scene, and/or the affordance (the implied possibility) of moving though the scene.

As a tentative first explanation, we propose that the globally consistent stimulus allowed for the interpretation as a natural scene, which could in turn have facilitated vection via two possible mechanisms: On the one hand, the pictorial depth cues contained in the depiction of the natural scene might have increased the perceived distance of the moving stimulus, such that it appeared as being further away than the projection screen and the surrounding VR setup. As Howard and Heckmann demonstrated, a moving central display results in stronger vection if it is physically placed behind a stationary surround instead of before [Howard and Heckmann 1989].

Hence, the pictorial depth cues contained in the globally consistent stimulus might have resulted in a *perceived* foreground-background separation and hence a *perceived* background motion, which in turn might have enhanced vection. If this was the case, a physical foreground-background depth separation would not be necessary to enhance vection, and pictorial depth cues would be sufficient. This extends findings by Ohmi et al., who showed that when observers are presented with two similar competing visual stimuli (one moving and one stationary) circular vection occurs only when the moving display is perceived as being behind the stationary display (even though it might physically be closer) [Ohmi et al. 1987].

On the other hand, we propose that interpretation of the globally consistent stimulus as a natural scene allowed for higher believability and presence in the simulated environment and thus provided observers with a more convincing, stable reference frame and primary rest frame with respect to which stimulus motion is being judged more easily as self-motion instead of object or image motion. The proposed mediating influence of presence for the self-motion illusion is in agreement with the "presence hypothesis" proposed by Prothero, which states that "the sense of presence in the environment reflects the degree to which that environment influences the selected rest frame" [Prothero 1998]. Even though further experiments are required to corroborate and further elucidate this phenomenon, the current experiment supports the notion that higher-level mechanisms do indeed affect the visually-induced self-motion illusion, a phenomenon that was traditionally believed to be mainly bottom-up driven. A similar higher-level or top-down influence was observed in a recent study on auditorily induced vection [Larsson et al. 2004]: Presenting sound sources that are normally associated with stationary objects (so-called "acoustic landmarks" like church bells) were more powerful in inducing circular vection than artificial sounds or sound typically generating from moving objects (e.g., foot steps). Hence, we propose that higher-level factors should be considered and further investigated both in self-motion simulation applications and in basic research – where they have been largely neglected apart from a few recent studies [Larsson et al. 2004; Lepecq et al. 1995; Palmisano and Chan 2004; Riecke et al. 2005c]. This could also be advantageous from a practical standpoint: Compared to other means of increasing the convincingness and effectiveness of self-motion simulations like increasing the visual field of view or using a motion platform, higher-level factors can often be manipulated rather easily and without much effort, such that they might be an important step towards a lean and elegant approach to effective ego-motion simulation. This could be achieved through enhancing overall believability and spatial presence in the simulated scene, for example by using a naturalistic stimulus and/or providing consistent multi-modal stimuli (e.g., adding acoustic landmarks to the visual scene [Larsson et al. 2004; Riecke et al. 2005a]). Furthermore, the current study suggests that the effectiveness of motion simulations could also be improved by increasing the perceived distance of the presented scene with respect to the VR setup using, for example, pictorial depth cues.

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

References

BECKER, W., RAAB, S., AND JÜRGENS, R. 2002. Circular vection during voluntary suppression of optokinetic reflex. *Experimental Brain Research 144*, 4 (June), 554–557.

BRANDT, T., DICHGANS, J., AND KOENIG, E. 1973. Differential effects of central versus peripheral vision on egocentric and exocentric motion perception. *Experimental Brain Research 16*, 476–491.

DICHGANS, J., AND BRANDT, T. 1978. Visual-vestibular interaction: Effects on self-motion perception and postural control. In *Perception*, R. Held, H. W. Leibowitz, and H.-L. Teuber, Eds., vol. VIII of *Handbook of Sensory Physiology*. Springer, 756–804.

DISTLER, H. K. 2003. *Wahrnehmung in Virtuellen Welten*. PhD thesis, Justus-Liebig-Universität Gießen.

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

FUSHIKI, H., TAKATA, S., AND WATANABE, Y. 2000. Influence of fixation on circular vection. *Journal of Vestibular Research-Equilibrium & Orientation 10*, 3, 151–155.

HETTINGER, L. J. 2002. Illusory self-motion in virtual environments. In *Handbook of Virtual Environments*, K. M. Stanney, Ed. Lawrence Erlbaum, ch. 23, 471–492.

HOWARD, I. P., AND HECKMANN, T. 1989. Circular vection as a function of the relative sizes, distances, and positions of 2 competing visual-displays. *Perception 18*, 5, 657–665.

HOWARD, I. P., AND HOWARD, A. 1994. Vection - the contributions of absolute and relative visual motion. *Perception 23*, 7, 745–751.

KENNEDY, R. S., LANE, N. E., LILIENTHAL, M. G., BERBAUM, K. S., AND LAWRENCE. 1992. Profile analysis of simulator sickness symptoms: Application to virtual environment systems. *Presence - Teleoperators and Virtual Environment 1*, 3, 295–301.

LARSSON, P., VÄSTFJÄLL, D., AND KLEINER, M. 2004. Perception of self-motion and presence in auditory virtual environments. In *Proceedings of Seventh Annual Workshop Presence 2004*, 252–258. Available: www.kyb.mpg.de/publication.html?publ=2953.

LEPECQ, J. C., JOUEN, F., AND DUBON, D. 1993. The effect of linear vection on manual aiming at memorized directions of stationary targets. *Perception 22*, 1, 49–60.

LEPECQ, J. C., GIANNOPULU, I., AND BAUDONNIERE, P. M. 1995. Cognitive effects on visually induced body motion in children. *Perception 24*, 4, 435–449.

NAKAMURA, S., AND SHIMOJO, S. 1999. Critical role of foreground stimuli in perceiving visually induced self-motion (vection). *Perception 28*, 7, 893–902.

OHMI, M., HOWARD, I. P., AND LANDOLT, J. P. 1987. Circular vection as a function of foreground-background relationships. *Perception 16*, 1, 17–22.

PALMISANO, S., AND CHAN, A. Y. C. 2004. Jitter and size effects on vection are immune to experimental instructions and demands. *Perception 33*, 8, 987–1000.

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

RIECKE, B. E., SCHULTE-PELKUM, J., AVRAAMIDES, M. N., AND BÜLTHOFF, H. H. 2004. Enhancing the visually induced self-motion illusion (vection) under natural viewing conditions in virtual reality. In *Proceedings of Seventh Annual Workshop Presence 2004*, 125–132. Available: www.kyb.mpg.de/publication.html?publ=2864.

RIECKE, B. E., SCHULTE-PELKUM, J., CANIARD, F., AND BÜLTHOFF, H. H. 2005. Influence of Auditory Cues on the visually-induced Self-Motion Illusion (Circular Vection) in Virtual Reality. In *Proceedings of Eigth Annual Workshop Presence 2005*. (submitted).

RIECKE, B. E., SCHULTE-PELKUM, J., CANIARD, F., AND BÜLTHOFF, H. H. 2005. Towards lean and elegant self-motion simulation in virtual reality. In *Proceedings of IEEE VR2005*, 131–138. www.vr2005.org.

RIECKE, B. E., VÄSTFJÄLL, D., LARSSON, P., AND SCHULTE-PELKUM, J. 2005. Top-down and multi-modal influences on self-motion perception in virtual reality. In *HCI international 2005 (accepted)*. www.hci-international.org.

SCHUBERT, T., FRIEDMANN, F., AND REGENBRECHT, H. 2001. The experience of presence: Factor analytic insights. *Presence - Teleoperators and Virtual Environments 10*, 3, 266–281.

WANN, J., AND RUSHTON, S. 1994. The illusion of self-motion in virtual-reality environments. *Behavioral and Brain Sciences 17*, 2 (June), 338–340.

WITMER, B. G., AND SINGER, M. J. 1998. Measuring presence in virtual environments: A presence questionnaire. *Presence - Teleoperators and Virtual Environments 7*, 3, 225–240.

ZACHARIAS, G. L., AND YOUNG, L. R. 1981. Influence of combined visual and vestibular cues on human perception and control of horizontal rotation. *Experimental Brain Research 41*, 159–171.