Max–Planck–Institut für biologische Kybernetik
Max Planck Institute for Biological Cybernetics

Technical Report No. 182

Scene scaling during simulated forward accelerations: Are explicit size cues used?

Daniel R. Berger[1]

January 2009

[1] Department Bülthoff, email: berger@tuebingen.mpg.de

Scene scaling during simulated forward accelerations: Are explicit size cues used?

Daniel R. Berger

Abstract. In many applications for motion simulators, movements which are considerably larger than the motion range of the actual setup are simulated. An important question is therefore how to design inertial movement envelopes and visual scenes so that the simulated movements are perceived as veridically as possible. In the study presented here, we investigated how participants derive an estimate of a visual forward acceleration and how they compare it to an inertially simulated forward acceleration (cross-modal matching).

We tested two possibilities: participants could either derive an *absolute* estimate of visual self-acceleration from objects of known size in the visual scene (images of people), which they compare to an estimate of absolute inertial acceleration; or they could resort to 'range matching', where they do not use absolute estimates but separate *relative* estimates of visual and inertial cues in their respective stimulus ranges, which they then compare.

We found that in our experiment subjects' responses were indeed largely consistent with a 'range matching' strategy. Explicit visual size cues were only used by 4 of the 12 participants, and across all participants, size cues did not have a significant effect on visual-inertial matching.

We conclude that when exposed to linear accelerations in a simulator environment, participants do not necessarily use visual size cues to derive an absolute estimate of the visual acceleration, but that most adapt to the range of presented stimuli, and perceive visual and inertial accelerations as 'matching' if they have similar relative magnitudes. This suggests that one should be careful when interpreting results of cross-modal magnitude matching experiments, since they might be strongly influenced by the ranges of stimuli presented in the experiment.

1 Introduction

In flight or driving simulators, movements of the observer through space can be mediated by different senses. Most frequently visual stimuli are provided to simulate self-motion. Many simulators also have the capability of physically moving the observer, thus also providing inertial information.[1] The physical movements are often not, however, identical to the visually simulated movement, because motion simulators typically have a restricted movement range. For example, continuously accelerating for a distance of more than a few meters is not feasible on a simulator platform. To achieve a good simulation of self-motion, stimuli in the different senses should be chosen so that participants perceive the simulated motion as veridical as possible, and do not notice conflicts between the different movement cues, within or across modalities.

If a forward acceleration is presented to an observer visually, the absolute magnitude of the acceleration may be undefined if the scale of the scene is unknown. Observers can, in principle, estimate the absolute acceleration $a = dv/dt$ if they know the absolute velocity, which can be estimated if distances are known, by $v = ds/dt$. Consequently, if no distance information is available, velocities and accelerations are also ambiguous. The same optic flow pattern on the retina can result from completely different absolute velocities and accelerations of the observer if the scales of the scenes are different. In an optic flow display, absolute distances can, for example, be available if stereo cues are provided, so that observers could use the vergence angle between the eyes to estimate the distance of a visual feature that they are fixating on. Stereo vergence cues are, however, not effective for larger distances and evidence has shown that humans are in fact not very good at estimating absolute distances from eye vergence or accommodation [1]. Another estimate of visual size can be derived from objects of known size in the scene. If such objects of known size, such as people, are standing on a horizontal ground plane which extends to infinity, the eye height of the observer above the ground plane is equal to the height of the intersection of the horizon with the people. We used such a visual stimulus for the study described here (see Figure 3). Observers

---

[1] In this paper, vestibular, somatosensory and proprioceptive modalities are treated together as 'inertial' modalities, because we stimulate them together and do not study them separately.

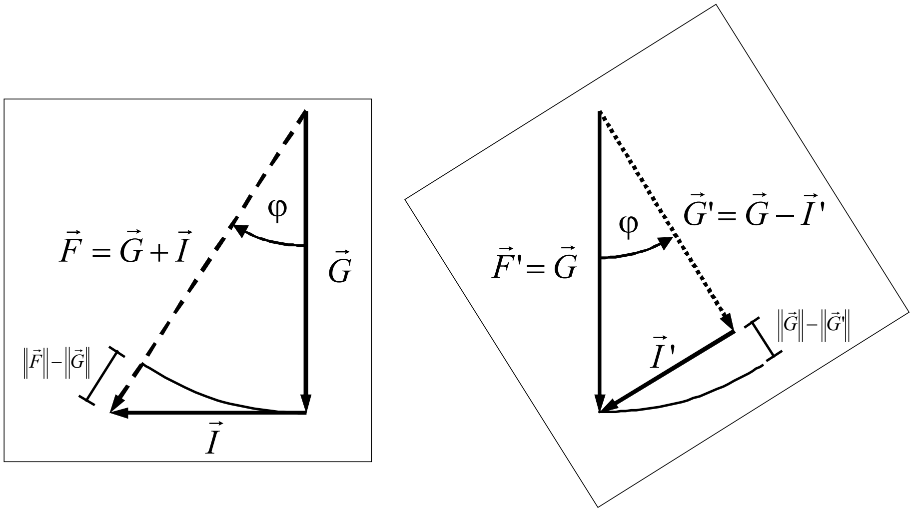


Figure 1: Explanation how a tilt φ of the platform can be used to simulate a linear acceleration with force vector $\vec{I}$ perpendicular to the direction of the gravitational force $\vec{G}$ ('tilt coordination'), by aligning the direction of the combined force vector $\vec{F}$ with the direction of gravity. Left: force diagram during an actual linear acceleration to the right. Right: simulation of the acceleration by tilting the platform.

can, in principle, use the visual information provided to derive an estimate of absolute distance to objects in the scene and from that, their absolute acceleration in the visual scene.

Many flight simulators also provide inertial cues to improve the impression of self-motion. Large inertial accelerations are simulated on a Stewart platform typically by pitching the observer backwards (a method called 'tilt coordination', see [2, 3]), and often combined with other movements of the platform such as brief forward translations or vibrations. During 'tilt coordination', the gravitational force of the earth is used to simulate a gravitational acceleration and a linear acceleration together. Figure 1 shows the basic principle. Intuitively, a pitch backwards and an inertial linear acceleration forwards are perceptually similar since, in both cases, the observer's back is pressed against the back of the seat. For each linear acceleration an equivalent tilt angle can be computed, so that the resulting force vector has the same direction with respect to the observer's body. This principle also allows us to mathematically compare inertial body tilts and visual linear accelerations:

$$\varphi = \arctan(\vec{I}/\vec{G}) \tag{1}$$

$$\vec{I} = \vec{G} \cdot \tan(\varphi) \tag{2}$$

Participants can presumably sense this force vector direction from signals of the otoliths of the vestibular system and from somatosensory and proprioceptive sensors in their skin and body. They could, in principle, use this direction of the gravitoinertial force vector with respect to their body to obtain an estimate of the simulated absolute inertial forward acceleration. It is assumed that the reduction of the magnitudes of the forces during tilt coordination does not play a critical role in the perception of the accelerations and that the tilting is performed in such a way that signals from the vestibular canals, which sense rotations of the head, do not interfere with the percept of the acceleration.

In this experiment we were interested in how participants compare simulated visual and inertial accelerations in a motion simulator and what determines which combination of stimuli in the two modalities are perceived as 'matching' (perceptually equivalent). We investigated two hypotheses. First, participants could derive *absolute* estimates of their acceleration visually and inertially and then compare these absolute estimates to each other when judging how well visual and inertial accelerations match. Second, they could alternatively form a *relative* estimate of the acceleration in each modality separately by comparing it to the range of stimuli presented in that modality during the experimental block, and then compare the two relative estimates across modalities. We call this latter strategy 'range matching'. While 'range matching' depends on the range of stimulus magnitudes presented in the experiment, an 'absolute' estimate is not affected by the sampled range of stimulus magnitudes – a given visual

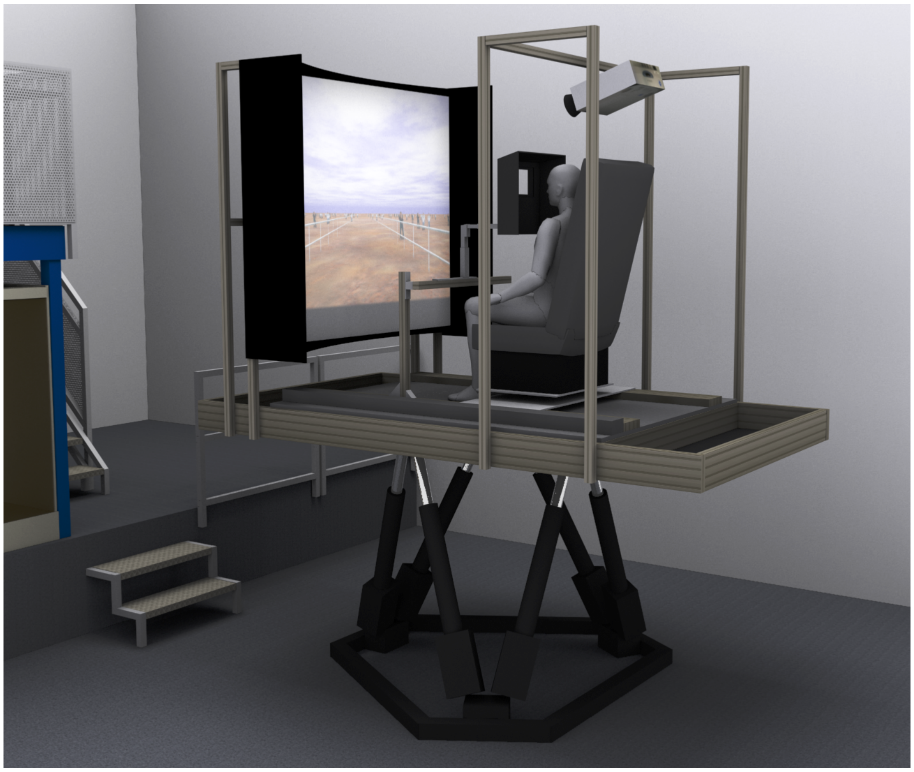


Figure 2: Setup used for this experiment. During the actual experiment, the cabin was completely enclosed by a black curtain

acceleration will always be matched to the same inertial stimulus and vice versa, independent of the other trials in the experiment.[2]

Comparison of stimulus magnitudes on an absolute scale would mean that participants know which magnitude in one modality corresponds to which magnitude in the other modality, either because the mapping is innate or has been learned over the participant's life experiences, and thus, this correspondence should not change as a function of the stimulus ranges that are sampled in the experiment. Comparison on a relative scale, on the other hand, means that such a fixed intermodal correspondence is not used, but that participants compare the stimulus magnitude to magnitudes of other stimuli in the same modality.

If participants use 'absolute' estimates, the question becomes which information do they base these estimates on. An absolute estimate of the inertial acceleration can be derived from the angle (or change of angle, for a change of acceleration) of the gravitoinertial force vector with respect to the body. To get an absolute estimate of the visual acceleration, knowledge about the absolute scale of the visual scene is needed. We were interested in whether participants use the available visual size cues to scale the visual scene, or whether they use some independent scaling. For example, they could ignore the visual size cues and assume a default eye height, such as a height which might be experienced in a car or on a bike.

To differentiate between these two hypotheses, we performed an experiment on a Stewart motion platform equipped with a projection screen (Figure 2), where we presented concurrently simulated visual and inertial linear forward accelerations. The visual forward accelerations were simulated by moving a virtual camera through a computer-generated scene that consisted of a ground plane with people and a fence (Figure 3). Based on parameters defined in a previous study [5, 6], the platform was moved in a manner that resulted in a realistic percept of forward accelerations. After each trial participants were asked to rate how believable the simulated acceleration was in terms of how well the inertially simulated acceleration matched the visual acceleration.

---

[2] What if participants had an absolute estimate of the stimulus in one of the modalities and a relative estimate in the other? Presumably, to make the stimuli comparable, either two relative or two absolute estimates are needed. Thus either the absolute estimate would have to be transformed into a relative one, by comparing it to the stimulus range in the experiment, or the relative estimate would have to be transformed into an absolute one, by 'anchoring' [4]. We do not explore this issue further in the present paper.

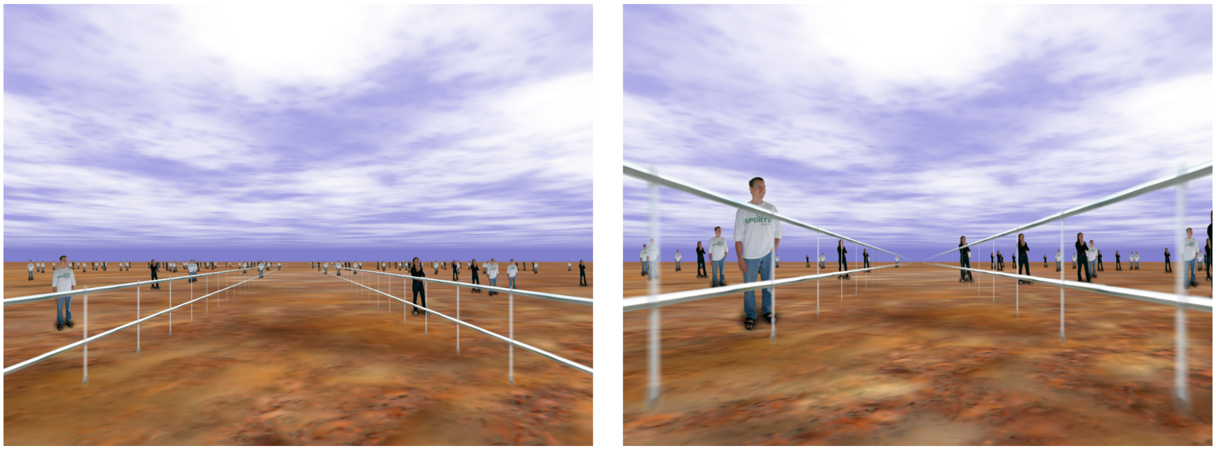


Figure 3: Examples of the visual scene, with small size cues (left; *'standard'* and *'low'* conditions) and big size cues (right; *'big'* condition).

We ran all participants in three different stimulus conditions, *'standard'*, *'low'* and *'big'*, on separate days. Inertial stimuli were equal in all three conditions, but visual stimuli were different. In the *'standard'* condition we presented visual accelerations of up to 2.5 $\mathrm{m/s^2}$ and small visual size cues. To investigate whether ratings depended of the range of visual accelerations in the experimental block, we compared the *'standard'* condition to a *'low'* condition which had the same visual size cues but a lower range of visual accelerations (up to 1.5 $\mathrm{m/s^2}$). If participants use 'range matching', we expect the responses to change as a function of the range of visual accelerations. To investigate whether the scaling of the explicit size cues has an influence on the responses of the participants, we compared the *'standard'* condition to a third condition (*'big'*), where we doubled the size of the explicit size cues while keeping the visual acceleration range equal to the *'standard'* condition. With big size cues the size of the visual scene is effectively doubled, which halves the acceleration indicated by the visual scene. Therefore participants should perceive lower tilt angles as matching, compared to the *'standard'* condition, if they use the size cues to scale the visual scene.

2 Methods

The methods of the experiment presented here largely follow the methods of a previous study [5, 6].

2.1 Subjects

Data was collected from 12 participants (7 female, 5 male, 18-30 years old, mean 24) with normal or corrected-to-normal eyesight and no known history of vestibular or neurological diseases. Participants were paid 8 EUR per hour for participation. They had to read an instruction sheet before the experiment and were also trained before the experiment by completing approximately 20 random practice trials.

2.2 Stimuli and Apparatus

Participants were seated on a Stewart motion platform equipped with a projection screen. An aperture in front of the participant's face covered the frame of the projection screen in their view (Figure 2). They also wore an eyepatch and watched the projection screen with one eye only so that they could not use the convergence angle between the eyes as a cue to estimate distance and, apart from eye accommodation, the only cue to distance was provided by the contents of the visual scene. The viewing distance to the screen was 1.19 m. Viewing height and distance were carefully controlled by adjusting the seat, so that the geometry of the visual scene was veridical (the participant's eye height was equal to the horizon height on the projection screen, and the visual angle of the virtual camera used for rendering the visual scene matched the actual visual angle of the participant).

The visual scene (see Figure 3) consisted of a ground plane and a sky with special random textures covering a large range of spatial frequencies [7]. To define an absolute scale of the scene, we put billboards[3] of people at

---

[3] 'Billboards' are front-facing rectangles which display a 2D-image in a 3D-scene. This rendering method is commonly used in real-time computer graphics.

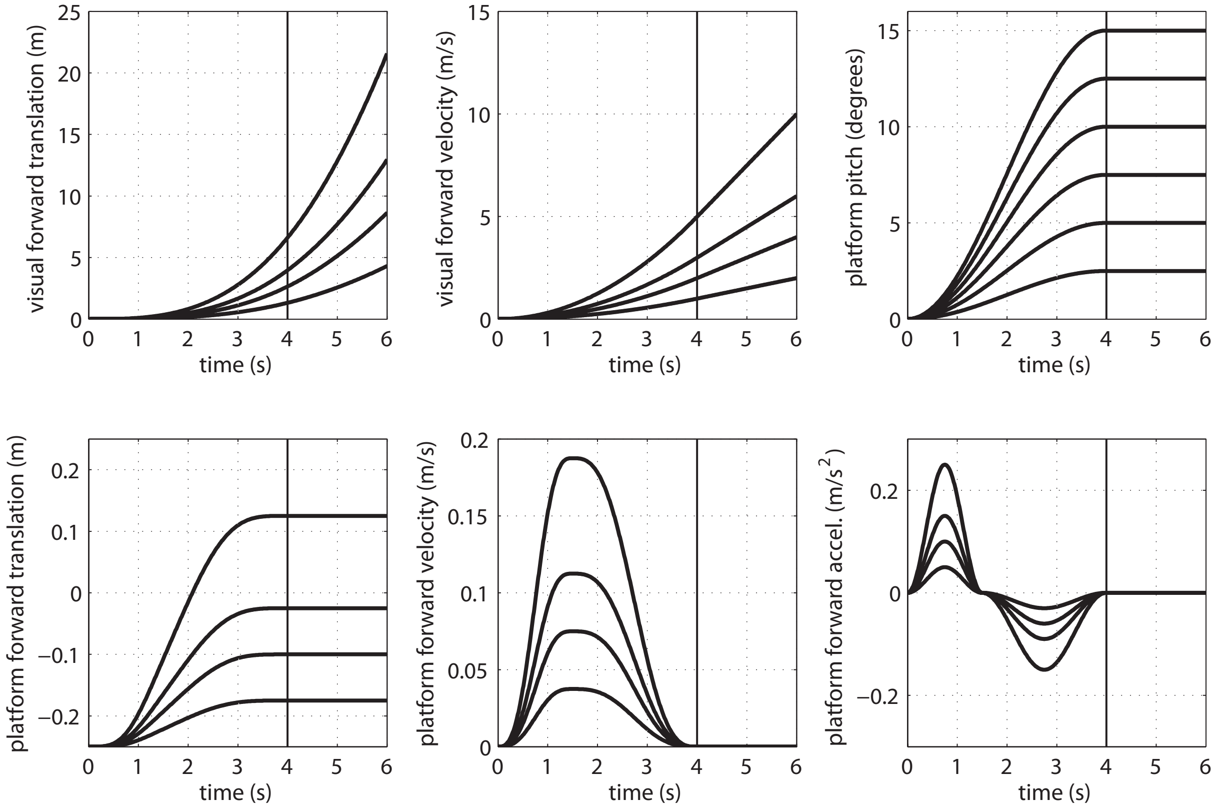


Figure 4: Movement trajectories used for visual forward movements, platform pitch, and platform forward (surge) translations

random locations on the ground plane, randomized for each trial. Also, the start location of the observer on the ground plane was randomized, so that participants could not use features of the ground or sky textures to estimate relative distances. During the forward acceleration, the camera moved horizontally forward over the ground plane on a straight path which was flanked by fences. In the *'big'* condition (Figure 3 right), the textures of the ground plane and the sky were not scaled differently than in the *'standard'* and *'low'* conditions (Figure 3 left), but the size of the billboards was doubled (double width and double height), and the number of the people reduced by a factor of four so that the amount of optic flow would be approximately equal for small and big billboard stimuli. In the *'standard'* and *'low'* conditions, the billboards indicated an eye height of roughly 1.7 m, in the *'big'* condition of roughly 0.85 m.

Inertial forward accelerations were simulated by moving the Stewart motion platform. We used a backwards pitch of the platform to simulate forward acceleration ('tilt-coordination'). In addition to the backwards pitch, the motion platform also performed a brief surge translation in each trial to increase the impression of a forward acceleration. The size of this surge translation was coupled to the visual acceleration, because in a previous experiment it was found that participants' acceleration believability ratings depended significantly on the combination of size of the surge movement and the visual acceleration (rather than on the combination of surge movement size and platform pitch angle) [5, 6]. We chose the size of these movements so that they would maximize the ratings, according to the previous experiment, as 0.15 times the value of the visual acceleration (0.075 m, 0.15 m, 0.225 m, and 0.375 m for 0.5 $\mathrm{m/s^2}$, 1.0 $\mathrm{m/s^2}$, 1.5 $\mathrm{m/s^2}$ and 2.5 $\mathrm{m/s^2}$ respectively). However, from the results of the previous experiment we expect that the surge translation has a smaller effect on the ratings than the platform pitch. Acceleration/deceleration timing of the surge translations (see methods of [5, 6]) was set to a constant value of 0.375.

The resulting trajectories are shown in Figure 4. Each trial consisted of a brief simulated forward acceleration (4 seconds fade-in of the acceleration and 2 seconds constant acceleration), simulated both visually and by platform movements. Then the screen faded to black within 0.5 s and the platform returned to the start position, which took about 4 s. Then the participant was asked to rate the simulated forward acceleration (not the return in darkness) on a continuous scale from 'not believable' to 'very believable', by using an adjustable bar displayed on the screen. Participants used a joystick to set the bar to the desired value and pressed a button. When released, the joystick re-centered automatically. After the button press and a brief pause the next acceleration was presented.

We measured responses in three conditions for all participants, *'standard'*, *'low'* and *'big'*, which were run in separate sessions on separate days. The differences between the three conditions were the range of visual accelerations presented and the size of the explicit visual size cues used. In all three conditions we used the same platform pitch angles to simulate inertial accelerations: $2.5°$, $5°$, $7.5°$, $10°$, $12.5°$ and $15°$. These were paired with three different visual acceleration magnitudes – 0.5 $\mathrm{m/s^2}$, 1.5 $\mathrm{m/s^2}$ and 2.5 $\mathrm{m/s^2}$ in *'standard'* and *'big'* conditions, and 0.5 $\mathrm{m/s^2}$, 1.0 $\mathrm{m/s^2}$ and 1.5 $\mathrm{m/s^2}$ in the *'low'* condition. *'Standard'* and *'big'* conditions differed only in the size of the visual size cues, which indicated an eye height of roughly 1.7 m in the *'standard'* and *'low'* conditions and 0.85 m in the *'big'* condition. Each condition was repeated 8 times, which yielded 6 (platform pitch angles) $\times$ 3 (visual accelerations) $\times$ 8 (repetitions) = 144 trials per person and condition, and therefore resulted in 432 trials per person in total. The 144 trials of each condition were measured in two experimental blocks of 72 trials each with a pause of roughly 15 minutes between them. The order of the trials in one session and the order of sessions across participants were randomized.

2.3 Instructions and Task

Participants were instructed to rate each simulated acceleration according to its 'believability'. We told them that they should always compare the inertial acceleration they feel to the visual acceleration they see (cross-modal matching). If they perceive that the accelerations in both modalities match, this indicates that they felt as if they were accelerated inertially in the same way as they experienced self-motion in the visual display, and the simulation of the acceleration is a 'believable' one. Participants were asked to rate trials as being particularly good if they had the impression that they were actually moving through the scene (i.e. they experienced linear vection), even if they were cognitively aware that the setup could not possibly move that far. On the other hand, we told them to rate trials as being particularly bad if they noticed conflicts between visual movement and inertial movement; for example if they noticed that they were tilted backwards physically instead of being accelerated forward (we explained to them explicitly that we would try to simulate a larger forward acceleration by moving the motion platform in a specific way, which might also involve tilts). They were also instructed to rate trials as being particularly bad if this conflict disturbed their impression of self-motion. Also, they were explicitly instructed not to simply report the amount of self-motion they perceived, but how well the stimulus they felt matched the stimulus they saw. In particular, a low visual acceleration presented together with a low inertial acceleration should also be rated 'good' if it was perceived as matching. After the experiment, participants generally reported that they were able to solve the task intuitively and without much cognitive effort.

Participants were naïve about the purpose of the experiment, and we did not tell them about the differences in the three conditions (different scaling of the size cues and different range of visual accelerations). Some participants reported that they noticed that something was different between the conditions, but most were not sure what the difference was.

3 Results

For each of the three experimental conditions (*'standard'*, *'low'*, *'big'*), we computed mean responses for each participant and for each level of visual acceleration and platform pitch. We also computed predictions for which combinations of visual acceleration and platform pitch should lead to the best ratings under different assumptions. Specifically, we considered whether explicit size cues are used to make a comparison of absolute acceleration magnitudes, or whether participants compare relative magnitudes across modalities (range matching).

For a 'range matching' strategy, we expect that participants rate those trials best in which the smallest visual acceleration is shown together with the smallest platform pitch, the intermediate visual acceleration is shown together with an intermediate platform pitch, or the largest visual acceleration is shown together with the largest platform pitch. We can also make predictions about which stimulus combinations should produce the best ratings if participants scale the scene correctly by using the visual size cues provided and compare the resulting absolute acceleration estimate to the absolute acceleration indicated by the platform pitch, as described by equations 1 and 2.

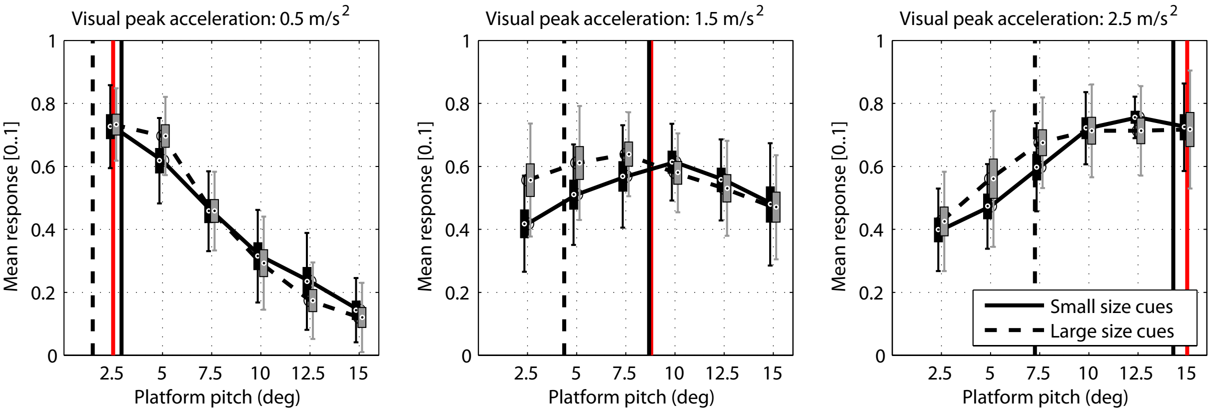


Figure 5: Comparison of the results with small and big size cues for all 12 participants (condition *'standard'* (solid) vs. *'big'* (dashed)). Black vertical lines represent the points at which the ratings should be best if participants compare absolute estimates of the acceleration and scale the visual scene by using the visual size cues. If participants do a simple linear 'range matching' the curves should peak at the red vertical lines. Error bars and whiskers show standard error and standard deviation of the mean responses over all 12 participants.

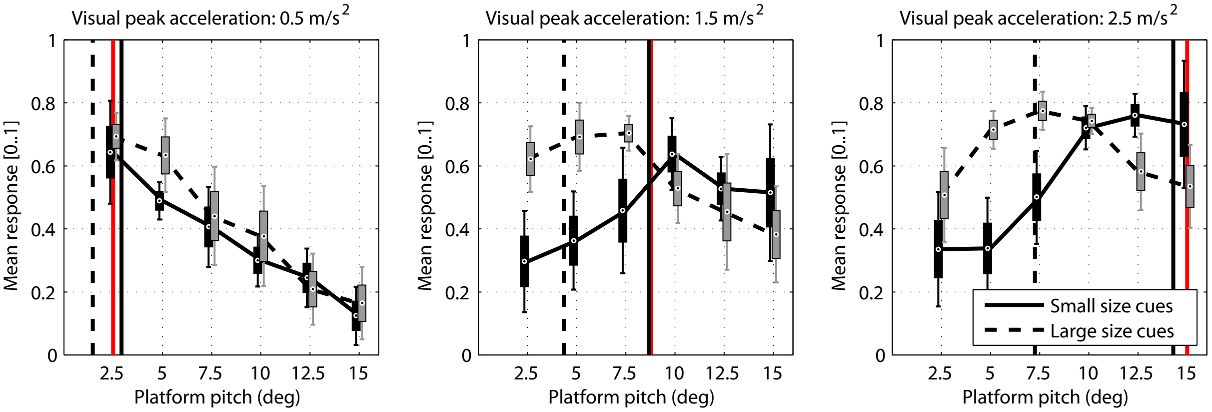


Figure 6: Responses of the four participants who showed the largest response differences as a function of the size cue scaling. See Figure 5 for more details.

3.1 Effect of the visual size cues

If participants scale the visual scene by using the provided explicit visual size cues, we expect to find a difference in the responses if we compare the *'standard'* and the *'big'* condition. The only difference between the stimuli in these two conditions was the scale of the visual size cues (people and fence). Combined results of all 12 participants are shown in Figure 5. As can be seen, the responses are almost identical for peak visual accelerations of 0.5 $\mathrm{m/s^2}$ and 2.5 $\mathrm{m/s^2}$. For visual accelerations of 1.5 $\mathrm{m/s^2}$ there seems to be a difference for small platform pitch angles. A two-way ANOVA with platform pitch and *'standard'* vs. *'big'* conditions as within-subject parameters did not, however, show any significant effects.

When looking at individual responses, it was apparent that four of the twelve participants did show considerable response differences as a function of the size cues. The responses of these four participants are shown in Figure 6. If the visual size cues are used to estimate absolute visual acceleration and the trials that are rated as the best are those in which the absolute visual acceleration matches the acceleration simulated by the platform pitch, the response curves should peak at the platform pitch values indicated by the black vertical lines in the plots of Figure 6. For these four participants, the predicted peaks match quite closely with to the actual peaks of the response curves. This suggests that at least these participants used the visual size cues to estimate an absolute visual acceleration and compared it to the simulated absolute inertial acceleration for their response.

3.2 Effect of the range of visual accelerations

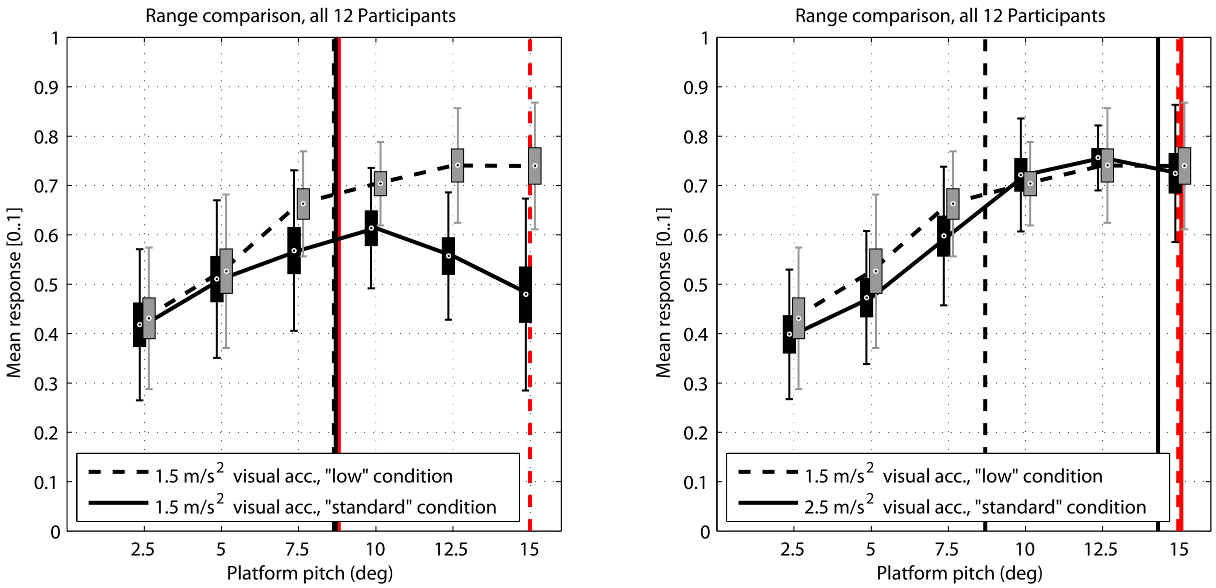
In the *'low'* condition, peak visual accelerations of 0.5 $\mathrm{m/s^2}$, 1.0 $\mathrm{m/s^2}$ and 1.5 $\mathrm{m/s^2}$ were presented, compared to 0.5 $\mathrm{m/s^2}$, 1.5 $\mathrm{m/s^2}$ and 2.5 $\mathrm{m/s^2}$ in the *'standard'* condition. The visual scene was equal in both conditions (i.e. small size cues). Comparing responses in these two conditions thus shows whether participants adapt their ratings to the stimulus range. If their ratings are independent of stimulus range, we should find the same responses for trials with 1.5 $\mathrm{m/s^2}$ peak acceleration in *'low'* and *'standard'* conditions. On the other hand, if we find that responses in trials with 1.5 $\mathrm{m/s^2}$ in the *'low'* condition and 2.5 $\mathrm{m/s^2}$ in the *'standard'* condition are the same, this would suggest that participants actually adjust their responses to the range of presented visual accelerations ('range matching').

Figure 7 shows the responses of all participants. The left plot compares responses to exactly the same stimuli (small size cues and 1.5 $\mathrm{m/s^2}$ visual peak acceleration) in conditions *'standard'* and *'low'*. The main effect of the visual acceleration range on the responses was significant in an ANOVA with 'visual acceleration range' (*'low'* vs. *'standard'*) and 'platform pitch' as within-subject factors: F(1,11)=15.7, p=0.002**. Thus, the range of visual accelerations presented in the experimental block has an influence on the responses, which suggests that participants do not compare absolute estimates of visual and inertial accelerations, but change their responses as a function of the stimulus range.

The right plot in Figure 7 compares responses in *'low'* and *'standard'* conditions at the respective top end of visual peak accelerations, which was 1.5 $\mathrm{m/s^2}$ in the *'low'* condition and 2.5 $\mathrm{m/s^2}$ in the *'standard'* condition. If participants compared estimates of the absolute visual and the inertial acceleration, these response curves should be different. An ANOVA with 'visual acceleration range' (*'low'* vs. *'standard'*) and 'platform pitch' as within-subject factors did not, however, reveal a significant effect of the 'visual acceleration range' factor in this comparison: F(1,11)=0.6, $p > 0.05$. This suggests that stimuli at the upper end of the range of visual accelerations in the experimental block are treated similarly. This again supports the hypothesis that participants respond by comparing relative estimates of visual and inertial accelerations, as a function of the stimulus ranges in the experimental block, not by comparing absolute estimates of the acceleration magnitudes.

Figure 8 shows responses of only those four participants who used visual size cues (same participants as in Figure 6). Even though they rate stimuli with visual peak accelerations of 1.5 $\mathrm{m/s^2}$ (Figure 8, left) generally better in the *'low'* than in the *'standard'* condition, the peaks of both curves coincide, showing that the visual accelerations are matched similarly to the platform pitches. In contrast to the right plot in Figure 7, the right plot in Figure 8 shows a clear difference in the response curves. This suggests again that these four participants are not relying on range matching, but use visual size cues to estimate absolute accelerations, which shifts the peaks of the curves towards the expected values.

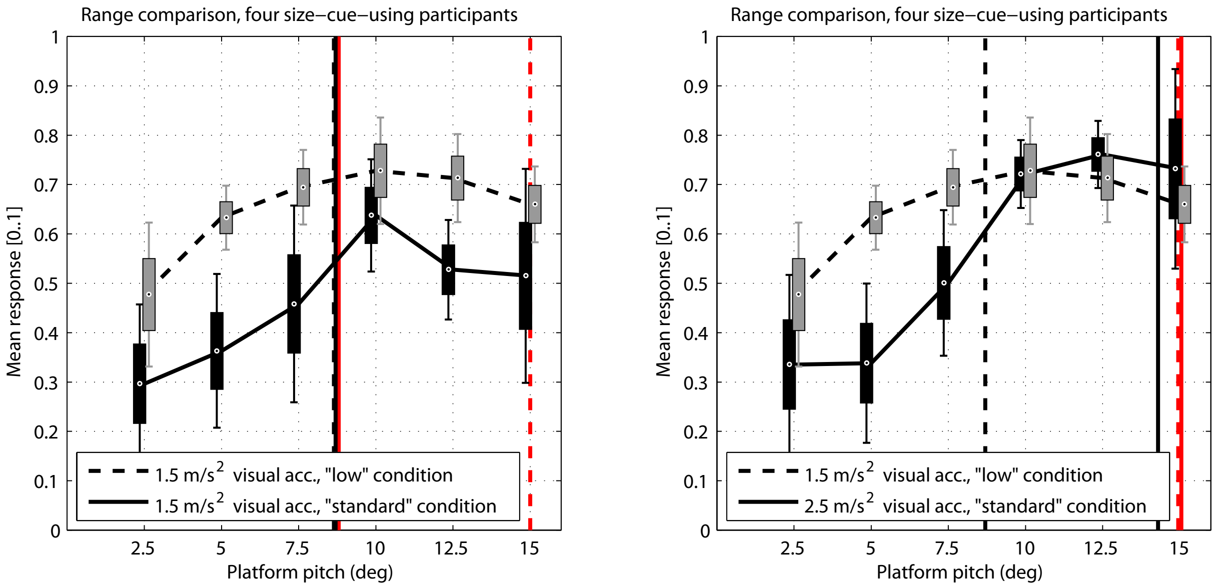
If participants use the range of stimuli in a given experimental session to assign estimates on a relative scale, their ratings might change with experience as knowledge about the range of stimuli occurring during the experiment is obtained. To investigate this, we evaluated the responses of the first time the participant responded to each stimulus combination, versus the following seven times (since we had eight repetitions of each stimulus). There was no apparent difference between first and later responses (see Figure 9). Also, comparing the first and second halves



Range comparison, all 12 Participants

Range comparison, all 12 Participants

Figure 7: Left: Comparison of responses for exactly the same stimuli (1.5 $\mathrm{m/s^2}$ peak visual acceleration) in *'low'* and *'standard'* conditions shows a dependency of the responses on the range of visual accelerations in the experimental block. Right: Comparing 1.5 $\mathrm{m/s^2}$ in *'low'* and 2.5 $\mathrm{m/s^2}$ in *'standard'* conditions shows that stimuli at the upper end of the stimulus range are treated the same, even though they are different. If participants used the visual size cues and compared absolute acceleration magnitude estimates, the curves should peak at the position of the respective black vertical lines (dashed and solid). Red vertical lines show where the dashed and solid curves should peak in the case of 'range matching'.



Range comparison, four size–cue–using participants

Range comparison, four size–cue–using participants

Figure 8: Same as Figure 7, but only showing the data of the four participants that apparently used the visual size cues for their response. See Figure 7 for more details.

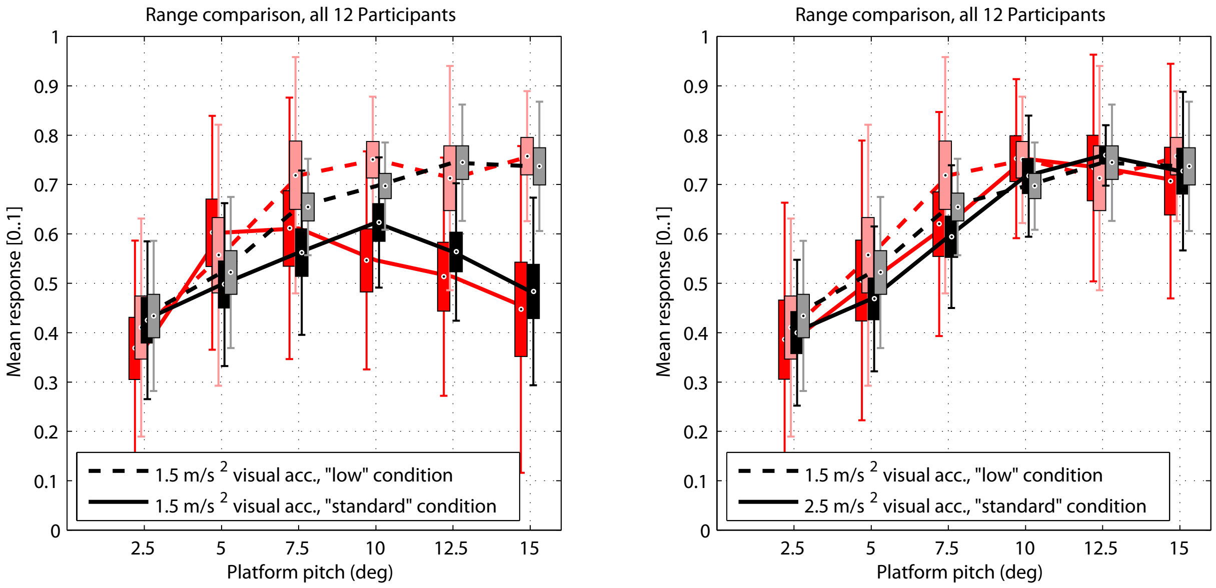


Figure 9: The effect of the visual acceleration range on first (red) and later responses (black) in each condition; compare to Figure 7.

of the experimental conditions (the first versus the second experimental block) did not yield any differences. It is likely that participants had acquired a rough estimate of the stimulus range already after the training trials, or during the first few trials of the experiment. Even though most participants did not receive further training trials before experimental sessions two and three, they would have seen all three visual accelerations in a given experimental block on average after only five to six trials. Since trials were randomized, the trials we used as 'first responses' for the comparison are not exactly the first trials of the experimental session, which might reduce the visibility of a learning effect in Figure 9 if there is one.

4 Discussion and Conclusion

Here we investigated how human participants compare simulated visual and inertial accelerations in a simulator setup. We were interested in whether participants are able to acquire absolute estimates of visual and inertial accelerations by the sensory information provided by each of the modalities, and compare these estimates in order to make their response. If so, we were interested whether participants would use explicit visual size cues to scale the visual scene in order to estimate the visual acceleration. We also considered the alternative hypothesis which was that participants actually use the relative magnitudes of the stimuli within their respective ranges for the cross-modal comparison.

Across all participants, we found that the responses were strongly influenced by the range of visual accelerations presented in the respective experimental block. This suggests that participants compared stimuli across modalities by using their magnitude relative to the range of stimuli presented in a given experimental block (range matching). At the same time, changing the size of the explicit visual size cues did not have a significant effect on the averaged responses. This suggests that, for the most part, participants did not use the visual size cues to scale the visual scene and to make an estimate of an absolute visual acceleration.

However, when looking at the responses of individual participants, we found that four of the twelve participants showed evidence for stimulus-range-independent estimates, and their responses were strongly influenced by the visual size cues. These results are consistent with the assumption that these participants compared absolute visual acceleration estimates (based on the explicit visual size cues), with absolute estimates of the inertial accelerations (based on the pitch angle). This is, however, not proof that these four participants perceived visual and inertial accelerations veridically, since a misestimation of both cues could lead to a similar result. If they would, for example, perceive visual distances as only half as far as simulated, leading to a visual acceleration estimate which is only half as large, and would also interpret the inertial cues as presenting a forward acceleration which is only half as large as intended, we would expect the same response pattern as with veridical perception of both simulated cues. A critical test would be to have participants compare the magnitudes of simulated and actual forward accelerations using, for example, a two-interval-forced-choice task.

Our results suggest that absolute estimates of the visual and inertial stimuli are not unconditionally and automatically used by the participants to make cross-modal comparisons of acceleration magnitude. Instead, many participants apparently use relative magnitude estimates by matching the ranges of stimulus magnitudes to each other. In a natural environment, when visual and inertial accelerations are largely consistent, such relative estimates (i.e. relative to the range of stimulus magnitudes occurring) can be sufficient to perform accurate cross-modal stimulus comparisons. The comparison of relative estimates is also possible when stimuli are compared that have no natural dependency and are only correlated in a specific situation; for example if the brightness of a light is to be compared with the loudness of a beep. A similar mechanism might be used if a participant has to report the magnitude of a stimulus on an arbitrary scale during an experiment (for example report a number between 1 and 10, or, like in this experiment, set a visual bar). When there is no clearly defined absolute match between stimuli that are being compared (or maybe if an absolute estimate is less reliable than a relative one), participants can then resort to a strategy for which they compare relative estimates instead of absolute ones.[4] After all, being passively accelerated is not a very natural stimulus. The fact that most participants use a comparison of relative magnitudes might indicate that our stimuli are not very 'natural', even though participants reported that the accelerations appeared more or less realistic.

We have to stress the point that the 'range matching' we found in this study is not the same as the 'range effect' which is often reported in psychophysical studies. The latter is an effect that is exhibited on the response axis. Specifically, this is the observation that, for certain tasks, participants tend to more frequently report values that are closer to the middle of a limited response scale. This can reduce the variance of the response values, but should not influence which combinations of stimuli are rated as matching best. In our study, a 'range effect' would exhibit itself as a tendency for the participants to rate trials closer to the mean of the believability response scale more often, and to avoid extreme responses. 'Range matching', however, is an effect which is found when two stimuli defined along separate dimensions have to be compared, so it concerns evaluating the matching of one stimulus to another and not the matching of a stimulus to a response.

Since we asked participants to report the perceived 'believability' of the accelerations, it appears that the presentation of visual and inertial accelerations which are matching in relative magnitude can be sufficient to induce a 'believable' percept of a forward acceleration in a simulator. But even though participants were instructed to base their ratings on a very specific criteria, we cannot be completely sure that they in fact truly reported 'believability' and not 'magnitude compatibility' across sensory modalities instead (i.e. simply how well they matched without the experience of true self-motion).

To our knowledge there is no previous study that has investigated the role of visual size cues for visual-inertial crossmodal matching during simulated forward accelerations. A few studies have investigated the simulation of forward accelerations with visual and inertial cues [8, 9, 2, 5, 6], but they did not address how the visual scene is scaled in perception and whether crossmodal comparisons in such an experiment are based on absolute or relative estimates. However, Groen et al. (2000, 2001) [8, 9] and Berger et al. (2007, 2009) [5, 6] investigated which platform pitch magnitude matched a given visual acceleration best. Groen et al. (2000, 2001) [8, 9] always presented the same visual acceleration (3.43 $\mathrm{m/s^2}$) and combined it with platform pitches with gains between 0 and 1 (simulating inertial accelerations between 0 and 3.43 $\mathrm{m/s^2}$). They found that participants rated the simulations best when an average platform pitch was presented (gain of 0.6). This is not really 'range matching' since there is no range of visual accelerations in the experiment, but matching the visual stimulus to approximately the average inertial stimulus seems to be a sensible thing to do when using relative estimates for the comparison.

The results from Groen et al. (2004) [2] are slightly more difficult to interpret because they did not use a complete experimental design with visual and inertial stimuli taken from fixed ranges, but combined each of three different visual accelerations with different ranges of platform pitches. For low sinusoidal movement frequencies, their results were more or less consistent with the usage of absolute acceleration estimates rather than range matching. Their visual scene consisted of a corridor, which could, in principle, provide more or less accurate visual scale information.

In Berger et al. (2007, 2009) [5, 6], we used the same setup and very similar stimuli to this experiment, and in that study we found that a relatively large range of platform pitch gains were rated almost equally well (approxi-

---

[4] Maybe one could even do an 'ideal observer' analysis to find out under which sensory conditions relative or absolute estimates should ideally be used.

mately 0.7 to 1.8), so that both the comparison of absolute estimates (expected best gain 1.0) and 'range matching' (expected best gain 1.72) are consistent with the experimental results.

The issue of absolute versus relative perception is also addressed in other areas of perception research. On example is the investigation of brightness perception, where the brightness of the brightest spot in a visual display has an effect on how bright all visual items in the display are perceived. This process, by which absolute brightness is derived from relative brightness, is called 'anchoring' (see Gilchrist et al. (1999) [4] for review).

Gogel (1969) [10] investigated the importance of explicit size cues for the disambiguation of absolute size and absolute distance of objects which are viewed in isolation in a monocular display. He found that participants can use the known size of familiar objects to disambiguate absolute size and distance under these conditions, but since objects of known size can be perceived as off-sized, also the retinal size plays a role in the estimate. In that study and in related studies it was found that if no distance cues are available, objects are perceived at a 'prior' distance of approximately 1.2 - 3.7 m. Concerning the issue of relative versus absolute estimates, Gogel (1963) [1] argued that humans do not perceive absolute distances, from which they would derive relative distances of different objects, but that the basic perceptual data is relative distances, from which absolute distances may be estimated. This is reminiscent of our results, which suggest that responses are dominated by relative estimates of the simulated accelerations.

If participants use relative rather than absolute estimates for cross-modal comparisons, the question arises what the implications are for multisensory integration. To be integrated, cues in different modalities must be brought into a commensurate representation, and whether two sensory cues should be integrated or not depends on whether or not they are in conflict ('robust integration'; see [11]). Recently there has been some recurring interest in this topic; see [12, 13, 14]. One could assume that during experience in natural environments, neural systems of multisensory integration should have learned which visual and inertial stimuli are equivalent so that the mapping of a stimulus in one modality to a stimulus in another should be constant and known. However there is evidence that the nervous system can adapt to sensory discrepancies, for example when participants wear prism glasses. This suggests that the matching of signals in different modalities in the brain is flexible. The system also needs to cope with effects of muscle fatigue, as well as with other forms of plasticity. Therefore, a mapping of relative magnitudes ('range matching') might often be a more robust mechanism for multisensory integration than relying on estimates of absolute stimulus magnitude. Still, the relationship of mechanisms of adaptation, conflict detection and integration needs further investigation. This work should also be extended to evaluate whether firstly range-relative estimates as described in this paper, secondly the classical response 'range effect', and thirdly effects of 'adaptation' to the stimulus situation are in fact different instantiations of the same underlying process.

Also, in this experiment, we only manipulated the upper end of the range of visual accelerations. It would be interesting to see whether the adaptation of the range for cross-modal matching also accounts for the lower end of the stimulus range, and whether a manipulation of the range of inertial stimuli yields equivalent results. Also it would be interesting to see whether 'range matching' takes place in other multisensory situations.

In conclusion, we found that when exposed to linear accelerations in a simulator environment, participants do not necessarily use visual size cues to derive an absolute estimate of the visual acceleration. Most participants adapt to the range of presented stimuli, and perceive visual and inertial accelerations as 'matching' if they have similar relative magnitudes (relative to the respective range of stimuli in the experimental block). This suggests that one should be careful when interpreting results of cross-modal magnitude matching experiments, since they might be strongly influenced by the ranges of stimuli presented in the experiment.

Acknowledgments. The author wants to particularly thank Jenny Campos for helpful discussions and comments on the manuscript.

References

[1] Walter C. Gogel. The visual perception of size and distance. *Vision Research*, 3:101–120, 1963.

[2] E. L. Groen and W. Bles. How to use body tilt for the simulation of linear self motion. *Journal of Vestibular Research*, 14(5):375–385, 2004.

[3] L. D. Reid and M. A. Nahon. Flight simulation motion-base drive algorithms. Part 1: Developing and testing the equations. Technical Report UTIAS report 296, University of Toronto, Institute for Aerospace Studies, 1985.

[4] Alan Gilchrist, Christos Kossyfidis, Frederick Bonato, Tiziano Agostini, Joseph Cataliotti, Xiaojun Li, and Branka Spehar. An anchoring theory of lightness perception. *Psychological Review*, 106(4):795–834, 1999.

[5] Daniel R. Berger, Jörg Schulte-Pelkum, and Heinrich H. Bülthoff. Simulating believable forward accelerations on a stewart motion platform. *Max Planck Institute for Biological Cybernetics, Technical Report*, 159, February 2007.

[6] Daniel R. Berger, Jörg Schulte-Pelkum, and Heinrich H. Bülthoff. Simulating believable forward accelerations on a stewart motion platform. *ACM Transactions on Applied Perception*. in press.

[7] Daniel R. Berger. Spectral texturing for real-time applications. In *Siggraph 2003 Sketches and Applications*. ACM Press, New York, 2003. DOI 10.1145/965400.965509.

[8] Eric L. Groen, Mario S. V. Valenti Clari, and Ruud J. A. W. Hosman. Psychophysical thresholds associated with the simulation of linear acceleration. In *Proceedings of the AIAA Modeling and Simulation Technologies Conference, Denver (CO), August 14-17, (paper number AIAA 2000-4294)*, pages 1–9. TNO Human Factors, American Institute of Aeronautics and Astronautics, Inc., 2000.

[9] Eric L. Groen, Mario S. V. Valenti Clari, and Ruud J. A. W. Hosman. Evaluation of perceived motion during a simulated takeoff run. *Journal of Aircraft*, 38:600–606, 2001.

[10] Walter C. Gogel. The effect of object familiarity on the perception of size and distance. *The Quarterly Journal of Experimental Psychology*, 21(3):239–247, 1969.

[11] Michael S. Landy, Laurence T. Maloney, Elizabeth B. Johnston, and Mark Young. Measurement and modeling of depth cue combination: in defense of weak fusion. *Vision Research*, 35(3):389–412, 1995.

[12] Neil W. Roach, James Heron, and Paul V. McGraw. Resolving multisensory conflict: a strategy for balancing the costs and benefits of audio-visual integration. *Proc. R. Soc. B*, 273:2159–2168, 2006.

[13] Marc O. Ernst. A bayesian view on multimodal cue integration. In G. Knoblich, I.M. Thornton, M. Grosjean, and M. Shiffrar, editors, *Human body perception from the inside out*, pages 105–131. Oxford University Press, New York, 2006.

[14] Konrad P. Körding, Ulrik Beierholm, Wei Ji Ma, Steven Quartz, Joshua B. Tenenbaum, and Ladan Shams. Causal inference in multisensory perception. *PLoS ONE*, 2(9):e943, 2007. doi:10.1371/journal.pone.0000943.