Revisiting Detection Thresholds for Redirected Walking: Combining Translation and Curvature Gains

Timofey Grechkin* Jerald Thomas* Mahdi Azmandian* Mark Bolas*† Evan Suma*

*USC Institute for Creative Technologies; †USC School of Cinematic Arts

Figure 1: *Physical and virtual setup for estimation of detection thresholds: Participant wearing head-mounted display walking with an experimenter (left) and test virtual environment (right).*

Abstract

Redirected walking enables the exploration of large virtual environments while requiring only a finite amount of physical space. Unfortunately, in living room sized tracked areas the effectiveness of common redirection algorithms such as Steer-to-Center is very limited. A potential solution is to increase redirection effectiveness by applying two types of perceptual manipulations (curvature and translation gains) simultaneously. This paper investigates how such combination may affect detection thresholds for curvature gain. To this end we analyze the estimation methodology and discuss selection process for a suitable estimation method. We then compare curvature detection thresholds obtained under different levels of translation gain using two different estimation methods: method of constant stimuli and Green's maximum likelihood procedure. The data from both experiments shows no evidence that curvature gain detection thresholds were affected by the presence of translation gain (with test levels spanning previously estimated interval of undetectable translation gain levels). This suggests that in practice currently used levels of translation and curvature gains can be safely applied simultaneously. Furthermore, we present some evidence that curvature detection thresholds may be lower that previously reported. Our estimates indicate that users can be redirected on a circular arc with radius of either 11.6m or 6.4m depending on the estimation method vs. the previously reported value of 22m. These results highlight that the detection threshold estimates vary significantly with the estimation method and suggest the need for further studies to define efficient and reliable estimation methodology.

Keywords: Virtual Reality, Locomotion, Perception

Concepts: •**Computing methodologies** → **Perception; Virtual reality;** •**Human-centered computing** → *Empirical studies in HCI;* User studies;

1 Introduction

Physical walking in virtual spaces can create truly compelling user experiences. However, the exploration of large virtual environments is constrained by the size of available physical space. *Redirected walking* (RDW) [Razzaque 2005] attempts to address this issue through subtle manipulation of the mapping between physical and virtual movement to decouple the user's virtual path from the real-world trajectory. As a result, users can explore relatively large virtual environments while being physically contained within the boundaries of the tracked space. RDW provides the numerous benefits of unconstrained physical walking in virtual environments [Usoh et al. 1999; Ruddle and Lessels 2009; Ruddle et al. 2011; Suma et al. 2010] at the limited cost of some cognitive load on the user [Bruder et al. 2015].

There exists a trade-off between redirection intensity and the likelihood of users noticing underlying perceptual manipulations. Typically, the intensity of each type of manipulation (reffered to as *gain*) is bounded by its corresponding *detection threshold*, which limits the effectiveness of the overall redirection. A recent study by Azmandian et al. [Azmandian et al. 2015] shows that there is little benefit in deploying common general-purpose RDW algorithms such as Steer-to-Center in living room scaled tracked areas smaller than 6×6 meters. At the same time this study suggests that the combination of Steer-to-Center algorithm and scaled translations similar to those used in the *Seven League Boots* metaphor [Interrante et al. 2007] can significantly improve performance in small tracked areas by reducing the number of contacts with the boundaries of tracked space. However, the effect of combined gains on detection thresh-

*e-mail: {grechkin, jthomas, mazmandian, bolas, suma}@ict.usc.edu

SAP '16, July 22 - 23, 2016, Anaheim, CA, USA
ISBN: 978-1-4503-4383-1/16/07
DOI: http://dx.doi.org/10.1145/2931002.2931018

olds is unknown since previous studies have examined the detection threshold for each gain independently, in the absence of other gains.

In this paper we revisit the problem of estimating detection thresholds to explore the effects of simultaneous application of translation and curvature gains. We discuss the issue of finding a suitable threshold estimation method and explore alternative options. We also present two experiments that employ two different estimation methods to obtain curvature detection thresholds in the presence of translation gain.

2 Background

The human perceptual system has naturally evolved to perceive the surrounding environment as stable during head movements that occur along with locomotion. Studies show that the environment can still be perceived as stable even when there is a certain amount of discrepancy between physical head movement and observed visual movement [Jaekl et al. 2005]. The prominence of visual information during locomotion can be further illustrated by the observation that humans tend to walk in circles when attempting to follow a straight line in an unfamiliar environment in the absence of prominent visual direction cues [Souman et al. 2009].

Redirected walking takes advantage of these natural properties of the perceptual system to imperceptibly decouple virtual and real-world user trajectories in order to condense the latter into available physical space. This goal is achieved using three primary types of perceptual manipulations [Suma et al. 2012]: *translation*, *rotation*, and *curvature* gains.

Translation gain scales virtual translations relative to the real world movement, resulting in faster or slower displacement in the virtual world. Translation gain is typically expressed as ratio of virtual to real displacement. When no translation gain is applied, this ratio equals to 1.0. Values above 1.0 indicate that virtual translation is faster than the actual displacement in the real world, and values below 1.0 indicate that virtual translations are smaller than the real ones.

Rotation gain applies scaling to rotations, effectively increasing or decreasing the amount of virtual rotation relative to user's real-world movement. Similar to translation gain, rotation gain is defined as ratio of virtual rotation to physical rotation.

Finally, *curvature* gain induces virtual rotations when the user is walking (i.e. primarily translating) in the real environment. These rotations are perceived by the user as error in maintaining the course towards the intended target and the resulting course correction introduces a curvature to the real-world trajectory. Therefore, a straight virtual path is realized as a curved path in the real-world environment. Curvature gain is expressed as ratio of $\frac{1}{r}$, where r is the radius of the arc-shaped path travelled in the real world. In the absence of curvature ($r = \infty$) curvature gain is equal to 0; gain values larger than zero correspond to greater curvature curvature intensities (i. e. tighter real-world arcs).

A redirection algorithm in its general form determines the most suitable type and intensity of gain to apply at each moment as part of a strategy to keep users within bounds of the available physical space. To ensure these gains are unnoticed by the user, it is important to know perceptual *detection thresholds* for each type of gain.

2.1 Detection thresholds for redirected walking

A *detection threshold* indicates the maximum intensity of a particular type of redirection gain that remains unnoticed by a user. The first informal assessment of detection thresholds for curvature gain was attempted by Razzaque [Razzaque 2005], in his PhD dissertation. However, these initial experiments were conducted with a very small number of subjects and were limited to only 20 trials, preventing him from obtaining reliable threshold estimates.

The most authoritative and comprehensive study on this subject to date was conducted by Steinicke et al. [Steinicke et al. 2010]. This study consists of three experiments designed to estimate psychometric functions and perceptual detection thresholds for rotation, translation, and curvature gains respectively. Detection thresholds for rotation gain have been estimated at 49% for positive (same direction as real world rotation) gain and 20% for negative (opposite to real world rotation) gain. For translation gain the thresholds were estimated at 26% for up-scaling and 14% for downscaling. Finally, according to their data for curvature gain to remain undetected, the circular walking arc needs to have a radius of at least 22m.

Interestingly, the estimated threshold values vary significantly from one study to another. For rotation gain, Jerald et. al [Jerald et al. 2008] estimated detection thresholds at 11.2% for positive (same direction as real world rotation) gain and at 5.2% for negative (going opposite to real world rotation) gain. For curvature gain, Hodgson and Bachmann [Hodgson and Bachmann 2013] successfully used an informally estimated curvature radius of 7.5 meters with no users reporting noticing redirection.

While several factors might have contributed to this variation, it is clear that the estimation methodology played a critical role. In an earlier version of their study Steinicke et. al [Steinicke et al. 2009] used a different version of the experimental task (which authors believed might have introduced a bias in participants' responses). This study yielded a very different set of threshold estimates for both rotation (41% and 10% respectively) and curvature (15 meter radius) gains. In the next section we will take a closer look at the alternative experimental methodologies that can be used to estimate perceptual detection thresholds as well as the key properties of the redirected walking domain that could affect the choice of the most suitable method, particularly for curvature gain.

2.2 Estimation methods for detection thresholds

The first step in designing a method for detection threshold estimation is choosing the type of experimental task to be used. Detection thresholds can be estimated using either a yes/no or a 2-alternatives-forced-choice (2AFC) task [Klein 2001]. In each trial of a yes/no task, a participant attempts to detect whether a signal was present and provide a binary yes/no response. The performance is summarized as the proportion of correct responses and the detection threshold is typically defined as the stimulus level corresponding to the 50% detection rate. In a 2AFC task participants are required to compare 2 alternatives. For instance, in auditory research participants might be asked to determine in which of two consecutive time intervals a sound tone was presented. In another version, participants may compare two alternative stimuli to determine which one possesses a more pronounced characteristic of interest (i.e. which sound tone is higher). Signal detection theory suggests that 2AFC tasks may be helpful to discourage response biases.

Steinecke et al. [Steinicke et al. 2010] describe their estimation methodology as a 2AFC task. This is based on the fact that participants are forced to choose among two possible answers. For example, when attempting to detect curvature gains participants were presented with a walking trial and then asked "Was your physical path curved left or right?". Participants had to choose one of the two answers even if they were not sure which response was correct. Although this task was designed to reduce bias in participants' responses one might argue that it can still induce bias, because both possible answers have an expectation that some type of curvature gain is present. On the other hand, participants' may scrutinize stimulus more intensely to detect gain in every trial, leading to more conservative threshold estimates. Importantly, this task does not actually meet the established definition of the 2AFC task used in the psychometric literature, because participants are not presented with 2 distinct alternatives in terms of stimulus levels. It would be more accurate to describe this as a special variation of the yes/no task, where the question is formulated to provide a different kind of binary response. We will refer to this task as a *pseudo-2AFC task* to distinguish it from a true 2AFC task.

Using a true 2AFC task to estimate detection threshold for curvature gain may not be practical. Curvature detection trials are relatively time-consuming and require physical effort because participants need to physically walk to experience curvature manipulation. Both time and physical effort can quickly add up to substantial levels over multiple trials. It is also impossible to present two different levels of curvature gains at the same time. Therefore, a true 2AFC task for curvature detection would require participants to complete two intervals of walking, then either decide in which of the walking intervals curvature gains were applied or determine in which of the two intervals their physical path had more extreme curvature. This would effectively double the required amount of walking compared to a simple yes/no task. We believe this makes a case for a yes/no task (or a pseudo-2AFC task) being more suitable for estimating curvature detection thresholds.

Another important component in threshold detection methods is the rule for selecting the level of stimuli to be tested in each trial. The established approach in redirected walking literature is to use the *method of constant stimuli*. Under this approach test levels are selected a priori to sample the entire range of stimulus values. Each test level is presented to the participant multiple times and the order of presentation is randomized. For each test level, the proportion of "correct" responses is computed. This allows estimating the entire psychometric function describing the likelihood of stimulus detection by the participant as a function of stimulus level. However, this procedure requires a high number of trials leading to significant time requirements when each trial is relatively lengthy. For example, in their study Steinicke et al. [Steinicke et al. 2010] used approximately 100 trials (10 test levels with 10 repetitions each) to estimate one psychometric function. The total number of trials per participant reached 300 and the required total time commitment was about 3 hours. It is desirable to have an experimental method that can use trials in a more efficient manner to reduce the total number of trials required.

In situations where individual trials are "expensive", *adaptive methods* represent a compelling alternative to method of constant stimuli (for an overview see [Leek 2001]). Adaptive methods rely on previously observed responses to determine which stimulus level should be tested next. For greater efficiency, the goal typically is to select the test level that is likely to provide the most information regarding the estimated threshold. In order to further reduce the required number of trials, most adaptive methods focus only on estimation of detection thresholds, not the full psychometric function. While adaptive methods are generally more efficient compared to the method of constant stimuli, they can also be more susceptible to bias.

There are relatively few examples of using adaptive methods in the redirected walking domain. Razzaque [Razzaque 2005] explored the use of adaptive 2-track staircase methods to estimate detection thresholds for rotation gain within approximately 20 trials. While this method was generally successful, high error rates for reported estimated thresholds suggest that the optimal number of trails should be higher. Nevertheless, this example demonstrates that adaptive methods can be useful for estimating detection thresholds for redirected walking using a limited amount of trials.

Figure 2: *Virtual environment used to reposition participants between trials.*

3 Research aims

Our primary goal in this study was to investigate how simultaneous application of translation and curvature gains may affect detection thresholds. Neth et al. [Neth et al. 2012] have shown that participants' real-world walking speed affects curvature detection thresholds. In particular, they found that faster walking participants are more sensitive to curvature gain. A similar effect was also observed for participants walking blindfolded in the real world [Kallie et al. 2007]. Translation gain scales virtual movement relative to the real world, effectively changing participants' virtual speed. One might expect that such change in virtual speed might also affect curvature detection thresholds.

To fully assess the interaction between translation and curvature gains one would need to explore a two-dimensional space of possible combinations of gains, making this hypothetical experiment prohibitively lengthy. As a first step towards this goal we opted for estimating curvature detection thresholds for three fixed levels of translation gain: no translation gain, down-scaled translation, and up-scaled translation.

Our first experiment closely followed an established estimation methodology with a pseudo-2AFC task coupled with a method of constant stimuli, which was previously used by both [Steinicke et al. 2010] and [Neth et al. 2012]. However, we were unable to recruit participants for similarly long experimental sessions (3 and 6 hours respectively) and had to reduce the total number of trials. Due to these time constraints we were also interested in using adaptive methods. In particular, we wanted to explore to what extent the estimated thresholds would be affected by the estimation method. To this end, we conducted a second experiment, this time using Green's maximum likelihood procedure [Green 1993]. In our opinion this variant of the adaptive procedure is particularly suitable for use in redirected walking, because it is specifically designed to rapidly estimate detection thresholds using a binary yes/no task. To achieve a better match with this procedure we also reformulated the experimental task as a simple yes/no task.

4 Experiment 1: Method of constant stimuli

The first experiment followed the established experimental methodology combining the method of constant stimuli with a pseudo-2AFC task. We defined three translation gain conditions by deliberately choosing gain levels to exceed detection thresholds reported in [Steinicke et al. 2010] for both up-scaling and down-scaling. In particular, the down-scaling translation gain of 0.75 was chosen to match the relative reduction in speed from "normal" (1m/s) to "slow" (0.75m/s) condition reported by [Neth et al. 2012]. For up-scaling we chose translation gain of 1.4, which is well above reported detection threshold (26%).

To limit the total number of trials we restricted the number of test levels of curvature gains to four. We wanted to explicitly test previously reported detection thresholds corresponding to curvature radii of 7.5m and 22m. To sample more extreme curvature we added 6m curvature level. Based on our preliminary tests we thought this curvature level was sufficiently obvious to be easily detected by most participants. To complete the design, we also added added a more shallow curvature of 56m.

Initially we considered implementing a method similar to that used by [Neth et al. 2012] to control participants' speed by asking them to follow behind a virtual sphere moving with fixed, pre-determined velocity. However, during pilot testing it became clear that one can easily detect the change in translation gain between trials by observing the movement of the virtual sphere. An alternative method of tending to a speed gauge while walking proved to impose high cognitive demands making it very hard to attend to the main perceptual task. As a result, we decided not to control participants' speed, but to check for its variability after the fact.

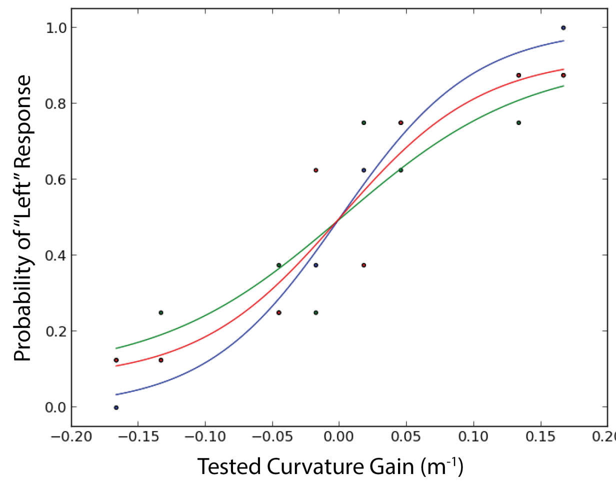

Figure 3: *Fitted psychometric functions for one of the participants. Red line corresponds to translation gains value of 1.4, blue line to gains of 0.75, and green line to 1.0.*

4.1 Method

Participants were immersed into an indoor, room-sized virtual environment facing a man's portrait on the wall opposite to their initial location. They were asked to walk straight towards the portrait on the opposite side on the room. In each trial participants were presented with of one of the 4 levels of curvature gain (corresponding to 6m, 7.5m, 22m, or 56m radius) and one of the 3 levels of translation gain (0.75, 1.0, 1.4). The randomization process was designed to ensure uniform normal distribution with 8 repetitions for each pair of gains (4 paths curving left and 4 paths curving right).

After covering 4 meters in the virtual environment (real-world walking distances varied due to translation and curvature gains) participants were asked the following question: "Did your real-world path curve left or right?". Responses were voiced aloud by participants and recorded by the experimenter via a button press. Once the response was recorded participants were presented with a second virtual environment (Figure 2) to physically reposition them in the real-world tracking area in preparation for the next trial. This second environment was intentionally designed to limit participants' ability to trace their movement relative to the real world from trial-to-trial. The starting position was always located at the opposite side of the tracking area relative to the starting location in the prior trial and chosen at random from two pre-defined options. When participants reached new starting position they were directed to rotate 180 degrees and then a new trial was initiated. Overall, each participant completed a total of 96 test trials.

Each experiment also contained 6 practice trials: two trial with no translation or curvature gains, two trials demonstrating two different levels of translation gain, and 2 trials combining translation gain with severe curvature gain (with corresponding curvature radius of 4m). To detect any potential effects of simulator sickness participants completed a pre- and post experiment Simulator Sickness Scale questionnaire [Kennedy et al. 1993]. At the end of the study they were also asked to provide basic demographics information, such as age and gender. On average the entire experiment took approximately 50 min including 10 min break.

4.2 Apparatus

During this experiment participants were wearing an Oculus Rift DK2 HMD. The DK2 HMD has a 960x1080 per eye resolution (1080p total resolution), 60Hz refresh rate, 110 degree nominal field of view (FOV), and an internal 9 degree of freedom (9DOF) inertial sensor. The participants were tracked in a roughly 3 meter by 4 meter space with a PhaseSpace Impulse X2 motion capture system. To enable 6DOF head tracking by the Impulse X2 system we attached a 5 LED, non-coplanar rig to the DK2. At runtime head orientation was tracked using DK2 internal inertial sensor, while positional information was supplied by the PhaseSpace system. We also performed periodic orientation drift correction between trials using orientation data from the Impulse X2 system. Sound cues were provided using a pair of "over the ear" Sennheiser headphones.

4.3 Participants

Eighteen participants (6 females and 13 males) were recruited for this study using online classifieds board (local Craigslist volunteers section). Participants were between 25 and 63 years old with a mean age of 39 years and a median age of 35 years (one participant did not report his age) and were required to have normal or corrected normal vision. We did not pre-screen participants for gaming or virtual reality experience.

Three of the participants were unable to complete the study due to simulator sickness and one additional participant dropped out due to physical fatigue. Furthermore, 2 participants exhibited highly inconsistent pattern of responses, suggesting that they did not understand the task, and were also excluded from the analysis. The total number of participants, who successfully completed the study was 12.

Participant recruitment and experimental procedures were approved by the local Institutional Review Board (IRB). Participants were paid USD $25 for their efforts.

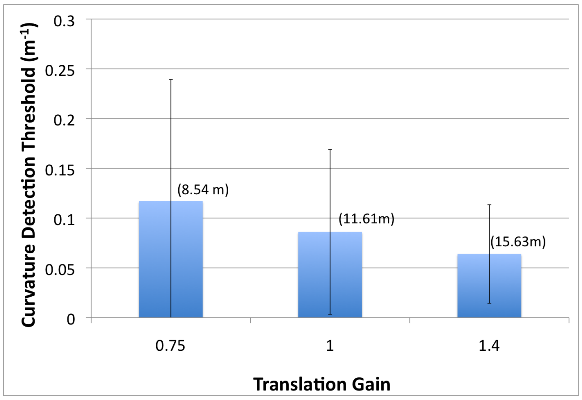

Figure 4: *Estimated curvature detection thresholds for Experiment 1. Data labels show corresponding radius in brackets. Error bars denote standard deviation.*

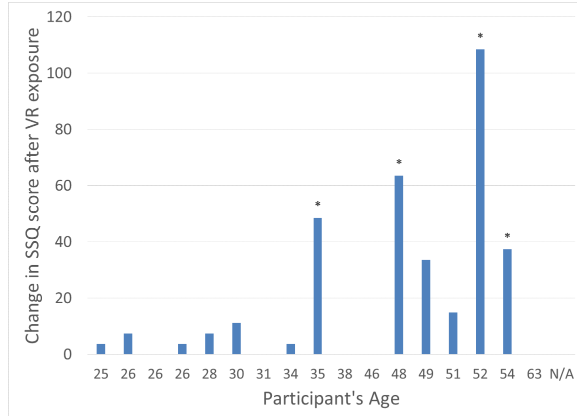

Figure 5: *Change in simulator sickness score (SSQ score) after the VR exposure. Stars indicate participants who were unable to complete the entire experiment.*

4.4 Results

For each participant we estimated 3 psychometric functions (one for each level of translation gain) based on the proportion of "left" responses using the Psignifit3 package [Fründ et al. 2011]. Figure 3 shows an example for one of the participants.

We used a standard logistic psychometric function with an a-b core in the following form:

$$F(x;\alpha;\beta) = \frac{1}{1 + exp(-\frac{x-\alpha}{\beta})}$$

Some participants systematically switched their "left" and "right" responses. This suggests that they relied on the virtual rather than the real-world frame of reference to estimate the direction of curvature. We re-mapped their responses accordingly.

We pooled together responses for left and right curvatures with the same level of curvature and translation gains. This is equivalent to the assumption that Point of Subjective Equality (PSE), the point with 50% probability of "left" responses coincides with zero curvature gain. While it is possible that participants might be slightly biased toward walking on a leftward or rightward curve due to physical asymmetries in their bodies, earlier studies suggest that such bias is minimal. Under these assumptions the detection threshold corresponds to 75% probability of "left" responses. For each participant psychometric curves yielded three estimates of the curvature gain detection threshold under each of the three translation gains conditions. Figure 4 shows mean curvature detection thresholds for each of the three levels of translation gain.

To compare estimated curvature gain detection thresholds across three translation gain conditions we fitted a one-way ANOVA model using Subject as a random effect to account for between-participant variability. We found no significant differences across conditions ($F(22, 2) = 1.59, p = 0.23$).

We also directly compared estimated detection thresholds with previously reported curvature values corresponding to curvature radii of 22m and 7.5m. Using a Bonferroni adjustment for multiple comparison the critical value α was set to 0.008. For a translation gain of 0.75, estimated threshold was significantly smaller than 22m radius ($t = -10.86, p < 0.001$), but did not differ significantly from 7.5m ($t(11) = -0.46, p = 0.65$). For translation gain of 1.0, estimated threshold was significantly smaller than 22m radius ($t(11) = -17.33, p < 0.001$), but not significantly larger than 7.5m radius ($t(11) = -1.97, p = 0.037$). Finally, for a translation gain of 1.4 estimated curvature detection threshold was significantly smaller than 22m radius ($t(11) = -30.5, p < 0.001$) and significantly larger than 7.5m radius ($t(11) = -4.85, p < 0.001$).

For each participant and each translation gain condition we computed mean physical speed. Since participants' speed was uncontrolled, it is possible that participants adjusted their physical walking speed in response to change in translation gain, which in turn would have affected estimated curvature detection thresholds. To explore the effect of translation gain on participants' physical speed we fitted a one-way ANOVA model, which also included participant as a random effect to account for between-subject variability. The model indicates that there is no sufficient evidence to reject the null hypothesis that participants' speeds were similar across all three translation gain conditions ($F(22, 2) = 1.21, p = 0.316$). On average, participants maintained a physical walking speed of 1.0m/s.

Based on [Neth et al. 2012] one might expect that participant's referred speed may be correlated with the estimated detection threshold. We examined the correlation between participant's average speed and estimated curvature detection threshold for each of the translation gain conditions. We found virtually no correlation for translation gain of 0.75 ($r = -0.077$) and translation gain of 1.4 ($r = -0.03$). However, there was a moderate positive correlation for translation gain of 1.0 ($r = 0.506$), which matched the findings in [Neth et al. 2012].

To analyze the effects of the VR exposure during the experiment on simulator sickness we computed the change in SSQ scores after the experiment for each participant (see Figure 5). On average participants aged 35 and older reported higher SSQ scores compared than younger participants. Statistically, SSQ scores tend to be highly skewed: participants affected by simulator sickness have very high scores, wheres unaffected participants have low scores. To compare SSQ scores across the two age groups we used a non-parametric Wilcoxon-Mann-Whitney test, which does not require normality assumptions for the underlying distribution. The results indicate that there is no sufficient evidence to support the hypothesis that the two age groups had different SSQ scores ($p = 0.29$). However, Pearson's χ^2 test indicated that younger participants had a significantly higher probability of completing the experiment compared to older ones ($\chi^2 = 4.11, p = 0.04$).

4.5 Discussion

Our data does not support the hypothesis that curvature detection thresholds are systematically affected by simultaneous application of translation gain. It is possible that we were simply unable to detect an underlying significant effect due to relatively low number of trials in our experiment. However, we also observed a high level of between-subject variability. In particular, the estimated detection thresholds for individual participants (see for example Figure 3) did not always follow the same trend as the population means, where curvature thresholds decreased with increase in translation gain (as shown on Figure 4). Furthermore, when checking the effects of participants' speed we did not find systematic correlation between participants' responses and their speed with the exception of no translation gain condition. Taken together these findings suggests that any underlying effect (if exists) is likely to be relatively small.

Interestingly, our estimates for curvature detection thresholds differ significantly from those reported by Steinicke et al.[Steinicke et al. 2010]. In particular, in the absence of translation gain (translation gain of 1.0) estimated detection threshold for curvature gain corresponded to radius as low as 11.61 meters, which was significantly smaller compared to the previously reported 22 meter radius. The reasons for this difference are unclear and should be more thoroughly investigated in future studies. One possibility is the difference between test virtual environments. Another potential factor is the difference in HMD characteristics. Both [Steinicke et al. 2010] and [Neth et al. 2012] used the same type of HMD - eMagin Z800 3D Visor with a limited field-of-view. In our experiment the field-of-view was significantly larger. Because peripheral vision plays an important role in motion detection, these technology differences could have affected the estimates. Finally, the difference in estimates might be explained by differences in the population of participants. Earlier studies primarily relied on a college students population to recruit participants. Our participants were significantly older and had more varied socio–economic backgrounds. The differences in both life experiences and physiological age-related factors could have affected participants' responses.

We also observed relatively high level of simulator sickness in this experiment. We believe that this was to a large extent due to the step-like change in translation gains from trial to trial. This issue was particularly acute for older participants, who were more likely to drop out due to simulator sickness. We conclude that a blocked design where trials are grouped by translation gain might be suitable for any future studies.

5 Experiment 2: Maximum likelihood procedure

The second experiment was designed to study the sensitivity of estimated thresholds to the differences in estimation methods. In this experiment we implemented Green's adaptive maximum likelihood procedure [Green 1993]. This procedure computes a fit of the psychometric function under a a range of hypotheses (possible detection threshold levels) and determines the most likely hypothesis based on observed responses. At each step the most likely psychometric function is used to to compute the next target stimulus level (the so called "sweetspot"), which provides the maximum amount of information about the location of the true threshold level. The participant is then presented with the next test level of stimulus and the new observation is added to the maximum likelihood computation. The procedure is repeated until the maximum number of trials is reached. The most likely psychometric function at the end of the procedure yields the threshold estimate.

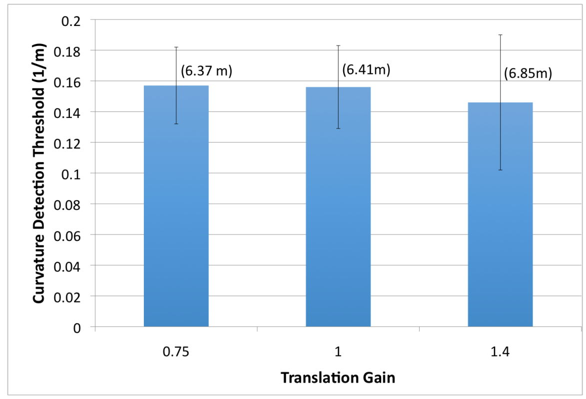

Figure 6: *Estimated curvature detection thresholds for Experiment 2. Data labels show corresponding radius in brackets. Error bars represent standard deviation.*

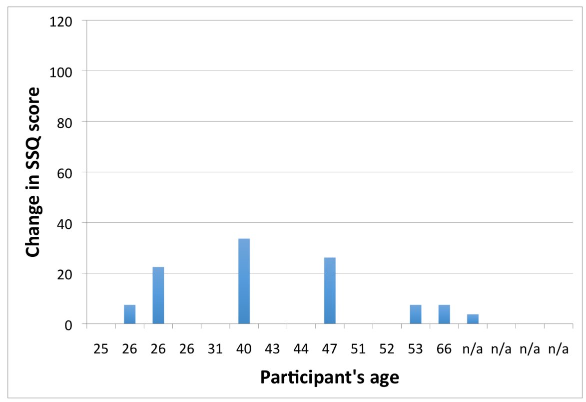

Figure 7: *Change in simulator sickness questionnaires (SSQ) scores after the VR exposure in experiment 2.*

We observed that some participants in experiment 1 were confused the question asked in Steinicke's pseudo-2AFC task. Considering that Green's method was designed for a simple yes/no task, we also reformulated our task to match.

5.1 Method

Participants were immersed in the same virtual environments as in Experiment 1 for both the main walking trial the following repositioning phase. At the end of each trial participants were asked and the following question: "Was your path curved?" and provided a yes/no answer.

The experiment was divided into three blocks, each block corresponding to a fixed translation gain value of 0.75, 1.0, or 1.4. Curvature gain levels were selected using Green's maximum likelihood adaptive procedure. The procedure assumed a zero false-positive detection rate. The slope of the psychometric function was estimated using the data reported in [Steinicke et al. 2010] and set to 34.0. The range of tested hypotheses was defined in terms of curvature values and covered values between 0 (straight line) and $0.16m^{-1}$ (corresponding to a 6.25m radius) with a step of $0.001m^{-1}$. The adaptive procedure terminated after 35 trials and the current most likely hypothesis was used as the final estimate.

The hardware setup was the same as in Experiment 1.

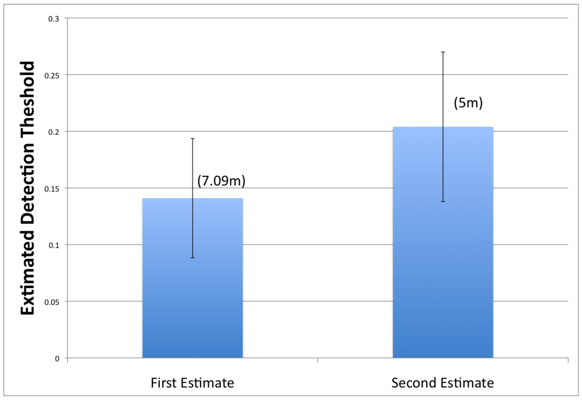

Figure 8: *Comparison of estimated curvature detection threshold for standard (first) and extended (second) range of hypothesis.*

5.2 Participants

Eighteen additional Craigslist volunteers were recruited to participate in this experiment. One participant appeared to be intoxicated during this study and was excluded from further analysis. In total, 17 participants (9 males and 8 females) completed this experiment. Participant's age ranged from 25 to 64 years (5 participants chose not to report their age). Mean reported age was 40 years and median reported age was 43 years.

5.3 Results

Figure 6 shows estimated curvature detection thresholds for each of the 3 translation gain conditions. To compare estimated curvature gain detection thresholds across three translation gain conditions we fitted a one-way ANOVA model using Subject as a random effect to account for between-participant variability. We found no significant differences across conditions ($F(31, 2) = 1.79, p = 0.18$).

The SSQ scores were noticeably lower compared to Experiment 1, suggesting that blocked design successfully mitigated the simulator sickness issues observed in Experiment 1.

We have observed that for several participants estimated thresholds corresponded to the boundaries of the tested range curvatures. This could mean that a larger range would result in different estimates. We explored this possibility by adding a fourth block to the end of the experiment for the last 6 recruited participants. In this block of trials the translation gain was set to 1.0 and the most extreme curvature value for the range of hypothesis was set at $0.25m^{-1}$ (corresponds to 4 m radius). The estimated detection threshold for the first and the second estimates are shown in Figure 8.

The responses exhibit high correlation between first and second estimate obtained for the same participant ($r = 0.978$). However, the second estimates were significantly higher ($t(6) = -9.14, p < 0.001$) (corresponding radius was smaller). These results suggest that participants were willing to accept more extreme curvature when the experimental procedure was modified to enable testing wide range of curvature values.

5.4 Discussion

As in Experiment 1 we found no evidence to support the hypothesis that translation gain has a significant effect on curvature detection thresholds. The estimated minimal radius values for the threshold curvatures were once again much lower compared to previously reported results and were also lower than in Experiment 1.For the no translation gain condition the estimated minimum curvature radius was 6.41m. Furthermore, when some participants were re-tested with an expanded range of possible threshold values, the estimated radius value was even lower at 5m.

These results suggest that when curvature gains are increasing from trial to trial in very small increments as can happen in our adaptive procedure after the initial burn-in period, participants are less likely to detect even significant curvature gain. While this effect introduces an undesirable bias for estimation of the true curvature detection threshold, from the application viewpoint the fact that a very gradual introduction of curvature gain might actually desensitize participants to this perceptual manipulation can be valuable.

To counter the potential estimation bias in maximum likelihood procedures for estimation of detection threshold, future studies should examine the possibility of using a 2-track adaptive procedure with two different starting values. This setup is likely to introduce higher contrast between successive trials and improve the accuracy of the estimation [Leek 2001].

6 Conclusion

Our primary goal in this study was to explore the effects of combined translation and curvature gains on curvature detection thresholds. Existing literature suggests that curvature detection thresholds can be significantly affected by the estimation method. In this paper we presented an analysis of experimental methodologies that can be suitable for estimating curvature detection thresholds. We discussed how the characteristics of a curvature detection task that can affect the applicability of a particular estimation method and considered the advantages and disadvantages of both the established experimental method of constant stimuli and the alternative methods based on adaptive procedures. We then described two experimental studies using these alternative methods to estimate curvature detection thresholds for situations where translation and curvature gains are applied at the same time.

Notwithstanding the significant differences in experimental methods, in both of our experiments we found no evidence to support the hypothesis that curvature detection thresholds are systematically changing with translation gain. This is particularly noteworthy, because we tested translation gain values that span the previously estimated range of undetectable translation gains. The practical implication of this is that currently used levels of translation and curvature gains can be safely combined to improve redirection effectiveness. This is particularly important for small, room-sized spaces that can be tracked by emerging consumer-level tracking devices, where generalized algorithms have limited effectiveness and can benefit from performance boost the most. Furthermore, the combination of translation and curvature gains can be used to develop improved specialized redirection algorithms that can be very effective in using the small physical space to support pre-planned virtual trajectories (for an example of a specialized algorithm, see “Near-Field VR” use-case study in [Azmandian et al. 2016]).

Interestingly, our estimates for curvature detection thresholds were also significantly smaller compared to levels previously established by Steinicke et al. [Steinicke et al. 2010]. Based on our estimates, participants are less sensitive to curvature gain than previously reported. The exact cause of the observed differences in estimated thresholds is unclear and can potentially be explained by several factors including hardware setup, methodological differences and population differences. Our data shows that detection threshold estimates are quite sensitive to changes in estimation method and underlines the importance of further research in this area.

Given large between-subject variability observed in this and previous studies, it would be particularly beneficial to be able to calibrate redirection algorithms for each user. We believe that this can be best achieved through development of a reliable estimation method based on an adaptive procedure. However, our data suggests that a simple implementation of an adaptive procedure might be susceptible to bias. One possibility to address this issue is to introduce of the 2-track Green's procedure. Another option could be the modification of Green's maximum likelihood procedure developed by Shen et al. [Shen and Richards 2012], which can be used to simultaneously estimate detection threshold, slope and error rate parameters of the psychometric function. We plan to explore these modifications in future studies.

Acknowledgements

This work is sponsored by the U.S. Army Research Laboratory (ARL) under contract number W911NF-14-D-0005. Statements and opinions expressed and content included do not necessarily reflect the position or the policy of the Government, and no official endorsement should be inferred.

References

AZMANDIAN, M., GRECHKIN, T., BOLAS, M., AND SUMA, E. 2015. Physical Space Requirements for Redirected Walking: How Size and Shape Affect Performance. In *ICAT-EGVE 2015 - International Conference on Artificial Reality and Telexistence and Eurographics Symposium on Virtual Environments*, M. Imura, P. Figueroa, and B. Mohler, Eds., The Eurographics Association.

AZMANDIAN, M., GRECHKIN, T., BOLAS, M., AND SUMA, E. 2016. The redirected walking toolkit: A unified development and deployment platform for exploring large virtual environments. In *2nd Workshop on Everyday Virtual Reality (WEVR)*, IEEE.

BRUDER, G., LUBAS, P., AND STEINICKE, F. 2015. Cognitive Resource Demands of Redirected Walking. *IEEE Transactions on Visualization and Computer Graphics 21*, 4, 539–544.

FRÜND, I., HAENEL, N. V., AND WICHMANN, F. A. 2011. Inference for psychometric functions in the presence of nonstationary behavior. *Journal of Vision 11*, 6, Article 16.

GREEN, D. M. 1993. A maximumlikelihood method for estimating thresholds in a yes-no task. *The Journal of the Acoustical Society of America 93*, 4, 2096–2105.

HODGSON, E., AND BACHMANN, E. 2013. Comparing Four Approaches to Generalized Redirected Walking: Simulation and Live User Data. *IEEE Transactions on Visualization and Computer Graphics 19*, 4, 634–643.

INTERRANTE, V., RIES, B., AND ANDERSON, L. 2007. Seven league boots: A new metaphor for augmented locomotion through moderately large scale immersive virtual environments. In *Symposium on 3D User Interfaces*, IEEE, 167–170.

JAEKL, P. M., JENKIN, M. R., AND HARRIS, L. 2005. Perceiving a stable world during active rotational and translational head movements. *Experimental Brain Research 163*, 3, 388–399.

JERALD, J., PECK, T., STEINICKE, F., AND WHITTON, M. 2008. Sensitivity to scene motion for phases of head yaws. In *Proceedings of the 5th Symposium on Applied Perception in Graphics and Visualization*, APGV '08, ACM, 155–162.

KALLIE, C. S., SCHRATER, P. R., AND LEGGE, G. E. 2007. Variability in stepping direction explains the veering behavior of blind walkers. *Journal of Experimental Psychology: Human Perception and Performance 33*, 1, 183–20.

KENNEDY, R. S., LANE, N. E., BERBAUM, K. S., AND LILIENTHAL, M. G. 1993. Simulator Sickness Questionnaire: An Enhanced Method for Quantifying Simulator Sickness. *The International Journal of Aviation Psychology 3*, 3, 203–220.

KLEIN, S. 2001. Measuring, estimating, and understanding the psychometric function: A commentary. *Perception & Psychophysics 63*, 8, 1421–1455.

LEEK, M. R. 2001. Adaptive procedures in psychophysical research. *Perception & Psychophysics 63*, 8, 1279–1292.

NETH, C. T., SOUMAN, J. L., ENGEL, D., KLOOS, U., BULTHOFF, H. H., AND MOHLER, B. J. 2012. Velocity-Dependent Dynamic Curvature Gain for Redirected Walking. *IEEE Transactions on Visualization and Computer Graphics 18*, 7, 1041–1052.

RAZZAQUE, S. 2005. *Redirected Walking*. PhD thesis, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA.

RUDDLE, R. A., AND LESSELS, S. 2009. The benefits of using a walking interface to navigate virtual environments. *ACM Transactions on Computer-Human Interaction (TOCHI) 16*, 1, 5:1–5:18.

RUDDLE, R. A., VOLKOVA, E., AND BÜLTHOFF, H. H. 2011. Walking improves your cognitive map in environments that are large-scale and large in extent. *ACM Transactions on Computer-Human Interaction (TOCHI) 18*, 2, 10:1–10:20.

SHEN, Y., AND RICHARDS, V. M. 2012. A maximum-likelihood procedure for estimating psychometric functions: Thresholds, slopes, and lapses of attention. *The Journal of the Acoustical Society of America 132*, 2, 957–967.

SOUMAN, J. L., FRISSEN, I., SREENIVASA, M. N., AND ERNST, M. O. 2009. Walking Straight into Circles. *Current Biology 19*, 18, 1538–1542.

STEINICKE, F., BRUDER, G., HINRICHS, K., JERALD, J., FRENZ, H., AND LAPPE., M. 2009. Real walking through virtual environments by redirection techniques. *Journal of Virtual Reality and Broadcasting 6*, Article 2.

STEINICKE, F., BRUDER, G., JERALD, J., FRENZ, H., AND LAPPE, M. 2010. Estimation of Detection Thresholds for Redirected Walking Techniques. *IEEE Transactions on Visualization and Computer Graphics 16*, 1, 17–27.

SUMA, E., FINKELSTEIN, S., REID, M., BABU, S., ULINSKI, A., AND HODGES, L. F. 2010. Evaluation of the cognitive effects of travel technique in complex real and virtual environments. *IEEE Transactions on Visualization and Computer Graphics 16*, 4, 690–702.

SUMA, E. A., BRUDER, G., STEINICKE, F., KRUM, D. M., AND BOLAS, M. 2012. A taxonomy for deploying redirection techniques in immersive virtual environments. In *2012 IEEE Virtual Reality Workshops (VRW)*, IEEE, 43–46.

USOH, M., ARTHUR, K., WHITTON, M. C., BASTOS, R., STEED, A., SLATER, M., AND BROOKS JR., F. P. 1999. Walking > Walking-in-place > Flying, in Virtual Environments. In *Proceedings of the 26th Annual Conference on Computer Graphics and Interactive Techniques*, SIGGRAPH '99, ACM Press/Addison-Wesley Publishing Co., 359–364.