Threefolded Motion Perception During Immersive Walkthroughs

Gerd Bruder*, Frank Steinicke†
Department of Computer Science
University of Hamburg

Abstract

Locomotion is one of the most fundamental processes in the real world, and its consideration in immersive virtual environments (IVEs) is of major importance for many application domains requiring immersive walkthroughs. From a simple physics perspective, such self-motion can be defined by the three components speed, distance, and time. Determining motions in the frame of reference of a human observer imposes a significant challenge to the perceptual processes in the human brain, and the resulting speed, distance, and time percepts are not always veridical. In previous work in the area of IVEs, these components were evaluated in separate experiments, i.e., using largely different hardware, software and protocols.

In this paper we analyze the perception of the three components of locomotion during immersive walkthroughs using the same setup and similar protocols. We conducted experiments in an Oculus Rift head-mounted display (HMD) environment which showed that subjects largely underestimated virtual distances, slightly underestimated virtual speed, and we observed that subjects slightly overestimated elapsed time.

CR Categories: H.5.1 [Information Interfaces and Presentation]: Multimedia Information Systems—Artificial, augmented, and virtual realities; I.3.7 [Computer Graphics]: Three-Dimensional Graphics and Realism—Virtual reality

Keywords: Locomotion, real walking, perception, immersive virtual environments.

1 Introduction

The motion of an observer or scene object in the real world or in a virtual environment (VE) is of great interest for many research and application fields. This includes computer-generated imagery, e.g., in movies or games, in which a sequence of individual images evoke the illusion of a moving picture [Thompson et al. 2011]. In the real world, humans move, for example, by walking or running, physical objects move and sometimes actuate each other, and, finally, the earth spins around itself as well as around the sun. From a simple physics perspective, such motions can be defined by three main components: (i) (linear or angular) speed and (ii) distances, as well as (iii) time. The interrelation between these components is given by the speed of motion, which is defined as the change in position or orientation of an object with respect to time:

\[ s = \frac{d}{t} \quad (1) \]

with speed \( s \), distance \( d \), and time \( t \). Motion can be observed by attaching a frame of reference to an object and measuring its change relative to another reference frame. As there is no absolute frame of reference, absolute motion cannot be determined; everything in the universe can be considered to be moving [Bottema and Roth 2012]. Determining motions in the frame of reference of a human observer imposes a significant challenge to the perceptual processes in the human brain, and the resulting percepts of motions are not always veridical [Berthoz 2000]. Misperception of speed, distances, and time has been observed for different forms of self-motion in the real world [Efron 1970; Gibb et al. 2010; Mao et al. 2010].

In the context of self-motion, walking is often regarded as the most basic and intuitive form of locomotion in an environment. Self-motion estimates from walking are often found to be more accurate than for other forms of motion [Ruddle and Lessels 2009]. This may be explained by adaptation and training since early childhood and evolutionary tuning of the human brain to the physical affordances for locomotion of the body [Choi and Bastian 2007]. While walking in the real world, sensory information such as vestibular, proprioceptive, and efferent copy signals as well as visual and auditory information create consistent multi-sensory cues that indicate one’s own motion [Wertheim 1994].

However, a large body of studies have shown that spatial perception in IVEs differs from the real world. Empirical data shows that static and walked distances are often systematically over- or underestimated in IVEs (cf. [Loomis and Knapp 2003] and [Renner et al. 2013] for thorough reviews) even if the displayed world is modeled as an exact replica of the real laboratory in which the experiments are conducted in [Interrante et al. 2006]. Less empirical data exists on speed misperception in IVEs, which shows a tendency that visual speed during walking is underestimated [Banton et al. 2005; Bruder et al. 2012; Durgin et al. 2005; Steinicke et al. 2010].

We are not aware of any empirical data which has been collected for time misperception in IVEs. However, there is evidence that time perception can deviate from veridical judgments due to visual or auditory stimulation related to motion misperception [Gromov 2008; Roussel et al. 2009; Sarrazin et al. 2004]. Different causes of motion misperception in IVEs have been identified, including hardware characteristics [Jones et al. 2011; Willemsen et al. 2009], rendering issues [Thompson et al. 2011], and miscalibration [Kuhl et al. 2009; Willemsen et al. 2008].

Considering that self-motion perception is affected by different hard- and software factors, it is unfortunate that no comprehensive analysis to this day exists in which the different components were tested using the same setup. Hence, it remains unclear whether or not there is a correlation or causal relation between speed, distance, and time misperception in current-state IVEs.

Our main contributions in this paper are:

- We compare self-motion estimation in the three components using the same setup and similar experimental designs.
- We introduce novel action-based two-alternative forced-choice methods for distance, speed, and time judgments.
2 Background

As stated above, research on distance, speed, and time perception in IVEs is divided in separate experiments using largely different hardware, software and experimental protocols. In this section we give an overview of the results for distance, speed, and time perception in IVEs.

2.1 Distance Perception

During self-motion in IVEs, different cues provide information about the travel distance with respect to the motion speed or time [Mohler 2007]. Humans can use these cues to keep track of the distance they traveled, the remaining distance to a goal, or discriminate travel distance intervals [Bremmer and Lappe 1999]. Although humans are considerably good at making distance judgments in the real world, experiments in IVEs show that characteristic estimation errors occur such that distances are often severely overestimated for very short distances and underestimated for longer distances [Loomis et al. 1993]. Different misperception effects were found over a large range of IVEs and experimental methods to measure distance estimation. While verbal estimates of the distance to a target can be used to assess distance perception, methods based on visually directed actions have been found to generally provide more accurate results [Loomis and Knapp 2003]. The most widely used action-based method is blind walking, in which subjects are asked to walk without vision to a previously seen target. Several experiments have shown that over medium range distances subjects can accurately indicate distances using the blind walking method [Rieser et al. 1990]. Other action-based methods include triangulated walking and timed imagined walking [Fukusima et al. 1997; Klein et al. 2009; Plumeri et al. 2004]. Moreover, perceptual matching methods can be used, in which subjects match the distance or size of a target to the distance or size of a reference object, respectively [Loomis and Philbeck 2008].

Although there is a large interest in solving the distance misperception problem, the reasons for this perceptual shift are still largely unknown, as are approaches to reduce such misperceptions. Kuhl et al. [Kuhl et al. 2009] observed that miscalibration of HMD optics is a main reason for distance misperception, although subjects underestimated distances even for correctly calibrated HMDs [Keller et al. 2012]. Willemsen et al. [Willemsen et al. 2009] compared HMD properties with natural viewing in the real world and observed that mechanical restrictions of HMDs can cause slight differences in distance judgments. Jones et al. [Jones et al. 2012; Jones et al. 2011] found that increasing the field of view of HMDs to approximate the visual angle of the human eyes helps alleviate distance misperception. Interrante et al. [Interrante et al. 2006] showed that the VE has an impact on distance judgments with underestimation being reduced if subjects are immersed in an accurate virtual replica of the real-world surroundings than in a different VE. Studies by Phillips et al. [Phillips et al. 2009] further show that the visual rendering style affects distance judgments. They observed that distance underestimation was increased for a non-photorealistic rendering style than in a photorealistic scene.

2.2 Speed Perception

Different sensory motion cues support the perception of the speed of walking in an IVE (cf. [Durgin et al. 2005]). Visual motion information is often estimated as most reliable by the perceptual system, but can cause incorrect motion percepts. For example, in the illusion of linear vection [Berthoz 2000] observers feel their body moving although they are physically stationary because they are presented with large-field visual motion that resembles the motion pattern normally experienced during self-motion. Humans use such optic flow patterns to determine the speed of movement, although the speed of retinal motion signals is not uniquely related to movement speed. For any translational motion the visual velocity of any point in the scene depends on the distance of this point from the eye, i.e., points farther away move slower over the retina than points closer to the eye [Bremmer and Lappe 1999; Warren, Jr. 1998]. Banton et al. [Banton et al. 2005] observed for subjects walking on a treadmill with an HMD that optic flow fields at the speed of the treadmill were estimated as approximately 53% slower than their walking speed. Durgin et al. [Durgin et al. 2005] reported on a series of experiments with subjects wearing HMDs while walking on a treadmill or over solid ground. Their results show that subjects often estimated subtracted speeds of displayed optic flow fields as matching their walking speed. Steinicke et al. [Steinicke et al. 2010] evaluated speed estimation of subjects in an HMD environment with a real walking user interface in which they manipulated subjects’ self-motion speed in the VE compared to their walking speed in the real world. Their results show that on average subjects underestimated their walking speed by approximately 7%. Bruder et al. [Bruder et al. 2012; Bruder et al. 2013] showed that visual illusions related to optic flow perception can change self-motion speed estimates in IVEs.

2.3 Time Perception

As discussed for distance and speed perception in IVEs, over- or underestimations have been well documented by researchers in different hard-, software and experimental protocols. In contrast, perception of time has not been extensively researched in IVEs so far. Experimental studies of time perception in the field of psychology have well established that estimates of stimulus duration do not always match its veridical time interval, and are affected by a variety of factors [Efron 1970]. Since time cannot be directly measured at a given moment, the brain is often assumed to estimate time based on internal biological or psychological events, or external signals [Gronvid 2008]. The effect of exogenous cues (i.e., zeitgebers) from the local environment on endogenous biological clocks (e.g., circadian rhythms) is studied in the field of chronobiology [Kramer and Merrow 2013]. It is possible that differences in exogenous time cues between the real world and IVEs have an effect on internal human time perception. In particular, system latency is known to change the perception of sensory synchronicity [Shi et al. 2010] and can degrade the perceptual stability of the environment [Allison et al. 2001].

Space and time are interdependent phenomena not only in physics, but also in human perception [Gronvid 2008]. Nelson keyed the term tau effect for the phenomenon that the variation of the time between spatial events can affect judgments of their spatial layout (cf. [Helson and King 1931; Jones and Huang 1982; Sarrazin et al. 2004]). For instance, Nelson and King [Helson and King 1931] observed for a tactile estimation experiment that stimulating three equidistant surface points $p_1$, $p_2$, and $p_3$ with $||p_2 - p_1|| = ||p_3 - p_2||$ at points in time $t_1$, $t_2$, and $t_3$ for different durations $|t_2 - t_1| > |t_3 - t_2|$ that subjects judge the distance between $p_1$ and $p_2$ as longer than between $p_2$ and $p_3$. 

The paper is structured as follows. In Section 2 we provide background information on self-motion perception in virtual reality (VR) setups. We describe the experiments in Section 3 and discuss the results in Section 4. Section 5 concludes the paper.
Conversely, Cohen et al. [Cohen et al. 1953] keyed the term *kappa effect* for the phenomenon that the variation of the spatial layout of events can affect judgments of their temporal layout (cf. [Grondin 2008; Roussel et al. 2009]). They observed for a visual bisection task that three successive flashes at spatial points \( p_1, p_2, \) and \( p_3 \) for different distances \( ||p_2 - p_1|| > ||p_3 - p_2|| \) with points in time \( t_1, t_2, \) and \( t_3 \) with \( ||t_2 - t_1|| = ||t_3 - t_2|| \) that subjects judge the duration between \( t_1 \) and \( t_2 \) as shorter than the duration between \( t_2 \) and \( t_3 \).

3 Psychophysical Experiments

In this section we evaluate misperception while walking in an IVE. We describe three experiments that we conducted to evaluate distance, speed, and time perception based on similar measurement protocols to obtain judgments using the same IVE.

3.1 Participants

18 subjects (6 female and 12 male, ages 19–38, \( M=25.0, \) SD=5.6) participated in the distance and speed experiment, while 10 of them participated in the time experiment, which was performed one day later. The subjects were students or members of the local departments of computer science, psychology, or human-computer media. Students obtained class credit for their participation. We confirmed stereoscopic vision of all subjects via anaglyphic random-dot stereograms before the experiment. All of our subjects had normal or corrected to normal vision. 2 subjects wore glasses and 7 subjects wore contact lenses during the experiment. We confirmed 20/20 visual acuity of all subjects with a vision test based on a Snellen chart before the experiment. None of our subjects reported a disorder of equilibrium. Two of our subjects reported a slight red-green weakness. No other vision disorders have been reported by our subjects.

10 subjects had prior experience with HMDs, and 8 of them had participated in an experiment involving HMDs before. All but two subjects reported experience with 3D video games, and 13 subjects reported experience with 3D stereoscopic cinema. The eye height of our subjects ranged between 1.59–1.84m (\( M=1.70m, \) SD=0.06m) above the ground. We used this value to adjust the height of target markers in the VE as shown in Figure 1. We measured the interpupillary distances (IPDs) of our subjects before the experiment [Willemsen et al. 2008]. The IPDs of our subjects ranged between 5.9–6.8cm (\( M=6.38cm, \) SD=0.24cm). We used the IPD of each subject to provide a correct perspective on the HMD. All subjects were naive to the experimental conditions. The order of the distance and speed experiments was randomized and counterbalanced between subjects. The total time per subject, including pre-questionnaires, instructions, experiment, breaks, post-questionnaires, and debriefing, was 2 hours. Subjects wore the HMD for approximately 1 hour. They were allowed to take breaks at any time; short breaks after every 30 experiment trials were mandatory.

3.2 General Materials

We performed the experiment in an 8m \( \times \) 14m darkened laboratory room. As illustrated in Figure 1, subjects wore an Oculus Rift DK1 HMD for the stimulus presentation, which provides a resolution of 640\( \times \)800 pixels per eye with a refresh rate of 60Hz and an approximately 110° diagonal field of view. We attached an active infrared marker to the HMD and tracked its position within the laboratory with a WorldViz Precision Position Tracking PPT-X4 active optical tracking system at an update rate of 60Hz. The head orientation was tracked with the inertial orientation tracker of the Oculus Rift HMD. We compensated for inertial orientation drift by incorporating the PPT optical heading plugin to improve the tracker output. The visual stimulus consisted of a virtual replica of our laboratory that we modeled with centimeter accuracy and textured with photos matching the look of the real laboratory (see Figure 1). For rendering, system control and logging we used an Intel computer with 3.4GHz Core i7 processor, 8GB of main memory and Nvidia Quadro 4000 graphics card. The stimuli were rendered with the Unity 3D Pro engine. In order to focus subjects on the task no communication between experimenter and subject was performed during the experiment. Task instructions were presented via slides on the HMD. Subjects judged their perceived self-motions via button presses on a Nintendo Wii remote controller.

3.3 General Methods

In the experiments we make use of two-alternative forced choice tasks (2-AFCT) to determine effects of gains on percepts in IVEs [Ferwerda 2008]. In this method, subjects are asked to move an eccentric visual target with the body. During the movement, motion gains \( g \in \mathbb{R}^+ \) between the movement of the subject and the sensory motion feedback are varied between trials in a within-subjects design. If \( g = 1 \), the sensory feedback matches the subject’s movement. For other gains, the sensory feedback is increased \( (g > 1) \) or decreased \( (g < 1) \) compared to the movement.

After the movement, the subject has to judge whether characteristics of the perceived motion were *slower* or *faster* \( / \) *longer* or *shorter* \( / \) *smaller* or *larger* than those of the subject’s movement in a 2-AFCT [Ferwerda 2008]. In these experiments we use an adaptive staircase design, which starts with a discrepancy which is easy to detect [Cornsweet 1962; Leek 2001]. For each following trial, the new gain is computed by the previous gain plus or minus a fixed step width (1-up-1-down) depending on the answer to the aforementioned 2-AFCT question. To eliminate response bias, we interleaved two staircases starting from a minimum and maximum gain [Macmillan and Creelman 2004]. By iterative refinement the interleaved staircase design converges on the point of subjective equality (PSE), i.e., the gain at which subjects judge the virtual motion as identical to the physical movement, which allows us to identify possible systematic over- or underestimation (i.e., \( g > 1 \) or \( g < 1 \)). All subjects had to complete 30 trials in total. In the interleaved staircase design we started with a minimum and maximum gain of \( g = 0.4 \) and \( g = 1.6 \) and a step width of 0.2. All subjects were able to identify the sensory discrepancies in the 2-AFCT experiment at these start gains. After two turns in response behavior (cf. [Leek 2001; Macmillan and Creelman 2004]) we halved the step width, and computed the PSE as the mean of the last 10 trials.
3.4 Experiment E1: Distance Estimation

In this section we describe the experiment that we conducted to evaluate the perception of walking distances in the IVE using a novel 2-AFCT approach for distance judgments.

3.4.1 Methods

The trials started with an initial stimulus phase in which subjects were presented with a virtual view of the replica model of our laboratory on the HMD. As shown in Figure 1 we placed a marker at eye height at a target distance of 3m, 4m, or 5m from the subject’s start position. After subjects felt positive of having memorized the target distance, they pressed a button on the Wii remote controller, and the view on the HMD went black.

With no visual feedback of their self-motion in the VE, subjects then walked in the direction of the previously seen target marker. Subjects walked a physical distance that was scaled relative to the virtual target distance. Therefore, we apply distance gains \( g_d \in \mathbb{R}_0^+ \) to determine a subject’s physical walking distance in each trial, describing the relation between the physical walking distance and virtual target distance. For \( g_y = 1 \) physical Euclidean walking distances along the floor plane \( D_y \in \mathbb{R}_0^+ \) match one-to-one the virtual target distance \( D_v \in \{3m, 4m, 5m\} \) with \( D_p = D_v \cdot g_d \). In contrast, gains \( g_d < 1 \) result in subjects walking shorter and \( g_y > 1 \) in longer distances in the real world. After walking the scaled physical distance, subjects were asked to answer a 2-AFCT question: “Did you move farther or shorter than the virtual target distance?” The subjects had to choose one of the two alternatives by pressing the up or down button on the Wii remote controller. If a subject cannot reliably discriminate between the virtual and real distance, the subject must guess, and will be correct on average in 50% of the trials. We applied gains in the range between \( g_d = 0.4 \) and \( g_d = 1.6 \), i.e., the walked distance differed by up to \( \pm 60\% \) from the target distance.

After answering the question, subjects were guided back to the start position of the next trial. Subjects did not receive feedback about their walked distance. The gains converged on the PSE following the interleaved staircase approach described in Section 3.3. We measured a PSE for each subject for each of the three target distances, i.e., the gains at which the subjects judge the physical and virtual distances as identical. The order in which the target distances were tested for each subject was given by Latin squares.

3.4.2 Results

Figure 2(a) shows the results for the three tested within-subjects target distances. The x-axis shows the target distances, and the y-axis shows the mean judged PSEs of the subjects. The vertical bars and colored regions illustrate the standard deviations.

Subjects judged a physical distance as matching the virtual target distance that was approximately 2.39m (\( g_d=0.795, 20.4\% \) decreased) for a target distance of 3m, 3.36m (\( g_d=0.839, 16.1\% \) decreased) for a target distance of 4m, and 3.98m (\( g_d=0.795, 20.5\% \) decreased) for a target distance of 5m. Over all target distances, subjects judged an approximately 19.0% decreased physical distance as identical to the virtual distance. Considering the PSEs of individual subjects, 15 of the 18 subjects consistently judged a decreased physical distance to match the virtual distance, whereas 1 subject consistently judged an increased physical distance as correct, and 2 subjects made mixed judgments.

Table 1 lists the judged PSEs and means over all subjects for the three virtual target distances. We analyzed the results with a repeated measure ANOVA and paired-samples t-test. The results were normally distributed according to a Shapiro-Wilk test at the 5% level. Mauchly’s test indicated that the assumption of sphericity had been violated (\( \chi^2(2)=9.44, p<.01 \)), therefore degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity (\( \eta_p^2=.36 \)). We found a significant difference between the judged PSEs over all subjects and veridical results with \( g_d = 1 \) \( t(17)=-6.25, p<.001 \). We found no significant main effect of target distance on the judged PSEs \( F(1,384, 23.52)=1.636, p<.22, \eta_p^2=.088 \).

3.5 Experiment E2: Speed Estimation

In this section we describe the experiment that we conducted to evaluate the perception of self-motion speed in the IVE.

3.5.1 Methods

The trials started with an initial startup phase in which subjects had to increase their walking speed to either 1m/s, 1.25m/s, or 1.5m/s. During this phase, subjects adjusted their self-motion to a target speed using a speedometer that was indicated on the HMD while no self-motion in the VE was shown. A walking speed of ca. 1.25m/s...
After the initial startup distance, subjects had to continue walking with the assumed pace for a duration of 3 seconds towards a virtual target marker that we displayed at eye height at 10m distance. During this time, subjects received visual feedback of their self-motion in the virtual replica of the laboratory, which we scaled relative to their physical motion. While a user moves in the real world, applying gains to a user’s virtual self-motion corresponds to changes in the mapping from physical to virtual self-motion speed. Such differences were implemented based on user-centric coordinates as introduced by Steinicke et al. [Steinicke et al. 2010]. Virtual self-motions can be scaled with speed gains \( g \in \mathbb{R}_0^+ \), describing the mapping from movements in the real world to motions of a head-referenced virtual camera. For \( g = 1 \) physical translations \( T_p \in \mathbb{R}^3 \) are mapped one-to-one to virtual translations \( T_v \in \mathbb{R}^3 \) with \( T_v = T_p \cdot g \), i.e., the virtual scene remains stable considering the physical position change, whereas \( g_t < 1 \) result in slower and \( g_t > 1 \) in faster virtual self-motion speed. We applied gains in the range between \( g_t = 0.4 \) and \( g_t = 1.6 \) (cf. Section 3.3), i.e., the virtual speed differed by up to \( \pm 60\% \) from the physical speed.

Following the stimulus phase, the subjects were asked to answer a 2-AFCT question: “Did you move faster or slower in the virtual world than in the real world?” The subjects had to choose one of the two alternatives by pressing the up or down button on the Wii remote controller. If subjects cannot reliably discriminate between the virtual and real speed, they have to guess, and will be correct on average in 50% of the trials. After answering the question, subjects were guided back to the start position of the next trial. The gains converged on the PSE following the interleaved staircase approach described in Section 3.3. We measured a PSE, i.e., the gain on which the interleaved staircase approach converged, for each subject for each of the three physical speeds. The order in which the physical speeds were tested was given by Latin squares.

### Table 1: PSEs of all subjects for the tested speeds and distances.

<table>
<thead>
<tr>
<th>subject</th>
<th>physical speed</th>
<th>virtual distance</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1m/s</td>
<td>1.25m/s</td>
</tr>
<tr>
<td>s01</td>
<td>0.49</td>
<td>0.72</td>
</tr>
<tr>
<td>s02</td>
<td>1.39</td>
<td>1.37</td>
</tr>
<tr>
<td>s03</td>
<td>1.30</td>
<td>1.31</td>
</tr>
<tr>
<td>s04</td>
<td>1.43</td>
<td>0.87</td>
</tr>
<tr>
<td>s05</td>
<td>0.56</td>
<td>0.93</td>
</tr>
<tr>
<td>s06</td>
<td>1.19</td>
<td>1.22</td>
</tr>
<tr>
<td>s07</td>
<td>0.80</td>
<td>0.88</td>
</tr>
<tr>
<td>s08</td>
<td>1.21</td>
<td>1.41</td>
</tr>
<tr>
<td>s09</td>
<td>0.77</td>
<td>0.74</td>
</tr>
<tr>
<td>s10</td>
<td>0.87</td>
<td>0.89</td>
</tr>
<tr>
<td>s11</td>
<td>1.36</td>
<td>1.42</td>
</tr>
<tr>
<td>s12</td>
<td>1.11</td>
<td>1.33</td>
</tr>
<tr>
<td>s13</td>
<td>1.15</td>
<td>1.28</td>
</tr>
<tr>
<td>s14</td>
<td>1.13</td>
<td>1.13</td>
</tr>
<tr>
<td>s15</td>
<td>0.83</td>
<td>0.80</td>
</tr>
<tr>
<td>s16</td>
<td>1.35</td>
<td>1.09</td>
</tr>
<tr>
<td>s17</td>
<td>0.61</td>
<td>0.75</td>
</tr>
<tr>
<td>s18</td>
<td>1.07</td>
<td>1.11</td>
</tr>
</tbody>
</table>

Mean: 1.034 1.069 1.087 0.796 0.839 0.795

correlates to an average walking speed of HMD users according to Mohler et al. [Mohler et al. 2007]. Accordingly, the speed of 1m/s correlates to slow walking, and 1.5m/s correlates to a brisk walking pace. If subjects were unable to walk at a steady pace for at least 1m with the target speed (±0.1m/s) after an initial 2m straight distance the trial was repeated. Less than 10% of all trials had to be repeated.

### 3.5.2 Results

Figure 2(b) shows the results for the three tested within-subjects walking speeds. The x-axis shows the physical walking speed, and the y-axis shows the mean virtual walking speed that the subjects judged as equivalent to their physical motion. The vertical bars and colored regions illustrate the standard deviations.

Subjects judged a virtual self-motion speed as matching their physical walking speed that was approximately 1.03m/s (\( g_v = 1.034, 3.4\% \) increased) for a physical walking speed of 1m/s, 1.34m/s (\( g_v = 1.069, 6.9\% \) increased) for a physical walking speed of 1.25m/s, and 1.60m/s (\( g_v = 1.067, 6.7\% \) increased) for a physical walking speed of 1.5m/s. Over all physical walking speeds, subjects judged an approximately 5.7% increased virtual speed as identical to their physical motion, which corresponds to an underestimation of virtual speeds. Considering the PSEs of individual subjects, 9 of the 18 subjects consistently judged increased virtual speeds as identical with their physical speed, whereas 6 subjects consistently judged decreased virtual speeds as correct, and 3 subjects made mixed judgments.

Table 1 lists the judged PSEs and means over all subjects for the three physical walking speeds. We analyzed the results with a repeated measure ANOVA and paired-samples t-test. The results were normally distributed according to a Shapiro-Wilk test at the 5% level. The sphericity assumption was supported by Mauchly’s test of sphericity at the 5% level. We did not find a significant difference between the judged PSEs over all subjects and veridical results with \( g_v = 1 \) (t(17)=1.03, p<.32). We found no significant main effect of physical walking speed on the judged PSEs (F(2, 34)=2.75, p<.77, \( \eta_p^2=0.16 \)).

### 3.6 Experiment E3: Time Estimation

In this section we describe the experiment that we conducted to evaluate the perception of time in the IVE.

#### 3.6.1 Methods

Since there is no clear notion of a discrepancy between elapsed time in the real and virtual world that could be compared simultaneously with a 2-AFCT, we measured values for time estimation separately in the IVE (i.e., with HMD) and in the real world (i.e., without HMD). The trials in the IVE started with an initial stimulus phase in which subjects were presented with a view on the HMD to the virtual replica model of our laboratory. When they felt ready to start, subjects pressed a button on the Wii remote controller, which caused a short acoustic signal to be heard, and then walked in the direction of a virtual target marker that we displayed at eye height at 10m distance. A second short acoustic signal was displayed after subjects had walked over a time interval that we varied between experiment trials. Thereafter, subjects answered a 2-AFCT question, and were guided back to the start position of the next trial. We replicated the same setup for the experiment trials in the real world without the HMD.

For each trial subjects walked over a time interval that was scaled relative to a reference time. Therefore, we apply time gains \( g_t \in \mathbb{R}_0^+ \) to determine a subject’s walking interval in each trial, describing the relation between the trial’s interval and reference time. For \( g_t = 1 \) the trial’s interval \( T_v \in \mathbb{R}_0^+ \) matches the reference time \( T_p \in \{2s, 3s, 4s, 5s\} \) with \( T_v = T_p \cdot g_t \). In contrast, gains \( g_t < 1 \) result in subjects walking over a shorter interval and \( g_t > 1 \) in a longer interval relative to the reference time. We chose the reference times based on the common durations in the distance and speed experiments. After walking the trial’s interval, subjects were asked to...
answer a 2-AFCT question: “Did you move longer or shorter than # seconds?” with the # replaced by the corresponding reference time. The subjects had to choose one of the two alternatives by pressing the up or down button on a Wii remote controller. If a subject cannot reliably discriminate between the elapsed and reference time, the subject must guess, and will be correct on average in 50% of the trials. We applied gains in the range between $g_l = 0.4$ and $g_r = 1.6$ (cf. Section 3.3), i.e., the walked interval differed by up to ±60% from the reference time.

After answering the question, subjects were guided back to the start position of the next trial, i.e., they did not receive feedback about the elapsed time. The gains converged on the PSE following the interleaved staircase approach described in Section 3.3. We measured a PSE for each subject for each of the reference times, i.e., the gains at which the subjects judge the elapsed and reference times as identical. PSEs that deviate from $g_r = 1$ indicate subjective time dilation or expansion for the different reference times. The order in which the reference times were tested was randomized for each subject.

### 3.6.2 Results

Figure 2(c) shows the results for the tested within-subjects reference times. The x-axis shows the actual reference times, and the y-axis shows the judged times with and without HMD. The vertical bars and colored regions illustrate the standard deviations.

In the real world without using the HMD subjects judged an elapsed time as matching the reference time that was approximately 2.10s ($g_l = 1.051$, 5.1% longer) for a reference time of 2s, 3.05s ($g_l = 1.017$, 1.7% longer) for a reference time of 3s, 4.06s ($g_l = 1.016$, 1.6% longer) for a reference time of 4s, and 5.04s ($g_l = 1.007$, 0.7% longer) for a reference time of 5s. In contrast, for the experiment trials in the IVE, subjects judged an elapsed time as matching the reference time that was approximately 2.16s ($g_r = 1.081$, 8.1% longer) for a reference time of 2s, 3.22s ($g_r = 1.073$, 7.3% longer) for a reference time of 3s, 4.24s ($g_r = 1.061$, 6.1% longer) for a reference time of 4s, and 5.23s ($g_r = 1.046$, 4.6% longer) for a reference time of 5s.

Over all reference times, subjects judged an elapsed time as identical to the reference time that was approximately 6.5% increased ($g_r = 1.065$) in the IVE, and approximately 2.3% increased ($g_r = 1.023$) in the real world. The difference between the PSEs in the real world condition and the IVE suggest that times were overestimated in the IVE by approximately 4.2% compared to estimates in the real world condition. Considering the PSEs of individual subjects, 7 of the 10 subjects judged a longer elapsed time in the IVE compared to the real world, whereas 3 subjects judged a shorter elapsed time.

We analyzed the results with a repeated measure ANOVA and paired-samples t-tests. The results were normally distributed according to a Shapiro-Wilk test at the 5% level. The sphericity assumption was supported by Mauchly’s test of sphericity at the 5% level. We found a significant main effect of reference time on the judged PSEs according to a Shapiro-Wilk test at the 5% level. The sphericity assumption was supported by Mauchly’s test of sphericity at the 5% level. The sphericity assumption was supported by Mauchly’s test of sphericity at the 5% level. The sphericity assumption was supported by Mauchly’s test of sphericity at the 5% level.

Figure 3 shows that the results for the motion components do not differ from their physical self-motion. The PSEs suggest that half of the subjects underestimated the visual speed cues from the virtual world compared to the proprioceptive-vestibular speed cues from the real world. This notion supports previous experimental results in the literature, which suggested a tendency towards underestimation of virtual walking speed in IVEs [Banton et al. 2005; Bruder et al. 2012; Durgin et al. 2005; Steinicke et al. 2010].

The time estimation experiment showed that subjects were able to estimate time in the real world quite reliably without large errors, but judged times in the IVE were approximately 4.2% increased over the estimates in the real world. This apparent time dilation in IVEs is an interesting observation, which warrants further investigation. In particular, an open question which effect hard- and software factors have on time perception, e.g., compared to traditional computer graphics or game environments.

We performed our experiments in a state-of-the-art IVE with an Oculus Rift HMD and a large real walking space. In this IVE we found different magnitudes of biases in self-motion estimates regarding the three motion components. Pooled over all subjects we observed a large underestimation of virtual distances, slight underestimation of virtual speed, and a slight overestimation of time. Figure 3 shows that the results for the motion components do not have to adhere to their mathematical relation described in Section 1. Further research is necessary to triangulate how this self-motion perception triple looks in different IVEs considering that other researchers found different magnitudes of biases as described in Section 2. Moreover, Figure 4 shows that the self-motion perception tringles of individual users may differ from the means computed for a specific IVE, which should be carefully considered. In particular, the results show that all subjects are rather accurate in judging
1.5
1.25
1
0.75
0.5
0.25
0
0.25
0.5
0.75
1
1.25
1.5
0
0.25
0.5
0.75
1
1.25
1.5
actual
estimated
time
distance speed

Figure 4: Individual self-motion perception triples for all subjects who participated in all three experiments.

elapsed time, whereas individual differences in distance and speed judgments can be observed. In particular, nine out of ten subjects underestimate distances (except for the subject whose results are depicted in Figure 4(c)). Furthermore, the results highlight that six subjects underestimate speed (cf. Figure 4(b),(f)-(j)), but that subjects in case they overestimate speed, tend to highly overestimate speed (cf. Figure 4(a),(c),(d),(e)). In future experiments, such correlations may be considered in more detail.

It is important to determine the magnitude of common misperception in IVEs, in particular with current-state HMDs like the Oculus Rift. Such misperception effects greatly limit the applications of VR technologies in domains that require spatial perception that matches the real world. However, it is a challenging task to remedy the causes of these effects. In contrast, different approaches may be used to compensate for misperception by alleviating the effects in one of the motion components. Examples are magnification based on the field of view [Kuhl et al. 2009], increased optic flow [Bruder et al. 2012], up-scaled travel distances [Interrante et al. 2007] etc. In particular, our PSEs suggest that distances and speed may be scaled such that users judge them as veridical, and exogenous cues may be scaled to provide a more accurate time perception. However, while these approaches provide advantages in the scope of a motion component, and may even lead to accurate impressions of self-motion, potential disadvantages may arise from side-effects of manipulations on the other motion components and cognitive processes, which should be considered.

5 Conclusion

In this paper we have analyzed the triple of self-motion speed, distance, and time perception in an IVE using the same setup and similar protocols. In psychophysical experiments in an Oculus Rift IVE we measured PSEs indicating a bias in the three components of self-motion perception. The results show that virtual walking distances on average were underestimated by 19.0%, virtual speeds on average were underestimated by 5.7%, and time was overestimated by 4.2% in the IVE.

As illustrated in Figure 3, differences in this self-motion triple indicate perceptual biases, which may be effected by VR hard- and software or individual differences, and should be triangulated over different IVEs and user groups. More research is necessary to understand the reasons, interrelations, and implications of such perceptual biases introduced by VR technologies in IVEs.

Acknowledgements

Authors of this work are supported by the German Research Foundation.

References


