GiAnt: stereoscopic-compliant multi-scale navigation in VEs

Ferran Argelaguet Sanz, Maignant Morgan

GiAnt: Stereoscopic-Compliant Multi-Scale Navigation in VEs

Ferran Argelaguet* Morgant Maignant
Inria, France École nationale d'ingénieurs de Brest, France

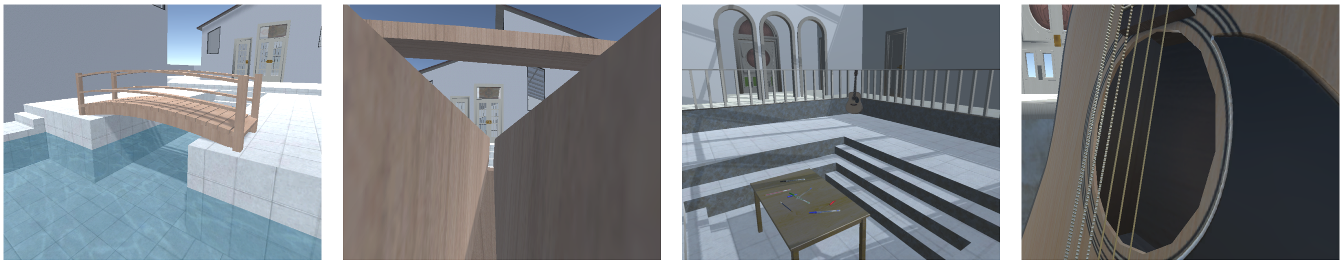

Figure 1: *This sequence presents four different camera locations during a multi-scale navigation path which requires not only the adaptation of the camera speed but also the stereoscopic rendering parameters. The proposed navigation technique, GiAnt, automatically adjusts the camera speed and the scale factor of the virtual environment in order to ensure a comfortable navigation experience.*

Abstract

Navigation in multi-scale virtual environments (MSVE) requires the adjustment of the navigation parameters to ensure optimal navigation experiences at each level of scale. In particular, in immersive stereoscopic systems, e.g. when performing zoom-in and zoom-out operations, the navigation speed and the stereoscopic rendering parameters have to be adjusted accordingly. Although this adjustment can be done manually by the user, it can be complex, tedious and strongly depends on the virtual environment. In this work we propose a new multi-scale navigation technique named GiAnt (GIant/ANT) which automatically and seamlessly adjusts the navigation speed and the scale factor of the virtual environment based on the user's perceived navigation speed. The adjustment ensures an almost-constant perceived navigation speed while avoiding diplopia effects or diminished depth perception due to improper stereoscopic rendering configurations. The results from the conducted user evaluation shows that GiAnt is an efficient multi-scale navigation which minimizes the changes of the scale factor of the virtual environment compared to state-of-the-art multi-scale navigation techniques.

Keywords: Multi-Scale, Navigation, 3DUI, Optical Flow

Concepts: **•Computing methodologies** → **Virtual reality;** **•Human-centered computing** → *User centered design;*

1 Introduction

Navigation is a key component of any 3D user interface [Kulik 2009] as it allows the user to control the virtual camera viewpoint enabling the exploration of the virtual environment. With the actual complexity of virtual environments, navigation interfaces need to support zoom-in and zoom-out operations in order to fit the desired area of the virtual environment with the correct amount of detail [Cho et al. 2014]. Zoom-in and zoom-out operations, when interactively navigating in the virtual environment, require a real-time adaptation of the navigation parameters, such as the navigation speed [McCrae et al. 2009] or the stereo display parameters [Ware et al. 1998], to ensure its usability. However, when the assumptions are violated, suboptimal navigation experiences can be generated. Fast camera motion might induce motion sickness [So et al. 2001a], slow motions might decrease user performance and engagement, excessive parallax might generate diplopia effects or not enough parallax might decrease depth perception. Thus, robust multi-scale navigation techniques are required to avoid such situations.

Current multi-scale navigation techniques adjust the navigation speed in order to ensure an optimal navigation speed, for example by analyzing the user's surroundings [McCrae et al. 2009] or the motion perception [Argelaguet 2014], focusing on either depth information [McCrae et al. 2009] or optical flow analysis [Argelaguet 2014]. However, there is a lack of study and evaluation of multi-scale navigation techniques compliant with stereoscopic content, being the works by Carvalho et al. [Carvalho et al. 2011] and Cho et al. [Cho et al. 2014] the only few exceptions. Multi-scale navigation techniques for stereoscopic content have also to adjust the stereo rendering parameters in order to (1) avoid diplopia effects (e.g. when approaching small objects in the virtual environment) and (2) maximize the user's depth perception (e.g. avoid the feeling that all the virtual environment is at the infinity). Diplopia effects are a critical issue as they generates strong visual discomfort [Lambooij et al. 2009].

In this work, we propose and evaluate a novel navigation technique for MSVE: GiAnt (GIant/ANT). GiAnt monitors the user's perceived navigation speed and automatically adjusts the navigation speed and the scale factor of the virtual environment to ensure an optimal navigation experience, both in terms of motion perception and stereoscopic comfort. The basis for this work are existing heuristic methods which dynamically adjust navigation speed taking into account the relationship between the user and the virtual environment [McCrae et al. 2009] and the user perception of the navigation speed [Argelaguet 2014]. The main contributions of this paper are (1) the design of a novel navigation technique which adapts the navigation speed and the scale of the virtual environment to ensure an optimal navigation experience in MSVE and (2) a formal evaluation of existing MSVE navigation techniques.

The remainder of the paper is structured as follows. Section 2 discusses related work on multi-scale navigation and motion sickness. Section 3 describes the GiAnt technique followed by its evaluation in Section 4. Finally, the conclusion is presented in Section 5.

*e-mail:fernando.argelaguet_sanz@inria.fr

2 Related Work

2.1 Multi-Scale Navigation

Multi-scale navigation techniques, in addition to provide mechanisms to allow the user to navigate through the virtual environment [Kulik 2009], they have to allow the user to seamlessly navigate at each level of the MSVE. In particular, multi-scale navigation techniques focus on the automatic adjustment of the navigation speed [McCrae et al. 2009; Argelaguet 2014] and the scale factor [Kopper et al. 2006; Cho et al. 2014]. According on how the different levels of scales are handled by the navigation technique we can distinguish discrete and continuous techniques.

Discrete techniques consider that the VE has independent levels of scale. Each level of scale has its particular navigation technique, predefined and discreet transitions are provided to navigate among levels of scale. In these solutions, the user does not have the control of the camera during the transitions. For example, the work presented in [Kopper et al. 2006] used a world in miniature representation of the virtual environment which allowed the user to navigate through the different levels of scale using a magnifying glass metaphor. The user could zoom in and zoom out using a virtual magnifying glass. Once in the new level of scale is loaded the user could navigate using standard navigation techniques. This solutions provide a measurable change of scale but has to be defined beforehand. In contrast, the user could lose context information since a teleportation is normally used, and landmarks determine the available destinations [Pierce and Pausch 2004]. [Kopper et al. 2006] minimized this issues through the world in miniature which allowed to track the current position of the user in the VE. Bacim et al. [Bacim et al. 2013] proposed another solution allowing the user to zoom in to desired location through progressive refinement.

Continuous Techniques allow for smooth transitions among different levels of scale by using adaptive navigation techniques. Such techniques monitor the user's viewpoint and actions in order to adapt the navigation parameters. The most common approaches use the distance between the user and the MSVE encoded either in the depth buffer [Ware and Fleet 1997] or in a six-sided depth CubeMap sampling the MSVE in all directions [McCrae et al. 2009; Trindade and Raposo 2011]. The depth information is afterwards processed, and typically, the minimum depth value is used to linearly adapt the navigation speed [Ware and Fleet 1997; McCrae et al. 2009]. Additional information can be extracted, such as the optical flow [Argelaguet 2014] or the viewpoint quality [Freitag et al. 2016]. Optical flow provides a perceptual measure about the perceived motion which can be used in combination with the depth information to ensure that the perceived navigation speed remains constant [Argelaguet 2014]. In contrast, the viewpoint quality provides an information about the relevance of the user's viewpoint, this information can be used to decrease the navigation speed when the viewpoint has a high visual quality and vice versa [Freitag et al. 2016].

2.2 Dynamic Adjustment of Stereo Parameters

An additional challenge when navigating in MSVE is the management of the depth range in order to reduce visual discomfort. Lambooij et al. [Lambooij et al. 2009] identified three main sources of visual discomfort in stereoscopic displays (1) excessive changes in accommodation and convergence linkage, (2) 3D artifacts resulting from insufficient depth information in the retinal images yielding spatial and temporal inconsistencies and (3) unnatural amounts of blur. However, current navigation techniques do not fully take into account these three sources of visual discomfort, while for desktop-based navigation the adaptation of the speed is enough, when stereoscopic rendering is considered this does not suffice. For example, when zooming in the virtual environment fusion problems will appear. On the contrary, when zooming out, due to the limitations on depth perception on stereoscopic content, there is the risk of perceiving the entire VE to be at the infinity (constant parallax). In such situations, one potential solution is to adjust the inter-pupil distance (IPD) [Ware et al. 1998; Carvalho et al. 2011] or adjust the depth range of the virtual environment [Cho et al. 2014] based on the distance between the user and the virtual environment. The goal in all cases is to maximize depth perception and avoid visual discomfort. However, existing solutions rely on depth-based heuristics or have strong a priori information. For example, in the work of Cho et al. [Cho et al. 2014] the considered virtual environment was a 3D model of the earth, which allowed authors to include additional rules when handling the depth adjustment. A priori information makes easier to avoid situations in which fuse problems might arise but it is not generalizable to arbitrary MSVE.

2.3 Motion Sickness and Navigation Control

Motion sickness can appear at different grades and different exposure times according to user individual differences. However, it has been acknowledged that the main cause is due to conflicting sensory information involving visual, vestibular and proprioceptive channels [Reason 1978; Stanney and Hash 1998]. These conflicts are typically manifested in three main forms: nausea, eyestrain (oculomotor disturbances) and disorientation [Kennedy et al. 1993].

Potential factors that have been studied for motion sickness in virtual environments (cybersickness) are diverse, from task-related factors (e.g. duration of exposure [Kolasinski 1995], habituation, navigation control [Stanney and Hash 1998], navigation speed [So et al. 2001b], viewing posture, the virtual reality setup (e.g. type of display, stereoscopy and latency [Sharples et al. 2008]) and user-related (e.g., age, gender [Sharples et al. 2008]). Although the findings in the literature vary from different studies, the majority of the studies rely on the Simulator Sickness Questionnaire (SSQ) [Kennedy et al. 1993] for the formal evaluation of motion sickness. The SSQ has been proven to provide a reliable measure of the degree of simulator sickness experienced by participants.

The feeling of being in control of the system (e.g. navigation) is known to have a positive impact on motion sickness symptoms [Rolnick and Lubow 1991; Stanney and Hash 1998]. If the user has no control over the system (e.g. automatic navigation) unexpected conflicts (the user is not able to predict the motion) will occur between different sensory inputs. In this situation, the participant is more likely to experience motion sickness.Similar results where obtained in [Sharples et al. 2008] when exploring active vs. passive navigation with different VR setups (e.g. HMD, projection systems).Conflicting situations results in a mismatch between the expected sensations and the experienced ones. This hypothesis was further supported by [Diels and Howarth 2011] which observed that motion profiles encountered in real life generate higher grades of motion sickness that unrealistic ones.

Recommendations by [Sharples et al. 2008] related to navigation control included (1) instruct participants with appropriate navigation strategies in order to minimize negative symptoms and (2) provide total control in terms of movement. Additionally, considering the results in [Stanney and Hash 1998], navigation control techniques have to be also constrained, decreasing their complexity and thus improving their ease of use. In line with these recommendations, in this work we propose an adaptive navigation technique in which the user fully controls the motion of the virtual camera, but the speed and the scale of the environment are automatically adapted in order to ensure a smooth navigation and an easy control.

3 Stereoscopic-Aware Multi-Scale Navigation

The proposed multi-scale navigation technique (GiAnt) extends the current state of the art of adaptive navigation techniques in order to enable multi-scale navigation in stereoscopic immersive displays. GiAnt is based on adaptive navigation techniques which monitor the perceived navigation speed and adjust the navigation parameters to ensure "optimal" navigation experiences [Argelaguet 2014]. The concept of this work is based on the fact that the perceived navigation speed is related both to the navigation speed and to the size of the virtual environment (level of scale). Thus, in addition to adapt the navigation speed, the adjustment of the perceived navigation speed can be achieved by modifying the scale of the virtual environment through a cyclopean scale transformation [Ware 1995]. For example, when the user is too close to a virtual object (e.g. a zoom-in operation) or reaching an empty space (e.g. getting out from a narrow tunnel) the navigation speed can reach low (e.g. $s < 1^{cm}/_{s}$) or fast speeds (e.g. $s > 1^{km}/_{s}$) which will strongly deviate from human walking speeds ($s \approx 1.4^{m}/_{s}$ [Marchal et al. 2011]). In this situations, we can assume that the level of scale is not adapted for user's viewpoint, thus an adaptation of the level of scale is required. However, as the change of the level of scale can be perceived by the user, instead of continuously adapting the scale factor, GiAnt (see Figure 2) follows a hybrid approach which adapts both the navigation speed and the scale factor of the virtual environment. Both adaptations, speed and scale, are computed according to the user's inputs (steering direction and user viewpoint) and to the user's perceived navigation speed. Furthermore, a speed correction step is introduced to ensure that the speed remains at a "human-level" scale and the scale correction aims to minimize diplopia.

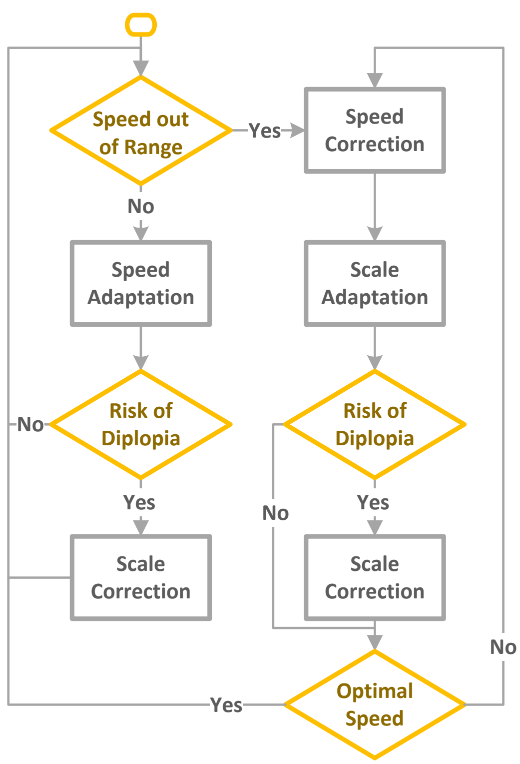

Figure 2: *Flow diagram of the GiAnt multi-scale navigation technique. The main processes are the speed adaptation and the scale adaptation which ensure that the perceived navigation speed and the level of scale are optimal.*

3.1 Computing the Perceived Navigation Speed

The computation of the user's perceived navigation speed follows the approach proposed in [Argelaguet 2014] which takes into account the relative distance of the user towards the VE based on the navigation speed and the optical flow. All computations are done using vertex and fragment shaders during the rendering process which decreases the computational foot-print and no pre-processing steps are required. During the rendering of the virtual environment, two image buffers ($n = 128 \times 128$ pixels) are computed:

- **Time-To-Collision (TTC) map**. Considering the current navigation speed, the TTC map stores the estimated time needed to reach each pixel. The computation only takes into account the distance in the direction parallel to the viewing direction.
- **Optical Flow (OF) map**. Stores the optical flow for each pixel. In contrast to previous work [Argelaguet 2014], the optical flow is estimated considering the current navigation speed (extrapolation). Using this estimation the rotational optical flow is easily discarded. Rotational optical flow can be introduced by head movements or camera rotations, which can distort the adaptation algorithm.

The data from the time-to-collision and optical flow maps is integrated to compute an estimation of the perceived navigation speed h (see Equation 1).

$$h = \sum_{i=0}^{n} \frac{TTC_i * (1 - OF_i) + OF_i}{n} \quad (1)$$

3.2 Speed Adaptation

The speed adaptation algorithm is inspired by previous works [Argelaguet 2014], the navigation speed is adapted to ensure that the perceived navigation speed (h) remains close to the optimal perceived navigation speed (h_{opt}). The speed adjustment algorithm considers that the relationship between h and h_{opt} is the same as the relationship between the current navigation speed v and the "optimal" navigation speed (v_{opt}) (see Equation 2). To avoid abrupt changes of speed, if $h < h_{opt}$, an hysteresis is used to smooth the acceleration. In contrast, when $h > h_{opt}$, a less smooth hysteresis is considered as the user is potentially navigating too fast, thus requiring smaller reaction times, additional details in this process can be found in [Argelaguet 2014].

$$v_{opt} = v \frac{h_{opt}}{h} \quad (2)$$

However, in GiAnt, the speed adaptation is only enabled when the navigation speed is close to the human walking speed $1.4^{m}/_{s}$ [Marchal et al. 2011], referred hereinafter as comfort speed, orange zone in Figure 3. The comfort speed provides a frame of reference to assess when the navigation speed is close to the average human walking speed. In other words, we assume that if the virtual environment is at the human scale, v_{opt} will be close to $1.4^{m}/_{s}$. When the speed adaptation is enabled, the navigation speed will oscillate according to the perceived navigation speed (see Figure 3). Yet, when the navigation speed diverges too much from the comfort speed, although the speed adaptation will ensure that the perceived navigation speed is optimal, we can assume that the level of scale is no longer at the human scale (it is either too big or to small).

We considered the acceptable range of speeds to be between $[0.7, 2.8]^{m}/_{s}$. The obtained results were not strongly dependent on this range. However, we consider that if the speed is half or twice the comfort speed the current level of scale is no longer valid and an update of the scale factor is required. The lower threshold is much more delicate that the upper one. A low speed means that either the user is approaching a virtual object or that it is navigating into a lower level of scale (e.g. getting inside an object), which can potentially increase the risk of diplopia. In contrast, the opposite scenario is less problematic. Increasing the speed means that either the user is navigating in an empty environment (e.g. crossing an empty factory) or moving towards a bigger level of scale (e.g. from street level to planetary level), which means that the virtual content is far away from the user (or the user is too small). Nevertheless, considering that stereoscopic depth perception is maximized at close distances, higher distances will result in decreased stereoscopic depth perception. Thus, there is a need to adjust the scale factor of the virtual environment not only to adjust the perceived navigation speed but to avoid the risk of diplopia and to improve depth perception.

3.3 Scale Adaptation

The scale adaptation algorithm (see Algorithm 1) uses the same rationale as the speed adaptation algorithm, however, it adjusts the perceived navigation speed by adjusting the scale factor (s) of the virtual environment. The adjustment is done computing the optimal scale factor (s_{opt}) according to the ratio between the perceived navigation speed (h) and the optimal perceived navigation speed (h_{opt}). For example, a ratio lower than one ($h < h_{opt}$) indicates that the perceived navigation speed is not fast enough, then, the perceived navigation speed can be increased by scaling the virtual environment down. However, (s_{opt}) cannot be applied directly as it might generate an abrupt and noticeable change of scale, instead, we apply an hysteresis to create a smooth transition which depends on the current perceived navigation speed. The convergence function ratio (see Algorithm 1, line 2) ensures that when the perceived navigation speed is closer to the optimal navigation speed the scale convergence increases. In our tests, specially when approaching virtual objects, there was a need to provide an asymmetrical convergence ratio. For example, when the perceived navigation speed is too high, there is a needed to provide a strong convergence ratio to minimize the risk of diplopia or to avoid potential collisions with the virtual environment. Once the new scale factor is computed (s_i), we apply a cyclopean scale which ensures that the user viewpoint remains unaltered. We have to notice that this algorithm is executed at each frame which ensures its stability and convergence.

In addition to adapt the scale factor to ensure an optimal perceived navigation speed, the scale adaptation has to adapt the scale factor of the virtual environment to ensure that the navigation speed converges towards the speed comfort zone. Intuitively, when the navigation speed is outside the comfort speed range, although the perceived navigation speed can be optimal, the level of scale of the virtual environment is no longer adapted to the human scale. However, once the scale mode is triggered, the navigation speed is no longer adjusted, which requires the addition of a speed correction step which will ensure the convergence towards the comfort zone.

Algorithm 1 Scale Adaptation Algorithm. The optimal scale factor s_{opt} is computed considering the ratio between the perceived navigation speed h and the optimal perceived navigation speed h_{opt}. An hysteresis (lines 2 and 3) is applied to compute the new scale factor to ensure a smooth transition. The convergence ratio parameters were determined empirically: $a = 2$, $b = 0.05$ and $c = 0.025$, in order to obtain the desired asymmetrical behavior based on the current value of h.

1: **Optimal Scale Factor**: $s_{opt} = s_{i-1} \cdot h_{opt}/h$
2: **Convergence Ratio**: $r = \Delta t/(a \cdot e^{(-b \cdot h)} + c)$
3: **Update Scale Factor**: $s_i = s_{i-1} + r(s_{opt} - s_{i-1})$
4: **Environment Scale**: CyclopeanScale($s_i - s_{i-1}$)

3.4 Speed Correction

The speed correction step ensures that the navigation speed v will converge towards the comfort speed ($v_{opt} = 1.4m/s$) in k seconds (see Equation 3). In our experiments k was 3s. An additional purpose of the speed correction step is to ensure that the speed adaptation is preferred over the scale adaptation. Although the scale adaptation is needed to ensure an optimal scale factor of the virtual environment, continuously adapting the scale of the environment can potentially disorient the user.

$$v_i = v_{i-1} + \frac{\Delta t}{k}(v_{opt} - v_{i-1}) \tag{3}$$

Finally, in order to avoid continuous mode changes between the scale and the speed adaptation and increase stability, once the scale adaptation is triggered, it will be active until the v_{opt} is reached again (Optimal Speed check in Figure 2). In Figure 3 we can observe how the scale adaptation algorithm is able to keep the perceived speed while strongly adapting the scale factor between $t = 2.5s$ and $t = 12s$.

3.5 Scale Correction

Diplopia provides a strong visual discomfort, thus it must be avoided at all costs. A normal person can focus an object without effort if this object is further than 20 cm [Acharya et al. 2012], objects closer than that distance increases the risk of diplopia effects as the brain will be not able to fuse both stereoscopic images. According to the nature of diplopia, when approaching virtual objects the risk of diplopia increases. We propose to monitor the risk of diplopia by analyzing the content of the depth buffer. Precisely, when more than 20% of the pixels of the depth buffer are closer than the comfort depth $d_c = 25cm$ to the viewer, we considered that there is a potential risk of diplopia. We consider $25cm$ for conservative reasons and a 20% coverage to ensure that the area of the screen which can lead to diplopia is relevant. When the risk of diplopia is detected, the environment is scaled up using a cyclopean scale in order to ensure that the objects that can create the diplopia effect are pushed away from the viewer, back to the comfort zone (see Equation 4). In order to keep the perceived navigation speed unchanged, the opposite scale factor is applied to the current navigation speed.

$$s_i = s_{i-1} \frac{d_c}{d_{min}} \tag{4}$$

Nevertheless, our tests showed that according to our speed range control, the risk of diplopia rarely occurs as the scale of the virtual environment is adapted to the human scale. When approaching a virtual object, the speed will be decreased due to an increased motion perception, which eventually will trigger a scale up of the environment before the risk of diplopia arises. Still, we included the test to ensure that the chances of diplopia remained minimal.

3.6 Results

We tested GiAnt in two different virtual environments: a virtual villa and the Menger's Sponge (see Figure 4). The villa environment provided a realistic virtual environment (in terms of scale) which only required the adaptation of the level of scale in particular

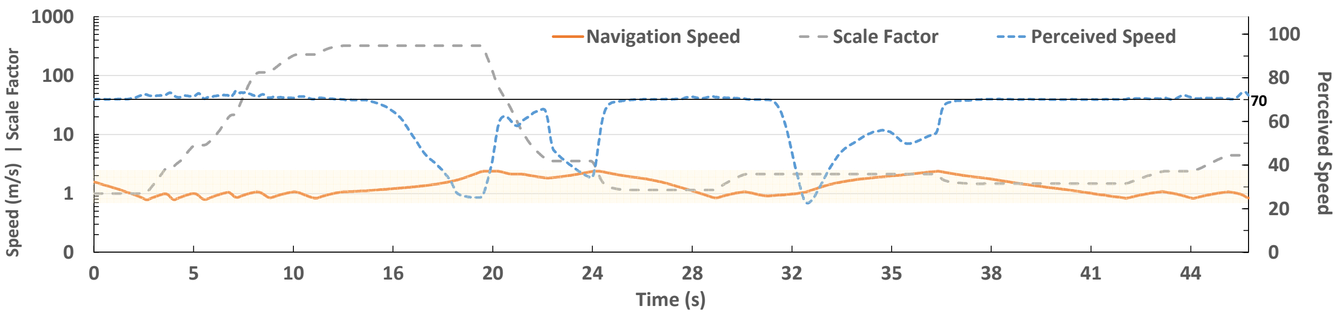

Figure 3: *Evolution of the navigation speed, the perceived navigation speed and the scale factor during the navigation depicted in Figure 1. When approaching to small cavities, the environment is scaled up to ensure the comfort speed range and avoid diplopia effects. Around $t = 3s$ the user stats to approach the bridge, notice that the strong change of the level of scale. Around $t = 42s$ the user approaches to the guitar, which also triggers a scale up of the environment.*

scenarios, e.g. while exploring small details or navigating through small passages (see Figure 5). In contrast, the Menger's Sponge is a procedural recursive virtual environment which allowed to stress the navigation technique as the navigation requires to continuously adjust the level of scale and the navigation speed.

Figure 3 shows the evolution of the navigation speed, the level of scale and the perceived navigation speed while performing the navigation depicted in Figure 1. The navigation required to approach to small cavities in the environment. The first on, at $t = 5s$, consisted in a close up view of the wooden bridge (first two images of Figure 1). From the plot we observe that the level of scale was increased on more than two levels of magnitude (from 1 to 300) in less than 10 seconds. After the user goes away from the bridge (between 24s and 41s) the user remains at a similar level of scale. Then, the user approach the second target, the guitar (last two images of Figure 1). When approaching the guitar, a new up scale of the environment is triggered allowing the user to get inside ($t = 45s$). During all the navigation, we observe that the navigation speed remains inside the comfort zone (orange rectangle). Regarding the evolution of the perceived navigation speed, we observe that the algorithm ensures that the value remains close to the optimal threshold ($h_{opt} = 70$). The drops of the perceived navigation speed (at $t = 16s$ and $t = 32$) match abrupt changes on the virtual environment: leaving the proximity of the bridge, and trespassing a window in order to get inside the house. The algorithm is conservative and avoids abrupt changes on the scale and the speed, unless the perceived navigation speed is higher than the optimal perceived navigation speed or there is a risk of diplopia.

We performed a simple pilot study asking participants to navigate in a VE and asked to find a set of eight hidden semitransparent spheres hidden at different locations (see additional material). In general participants (n=12) found the technique well suited for the task and surprisingly, they rarely noticed the changes in the scale factor of the environment. Still, considering that they had to explore small hidden locations, such as the interior of a pen (see Figure 5 right), participants could guess that the scale factor was adjusted. This promising results encouraged us to perform a formal user evaluation which is described in the next section.

4 User Evaluation

The goal of the conducted evaluation was twofold, first, compare GiAnt against state-of-the-art multi-scale navigation techniques, and second, evaluate the impact of dynamically adjusting the level of scale in terms of spatial perception and simulation sickness. We

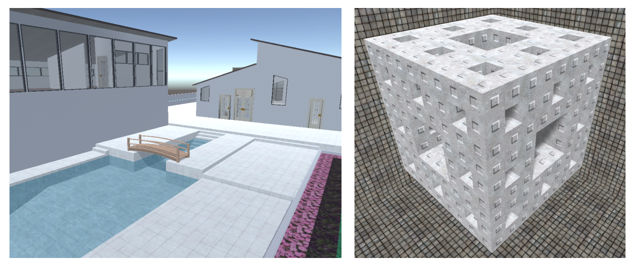

Figure 4: *Virtual environments used for testing. (Left) The villa environment, although not being a real multi scale environment, required the user to search for hidden objects inside a different hidden locations. (Right) The Menger Sponge is a recursive 3D fractal environment, due to rendering performance limitations, we restricted the navigation up to 6 levels of recursion.*

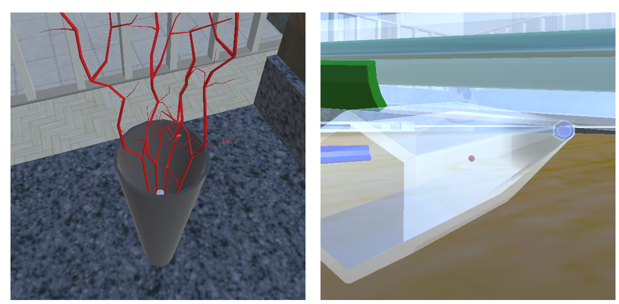

Figure 5: *In the villa virtual environment, participants were instructed to search for small hidden targets (semi-transparent spheres), hidden at different locations. (Left) Inside a vase. (Right) Inside a transparent pen.*

have to note that existing multi-scale navigation techniques have never been exhaustively evaluated [Carvalho et al. 2011; Ware et al. 1998]. Participants performed a path following task in a procedural (fractal) MSVE (see Figure 4 right) which allowed to stress the adaptation process.

4.1 Apparatus and Participants

Thirteen participants (age $M = 27$; $SD = 5.6$, 12 male and 1 female) from our lab took part in the experiment. One participant experienced moderate simulator sickness during the experiment and was discarded. All participants had a moderate gaming experience ($M = 5.6$; $SD = 1.3$) and VR experience ($M = 5$; $SD = 1$) (scale from 1 to 7). Considering the risk of simulation sickness, we preferred to recruit people used to VR equipment.

The virtual environment was displayed using an Oculus DK2 head-mounted display and participants could drive the navigation through a standard keyboard. Participants were seated during the experiment. The prototype was implement using the Unity game engine and provided a constant frame rate of 75fps, which ensured an optimal rendering for the Oculus Rift. We used a standard desktop computer to drive the system.

4.2 Experimental Protocol

At the beginning of the experiment, participants were provided with written instructions about the experiment and the user's consent form. After reading and signing the user's consent form, the experimenter asked the participant to fill a demographics questionnaire and were introduced to the equipment used during the experiment.

The navigation interface followed a flying metaphor. The steering direction was determined by the viewing direction and the forward motion was triggered when pressing the forward key arrow. Users had no control on the camera speed as the navigation technique was in charge of the speed adaptation. Additionally, if needed, users were allowed to rotate the camera towards the left and towards the right using the left and right arrow keys. The rotation speed was fixed at $1 ^{rad}/_{s}$.

Participants were asked to navigate through a procedural MSVE (see Figure 4 right) following a predefined path. The path required them to reach the deepest level of the Menger's sponge, which was constrained to six iterations. The path was nonlinear, in terms of scale factor, and allowed to measure the efficiency of the scale adaptation through many different scale transitions. The path was indicated through a set of 53 semitransparent spheres (see Figure 6) ensuring that participants followed a similar path. In addition, to avoid any possible ambiguity a 3D arrow was placed at the top center position of the user viewpoint pointing at the next waypoint. The arrow did not occlude their view, participants had to look up (without turning their head) in order to look at it.

The experiment was divided in three blocks, one for each tested navigation technique (see Section 4.3). Each block consisted in two traversals of the Menger's sponge, which required approximately 10 minutes to complete in total. After each block, a two minutes break was enforced in which participants removed the HMD. The break avoided a long and continuous exposure to the virtual environment. During the break participants filled the Simulation Sickness Questionnaire [Kennedy et al. 1993] and a subjective questionnaire (see Table 1).

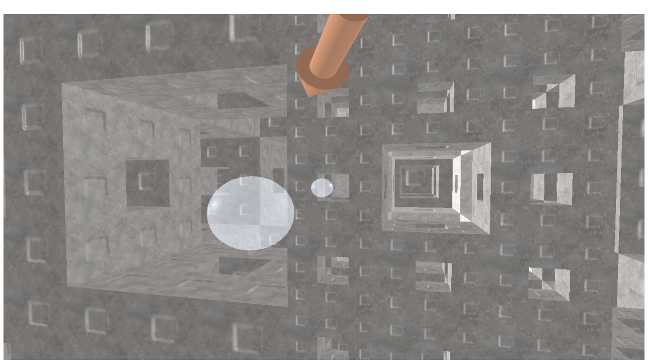

Figure 6: *The trajectory followed by participants during the experiment indicated by semitransparent spheres. Additionally an arrow placed at the center top of the user's viewport pointed to the next waypoint.*

4.3 Multi-Scale Navigation Techniques

In addition to GiAnt, two additional navigation techniques were considered. First, a multi-scale navigation technique based on the works presented in [McCrae et al. 2009] and in [Carvalho et al. 2011], which will be referred as *Depth-based*. The Depth-based technique provides a speed and an interpupillary distance (IPD) adaptation based on the minimum value of a depth cubemap (d_{min}) centered at the user's location. The speed (v) and the IPD are computed as shown in Equations 5 and 6. Changes in the IPD can be seen as scaling up and down the user which have the same effect as applying a cyclopean scale of $s = IPD_0/IPD$, with $IPD_0 = 0.064mm$ being the default IPD value:

$$v = d_{min} \cdot k_v \quad (5)$$

$$IPD = d_{min} \cdot k_{IPD} \quad (6)$$

The second technique was a combination between the speed adaptation algorithm proposed in [Argelaguet 2014] and the IPD adjustment proposed in [Carvalho et al. 2011]. The second technique which has never been evaluated before will be referred as *Perception-based*. The Perception-based technique and the GiAnt technique are driven by the perceived navigation speed (h) which is adapted to ensure an optimal perceived navigation speed (h_{opt}).

In order to minimize potential bias among techniques, we adjusted the parameters of each navigation technique to ensure that participants required in average the same time to perform the task and that the level of scale did not strongly deviate. First, we experimentally determined the optimal value for h_{opt} in a pilot study. We evaluated different values of h_{opt} which resulted of an optimal value of $h_{opt} = 70u$. With this configuration, we adjusted the different constants required for the depth- and perception-based approach in order to achieve a similar behavior: $k_v = 0.75s^{-1}$ and $k_{IPD} = 0.04$. Previous works used different adjustments for the depth-based approach [Carvalho et al. 2011] as they can be dependent on the virtual environment.

4.4 Experimental Design and Hypotheses

The independent variable was the navigation technique, which had three levels: Depth-based, Perception-based and GiAnt. We followed a repeated-measures within-subjects design, in which the order of the technique was counterbalanced using a Latin square design to minimize ordering effects. Two repetitions were done for each navigation technique, resulting in 6 traversals of the Menger's sponge for each participant. In total, the experiment lasted around forty-five minutes.

The dependent variables were the total task completion time and the accumulated scale change along the entire path. In addition, the time between to consecutive waypoints (53 waypoints) and the scale factor at each waypoint was also measured. For the depth-based and the perception-based techniques which adjusted the IPD, in order to allow a direct comparison with the GiAnt technique, the scale factor was computed as $s = IPD_0/IPD$. Finally, after each block, participants filled the Simulator Sickness Questionnaire and a subjective questionnaire with questions related to the navigation control, speed perception and scale perception (7-Likert scale). According to our experimental design, our hypotheses were:

- **[H1]** Faster task completion time for the perception and the GiAnt techniques,
- **[H2]** Smaller accumulated scale factor change for the GiAnt technique and
- **[H3]** Increased subjective preferences towards the perception and the GiAnt techniques.

4.5 Results and Analysis

The analysis of parametric data was done using repeated-measures ANOVA, when needed, Bonferroni multiple comparisons were performed ($\alpha = 0.05\%$). Friedman ANOVA and post-hoc Wilcoxon pairwise comparisons were used for the analysis non-parametric data.

4.5.1 Task Completion Time

The one-way ANOVA navigation technique vs. the total task-completion time showed a significant main effect ($F_{2,22}$=86.67; $p < 0.001$; η_p^2=0.44). Bonferroni post-hoc tests showed significant pairwise differences $p < 0.05$) between the Depth ($M = 237.69s$; $SD = 15.33s$) compared to the Perception technique ($M = 200.82s$; $SD = 4.45s$) and the GiAnt technique ($M = 198.58s$; $SD = 6.69s$). The increased task-completion time for the Depth technique supports **[H1]**.

Moreover, we also performed an ANOVA analysis considering the time required to traverse two consecutive waypoints (see Figure 7). For simplicity, we only report the significant pairwise differences between each navigation technique when the one-way ANOVA showed a significant main effect. The results, show that in 7 waypoints there was no significant differences among techniques, in 23 waypoints participants spend significantly more time with the Depth technique and in 23 waypoints participants spend significantly less time with Depth technique. In particular, the waypoints showing stronger differences (waypoints 10, 14, 23, 24, 28, 36, 44 and 51 in Figure 7) correspond to trajectories which involved a decrease of the level of scale (zoom-out) of the virtual environment, the changes of the level of scale are presented in Figure 8. The differences between the Perception technique and the Giant technique were minimal.

4.5.2 Scale Adaptation

In order to analyze the evolution of the scale factor along the trajectory, we considered the accumulated change of the level of scale along the entire trajectory. Before the analysis, and due to the range of the scale factor, a logarithmic transformation was applied to the data (see Figure 8). The one-way ANOVA navigation techniques vs. the accumulated scale factor, showed a significant main effect on technique ($F_{2,22}$=478.06; $p < 0.001$; η_p^2=0.97). Bonferroni post-hoc tests showed significant differences for all pair-wise comparisons (all $p < 0.05$). The Perception technique accumulated the highest scale factor ($M = 171.6$; $SD = 2,17$), followed by the Depth technique ($M = 166,8$; $SD = 4,67$) and the GiAnt technique ($M = 135,9$; $SD = 1,86$). This results supports **[H2]**.

Moreover, we performed an ANOVA analysis of the scale factor at each waypoint (see Figure 8) between the Depth technique and the GiAnt technique. We found a main effect in technique in 33 waypoints. Bonferroni post-hoc pairwise comparisons showed that in 13 waypoints the scale factor was significantly higher for the Depth technique and that in 20 waypoints it was the opposite.

Table 1: *At the end of each block participants answered the following questionnaire. The questions were grouped considering (Q1-Q2) the camera control, (Q3-Q5) the perception of the camera motion and (Q6-Q9) the perception of the scale factor of the virtual environment.*

Q1: The control of the camera was: (1) Difficult, (7) Easy
Q2: I wanted to change the speed of the camera. (1) Never, (7) Always
Q3: The motion of the camera was too fast. (1) Never, (7) Always
Q4: The motion of the camera was too slow. (1) Never, (7) Always
Q5: I felt that the camera motion was: (1) Variable, (7) Constant
Q6: The scale of the virtual environment changed. (1) Never, (7) Always
Q7: I experienced diplopia effects. (1) Never, (7) Always
Q8: Changes of the level of scale were: (1) Unnoticeable, (7) Annoying
Q9: The environment was at arms reach. (1) Never, (7) Always

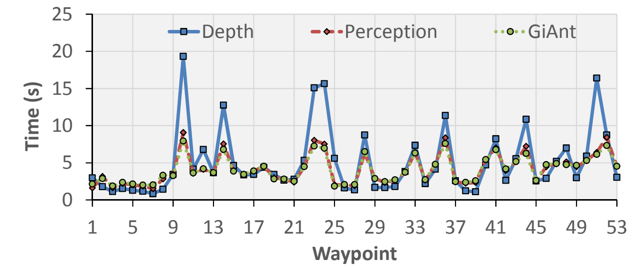

Figure 7: *Plot displaying the average time required to traverse two consecutive waypoints. We observe that in most situations all three techniques have a similar behavior, but the Depth technique presents some increased traversal times at several waypoints.*

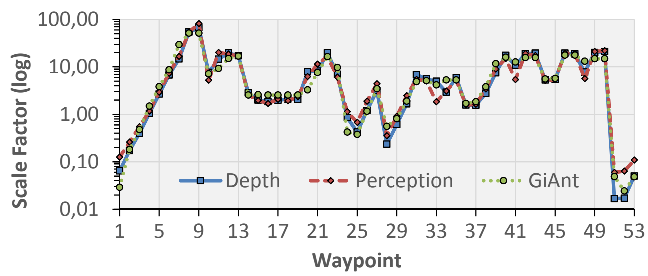

Figure 8: *Plot displaying the average scale factor at each waypoint for each tested technique. The navigation in the MSVE required constant adjustment of the scale factor. The results shows that both techniques had a similar behavior.*

4.5.3 Subjective Results

The analysis of the questionnaire data, only showed a main effect for Q3 (*The motion of the camera was too fast.*). Participants felt that the camera for the Depth technique was almost never too fast ($M = 1.66$; $SD = 0.49$) compared to the Perception technique ($M = 2.75$; $SD = 1.13$), no significant differences were found between the GiAnt technique and the others. For the remaining questions only the average for all techniques is discussed. Participants considered that the camera was easy to control (Q1, $M = 5.8$; $SD = 0.9$) although they moderately felt the need to change the speed (Q2, $M = 4.3$; $SD = 1.35$), mainly participants felt that the camera was generally too slow (Q4, $M=4.55$; $SD=1.46$). Regarding the speed perception, participants had a moderate perception of changes in the speed (Q5, $M=2.88$; $SD=1.36$). Finally, participants were moderately aware of the changes of the scale (Q6, $M=3.83$; $SD=1.59$), they rarely experienced diplopia (Q7, $M=1.36$; $SD=0.72$), the changes in the scale did not bother them (Q8, $M=1.66$; $SD=0.96$) and in general they perceived the environment to be close to them (Q9, $M=5.02$; $SD=1.42$). This results did not show a strong subjective preference for any the techniques, which do not support **[H3]** nor **[H4]**.

Regarding the SSQ ratings, the one-way ANOVA order vs. SSQ ratings showed a significant main effect ($F_{3,33}$=5.08; $p<0.01$; η_p^2=0.32). Post-hoc tests only showed a significant difference between the SSQ at the start of the experiment $M=1.5$; $SD=1.508$ and at the end $M=5.58$; $SD=6.13$. From all the participants which finished the experiment (12 out of 13) we did not observe any noticeable effect of simulation sickness.

4.6 Discussion

The navigation task inside the Merger's Sponges showed the need to adjust both the navigation speed and the level of scale of the virtual environment (or the IPD) in order to allow the user to freely navigate in multi-scale virtual environments. Speed and scale adaptation is needed in order to ensure an optimal navigation speed and a correct depth perception of the virtual environment. The results of the evaluation are discussed focusing on the management of the navigation speed, the management of the scale factor and the user comfort.

4.6.1 Speed Management

The Depth technique presented the highest task-completion time compared to the Perception and GiAnt techniques. This result contrasts with existing results in the literature [Argelaguet 2014] in which no differences were found between the Perception and Depth techniques (monoscopic content). Our hypothesis is that this discrepancy is explained by the fact that the navigation in the Merger's Sponge required a continuous adaptation of the scale factor which was not the case in the previous study. In the study from [Argelaguet 2014] the changes of level of scale were not abrupt and always consisted on zoom-in operations. This is partially supported by the analysis of the time required to traverse each waypoint which showed that participants required significantly more time to cross waypoints representing zoom-out operations (e.g. waypoint 10 or 23 in Figure 7 and 8). When the user performs a zoom-out operation, in most situations, the user is backwards to the geometry of the smaller level of scale (e.g. leaving a narrow tunnel into an open space). However, as the Depth technique adapts the navigation parameters considering all the information available in the surroundings of the user (depth cubemap), the content at the back of the user will still be considered for the navigation adaptation. In contrast, for the Perception and GiAnt techniques only the information in the viewing frustum is considered, which allows for faster, but unperceptive, accelerations during zoom-out operations.

4.6.2 Scale Management

GiAnt was the technique which had the minimum accumulated scale factor change (22% less). This shows, that the hybrid adaptation between the speed and the scale factor reduces the need of adjusting the scale factor, and it is only adjusted when necessary. We believe that this is the key feature of the GiAnt technique as changes in the scale factor can potentially alter the spatial perception of the user.

However, the similarity on the scale factor at each waypoint for all three techniques was a surprising and unexpected result (see Figure 8). This similarity shows, that although both techniques use a relative different method to compute the optimal scale factor, obtain comparable results. The explanation is that the Depth technique tries to enforce a relative navigation speed close to the average human walking speed ($1.4m/s$). Considering the speed adaptation for the Depth technique, we compute the relative speed ($v_r = s \cdot v$) from Equation 5, Equation 6 and the user defined constants k_v, k_{IPD} and IPD_0. It turns out, that the relative speed is constant and equals to $v_r = (IPD_0/k_{IPD}) \cdot k_v = 1.2m/s$. Considering that the Perception and GiAnt techniques were adjusted in order to ensure a perceived navigation speed close to the average human walking speed, this suggests, that all three techniques optimize the scale factor to ensure a similar relative navigation speed. This resemblance can partially explain the similarities of the scale factor at each waypoint.

In addition, although not considered in the user evaluation, the adaptation of the level of scale can allow for a hybrid navigation mechanism including head-tracking exploration. As GiAnt adjusts the level of scale to ensure an average human walking speed, the resulting scale factor will be always enable the user to directly explore the virtual environment by head-tracking exploration which is not directly supported by techniques only adjusting the IPD.

4.6.3 User Comfort

Although the subjective results were inconclusive, they show that participants were extremely tolerant to changes in the scale factor (Q8) and that they were mainly able to detect them due to contextual cues and not directly to the changes in parallax (Q6). This tolerance was also visible on the SSQ ratings. Although the SSQ score significantly increased along the experiment, the average SSQ score at the end of the experiment stayed within a safety margin. Only one participant experienced severe motion sickness, but he dropped the experiment at the end of the first block.

In addition, we observe that participants expressed a moderate need to manually adjust the navigation speed (Q2), mainly in situations in which the navigation speed was perceived to be too slow (Q4). This result was somehow expected due to the maximum accelerations allowed (e.g. convergence ratio from Algorithm 1). The filtering of the acceleration avoids extreme accelerations that could destabilize the system at the expense of potentially slow speed values. This behavior can be observed in Figure 3: due to abrupt changes of the MSVE the perceived speed navigation can be go below the optimal perceived navigation speed threshold. Some participants suggested the increase of the user control, for example by providing a "Boost mode", allowing to speed up the camera during such a situation. We envisioned to provide additional control to the user, but in order to keep the control as simple as possible we did not provided any additional speed control mechanism.

5 Conclusion

In this paper we have proposed and evaluated a novel multi-scale navigation techniques: GiAnt. GiAnt dynamically adapts the navigation speed and the scale factor of the virtual environment according to the user's perceived navigation speed. In contrast to existing techniques, the main novelty of our work is to provide a stereoscopic compliant technique which accounts for the user's perceived navigation speed and not only on spatial information as previous stereoscopic compliant techniques. In particular, the GiAnt technique ensures that the depth range of the virtual environment avoids diplopia effects and ensures that elements of the virtual environment remain close to the user, thus improving depth perception. A user evaluation was conducted in order to evaluate different multi-scale navigation techniques on two different virtual environments: a realistic and a procedural virtual environment. The results of the user evaluation showed that GiAnt allowed for an significant more efficient navigation while significantly minimizing the required changes of the level of scale. Taken together, these results indicate that GiAnt provides an efficient solution for multi-scale navigation in virtual environments. Future works should explore control methods providing more flexibility to "expert" users and explore the impact of adding additional degrees of freedom to the navigation control to the cognitive workload of the user.

References

ACHARYA, R., NQ, E. Y. K., AND SURI, J. S., Eds. 2012. *Image modeling of the Human Eye*. CRC Press, ch. 2.7.3, 61.

ARGELAGUET, F. 2014. Adaptive navigation for virtual environments. In *IEEE Symposium on 3D User Interfaces*, 91–94.

BACIM, F., KOPPER, R., AND BOWMAN, D. A. 2013. Design and evaluation of 3D selection techniques based on progressive refinement. *International Journal of Human-Computer Studies 71*, 7-8, 785–802.

CARVALHO, F., TRINDADE, D. R., DAM, P. F., RAPOSO, A., AND DOS SANTOS, I. H. F. 2011. Dynamic adjustment of stereo parameters for virtual reality tools. In *IEEE XIII Symposium on Virtual Reality (SVR)*, 66–72.

CHO, I., LI, J., AND WARTELL, Z. 2014. Evaluating dynamic-adjustment of stereo view parameters in a multi-scale virtual environment. 91–98.

DIELS, C., AND HOWARTH, P. A. 2011. Visually induced motion sickness: Single- versus dual-axis motion. *Displays 32*, 4, 175–180.

FREITAG, S., WEYERS, B., AND KUHLEN, T. W. 2016. Automatic speed adjustment for travel through immersive virtual environments based on viewpoint quality. In *IEEE Symposium on 3D User Interfaces*, 67–70.

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.

KOLASINSKI, E. M. 1995. Simulator sickness in virtual environments. Tech. rep., US Army Research Institute for the Behavioral and Social Sciences, (ARI Tech. rep. 1027). Alexandria, VA.

KOPPER, R., NI, T., BOWMAN, D. A., AND PINHO, M. 2006. Design and evaluation of navigation techniques for multiscale virtual environments. Ieee, 175–182.

KULIK, A. 2009. Building on realism and magic for designing 3D interaction techniques. *IEEE Computer Graphics and Applications 29*, 6, 22–33.

LAMBOOIJ, M., IJSSELSTEIJN, W., FORTUIN, M., AND HEYNDERICKX, I. 2009. Visual discomfort and visual fatigue of stereoscopic displays: A review. *Journal of Imaging Science and Technology 53*, 3, 030201.

MARCHAL, M., PETTRÉ, J., AND LÉCUYER, A. 2011. Joyman: A human-scale joystick for navigating in virtual worlds. In *IEEE Symposium on 3D User Interfaces*, 19–26.

MCCRAE, J., MORDATCH, I., GLUECK, M., AND KHAN, A. 2009. Multiscale 3D navigation. In *Symposium on Interactive 3D Graphics and Games*, 7–14.

PIERCE, J., AND PAUSCH, R. 2004. Navigation with place representations and visible landmarks. In *IEEE Virtual Reality*, 173–288.

REASON, J. T. 1978. Motion sickness adaptation: a neural mismatch model. *Journal of the Royal Society of Medicine 71*, 11, 819–29.

ROLNICK, A., AND LUBOW, R. E. 1991. Why is the driver rarely motion sick? The role of controllability in motion sickness. *Ergonomics 34*, 7, 867–879.

SHARPLES, S., COBB, S., MOODY, A., AND WILSON, J. R. 2008. Virtual reality induced symptoms and effects (VRISE): Comparison of head mounted display (HMD), desktop and projection display systems. *Displays 29*, 2, 58–69.

SO, R. H. Y., HO, A. T. K., AND LO, W. T. 2001. A metric to quantify virtual scene movement for the study of cybersickness: Definition, implementation, and verification. *Presence: Teleoperators and Virtual Environments 10*, 2, 193–215.

SO, R. H. Y., LO, W. T., AND HO, A. T. K. 2001. Effects of Navigation Speed on Motion Sickness Caused by an Immersive Virtual Environment. *Human Factors: The Journal of the Human Factors and Ergonomics Society 43*, 3, 452–461.

STANNEY, K. M., AND HASH, P. 1998. Locus of user-initiated control in virtual environments: Influences on cybersickness. *Presence: Teleoperators and Virtual Environments 7*, 5, 447–459.

TRINDADE, D. R., AND RAPOSO, A. B. 2011. Improving 3D navigation in multiscale environments using cubemap-based techniques. In *ACM Symposium on Applied Computing*, 1215.

WARE, C., AND FLEET, D. 1997. Context sensitive flying interface. In *Symposium on Interactive 3D graphics (SI3D)*, 127–ff.

WARE, C., GOBRECHT, C., AND PATON, M. 1998. Dynamic adjustment of stereo display parameters. *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans 28*, 1, 56–65.

WARE, C. 1995. Dynamic stereo displays. In *ACM SIGCHI Conference on Human Factors in Computing Systems (CHI)*, 310–316.