The “Infinite Loop” Towards Adaptive 3D User Interfaces

Ferran Argelaguet

To cite this version:

Ferran Argelaguet. The “Infinite Loop” Towards Adaptive 3D User Interfaces. Graphics [cs.GR]. Université Rennes I, 2021. tel-03903062

HAL Id: tel-03903062
https://inria.hal.science/tel-03903062v1

Submitted on 16 Dec 2022

HABILITATION
À DIRIGER DES RECHERCHES

présentée par

Ferran Argelaguet

Discipline: Informatique

Spécialité: Realité Virtuelle

The “Infinite Loop”
Towards Adaptive 3D User Interfaces

November 9, 2021

Jury

Victoria Interrante	Professeur, University of Minnesota	Rapporteur
Michel Beaudouin-Lafon	Professeur, Université Paris-Saclay	Rapporteur
Doug Bowman	Professeur, Virginia Tech	Rapporteur
Antony Steed	Proffeseur, University College London	Examinateur
Daniel Mestre	Directeur de Recherche, CNRS	Examinateur
François Chaumette	Directeur de Recherche, Inria	Examinateur

Habilitation à Diriger des Recherches préparée au sein de l’équipe Hybrid commune à l’Inria et au CNRS

Contents

The Infinite Loop

> *"There is something so familiar about this. Do you ever have déjà vu?"*
> *"Didn't you just ask me that?"*
>
> — Rita and Phil, *Groundhog Day*

While interacting with our environment, humans are driven by an infinite loop: the perception-action loop. This loop runs endlessly all day long, the brain receives and process external stimuli, determines which actions wants/needs to perform and executes them, such actions generate additional stimuli closing the loop. The perception action loop models a complex process which is bounded by the perceptual, cognitive and motor skills of every one of us. By iterating in this loop over and over, and thanks to the plasticity of our brains, humans are able to learn and adapt during our entire life. Indeed, this process allows a human to learn how to walk but also how to use any tool, from a physical hammer to a virtual flying interface while being immersed in a virtual environment. I am bounded to this infinite loop while writing these lines but also the reader while reading these same lines.

In real life, this loop is non-mediated, we can directly perceive the real world and act on it. Yet, when immersed on a virtual or augmented reality we perceive and act indirectly. The virtual world is perceived thanks to a number of output devices (e.g. screen, headphones, haptic devices) and we are able to act thanks to a number of input devices (e.g. tracking system, buttons, joysticks). The perception-action loop in virtual and augmented reality (see Figure 1) can be decomposed in: (1) the user receives multi-sensory feedback from the virtual environment (perception), (2) decides and plans the action he/she wants to perform (cognition), (3) executes the planed actions (action) and (4) the system interprets and executes the user's actions (commands). The execution of the commands will generate additional feedback (5), closing the loop. The user interface, commonly refereed as 3D user interface or 3DUI, becomes the tool that enables the user to interact and perceive the virtual environment, it translates the user's actions into commands and generates feedback that can be perceived by the user. However, issues in any of these processes will hinder user interaction, having a non-negligible impact on the overall usability of the system.

Using the perception-action loop as a canvas, this manuscripts presents a selection of my research activity since I started my PhD in 2006 in order to enhance such loop. My main goal, as stated in the title if this manuscript, is the design of adaptive 3D user interfaces. Interfaces that are aware of the perception and interaction capabilities of the users. Interfaces that are

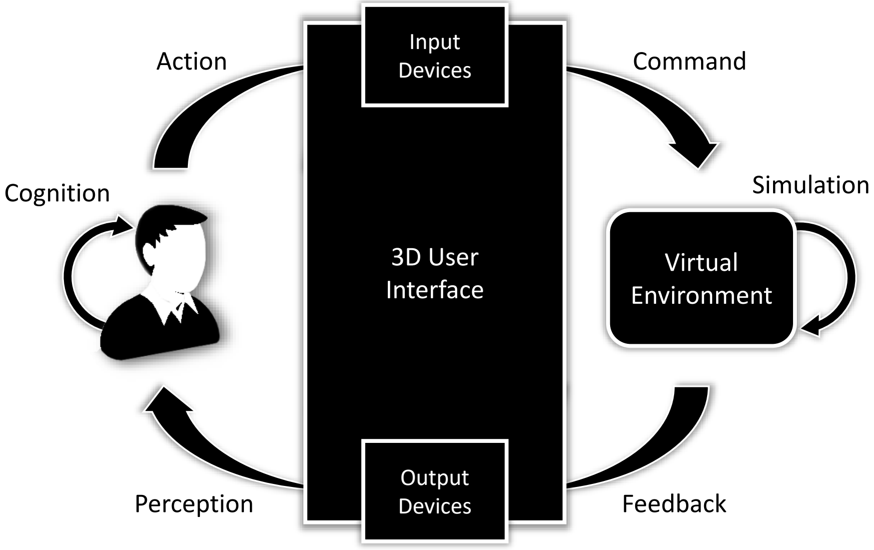


Figure 1: Perception-action loop in 3D User Interfaces

able to efficiently support the user while performing 3D interaction tasks. Compared to real life interactions which are bounded by the laws of physics, interactions in virtual environments are only bounded by our imagination.

Research Axes

In order to achieve my main goal, design interfaces better adapted to the user, I identify three major research axes, focusing on the understanding, improvement and enrichment of the "infinite loop".

Objective 1: Assess user perception in virtual environments. The way humans perceive the environment and ourselves is determined by our senses and previous experiences. However, when immersed in a virtual or augmented environment perception can be disrupted. In order to ensure efficient interaction, it is needed to understand how perception and the interaction process is altered. My research has been focused in the perception of distances, the perception of shapes and materials and the perception of the self.

Objective 2: Design of user-compliant 3D user interfaces. User interaction is driven by perception and users' previous knowledge. Thus, 3DUIs that better support the user will improve the interaction process. In this context, I have mainly focused in two interaction tasks: 3D object selection and navigation. While object selection is strongly driven by distance perception, navigation is influenced by motion perception.

Objective 3: Assess and exploit user's cognitive and mental states. Having precise knowledge of the user's cognitive and mental state could enable the design of interfaces that can adapt based on this additional input modality. We started to explore several potential usages, such as the design of virtual training applications that target specific cognitive states, the proposal of new evaluation methodologies and the usage of brain-computer interfaces.

Overview of the manuscript

The reminder of the manuscript is structured as follows. Part I focuses on perceptual aspects of the interaction process, describing relevant works we have conducted on assessing user perception in the scope of depth perception, pseudo-haptics and avatar perception. Then, Part II focuses on relevant 3D interaction techniques which we have proposed for 3D object selection and navigation interaction tasks. Finally, Part III presents research conducted to assess and to exploit the user's cognitive and mental state, either to design virtual training systems, monitor the user or to enable direct interactions. The manuscript concludes with a global discussion on the conducted research and future perspectives.

Part I

Understanding the loop - Perception in Virtual Environments

Introduction

> *"The environment of animals and men is what they perceive. The environment is not the same as the physical world, if one means by that the world described by physics."*
>
> — James J. Gibson*, The Ecological Approach to Visual Perception*

Humans are surrounded by a complex reality, which is sensed by our highly specialized senses (sight, smell, touch, taste, and hearing). All the sensory information captured, allows us to build a representation of our reality which will be the starting point for any interaction. For example, when grabbing a glass of water, we need to know where the glass is located in 3D space to plan the grabbing motion. Then, when the grabbing motion is performed, the information sensed by the skin in contact with the glass will determine the required force to actually grab it. If the force exerted is not enough, the glass will slip from our fingers but if it is too much the glass might break. While the task in itself serves a purpose, I want to drink because I am thirsty, all the actions performed are mostly unconscious, we have not though on how to grasp and lift the glass, previous experience and muscle memory have driven the interaction. As we are continuously perceiving our reality, our senses are adapted and calibrated for it. However, what happens when our perception is altered?

Due to current technical limitations, our perception is altered when we are immersed in a virtual environment. For example, although virtual reality displays are advancing at high pace, they are not still able to reproduce how light from a real environment reaches our eyes. Thus, as the visual stimulation coming from the virtual environment does not match the visual stimulation coming from the real world, a number of potential conflicts and ambiguities have to be resolved by our brain, i.e. the brain has to "interpret" the sensory data. As discussed in Section 1, one example is depth perception. While eye convergence and accommodation are depth cues, VR displays are still not able to provide a full convergence-accommodation coupling. While VR displays are able to fully support convergence using stereoscopic rendering, most VR displays only provide one focal plane, which does not allow to fully simulate accommodation depth cues. Known as convergence-accommodation conflict, it can potentially distort depth perception and increase ocular fatigue.

Another example is haptic perception. Haptic perception encloses tactile (skin and mechanoreceptors) and kinesthetic (muscles and tendons), which allows us to feel surfaces and external forces. However, current haptic systems are far away from being able to deliver precise and realistic haptic feedback. The main technical constraints relate to the high frequency features of surfaces and the ranges of forces that can be generated using mechanical systems. Approximations can be generated, such as using vibrotactile feedback to simulate the roughness of a surface or using single point force feedback to simulate contact. Nevertheless, haptic perception is partially driven by visual perception. One example is pseudo-haptic feedback, which leverages the dominance of vision over haptic feedback in order to increase the range of haptic sensations that can be generated. An isometric device, such as spring, coupled with the appropriate visual feedback can enable the simulation of wide range of haptic sensations (e.g. different stiffness values). Section 2 provides examples on how vision can be used to elicit stiffness while interacting with a 2D texture. This example illustrates how the brain tries to "make sense" of the perceived stimuli that is strongly influenced by our previous experiences.

In order to achieve the desired goal, and driven by perception, the user will determine the actions to perform. Once these actions are planed, the brain will initiate and monitor the realization of the task. For example, a simple arm reaching task (e.g. pushing a button), the realization can be decomposed in two phases. A ballistic phase, fast motion with low accuracy, in which the majority of the motion is performed, and a correction phase, lower speed with higher accuracy, in which the objective is reached. Such decomposition has been widely studied in 2D and 3D reaching tasks, and its virtual counterparts (e.g. clicking a button with a mouse or reaching a 3D virtual object with a hand-held device). While the ballistic phase is mainly driven by proprioception, the correction phase is mainly driven by visual information (e.g. eye-hand coordination). The brain, following a perception action loop, will assess whether the performed motion (vision and proprioception) is suited to reach the final goal and correct the motion accordingly. Reaching tasks have an intrinsic complexity, as firstly modeled by Fitts, they are defined by the amplitude of the motion and the size of the target. Intuitively, Fitts' law, states that the mean time to reach a target is higher the further away or smaller the target is. Fitts' law has been successfully used to improve reaching tasks in virtual reality interaction contexts. However, in addition to the intrinsic complexity of the reaching task, the accuracy of our perception also plays a key role on the overall performance. For example, a poor perception of our environment will generate inaccurate ballistic motions requiring additional slower corrective movements. Section 1.2 describes the Eye-Hand visibility mismatch which hinders rapid pointing gestures due to a limited spatial perception of the virtual environment. Thus, in the context of virtual reality, in which perception is distorted, motion planning could become inaccurate requiring additional, slower, corrective vision-driven loops. Furthermore, virtual reality involves mid-air interactions, which can increase the difficulty of corrective motions due to lack of physical constraints. How would you grasp a glass if you do not feel the glass under your fingers? Although haptic feedback could be considered, current systems tend to rely mainly on visual feedback. Finally, an additional aspect that can alter the interaction process is the perception of our body in the virtual environment.

With the democratization of head mounted displays, the representation of the user in the virtual environment has become an active research field. Once immersed in a virtual environment, using a head mounted display, our physical body is no longer visible, and thus, a virtual representation is needed. However, the design and conception of the user's virtual representation is tailored by a number of technical (e.g. motion capture capabilities), data (e.g. 3D model acquisition) and algorithmic (e.g. animation, 3D reconstruction) constraints (see Section 3.1). Although the virtual representation of the user is only bounded by imagination, we will only focus on anthropomorphic representations, and we will refer it as "the avatar". However, which are the mechanisms that enable us to become embodied in an avatar? Is the avatar perceived as being our real body? Referred as the *Sense of Embodiment*, numerous studies have been interested in this phenomena, which in the context of virtual reality, is decomposed in three main dimensions: self-location, agency, and ownership. While agency can be directly linked to the ability to interact with a system, the effects of self-location and ownership in interaction are less clear, as they relate on where the users perceive their body and whether they feel that they own the virtual body, i.e. the virtual body becomes their real body. Nevertheless, as discussed in Section 3.1 when users have to perform a specific task while being embodied in an avatar they tend to value the control of the avatar over their appearance. Yet, in situations in which the virtual body is threatened, their behavior can be altered depending on their sense of ownership.

In addition to the sense of embodiment, in an interaction context, the avatar plays a strong role on the perception action loop. First, our body is our reference frame. Although we can perceive our body thanks to proprioception, being embodied in a virtual body can modulate our perception. For example, being embodied in lower or taller avatar will alter how we perceive the size of objects. We will tend to perceive things bigger if we are embodied in a small virtual body. Second, we will judge the actions performed by our virtual body based on our actual motions, am I agent of the actions of my virtual body? If I move my right arm and the avatar moves the left arm I will feel disturbed. In both cases, when the shape and the control of our virtual body do not match our real body, the perception of our own actions will not match our expectations. This is illustrated in Section 3.2 in which we show that users tend to have high expectations when using human-like avatars compared to robot-like avatars feeling less agent of their actions with the human like avatar. Nevertheless, due to the fast adapting nature of our brain, the brain can integrate the new shape and control of our virtual body in order to recalibrate its motor actions and regain their "control" of the avatar.

Overview

The following chapters describe the works we have conducted in order to better understand distance estimation in virtual environments (Chapter 1), how haptic perception can be modulated using vision (Chapter 2) and how we perceive our virtual self in virtual reality (Chapter 3).

The works presented in Chapter 1 have been conducted in collaboration with (alphabetic order) Gerd Bruder, Yuta Itoh, Anatole Lécuyer, Jean-Marie Normand, Guillaume Moreau, Anne-Hélène Olivier, **Etienne Peillard** (PhD Student), Julien Pettré and Thomas Thebaud (Master Student).

The works presented in Chapter 2 have been conducted in collaboration with (alphabetic order) **Antoine Costes** (PhD Student), Fabien Danieau, Anatole Lécuyer, David Gomez Jauregui, Philippe Guillotel and Maud Marchal.

The works presented in Chapter 3 have been conducted in collaboration with (alphabetic order) **Diane Dewez** (PhD Student), Anatole Lécuyer, **Rebecca Fribourg** (PhD student), Ludovic Hoyet, Daniel Mestre, Corentin Nicole (Master Student), Mel Slater and Michael Trico (Master Student).

Chapter 1

Perception of Virtual Environments: Biases and Misperceptions

In this Chapter, I review the works focused on understand perceptual issues in mixed reality environments. As previously said, perception drives the interaction process, thus it is paramount to understand human perception to be able to propose methods capable of overcoming these perceptual issues.

1.1 Distance Perception in Mixed Reality

Understanding how humans perceive their surroundings has been, and still is, and active topic in human perception research [1]. However, in virtual and augmented realities, as they are mediated environments, the perception of distances can be distorted due to this mediation [2]. In particular, depth cues can be distorted due to intrinsic and extrinsic limitations of the visual rendering systems potentially biasing distance estimations. This is a major issue while interacting as poor distance estimations (underestimation or overestimation) will hinder the interaction process. For example, users will not be able to accurately plan their motions to reach virtual objects [3] or they will have a distorted perception of the virtual spaces.

In that respect, we conducted a number of experiments in order to explore different aspects of distance perception, considering different display systems (large projection displays [4] and retinal projection displays [5]) and different tasks such as exocentric distance estimation [6] or locomotion [7]). The different experiments have contributed to the better understanding of distance perception in virtual and augmented reality environments.

In the context of virtual reality, we have explored the role of the relative position between the user and a large immersive projection system [4] (see Figure 2). Our results showed that the screen distance and parallax have a strong asymmetric effect on distance judgments while screen resolution had a minimal effect. In particular, we observed increased underestimation for positive parallax conditions and slight distance overestimation for negative and zero parallax

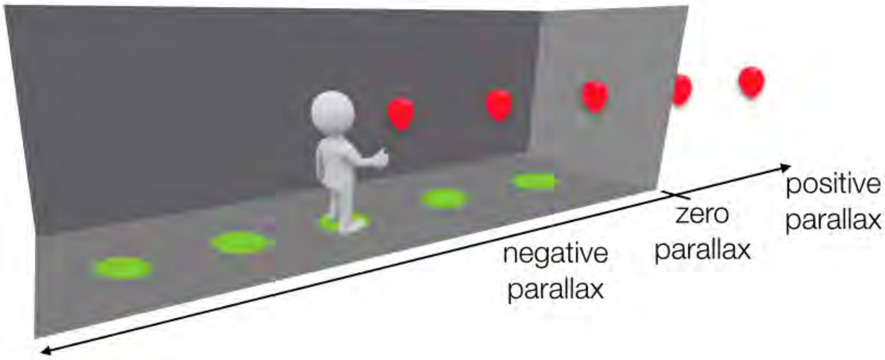


Figure 2: Experimental conditions for the egocentric distance perception on large immersive projection systems. The green dots represents the positions in which the user could stand. For each position five distances were tested (1m, 3m, 5m, 7m, 9m).

conditions. These results showed that distance perception was more accurate for objects placed between the user and the projection screens. In a different study [8], we explored the impact of the relative position of virtual objects with respect to the user. Most distance perception experiments consider that the stimuli is always placed in front of the user, however, what happens when the observer looks on the side? Confirmed by four controlled experiments, we discovered that the orientation of virtual stimuli with respect to the user introduces a distance perception bias: objects placed on the sides are systematically perceived farther away than objects in front. In addition, we could observe that this bias increases along with the angle, and appears to be independent of both the position of the object in the field of view as well as the quality of the virtual scene. These two examples show that nowadays we still lack knowledge on the impact of the visual display on distance perception and that there are perceptual biases that could be independent on the display system that are yet to be discovered.

In the context of augmented reality, the cohabitation between real and virtual information makes the study of distance perception even more challenging. First, distance judgments in VR are driven only by the perception of virtual content, however, in AR, the perception of real and virtual objects is intertangled. Second, due to the need of mixing virtual and real content, the degree of complexity of AR displays can be argued to be more complex than VR displays. See-through displays have to minimize the distortion of the real environment, the user's location have to be tracked precisely and accurately, and real and virtual content have to be blended seamlessly [9]. These challenges introduced more than 20 years ago are still partially solved. We conducted two perceptual studies [6, 5] to better understand potential biases that could alter distance perception in AR systems.

Distance perception studies (in MR) mainly focus on the estimation of objects placed in front of the observer. However, in the context of AR, the spatial relationships between objects plays a key role, especially when mixing virtual and real content. To explore this, we conducted a distance estimation experiment focusing on exocentric distance perception (see Figure 3).

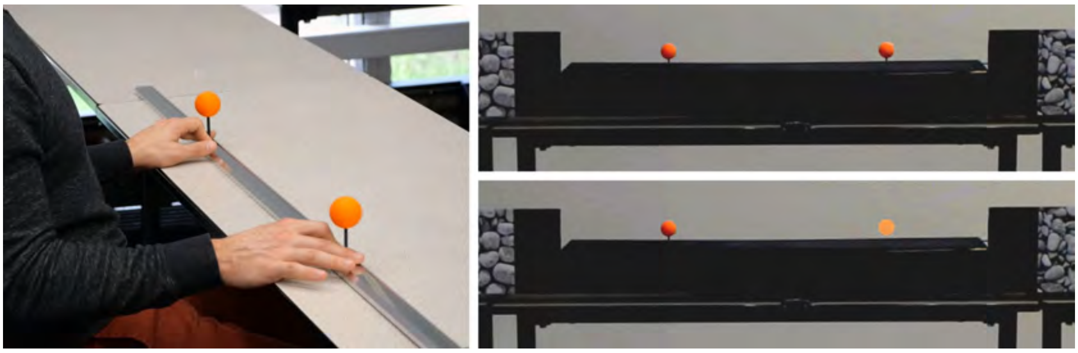


Figure 3: Exocentric distance perception in AR. Left, response protocol, the user had to match the distance between the spheres with the perceived exocentric distance. Right, experimental stimuli, real-real condition (top) and real-virtual condition (bottom).

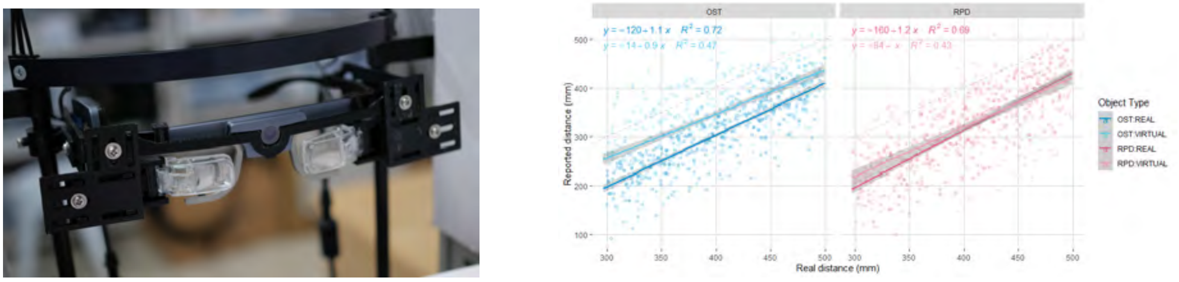


Figure 4: Egocentric distance perception. Left, binocular RPD device used, composed of two QD Laser RETISSA Displays. Each display module had a resolution of 1024 × 600 pixels and an approximate horizontal FoV of 20°. Right, regression analysis of the egocentric distance estimates. The plot shows that in the RPD condition (red) the distance estimates were closer to the real estimates compared to the OST condition (blue).

In this particular experiment, experiment we assessed how the exocentric distance estimates between two objects presented in a frontoparallel plane were influenced by the nature of stimuli (Real/Real, Real/Virtual, Virtual/Virtual). Concisely, the results showed that participants had an average of 20% overestimation in all conditions, and surprisingly the overestimation was higher for the real/real condition (4%).

In a different study [5], we explored the role that the mismatch between vergence and accommodation cues could have in egocentric distance estimates, which has been appointed a potential source of distance perception biases in virtual and augmented reality [10]. In this study, we compared a fixed-focus optical-see through device, Epson Moverio BT-30C AR glasses, with an infinite-focus optical-see through device, custom binocular retinal projector display (see Figure 4, left). During the study, participants had to provide distance estimates within the personal space (between 30cm and 50cm) of objects placed in front of them. The results showed that distance estimates with the retinal projection device (RPD) were significantly closer to real estimates (see Figure 4, right). Although it remains unclear the particular reason that generated this result, it shows that an infinite-focus device provided virtual distance estimates that better resembled real estimates. This work will potentially pave the way to numerous studies exploring the usages of binocular RPD in augmented reality applications.

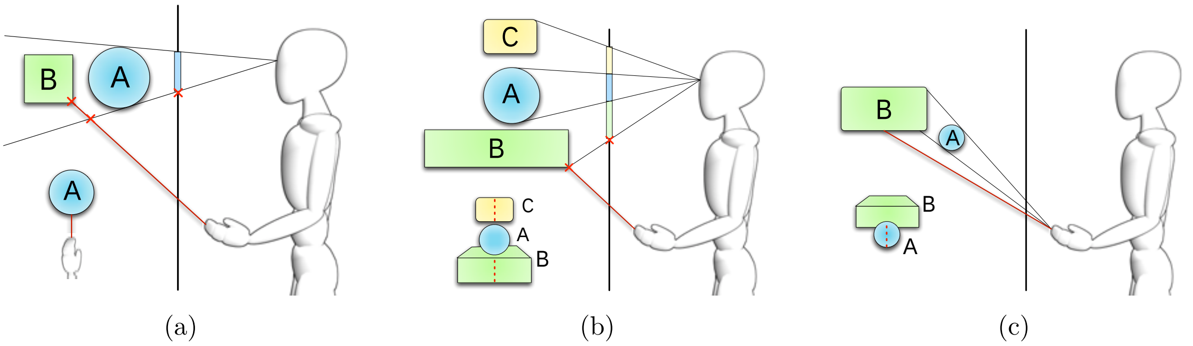
1.2 Eye-Hand Visibility Mismatch

In addition to absolute and relative distance perception, having a precise knowledge on the spatial distribution of virtual/real environment could also play a key role during the interaction process. Have you ever said to someone “Look at that!” while pointing with your index finger and realized that that person is not able to guess what are you pointing at? A similar issue can be found when selecting virtual objects using raycasting-based methods. Object selection techniques are a key component in any virtual reality application [11], and in particular, raycasting-based methods, which consider a virtual ray, cast from the user’s hand, the tool to designate virtual objects, are widely used techniques. In the paper “Overcoming Eye-Hand Visibility Mismatch in 3D Pointing Selection” [12] we discussed perceptual issues raised in object selection tasks when using raycasting-based methods, which arise due to the fact that the users’ viewpoint differs from the origin of the selection ray (i.e. the hand).

Lets consider a virtual object S. S is visible if at least one point in its boundary surface (∂S) is visible from the user’s viewpoint and is selectable if at least one point is visible from the user’s hand viewpoint. Let ν_E be the set of visible objects and ν_H the set of selectable objects. The set of objects $\nu_E \cup \nu_H$ can be decomposed into three disjoint sets: $\nu_E \cap \nu_H$, $\nu_E - \nu_H$ and $\nu_H - \nu_E$.

Let us first discuss the behavior of raycasting selection for objects in $\nu_H - \nu_E$. This correspond to objects which are hidden to the user’s eyes but are reachable from a ray emanating from the user’s hand (see Figure 5a). Object B is occluded from the eye position but not from the hand position. Therefore it is currently selectable even though it is not visible. This phenomenon is particularly distracting to the user. The user might think that the currently selected object is A, as there is an apparent intersection of the ray with object A (the screen projection of the last visible point of the ray is on the silhouette of the screen projection of A). In the absence of additional feedback, if the user triggers the selection confirmation at this point, the hidden object B would be erroneously selected.

Let us now examine the selection behavior for objects in $\nu_E - \nu_H$. This corresponds to objects which are visible but which are not reachable from a ray emanating from the hand (see Figure 5b). Object A is visible from the eye position but it is completely obscured from the hand position. Therefore, although being visible, object A cannot be selected while preserving the hand position. Unlike the above case, the screen projection of endpoint of the virtual ray Q appears over the projection of the currently selected object B, so there is no room for misinterpretation of the selected object. However, the user would be unable to select object A unless he changes the origin of the pointing tool. If the user only moves the ray upwards, increasing the elevation of the ray trying to bring Q to the screen projection of A, Q seen



(a) (b) (c)

Figure 5: Eye-hand visibility mismatch issues. (a) The user can selected an object which is hidden by another object. The last visible point on the ray, is projected over the screen projection of the occluding object, leading to misinterpretation: the ray appears to intersect object A, although the intersected object is behind. (b) A visible object cannot be selected because it cannot be reached by a ray emanating from the user's hand. The dotted line shows the path followed by the ray-scene intersection as seen on the screen, which skips object A. (c) Object A is visible from both E and H, but no point on its boundary is simultaneously visible from E and H.

would jump from object B to object C. The discontinuous path followed by Q on the screen is shown in Figure 5b. Note that Q skips the regions occupied by the screen projection of A. Some users might perceive this unexpected effect as an anomalous behavior, once they realize that the object is not accessible, they have to move their hand to an unoccluded position.

In the last group $\nu_E \cap \nu_H$, a more accurate approach to predict potential selection problems is to consider the solid angles $\Omega_E(S)$ and $\Omega_H(S)$. An object S is a potentially difficult target whenever $\Omega_E(S)$ or $\Omega_H(S)$ is below a threshold. But, as Figure 5c depicts, this approach is still inaccurate, as an object S with large $\Omega_E(S)$ and $\Omega_H(S)$ can still be difficult to select. Now, both $\Omega_E(A)$ and $\Omega_H(A)$ are large, but object A is still difficult to select because no point in the boundary of A is simultaneously visible from E and H. As a consequence, the user can intersect object A with the ray, but the intersection point is hidden by object B, keeping the user from having visual feedback. Therefore, a more accurate measure of the difficulty/accuracy required to select an object S must be defined in terms of its simultaneous visibility. Given an object S, we define $\Omega_{E \times H}(S)$ as the solid angle subtended by the points of S which are simultaneously visible from E and H.

In addition to the theoretical implications that the eye-hand visibility mismatch could generate, we also explored its potential practical effects in two different user studies [12, 13]. Both studies showed that virtual objects exhibiting higher eye-hand visibility mismatch were more difficult to select. In the second study, we also explored how the eye-hand visibility mismatch could disrupt Fitts' law predictions [14, 15]. As expected, objects exhibiting a high eye-hand visibility mismatch do not comply with Fitts' law estimations. Figure 6 shows in the left the results from a selection task in which all virtual objects were not exhibiting

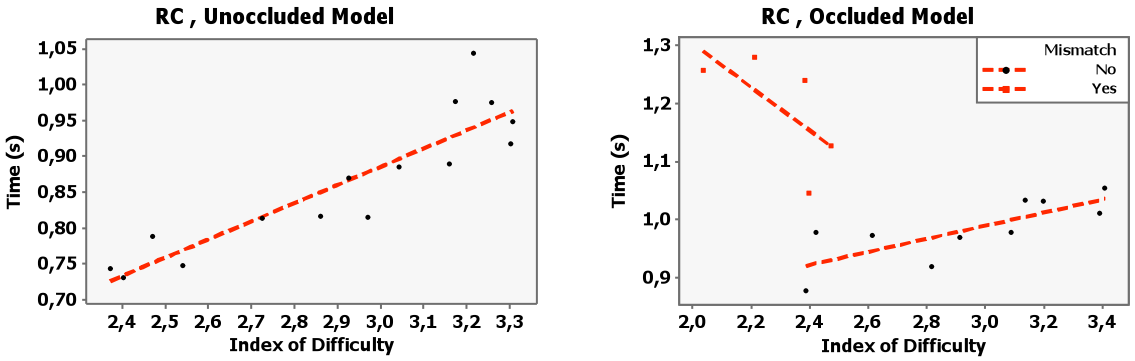


Figure 6: Scatterplot of mean object selection time vs index of difficulty for two selection tasks. (Left) All virtual objects did not exhibit eye-hand visibility mismatch ($r = 0.898$). (Right) A number of objects exhibited eye-hand visibility mismatch ($r = 0.755$, ignoring the objects with mismatch).

the eye-hand visibility mismatch, while the plot from the right shows that virtual objects exhibiting a eye-hand visibility mismatch required higher selection times compared to their theoretical index of difficulty level. The increase of the selection time can be easily explained as the ballistic motions become inefficient to reach the desired target, which require additional corrective motions in order to reach the target.

These results encouraged us to propose selection methods based on raycasting which would overcome the highlighted issues. Chapter 4 presents two selection techniques, which we proposed in order to overcome the eye-hand visibility mismatch. Furthermore, this work was also the initial idea which crystalized with a collaboration with the team of Pr. Bernd Froehlich at the Weimar University (Germany), which extended the analysis to multi-user interaction [16, 17].

1.3 Conclusion

The works presented in this Chapter highlighted perceptual issues that are found in VR/AR systems, which could alter the interaction process. First, we detailed a number of experiments focusing on the better understanding on how users perceive distances in VR and AR. The results support most of the works exploring distance perception in VR and AR studies: distance perception is distorted. However, the results of the AR experiments, also showed that distance perception in the real world is also biased. In a more practical way, the eye-hand visibility mismatch study showed how misperceptions of the layout of the virtual environment could hinder the selection process. The results showed that task difficulty increased when selecting objects suffering from the eye-hand visibility mismatch as participants had more difficulty on planning their selection motions due to a limited understanding of the environment.

Chapter 2

Perceptual Illusions in the Loop: Pseudo-Haptic Feedback

Perception is a complex process, which integrates the different sensory modalities to enable us to understand our surroundings. As evoked in the previous chapter, vision is an important sensory modality to understand the spatial layout of the environment. However, when interacting with real objects, haptic feedback enriches the interaction process by providing additional information (e.g. stiffness, roughness). Although haptic devices are rapidly evolving, they still present several limitations in terms of the haptic feedback that can render and bulkiness [18]. In order to overcome these limitations, our research has focused on the generation of visuo-haptic illusions capable of eliciting haptic properties, mainly on the physicalization of 2D interfaces (see Section 2.1) and avatars (see Section 2.2). In particular, we followed a pseudo-haptic approach [19] in which the user's actions are visually-modulated in order to elicit haptic properties. Pseudo-haptic feedback exploits the dominance of visual perception over haptic perception in situations where the haptic feedback is of low-quality or it is nonexistent. For example, in the context of stiffness, visual dominance allows the simulation of a wider amount of stiffness levels by the combination of haptic and visual feedback [20] and even without haptic feedback, stiffness can be visually discerned [21].

2.1 Eliciting Haptic Properties on 2D Images

When interacting with images using a computer mouse of a touch screen, the image and the cursor can be augmented in order to elicit haptic properties. A first example is the concept of "Elastic Images" [22], in which we proposed a pseudo-haptic approach to elicit local stiffness information in 2D images. In a nutshell, when the user clicks on an Elastic Image, the area surrounding the mouse cursor is deformed simulating that the user is pushing the image. The effect of pressing an image towards the screen was achieved by displacing the pixels towards the mouse cursor. Once the user clicks (i.e. presses) an image, a smooth animation at pixel

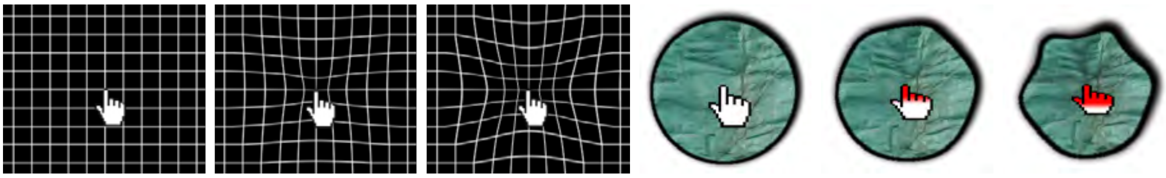


Figure 7: Elastic Image simulation. Left, animation steps for the proposed image-based deformation as the virtual pressure exerted by the user increases. Right, animation steps of the simulation of an Elastic Image with additional visual feedback.

level is generated. The deformation of the image was defined by the predefined stiffness of the image and the time the user keeps the mouse down. Furthermore, we proposed additional effects to reinforce the pseudo-haptic effect, such as procedural shadows, non-uniform deformations and the animation of the mouse cursor (see Figure 7). The psychophysical evaluation showed that users were able to recognize up to eight different stiffness values with our proposed method confirming that it provided a perceivable and exploitable sensation of elasticity. In this approach, the fact that users had the control of the interaction, and were not just passively looking at the image while it was being deformed, clearly reinforced the effect.

Inspired by this work, we explored other effects that could be generated while exploring 2D images, but this time on touch surfaces [23]. In touch surfaces, the user's finger is co-located with the 2D content, which creates an additional challenge. Precisely, the user's finger can occlude the content and its motions cannot be adapted as it is done with classical pseudo-haptic approaches [24]. Our solution, named "Touchy" aimed at enhancing touchscreen interactions without the need of any mechanical actuator, through a variety of pseudo-haptics effects (see Figure 8). When the user touches the screen, a cursor appears under the finger and follows it as it strokes the screen, before disappearing on release. As the finger hovers an area with haptic content, the cursor's motion and shape are altered in order to express the relevant haptic properties. For instance, the cursor might vibrate according to roughness, deform according to stickiness or slipperiness, or change its size according to stiffness. We showcased seven different pseudo-haptic effects that address five haptic properties: stiffness, (fine) roughness, reliefs, slipperiness, and stickiness. These five haptic properties are related to the four "tactile primary colors" [25] and can be organized along the corresponding perceptual dimensions: hardness, fine roughness, macro roughness, and friction. Two different user studies allowed to calibrate the different effects and to investigate which precision sensations were evoked with the different proposed effects. Overall, the results were mitigated, and while some effects elicited the desired sensations, some cross-effects were found due to the diffuse boundary among the different explored sensations. For example, effects designed to elicit fine roughness elicited mainly macro roughness sensations. Nevertheless, the use of a proxy object, the cursor, enabled the generation of different sensations by just altering its control mapping and shape.

Figure 8: Touchy is a co-localized cursor for tactile displays that deforms and/or moves to evoke a variety of haptic properties, covering four different perceptual dimensions: fine roughness, macro roughness, hardness, and friction.

2.2 Towards Pseudo-Haptic Avatars

In avatar-based interactions, the most common scenario is to map users actions directly to the avatar actions in order to enable the control of the avatar. However, such configuration, unless using haptic devices, results on avatar actions that are perceived as effortless. In the paper “Toward Pseudo-Haptic Avatars: Modifying the Visual Animation of Self-Avatar Can Simulate the Perception of Weight Lifting” [26], we explored how the mapping between the user’s gestures and the self-avatar animations could be altered in order to generate a change of the perceived effort. In addition to the simulation of surface properties, pseudo-haptic feedback has been also used for the simulation of forces, being the modulation of object weight using control/display adaptation the most common approach [27]. Thus, we explored how pseudo-haptic feedback could be used to change the perceived effort of the avatar actions when the avatar had to perform a weightlifting task using different weights. The user controlled the avatar by also performing a weightlifting task, but without weight (see Figure 9). In particular, we explored three different adaptations of the avatar control according to the virtual effort delivered: 1) changing the control/display mapping between the user’s and the avatar gestures, 2) selecting from a motion capture database a motion which matched the weight lifted, and 3) adapting the posture of the avatar (upper-body forward inclination).

For the first method, the principle was similar to classical pseudo-haptic feedback approach where the visual motion of a manipulated virtual object is amplified or reduced when compared to the motion of the user. Thus, in order the lift a heavier object, the user had to perform wider motions and vice-versa. In the second method, the elevation of the user’s wrist was used to play a pre-recorded animation of another subject which performed the same task. Using a full-body tracking system, we pre-recorded subjects lifting dumbbells with different weights. This approach aimed to reproduce subtle motions that arise when performing such a task. Finally, the third method consisted in modifying artificially the inclination of the avatar in order to simulate different lifting efforts. As observed in motion capture data, subjects had the tendency to lean more when the dumbbell was heavier. This visual feedback aimed to show

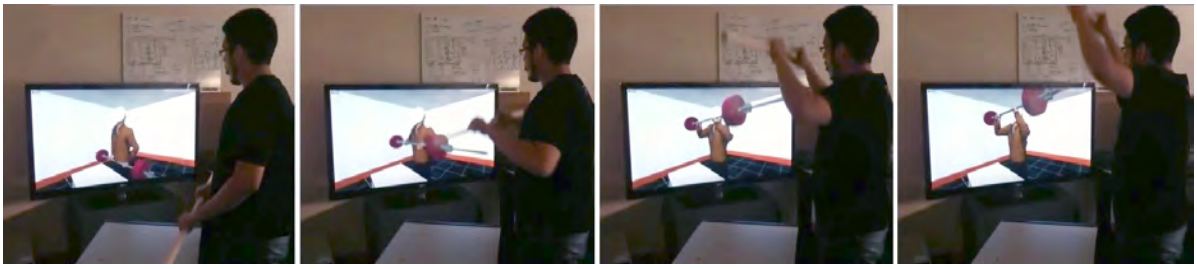


Figure 9: Virtual weightlifting task. The animation of the avatar is controlled by the user's actions and it is modulated according to the weight of the virtual dumbbell.

the effort via the avatar posture. We associated the lifting effort to the angle of inclination of the avatar.

In order to assess the different methods, we performed a user study in which participants had to sort different dumbbells relying solely on the visual feedback provided by the avatar. Overall, the results of the study showed that participants were able to rank the different weights of dumbbells by exclusively relying on the avatar animation. Yet, the "traditional" approach was the one that enabled better discrimination, followed by the avatar inclination and the pre-recorded animations.

2.3 Conclusion

The different works presented in this chapter highlight the potential usages of pseudo-haptic feedback when simulating tactile and kinesthetic feedback, the two dimensions of haptic feedback. While Touchy mainly focused on the simulation of tactile properties, the elastic images and the pseudo-haptic avatar approaches focused on the simulation of kinesthetic feedback. Yet, there are still a number of unanswered questions. While most of the studies showed that participants were able to discriminate the visual feedback, it still remains unclear whether the actual haptic perception is elicited and which is the association between the haptic and visual feedback made. One potential solution might be to enhance low-quality haptic feedback devices [28] to easily elicit the desired haptic feedback. In such scenarios, as haptic feedback is provided, the association between haptic and visual feedback is more straightforward. Moreover, all the conducted works focused on the simulation of isolated haptic properties. The question still remains whether and how different pseudo-haptic modalities can be combined and still enable the elicitation of haptic properties. Finally, in the context of avatar-based interactions, the work on the pseudo-haptic avatar raised the awareness of the relationship between the user and their avatar. In the conducted experiment, the user could see his/her avatar in a TV screen resembling a mirror, yet, as we will see in the following chapter, there are a wide range of factors that can influence such relationship.

Chapter 3

Perception of the Self: The Avatar in the Loop

When immersed in a virtual reality environment using a head-mounted display (HMD), we are unable to see our own body. Although we have the sense of proprioception than allows us to know where our body parts are located, for some tasks it becomes relevant to provide the user with a virtual body, hereinafter referred as "avatar". However, when embodied in a virtual avatar, one can wonder how the user perceives his/her avatar with respect to his/her own real body. Although this question might seem anecdotal, the avatar plays a key role in the

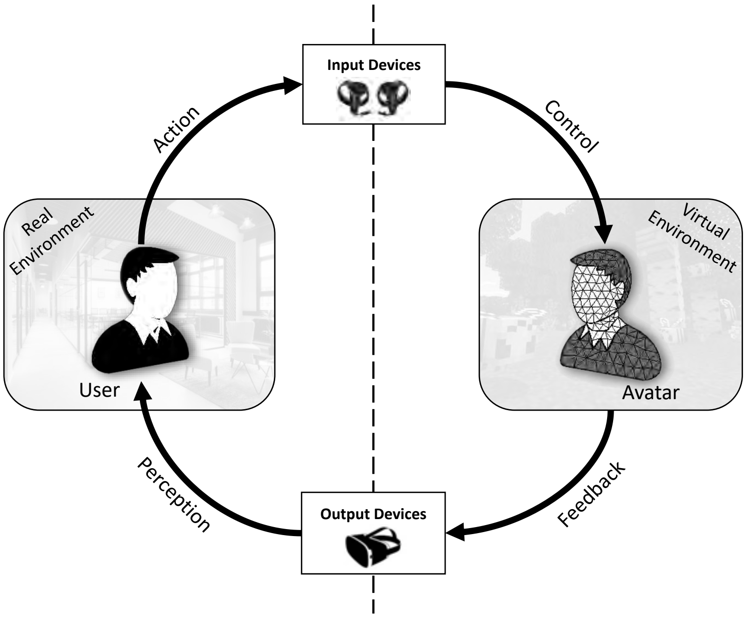


Figure 10: The avatar in the perception-action loop. The avatar, as it becomes the representation of the user in the virtual environment, provides feedback of the user's actions and can determines the actions that the user can perform.

interaction process, as user actions are driven by its perception (see Figure 10). For example, a number of works have shown that distance [29] or size [30] estimates can be influenced by the avatar appearance. The impact of the avatar has been linked with how well the user "appropriates" his/her virtual body. However, the assessment of such appropriation is complex as it is mainly subjective (i.e. a *qualia*), normally assessed using questionnaires [31]. Our work has been strongly inspired on virtual embodiment [32], which provided the following definition: *"the sense of embodiment (SoE) toward a body B as the sense that emerges when B's properties are processed as if they were the properties of one's own biological body"*. In this Chapter, I present the different research works focused on understanding the role of the avatar in the interaction process and to better understand the factors influencing virtual embodiment.

3.1 Factors Influencing Virtual Embodiment

Many studies have tried to better understand how users perceive their avatar in VR by evaluating their Sense of Embodiment (SoE). More precisely, they focused on three subcomponents of the SoE [32]: the Sense of Self-Location, the Sense of Ownership and the Sense of Agency. From those researches emerged different "factors of influence" towards these three subcomponents, e.g. the avatar's appearance [33] or the user's point of view [34].

In a recent work, we explored the relative preference for these three subcomponents, exploring how users prioritize them [35]. To cope with the difficulty of handling the numerous "factors of influence", we proposed to use a subjective matching method [36], which was been used for a presence study, to assess the relative preference of each factor. In particular, we considered three factors: the avatar appearance, the user's control of the avatar, and the user's point of view. Moreover, participants had to perform four different tasks (see Figure 11). In the main experimental task, participants had to match a given "high" SoE avatar configuration (realistic avatar, full-body motion capture, first-person point of view), starting by a "low" SoE configuration (minimal avatar, no control, third-person point of view), by increasing iteratively the level of each factor. As participants increased one level at a time, the order of their choices provided us insights with the relative preference of each factor for each task.

Figure 12 presents some of the results of the experiment in the walking and soccer task, represented as a Markov chain. Such representation enables the computation of the most likely path, which provides insights of the factors which were perceived by participants to contribute more to their sense of embodiment. In particular, the results showed that participants consistently increased first the control and the point of view levels before the appearance levels. This is particularly visible in their first choice (transition from the initial node {000}), in which the second and third digit is chosen with a higher probability (first digit being the appearance, the second the control and the third the point of view). Another interesting aspect observed in Figure 12, is that for some of the tasks the most probable sequence differs. For example, probabilities for reaching the nodes {001} and {010} were inversed between soccer and walk-

Figure 11: Tasks considered in the subjective matching experiment to assess the preference of different factors of influence of the sense of embodiment. From left to right, boxing, football, fitness and walking task.

ing tasks and in the soccer task participants the probability of reaching the node {021}. We hypothesized that these results are partially explained by the need of the user to be in control of the interaction, in which control and point of view play a more important role than appearance. Although a wide range of experiments in the literature have shown the impact of the different factors on the sense of embodiment, this experiment highlighted the potential impact between the interaction process and the different factors influencing the sense of embodiment.

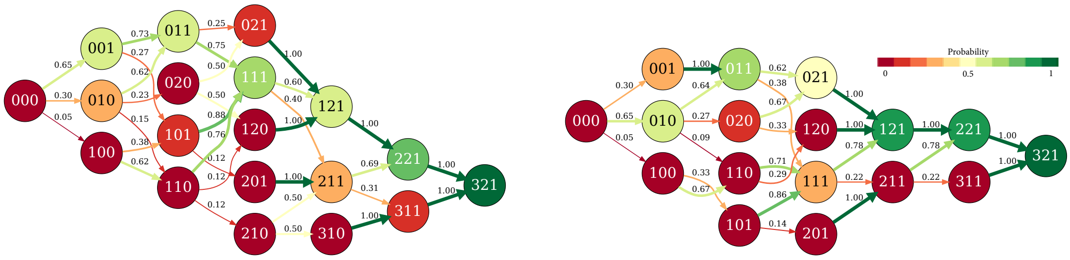


Figure 12: Markov chains representing the transition matrix probability for the walking (left) and soccer (right) tasks. The digits for each node represents the level of the appearance (0 to 3), control (0 to 2) and point of view (0 to 1). The color and the thickness of the edges represent the transition probability from a given node, while the color of the node encodes the probability that the node is reached.

Complementary to this experiment, which focused on factors directly linked with the relationship between the user and the avatar (avatar-related factors), other factors could also influence the sense of embodiment, such as external factors linked with the virtual environment (e.g. shared or collaborative experiences) and internal factors linked with the user's psychological or cognitive state. In this respect, we have started to explore other factors less tackled in the literature, such as the impact of collaborative experiences in the sense of embodiment [37, 38] (see Figure 13), the impact of personality traits and body awareness [39], haptics [40] or the ability to embody a virtual avatar with a different hand morphology [41]. Furthermore, in the particular context of the interaction loop, we further explored the role of the interaction in the sense of embodiment.

Figure 13: Collaborative experiences. (Left) Experimental task considered for the assessment of the sense of embodiment when sharing the virtual environment with another user [37]. (Right) Physical setup used in the virtual co-embodiment experiment [38] in which two users were embodied in the same avatar and share its control.

3.2 The Role of Interaction in Virtual Embodiment

The avatar, when interacting in virtual environments, provides feedback regarding all the actions that the user is performing. However, when interacting in virtual environments, the coherence between the avatar actions and the outcome of the interaction can play an important role. The concept of coherence has already been defined to the notion of presence [42]: *"the set of reasonable circumstances that can be demonstrated by the scenario without introducing unreasonable circumstances, where a reasonable circumstance is a state of affairs in a virtual scenario that is self-evident given prior knowledge"*. If we apply this concept of coherence to virtual embodiment, the avatar actions and interaction outcomes will be judged by the user's prior knowledge. Yet, as discussed in the previous section, most of the experiments regarding the sense of embodiment have mainly focused on the impact of the different factors of influence on the three subcomponents of the virtual environment. Yet, these experiments tend to consider passive or simple interactions. However, if we consider the user's prior knowledge on real world interactions, the interaction component can play a key role. Moreover, this process is dual, while the interaction process can influence the sense of embodiment, the sense of embodiment can also modulate the interaction process.

In the paper "The role of interaction on virtual embodiment: effects of virtual hand representation" [43], we explored the impact of having different virtual representations while performing simple pick-and-place tasks (see Figure 14). The results of this study showed that while the visual component have a direct impact on the sense of ownership, participants experienced a higher degree of ownership with the realistic representation. The sense of agency was also linked to the visual realism but inversely, participants experienced a higher degree of agency with the most iconic representation. The analysis of the objective and subjective responses showed that participants had higher "expectations" for the realistic representation in terms of interaction control compared to the iconic representations. Combined with the fact that the interaction just followed a plausible physically-based interaction and the lack of tactile

Figure 14: (Left) Visual hand conditions used in the pick-and-place task (right). The visual conditions considered from abstract to realistic representations while keeping the same pick-and-release mechanism.

feedback (participants could not feel the cube under their virtual hand), the coherence was "weaker" compared to an iconic representation in which prior knowledge might play a lower role.

While manipulating virtual objects is a particular technique to manipulate virtual objects, when interacting in virtual environments, other tasks can be considered [44], such as navigation tasks. A particular characteristic of navigation is that in real life requires full body motions (e.g. walking). However, in virtual reality environments, walking is discouraged due to the current limitations of tracking methods and more importantly to the available working space (typically several squared meters) [45]. Such limitations have been addressed by the design of virtual locomotion techniques such as walking-in-place or virtual steering methods in order to enable the user to navigate infinitely. However, in the context of a VR application in which a virtual avatar is provided, we wondered the impact of the coherence between the avatar animations and the virtual locomotion method [46]. In this particular work, we designed an experiment in which participants, using different virtual locomotion methods with different physical engagements, had to perform different tasks with different awareness of the virtual avatar (see Figure 15). As we considered a full-body tracking in order to control the avatar, the avatar animation of the avatar was consistent with walking when the user had to physically walk in the environment, moderately when the user was walking-in-place and unrelated when the user used just a joystick to navigate (e.g. the virtual floor slid under the virtual avatar feet). In a nutshell, the results showed that the sense of embodiment was not significantly impacted by the coherency between the avatar animation and the virtual locomotion method. The fact that participants kept direct control of the avatar all the time is our main explanation. However, we could observe that the quality of the tracking played a key role in the sense of embodiment. Finally, in terms of acceptability, participants tended to better accept the conditions in which the avatar motions were closer to the real walking animation.

Figure 15: Interrelation between avatar and virtual locomotion. From left to right, training task, corridor task, path-following task and avoidance task. Each participant performed the tasks with different locomotion techniques and with the absence or presence of an avatar.

3.3 Conclusion

This Chapter presented our works in the scope of understanding the role of the user's virtual representation on the sense of embodiment but also our first works exploring the interconnections between the interaction process and user's virtual representation. The first part mainly covered works aiming to better understand the different factors that influence the sense of embodiment, not only exploring the classical avatar-related factors (e.g. appearance, control, point of view), but also exploring less researched factors such as internal (e.g. personality traits) and environment-related (e.g. shared experiences). The different works highlight the challenges on modeling the impact of the different factors on the sense of embodiment. The second part aims to go further, but including the interaction process. How the interaction process alters the sense of embodiment? How the feeling of embodiment can influence the interaction process? These questions remain mostly unanswered and further research is still needed. To start bring the gap between the sense of embodiment and the interaction process in VR, in a recent paper [47], we presented a thorough analysis of the potential interaction between the sense of embodiment and 3D manipulation techniques in VR. Highlighting that the sense of embodiment is typically disregarded in interaction studies and that it should be considered as a user experience measure, such as presence and usability.

Conclusion

Through this first Part of the manuscript, I have detailed the different research works conducted in which the perceptual process is the main source of study. The main goal of the presented studies were to deepen the understanding of the user's perception during the interaction process in order to improve current mixed reality systems. The covered topics ranged from distance perception, how perceptual illusions can be generated, to how users perceive their virtual body while immersed in VR. Nevertheless, the impact of the different works presented differ with respect to their direct applicability on the interaction process. For example, distance perception studies in VR and AR enable to characterize the user's perception, while the eye-hand visibility mismatch highlights how a perceptual limitation can hinder eye-hand coordination during a remote pointing task. The characterization of this process is paramount to design interaction techniques better adapted to the user, which overcome/avoid such perceptual limitations. Finally, the studies on the role of the users' avatar and the sense of embodiment are still in its initial phases, and the conducted research has only scratched the surface. Interacting with an avatar in virtual reality is extremely challenging as a wide range of factors can be considered while using avatars (e.g. appearance and control methods), and the precise impact of the sense of embodiment on the interaction process still requires further research.

Another element worth mentioning is the impact of the technological-related factors (e.g. display technology) in the perceptual process. Over the years, VR and AR technology has greatly evolved and perceptual limitations have been improved (e.g. better resolution, better optical systems, improved rendering quality). This improvement of VR/AR technology had increased the fidelity of the perceived visual content, significantly reducing distance perception mismatches. Nevertheless, other biases are less dependent on the technology, such as the eye-hand visibility mismatch or the potential role of the users' virtual representation in the interaction process.

The next Part of the manuscript will cover research that explored how to apply results on perceptual and behavioral studies in order to propose interaction methods that better support the user.

Part II

Improving the Loop - Interaction in Virtual Environments

Introduction

> *"Design is concerned with how things work, how they are controlled, and the nature of the interaction between people and technology. When done well, the results are brilliant, pleasurable products. When done badly, the products are unusable, leading to great frustration and irritation. Or they might be usable, but force us to behave the way the product wishes rather than as we wish."*
>
> — Don Norman, *The Design of Everyday Things*

So far, I have discussed how perception drives our interactions and how when immersed in a virtual environment the perceptual limitations can alter and/or hinder interaction. In addition to our perception and interaction capabilities, we are bounded to the 3D user interface as it dictates the actions that can be performed in the virtual environment. A well designed 3DUI has to take into account both perceptual and motor limitations of the user. The feedback provided by the interface should take into account the perceptual limitations of users while the control of the interface should consider the motor limitations. In this part of the manuscript I will focus in two major interaction tasks: selection and navigation.

Object selection is a primary interaction task that any VR application has to provide. From selecting a 2D button in a 3D graphical user interface, selecting a 3D object, or defining a location in the environment to teleport to. The most common goal of an object selection technique is to determine a 3D position in the virtual environment. Although numerous object selection techniques exist [11], they are mainly characterized by the control scheme, i.e. how the user defines the 3D position, and the feedback, i.e. how the system displays the state of the interaction. In the case of Raycasting selection, the 3D position is defined by the intersection of the selection ray with the virtual environment in which the ray has five degrees of freedom (3DoF for the origin and 2 DoF for the orientation). The selection ray is controlled typically with a 6 DoF device and user performance can be modeled using Fitts' law [13]. Regarding the feedback, the ray and the intersection point should be displayed in order to provide information regarding the interaction state to the user. However, the perceptual limitations of users while immersed in virtual environments can degrade corrective movements. As described in Section 4, numerous improvements in Raycasting selection can be provided to better support the user in both the control and the visual feedback. One of these limitations is the Eye-Hand visibility mismatch already discussed in Section 1.2. In this respect, Section 4.1 describes two selection techniques which aimed at minimizing its negative effects by providing improved control and feedback mechanisms.

Moreover, navigation is another primary interaction task; users should be able to freely explore the virtual environment. While real walking has been acknowledged to be the most ecological approach to navigate in a VE as it better matches real locomotion tasks, it also requires a large physical workspace that is generally not available in most VR setups. Thus, since the beginning of VR systems, alternative navigation techniques have been explored in order to enable the user to navigate infinitely regardless of the size of the physical workspace. As such, virtual steering techniques remain the most used techniques as they require a minimal physical workspace, thus, being suitable for most applications. However, virtual steering techniques presents a major perceptual conflict: users can experience visually induced motion (i.e. vection) while their vestibular sense does not perceive any motion. Although the implications of the conflict between the visual and vestibular senses are yet to be fully understood, our research has focused on improving virtual steering techniques considering two orthogonal approaches understanding. First, on how users perceive motion during virtual steering 5.1 and second, on studying the impact of the steering technique on the user behavior 5.2. First, while virtually navigating, motion is mainly perceived due to the optical flow, i.e. the motion of the projected environment on the retina. We have proposed real time models to estimate the perceived motion and adjust the virtual navigation parameters accordingly. In addition to ease the user control, if the perceived motion remains constant, accelerations are less perceived, and considering that the vestibular sense is mainly sensible to accelerations, the visio-vestibular conflict should be reduced. Nevertheless, as previously stated, the implications are still unclear and additional works will be required to better understand the phenomenon.

Finally, human locomotion is a complex task involving motor and cognitive controls. For example, studies in the field of neuroscience have shown the importance of the head to control locomotion, acting as an inertial platform and a frame of reference to help the coordination of body segments. As such, while walking along a curved path a top-down reorientation strategy is observed: the gaze anticipates the future direction of the movement, followed by the head and then the shoulders. We have explored whether such reorientation behavior also happens while virtually navigating (see Section 5.2), and the potential impact that the navigation control law can have on other behavioral measures, such as the stereotypy of the trajectories or the unintentional positional drift that can happen while virtually navigating. These series of studies illustrate how the knowledge coming from the biomechanics research field can be taken into account while evaluating and designing navigation techniques.

Overview

The following chapters describe the work we have conducted in order to better understand and improve 3D user interfaces, and in particular object selection and navigation tasks. Chapter 4 describes works on 3D object selection, regarding the impact of perceptual limitations and how they can be overcome. Then, Chapter 5 describes the works conducted in order to assess motion perception while navigating in virtual worlds and the impact of virtual steering techniques on behavioral indicators.

The works presented in Chapter 4 are works done during my PhD under the supervision of Carlos Andujar.

The works presented in Chapter 5 have been conducted in collaboration with (alphabetic order) Carlos Andujar, Gerd Bruder, **Hugo Brument** (PhD Student), Hannes Kaufmann, Morgan Maignant (Master Student), Maud Marchal, Anne-Hélène Olivier and Iana Podkosova.

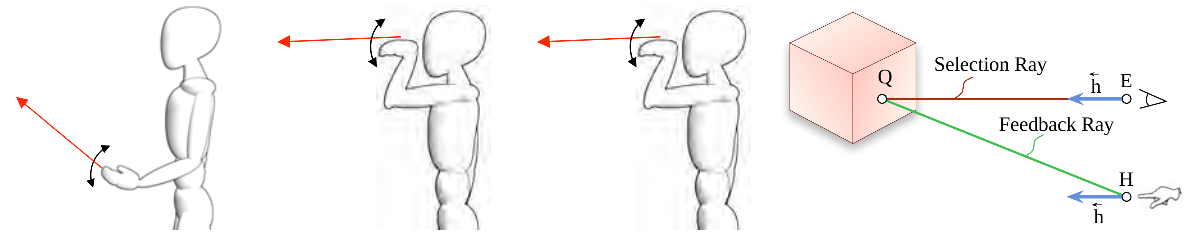
Chapter 4

Selecting Virtual Objects

Object selection is one of the fundamental tasks in a 3D user interface and the initial task for most common user interactions in a VE [11]. For example, manipulation tasks often depend on (and are preceded by) selection tasks. As a consequence, poorly designed selection techniques often have a significant negative impact on the overall user performance. From the plethora of the object selection methods, Raycasting is arguably the most used method due to its simplicity (e.g. virtual ray emanating from the user's dominant hand) and due to its ability to reach objects far away from the user. A number of previous research have explored variants of raycasting in order to increase selection accuracy, as the selection of small targets is a major challenge for selection tasks. In our research, we explored additional variants of raycasting selection either to minimize the eye-hand visibility mismatch (Section 4.1) and to support the selection of small targets for 2D graphical user interfaces in VEs (Section 4.2).

4.1 Overcoming the Eye-Hand Visibility Mismatch

In order to overcome the eye-hand visibility mismatch we proposed two different selection techniques, the Raycasting-from-the-Eye (RCE) [13] and the ViewFinder (VF) [48]. The two techniques share the same control scheme, although proposed two different visual feedback solutions. The goal of both methods was to provide an unambiguous selection gesture, avoiding the corrective actions that are required when there is eye-hand visibility mismatch. For this reason, to avoid any potential mismatch, the used selection ray emanates from the user's eye position (i.e. average position of the two eyes). Then, the orientation of the selection ray was directly determined by the dominant hand orientation (see Figure 16), which provides a better user comfort, as the user can keep his/her hand in a resting position. In a nutshell, the proposed control scheme combines the benefits of image-plane techniques (absence of visibility mismatch and continuity of the ray movement in screen-space) with the benefits of ray control through hand rotation (requiring less physical hand movement from the user). In this sense, it can be viewed as a hybrid technique between raycasting and image-plane techniques.



(a) (b) (c) (d)

Figure 16: In classic raycasting (a) the selection ray is cast from the user's hand, thus potentially suffering from eye-hand visibility mismatch. This problem persists unless users align their hand with the pointing direction (b), which results in an uncomfortable position. Our approach (c) uses a selection ray cast from the eye, whose direction is controlled with the hand orientation. Since the selection ray is insensitive to hand position, users can select objects as in (b) but with less physical effort. As the selection ray is not visible to the user, visual feedback is provided by drawing a ray cast from the hand to the first intersection of the selection ray with the scene (d).

However, a major challenge that was raised was the assessment of the visual feedback that will better support the selection process with such a method. As the ray emanates at the eye position, the selection ray will project into a single point. The naive solution was to display a 3D cursor at the intersection point between the ray and the virtual environment, yet, this resulted in visual discomfort due to diplopia effects, especially when selecting two close objects at different depth positions. One potential solution might require the use of one-eyed cursors, yet this solution was discarded as it might provide additional side effects, such as rivalry [48].

The first feedback solution implemented in the RCE method was to display a *feedback ray* defined by the hand position (H) and the current intersection point (Q) as depicted in Figure 16d. Although the depth position of the endpoint of the ray Q could change abruptly, the replacement of a 3D cursor by a ray notably alleviated the problem. The evaluation of the RCE approach showed that it was superior in terms of selection performance (time and accuracy), even for selecting objects which did not exhibited the eye-hand visibility mismatch. These results were explained by the continuity of the ray endpoint, which facilitated eye-hand coordination.

Nevertheless, the RCE feedback solution generated some undesired situations as the feedback ray could visually intersect other objects. This motivated us to find another potential visual feedback solution for the proposed selection control method: the ViewFiner. The key idea for the VF was to locally flatten potential targets in the vicinity of the pointing direction by projecting them onto a small virtual screen attached to the pointing direction itself (see Figure 17). We named this technique ViewFinder because the resulting effect is similar to looking a small part of the scene thought an LCD digital camera display. The viewfinder itself is rendered as a textured quad perpendicular to the pointing direction and 2D cursor on its center represents the pointing direction. Since the cursor and the objects displayed on

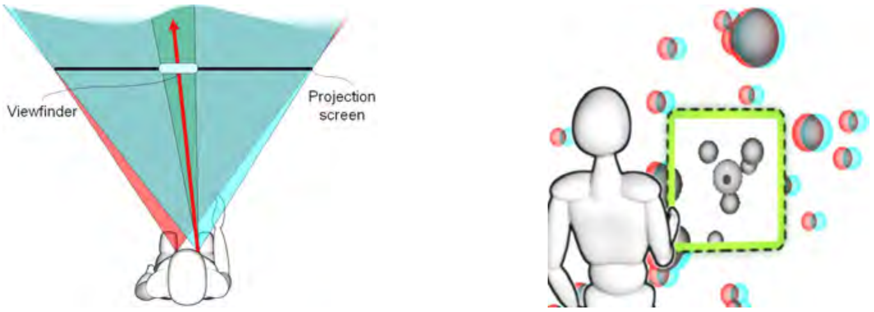


Figure 17: Viewfinder metaphor: Viewing pyramids of L/R eyes and the viewfinder (left). User view of the viewfinder (not to scale).

the viewfinder are drawn at fixed parallax, we avoid selection ambiguity problems that have discouraged image-plane techniques in stereoscopic displays.

To assess the advantages and disadvantages of the proposed methods, we conducted a user study, exploring impact of the selection control and visual feedback on selection performance. The results presented in [48], showed that the raycasting-from-the eye control resulted in the fastest and the most accurate method. Surprisingly, when combining the RCE control with a 2D cursor, the results were comparable to the results obtained with the RCE and VF feedback, yet, in conditions in which participants had to select virtual objects exhibiting great depth disparities, the cursor approaches were more prone to errors and less appreciated by the participants.

4.2 Interaction with 2D UI Elements

When interacting with Graphical User Interfaces (GUIs), selection techniques determine the available interactions. GUIs based on the WIMP (Windows, Icons, Menus and Pointer) paradigm are still a standard the facto for virtual reality applications. However, if not well designed, two-dimensional windows in 3D environments can include small widgets which can be difficult to select and manipulate. In comparison to the works presented in the previous Section, the selection of 2D elements laid out in a 3D plane are less prone to perceptual issues as all the elements are placed at similar depths and the eye-hand visibility mismatch is not relevant. In this scenario, selection performance is mainly constrained by the motor skills of the user. For example, if using raycasting, small rotations of the wrist sweep out large arcs at the end of the selection ray. Hand trembling and tracking errors are amplified with increasing distance. Accurate selection is also compromised by the hand instability caused by the absence of constraints on the hand movements (lack of physical support for manipulation). As a result, users attempting to select small widgets with raycasting have to make a considerable effort to stabilize their wrist.

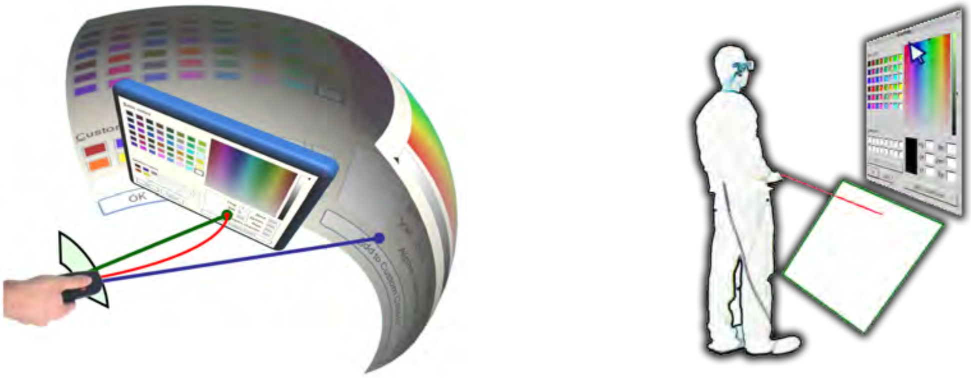


Figure 18: Left, the interaction and visual space of the Friction surfaces technique [49]. The interaction space (blue ray) is increased with respect to the visual space (green ray). Only the feedback ray (red) is displayed. Right, the interaction and visual space of the Virtual pads technique [51]. While the visual space remains unaltered, the interaction space is decoupled by the Virtual Pad, which can be customized (size and position) by the user.

In order to alleviate such limitations, we explored different alternatives to improve the user's precision and comfort while interacting with 2D GUIs embedded in virtual reality environments. First, we explored how precision could be increased while interacting with 2D GUIs. In that respect, we proposed the Friction Surfaces technique [49], which is a modified version of the raycasting technique which adapts the Control Display (CD) ratio according to the size and position of the virtual window. Inspired by 2D pointing facilitation techniques which modulate the Control-Display (CD) ratio in order to increase precision, we proposed to apply a CD gain higher than 1 to the hand rotations. However, while the manipulation of the CD ratio for mouse based interactions is transparent for the user, as it is a non-collocated interaction, for raycasting this will be easily noticeable for the user, as the input (user's hand) and the feedback (virtual ray) are co-located. To alleviate this effect, the modified CD ratio was only applied for each individual 2D window, and a bended ray was displayed to reinforce the effect (see Figure 18, left). The computation of the CD ratio enforced that the entire surface of the window was mapped to a fixed 90 angle. This ensured that the placement of the window (e.g. depth position) did not alter the user's precision while interacting with the window [50]. Overall, the results of the experiments showed that users rapidly integrated the change of the CD ratio as no performance drop was observed when interacting with large widgets. Moreover, selection time and accuracy was significantly improved with small widgets, which was the main objective.

Following a different approach, we further proposed the Virtual Pad metaphor [51]. The design rationale for the Virtual Pad was to decouple the visual (where the 2D window is displayed) and the motor space (where the user points to). This decoupling provided two interesting features, first, the placement of the Virtual Pad could be placed at a more comfortable position (see Figure 18, right). Second, the size of the Virtual Pad could be adjusted to alter the CD ratio, the larger the virtual pad, the higher the CD ratio. The virtual pad could be manually configured by the users to support their need. The evaluation of the Virtual Pad showed that users achieved similar performance than using direct methods, but with a significant increase in comfort.

All the experiments and the conception of the selection techniques were supported by a GUI framework that we proposed [52] to easily embed and author 2D GUIs in virtual reality leveraging existing 2D frameworks such as Qt. The proposed framework enabled to display 2D GUIs as 2D textured surfaces that could be manipulated using raycasting selection methods.

4.3 Conclusion

This Chapter has summarized several of my early works conducted in the scope of my PhD entitled “Pointing Facilitation Techniques for 3D Object Selection in Virtual Environments”. The works highlighted in this Chapter focus on how intrinsic limitations of humans can be mitigated when designing interaction techniques adapted to them. For example, the RayCasting-from-the-Eye and the ViewFinder techniques show how the control scheme can be adapted to minimize a lack of awareness of the spatial relationships between the virtual environment and the user. However, it also highlighted the need of adapted visual feedback techniques, which provide unambiguous feedback and are compliant with stereoscopic devices. Moreover, other interaction criteria such as accuracy and comfort should be taken also into account. Although conceived for the interaction of 2D elements, the Friction Surfaces and the Virtual Pad techniques highlighted how control methods conceived to increase user accuracy could also be beneficial both for the performance and comfort point of view. However, the “perfect” selection technique does not exist yet. Furthermore, the diversity of applications and layouts that can be found in virtual reality applications might require the co-existence of different selection methods. This is potentially a major drawback of the majority of studies considering selection methods, the difficulty to assess the generalization of current selection methods to the plethora of VR applications in the wild.

Chapter 5

Navigating in Virtual Environments

Navigation is another fundamental task in virtual reality applications, as it allows the user to explore the virtual environment. In contrast to selection, navigation is a task that we perform in a daily basis, either using our body (i.e. walking, running) or other transportation methods (e.g. bike, car). In comparison with real live navigation, virtual reality is not bounded by the physical limits of the real world, in VR we can fly, instantaneously teleport to far away locations and navigate through multi-scale environments (e.g. from the atom scale to the planetary level). However, the most common navigation technique in real life, locomotion, is typically unfeasible due to technical constraints. Although a wide range of methods to overcome these limitations exist, research is still being conducted to provide navigation/locomotion techniques which better support the user.

In this context, my research has been mainly focused on the study of virtual steering techniques, either to provide seamless navigation while virtually navigating through multi-scale environment (Section 5.1 or to propose novel evaluation methodologies (Section 5.2). Virtual steering techniques present a number of notable drawbacks, such as the lack of vestibular cues and potentially generate higher levels of cybersickness as methods requiring the user to physically walk. However, virtual steering techniques are the most flexible methods, as it can be used even with small tracking setups, generate less physical fatigue and are most suited for non-realistic (e.g. multi-scale) navigation.

5.1 Optical Flow-Based Multi-Scale Navigation

With the actual complexity of virtual environments, and in particular with multi-scale virtual environments, navigation interfaces have 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 [53]. Zoom-in and zoom-out operations, when interactively navigating in the virtual environment, require a real-time adaptation of additional navigation parameters, such as the range of the navigation speed [54] or the stereo display parameters [55] in order to ensure its usability. However, if not adapted properly, suboptimal navigation experiences can be generated. For example, fast camera motions might induce motion sickness [56], slow motions might decrease user performance and engagement, excessive parallax might generate diplopia effects or not enough parallax might decrease depth perception.

Our objective was to provide automatic navigation methods capable of adjusting the navigation speed and the scale factor of the virtual environment in order to limit the potential adverse effects previously listed. The main approach was to analyze in real-time the camera motions, in order to identify such potential undesired effects, with a particular analysis of the perceived navigation speed. In an initial work that did not focus in real-time navigation [57], we proposed an optimization method to automatically adjust the camera speed in order to generate trajectories which limited the amount of optical flow, avoided abrupt speed oscillations, and considered the relevance of the displayed content (salient features and new content). In this work, in addition to the motion constraints, we also considered aesthetic components. Its main purpose was to optimize the speed along a 3D trajectory to provide a pleasing and informative path, and in particular avoid strong acceleration and deceleration motions.

In a follow up work [58], we proposed a real-time version of the optimization method only focusing on the perceived navigation speed, disregarding the aesthetic and novelty aspects. In this version, the user was able to freely navigate in 3D virtual environments using a desktop based system. The estimation of the perceived speed was computed from (1) the state of the system (the relative distance between the 3D environment and the user, and the current navigation speed) and (2) the change of optical flow between frames. The heuristic employed operated in image space and only took into account the visible part of the scene. All computations were done directly in the GPU and stored in 2D textures (see Figure 19) for optimized processing:

- **Time-To-Collision (TTC) map**. The TTC map was meant to gather information about the relative distance between the user and the virtual environment, proving an estimate of potential collisions with the environment. For each pixel, we computed the relative speed and estimated the time when a potential collision might occur. The computation only took into account the minimum distance between each pixel and the viewing plane (z axis).
- **Optical Flow (OF) map**. The OF map stored the optical flow for each pixel and it was meant to gather information relative to the perceived motion of the user. The optical flow was estimated considering the current navigation speed (extrapolation). Using this estimation the rotational optical flow was easily discarded. Rotational optical flow can be introduced by head movements or camera rotations, which could distort the adaptation algorithm.

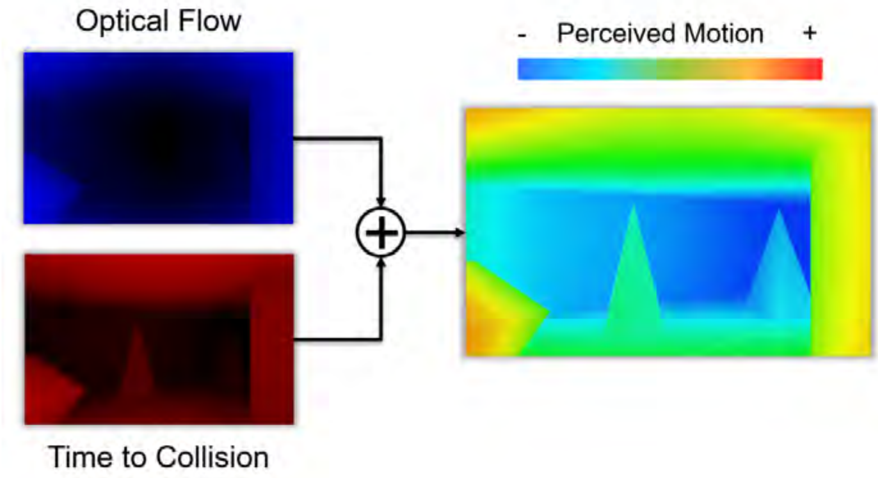


Figure 19: Information used to compute the perceived user's speed during a virtual steering task. Two main textures were computed in real time, the optical flow map which encodes visual motion, and the time-to-collision map which encodes the risk of collision.

The data from the time-to-collision and optical flow maps (normalized, see [58] for details) was integrated to compute an estimation of the perceived navigation speed h (see Equation 5.1).

$$h = \sum_{i=0}^{n} \frac{TTC_i * (1 - OF_i) + OF_i}{n} \tag{5.1}$$

Using these computed data, the navigation speed was adapted by updating the current acceleration of ensuring that an optimal perceived speed (h_{opt}) was achieved without abrupt speed changes. This approach was evaluated in [58] showing that it efficiently and effectively increased user control while navigating in multi-scale environments. However, this approach was not directly exploitable in virtual reality as the scale of the virtual environment was not adjusted.

In a follow up work, we proposed GiAnt [59]. The concept of GiAnt was 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 [60]. 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. $s < 1\,\mathrm{cm/s}$) or fast speeds (e.g. $s > 1\,\mathrm{km/s}$) which will strongly deviate from human walking speeds ($s \approx 1.4\,\mathrm{m/s}$ [61]). In these situations, we can assume that the level of scale is not adapted, 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 followed a hybrid approach which adapts both the navigation speed and the scale factor of the virtual environment. Both adaptations, speed and scale, were computed according to the

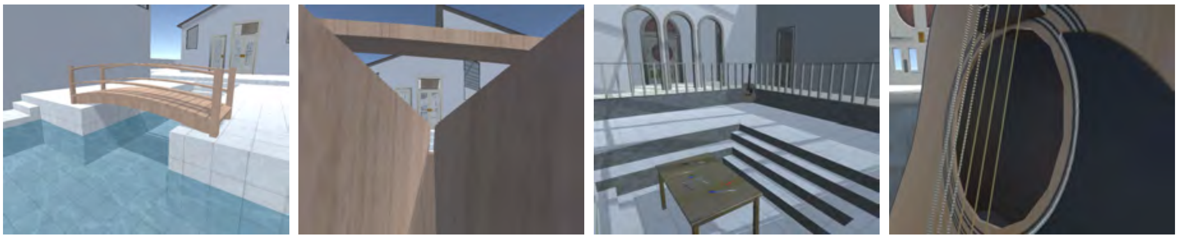


Figure 20: Multi-scale navigation sequence requiring not only the adaptation of the camera speed but also the stereoscopic rendering parameters. The proposed multi-scale navigation technique (GiAnt) ensured that the navigation speed and the scale factor of the virtual environment were adjusted in order to ensure a comfortable navigation experience.

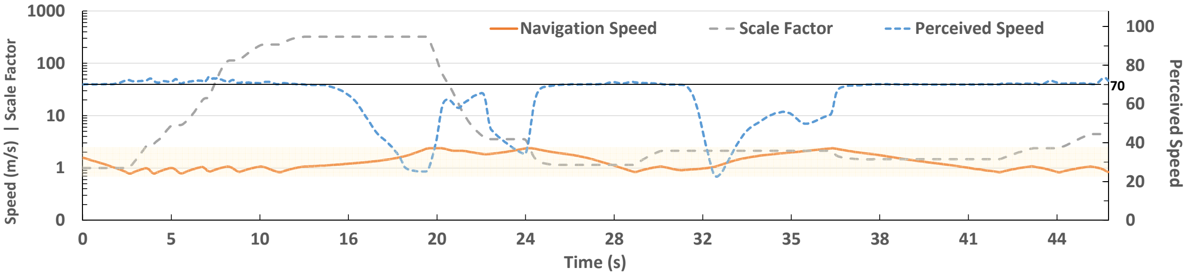


Figure 21: Evolution of the navigation speed, the perceived navigation speed and the scale factor during the navigation depicted in Figure 20. When approaching to small cavities, the environment is scaled up to ensure the comfort speed range and avoid diplopia effects. Around $t = 3\,\text{s}$ the user stats to approach the bridge which generated a strong change of the level of scale. Around $t = 42\,\text{s}$ the user approaches to the guitar, which also triggers a scale up of the environment.

user's inputs (steering direction and user viewpoint) and to the user's perceived navigation speed. Furthermore, a speed correction step was introduced to ensure that the speed remains at a "human-level" scale and the scale correction aimed to minimize diplopia.

Figures 20 and 21 provide an example of the behavior of GiAnt. Figure 21 shows the evolution of the navigation speed, the level of scale and the perceived navigation speed while performing the navigation depicted in Figure 21. The navigation required to approach small cavities in the environment. The first, at $t = 5\,\text{s}$, consisted in a close up view of the wooden bridge (first two images of Figure 20). 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 $t = 24\,\text{s}$ and $t = 41\,\text{s}$) the user remains at a similar level of scale. Then, the user approaches the second target, the guitar (last two images of Figure 20). When approaching the guitar, a new up scale of the environment is triggered allowing the user to get inside ($t = 45\,\text{s}$). During all the navigation, we observe that the navigation speed remains inside the comfort zone (orange rectangle, $s = [0{:}7;2{:}8]$ m/s.). 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 $t = 16\,\mathrm{s}$ and $t = 32\,\mathrm{s}$) 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.

Experimental results (see [59] for details) showed that GiAnt allowed for an significant more efficient navigation while significantly minimizing the required changes of the level of scale. These results indicated that GiAnt provided an efficient solution for multi-scale navigation in virtual environments. Furthermore, with GiAnt, users could navigate in arbitrary virtual environments which did not match real life constraints. Yet, most applications VR consider virtual environment that are the human scale, and thus, locomotion techniques which aim at matching real locomotion are typically preferred.

5.2 Studying the Users' Behavior During Virtual Locomotion

Human locomotion is a complex task involving motor and cognitive controls. For example, studies in the field of Neuroscience have shown the importance of the head to control locomotion, acting as an inertial platform and a frame of reference to help the coordination of body segments [62]. Through a series of studies, we have explored and analyzed the potential impact of virtual locomotion on the user's behavior, mainly on the reorientation strategy of body segments (eye, head, shoulders, torso) [63], the unintentional positional drift that can arise during virtual navigations [64] and the impact of rotational and translational motions on the detection of head rotation gains [65]. These studies aimed at better characterize the user's behavior while virtual navigating, propose novel evaluation metrics and eventually navigation techniques better adapted to the user.

In a first study [63], we explored the impact of the virtual locomotion techniques in the top-down reorientation strategy of the body while performing curve trajectories. While performing a curved path in a Real Environment (RE), a top-down reorientation strategy is consistently observed [66, 67, 68]: the gaze anticipates the future direction of the movement, followed by the head and then the shoulders. Our main hypothesis was that navigation techniques with higher fidelity to real walking would lead to synergies between body segments closer to the ones observed during real locomotion tasks. In particular, we chose five navigation techniques considering its relevance and the provided degree of user control. In particular, we focused on real walking, three virtual steering techniques (head, hand and torso steering) and one passive navigation technique. Regarding the control law, we proposed a new control law based on the relationship between speed and curvature [69, 70] which is found in real walking conditions. During a continuous trajectory, the instantaneous speed varies according to the local radius of the curvature following a power law:

$$R(t) = \frac{(\dot{x}^2 + \dot{y}^2)^{\frac{3}{2}}}{\dot{x}.\ddot{y} - \ddot{x}.\dot{y}} \tag{5.2}$$

where $\dot{x}$, $\dot{y}$, $\ddot{x}$ and $\ddot{y}$ are respectively the first and second derivatives of x and y coordinates of the user's position in the environment. In the case of walking trajectories, the speed of locomotion is proportional to the cubic root of the radius of curvature [69, 70]:

$$S(t) = K.R(t)^{\frac{1}{3}} \tag{5.3}$$

where $S(t)$ is the horizontal speed at time t, K is a gain speed coefficient and $R(t)$ is the radius of local curvature of the trajectory at time t. Intuitively, the higher the curvature the lowest the speed.

In order to assess the different techniques and its impact on the top-down reorientation behavior, we designed an experimental task inspired from paradigms used in real locomotion studies for assessing gaze anticipation during curved trajectories in RE [67, 71] in which participants perform lemniscate trajectories (see Figure 22, left).

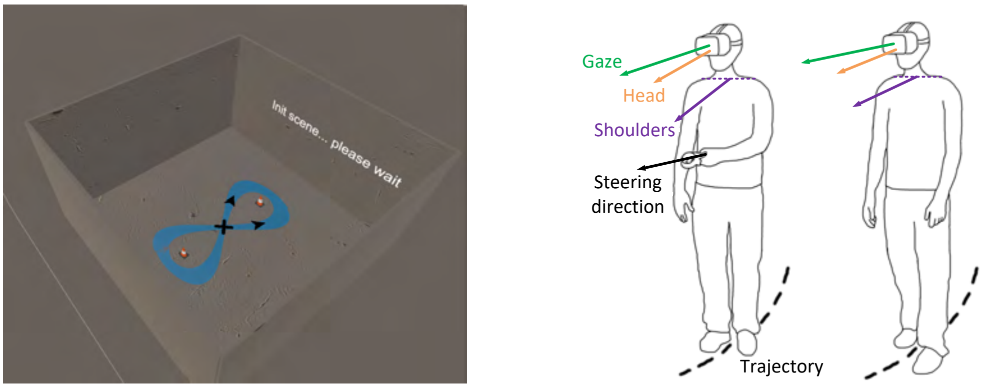


Figure 22: (Left) Virtual environment depicting the lemniscate trajectory that users had to perform. During the actual task all the virtual reference elements were hidden (trajectory, props). (Right) Depiction of the different body segment direction considered in our analysis.

The experimental results showed that the navigation technique had an effect on the user's temporal anticipation when considering the gaze and the head-heading directions and the three gaze-head, head-shoulders and gaze-shoulders body segments, as depicted in Figure 23. A similar ordered top-down sequence of reorientation of the gaze, head and shoulders during curved trajectories between walking in REs and in VEs (for all the evaluated techniques) was observed. However, the anticipation mechanism was significantly higher for the walking condition compared to the others. Torso steering resembled the most the top-down reorientation when walking in RE. For the walking and torso steering conditions, we could notice the coupling between the heading and the shoulders angles (both curves are overlapped) whereas for

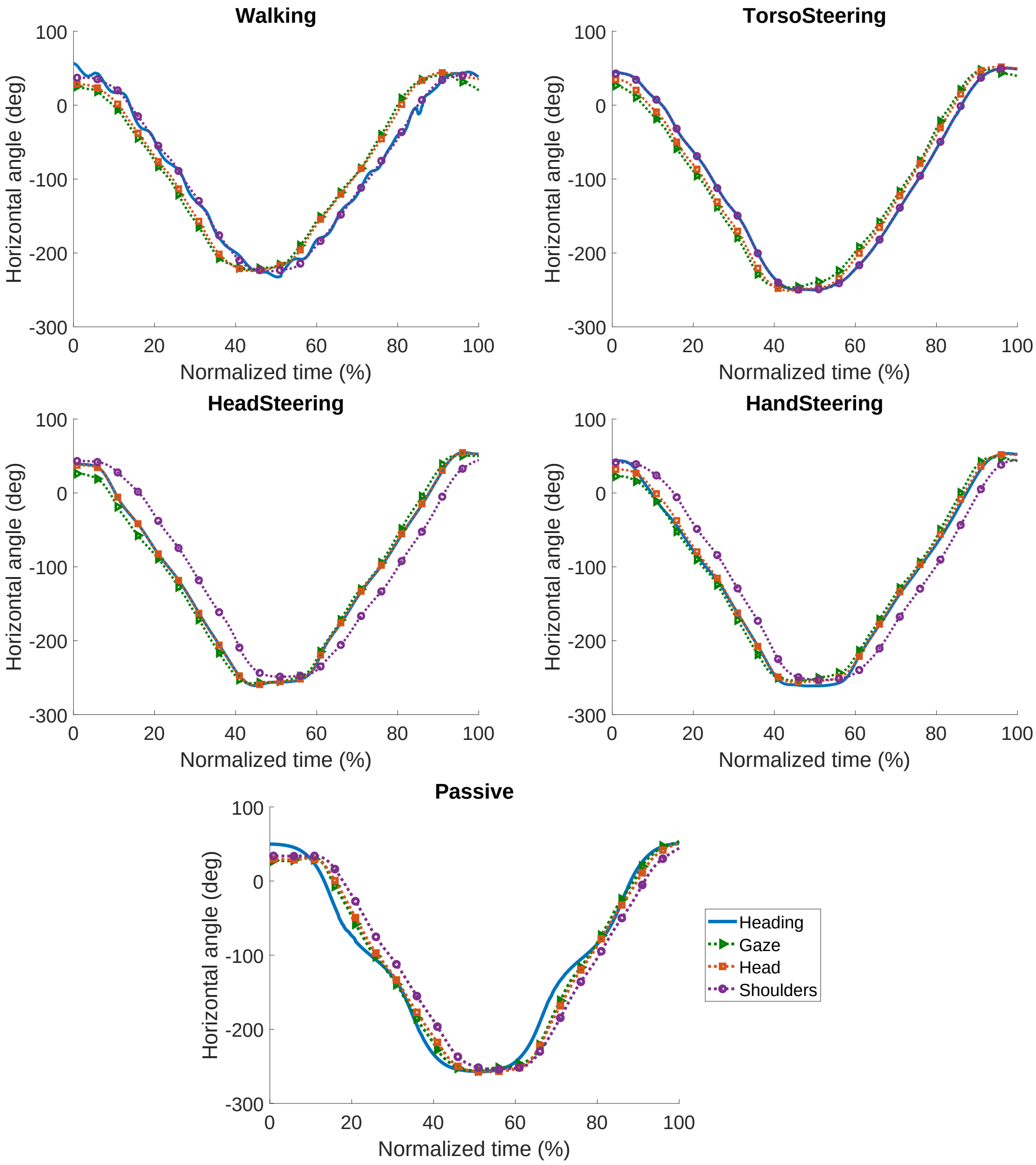


Figure 23: Typical evolution of the average angle of gaze (green), head (orange), shoulders (violet) and heading (blue) while performing the trajectory per condition.

hand steering and head steering, the coupling was between the heading and the head angles. Passive condition showed no particular coupling between horizontal angle of body segments. A body segment anticipation could be noticed if the evolution of its horizontal angle is shifted to the left with respect to the heading angle evolution.

One hypothesis was that steering direction provided by the torso would be more natural and participants would anticipate more easily thanks to the decoupling between the head direction and the heading direction which allowed the head not being involved in the steering direction.

Although further work is required, considering that the top-down reorientation behavior is a locomotion invariant, we argue that techniques which generate a similar behavior as real walking should be prioritized.

In addition to the input mechanism (e.g. head, torso steering), other factors could influence the user's behavior while virtually navigating. One of this additional factors is the control law, which models how the user translation and rotation viewpoint are updated in the VE considering the state of the system. The control law takes as input the state of the system, which encompasses the current navigation state and the user's input, but it can also consider other parameters such as the curvature of the current trajectory [63], the scale of the environment [54], the viewpoint quality [72] or the user's perceived motion [59]. To enable speed changes, control laws determine the navigation speed given the user's inputs (discrete, e.g. a button press or continuous, e.g. a joystick) and a transfer function that defines the mapping between each input data and transforms it into an output value [73].

We conducted an experiment [74] to investigate the effects of control laws on users' behavior and preferences during a navigation task with curved trajectories. We wanted to study the way users perform different types of turns (by varying the trajectory's curvature) while using virtual steering techniques, in particular, we assessed speed, acceleration and rotation profile while performing the task (see Figure 24). To this end, we designed a slalom task (sinusoidal trajectory) as it involves a continuous navigation with several turns, therefore potentially inducing speed and orientation adaptations. Three control laws were tested, a constant law in which navigation speed was always constant, a linear control law in which the user could modulate linearly the tangential speed using a hand-held controller (the joystick of an HTC Vive controller) and the control law presented previously, which takes into account the curvature of the trajectory. For all three cases, a torso steering navigation was considered as it showed the closest anticipation behavior as real walking.

The main analysis focused of the user speed profile during turns, which was defined a as the trajectory between two inflexion points of the sinusoidal-like trajectory. Figure 24 shows the user profiles for the medium curvature trajectories. Similar profiles where found with small and high curvature. The statistical analysis performed showed an effect of the control law on these time-series during the turn. Post-hoc tests demonstrated that the linear speed, acceleration and angular speed were smaller for the adaptive control law than the constant or linear ones during the turn. However, no difference was observed between constant and linear. One of the main results of this experiment was the impact that the control law had on the turn strategy for each turn. As seen in Figure 24, the maximum speed was approximately 25% lower for the adaptive law. This decrease in rotation speed could be linked with the comfort of the user while performing the task, which was partially supported by lower effort scores for the adaptive technique. Although participants were able to perform the task with all techniques, and the adaptive law had the highest task completion times, we argue that the proposed adaptive law is better suited for the evaluated task as it allowed to generate as it generated similar linear speed

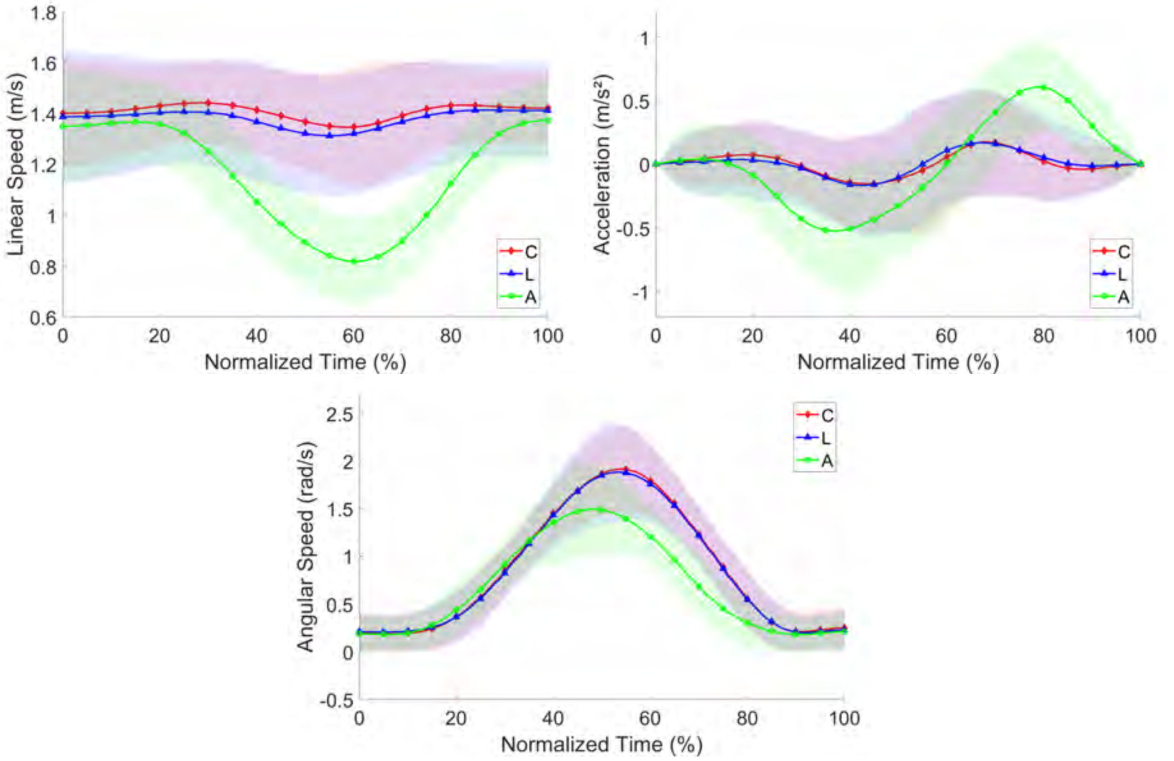


Figure 24: The averaged temporal evolution (mean and standard deviation) for the linear speed, the acceleration and the angular speed for each control law (constant (C) in red, linear (L) in blue and adaptive (A) in green) for the medium curvature trials. We can observe that there is an effect of the control law on the variables during most of the turn duration. The maximum curvature of the path was achieved at the 50% of the trajectory.

profiles than walking, allowing users to keep a consistent behavior even across trajectories with varying curvature. Yet, further research is required to validate these assumptions.

Another interesting observation of the experiment was that participants had the tendency to drift from their initial position in the real environment while virtually navigating (see Figure 25). This phenomenon, named unintentional positional drift (UPD), is problematic for prolonged VR sessions, as participants could reach the boundaries of their workspace even using virtual navigation techniques that do not require locomotion. While previous works have assessed UPD while using Walking-In-Place (WIP) techniques [75], little is known about UPD during steering navigation in VEs. The characterization of the UPD could lead to adapted techniques capable of reducing or compensating for this UPD. Thus, we worked for proposing an initial methodology for its characterization [64]. Based on the positional data gathered from the previous experiment (slalom task), we proposed two models (linear regression and Gaussian mixture) in order to encode the UPD based on the behavior of the users. The analysis performed showed that drift was consistent during turns, and the amplitude of the rotation was a promising predictor to model the UPD. However, the proposed model clearly had limitations

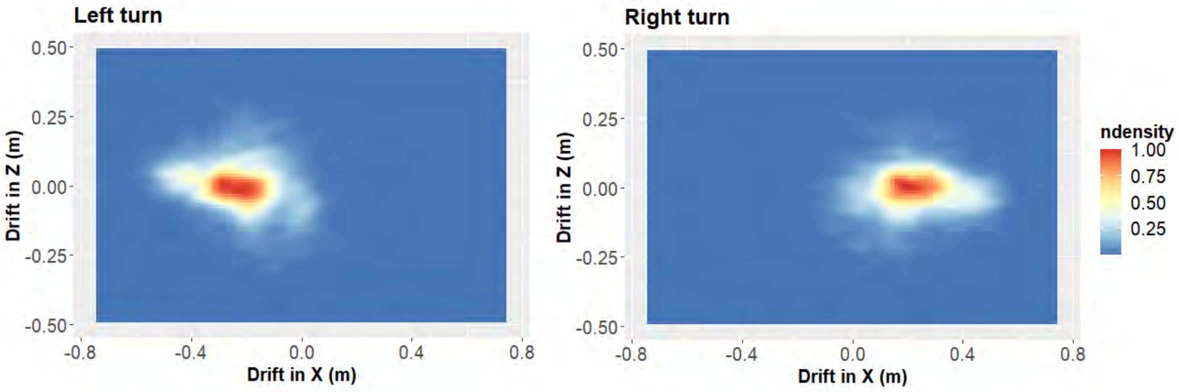


Figure 25: Density plots for the drift in the mediolateral (x) and anteroposterior (z) axis for the left and right turns.

as they were only considering a type of turn, which means that other temporal aspects of the rotation could also induce additional UPD. The long-term goal of this work is to propose UPD simulators capable of predicting the UPD during long sessions and test mitigation strategies to reduce the UPD. One promising approach being redirection techniques.

5.3 Conclusion

This Chapter has presented our main research works conducted in the context of virtual navigation in virtual environments. As our main goal is to propose navigation methods better adapted to the user, the different works presented were strongly grounded in perceptual and biomechanics research as they can drive the design process.

In Section 5.1, the proposed control law to enable multi-scale navigation is mainly grounded in visual perception as the main objective was to achieve a comfortable navigation along different levels of scale of a virtual environment. In particular, we considered optical-flow analysis which is the main information that the user has during virtual steering navigation in VR. The ability to keep an almost-constant optical flow during navigation enabled a comfortable navigation even with environments supporting changes on scale of several magnitudes. Proposing all-purpose navigation methods in multi-scale environments is challenging due to the potential heterogeneity among different virtual environments. For example, dramatic changes of the level of scale are hard to handle as the reactivity of the proposed method remains limited. Yet, these limitations could be overcome by well designed multi-scale environments. Further research directions could explore other input mechanisms as only head steering was considered. In addition, as it remains unclear the impact in terms of spatial awareness when changes on the level of scale are applied.

In contrast, Section 5.2 focused on the potential impact that navigation techniques could have in the user's behavior, specifically during locomotion tasks. Locomotion in real environment has been extensively studied in the biomechanics and neuroscience research fields, and in particular a number of locomotion invariants have been identified. Our objective was to assess whether these invariants where also preserved during virtual locomotion, and propose novel methodologies for the evaluation of virtual locomotion techniques. We believe that through this work we have mainly scratched the surface and a large number of follow up studies are still required. For example, the analysis of the top-down reorientation could be used as a new metric to evaluate locomotion techniques. Yet, the implications of reorientation strategies that do not match the ones observed during real walking remain uncertain. In addition, we believe that the analysis of the UPD could be considered while assessing navigation techniques, and that strategies to mitigate and reduce the UPD are worth exploring. Redirection techniques have been widely studied in the literature, but they have been applied in locomotion techniques which require the user to physically walk. We believe that redirection techniques for virtual steering techniques leveraging rotation gains could be a promising direction in the future.

Conclusion

Through this second Part of the manuscript, I have presented the main conducted works which aimed directly at improving interaction on virtual environments, in particular selection and navigation techniques. The presented work leverages action (motor, biomechanics) and perception (perceptual, neuroscience) research in order to design and evaluate novel interaction techniques. In addition to the contribution related to the new interaction methods, the conducted research also aimed to propose novel evaluation methodologies in order to better understand and model how humans interact in virtual environments.

Traditional 3DUI research has been strongly focused on performance and subjective analysis, which tends to overlook the subtleties of the interaction process and how the perception-action loop is unrolled. Additional assessment on user behavior while interacting is still needed, and it could be the cornerstone to enable the design of more adapted techniques and individual personalization could be envisioned. For example, perceptual thresholds of redirected walking techniques have been largely explored, but inter-personal variability is not negligible. The ability to predict the perceptual thresholds for a given individual, could enable personalized redirected walking controllers. In particular, we recently started to explore the potential role of anticipation and gaze behaviors when rotation gains are introduced [65], which could be assessed in the context of the UPD, to proposed redirection controllers which reduce/minimize the UPD.

The next Part of the manuscript will cover research which aimed to leverage the user's mental and cognitive state to the interaction process, with the long term goal of using that information for real time interaction and evaluation.

Part III

Enriching the Loop - Exploiting User's Cognitive and Mental State

Introduction

> *"Check the heart, check the, check the, is it the brain?"*
> *"No sign of cardiac analomy or unusual brain activity."*
> *"Okay, so I was poisoned?*
> *"My diagnosis is that you've experienced a severe anxiety attack."*
> *"Me?"*
>
> — Tony Stark and J.A.R.V.I.S., *Iron Man 3*

Interaction techniques leverage knowledge on user perception and user behavior in order to improve the interaction process. But, can we feed additional information to the loop? One potential solution relies on physiology computing, using real-time information from the current user's physiological state in order to enable and/or adapt the interaction. Changes in user physiology can reflect processes happening in the Central Nervous System (CNS) and in the Autonomic Nervous System (ANS), providing information regarding the user state. The most used measures of the ANS activity include measures of cardiovascular, electro-dermal, and respiratory activities, while measures of the CNS rely on brain imaging techniques such as ElectroEncephalography (EEG) or functional Magnetic Resonance Imaging (fMRI). However, physiology computing raises a wide range of research challenges, deciding which physiological data is relevant, identifying the state of the user from the user's physiological state and defining how to exploit this information to drive the interaction process. Moreover, physiological data is noisy, with strong user inter and intra-variability and dependent on the measurement equipment. Furthermore, changes in users' physiology can be due to numerous body and mental processes, which can generate similar physiological states.

In the context of the ANS, physiological data has been used to detect arousal, stress, comparing virtual and real situations, or measuring the human response when interacting with virtual characters. For example, in the "pit" experiment, the evidence strongly suggested that heart rate measured as beats per minute (bpm) increased when the stress was induced. In addition, a number of studies have focused on video games involving cardiac measurement as an index of valence, arousal, attention, cognitive effort, and stress. Yet, the heart and the circulatory system are regulated by a number of body processes, which are strongly driven by exercise. In our research, we have mainly focused in the monitoring of the mental workload. In particular, we have explored the potential influence of virtual reality (see Section 6.1) and how the monitoring of the mental workload can be exploited in training procedures (see Section 6.2).

Furthermore, the monitoring of the CNS provides additional insights with respect to the user state, as measures of the CNS can be directly localized at independent brain regions. Yet, the information that can be extracted highly depends on the brain imaging technique used. For instance, EEG data has a high temporal resolution but a low spatial resolution (a number of electrodes placed on the scalp surface). In contrast, fMRI has a low temporal resolution but a high spatial resolution (3D data). In our research, we have mainly focused on EEG as its temporal resolution is more adapted for real-time interaction and could be more easily coupled with VR equipment. In addition to the monitoring of the user's mental state (see Section 7.1), neurophysiological data, in the context of Brain Computer Interfaces (BCI), can be used as a direct input modality. Our research has been mainly focused on the use of Reactive BCIs, which monitor voluntary but non-spontaneous modulations of the brain activity. Precisely, we mainly focused on Steady State Visual Evoked Potentials (SSVEP) in which neurophysiological data is modulated using flickering stimuli (see Section 7.2). Nevertheless, other BCI applications have been explored, such as Active BCIs, which monitor voluntary and spontaneous modulations of the brain activity (Active BCIs) and Passive BCI in which non-voluntary, non-spontaneous modulations are monitored.

Overview

The following chapters describe the work we have conducted in order to assess the user cognitive and mental states using physiological and neurophysiological data and how it can be used to modulate or enable direct interactions. Chapter 6 focuses on the user's mental workload and how it can be used in the context of VR training scenarios. Then, Chapter 7 presents two studies with the objective to extract neurophysiological markers using EEG, and concludes with a study exploring the combination of BCIs and AR.

The works presented in Chapter 6 have been conducted in collaboration with (alphabetic order): Jean-Marc Diverrez, Anatole Lécuyer, **Tiffany Luong** (PhD student), Nicolas Martin and Anais Raison.

The works presented in Chapter 7 have been conducted in collaboration with (alphabetic order): Louis Albert (Master Student), Géry Casiez, Maria Duarte, **Andéol Évain** (PhD Student), Anatole Lécuyer, Catarina Lopes-Dias, Camille Jeunet, Gernot R. Müller-Putz, Reinhold Scherer and **Hakim Si-Mohammed** (PhD Student).

Chapter 6

Assessing the Users' Cognitive State

The measure of the user's cognitive state while interacting has a strong potential for the assessment of interactive systems and its adaptation in real time to support the users' cognitive resources. However, traditional assessment methods relying on questionnaires, such as the NASA-TLX, are administered at discrete intervals. These assessments provide a general overview of the state of the user during the entire exposure, or between pre-defined intervals. In order to increase the temporal precision, the use of physiological and behavioral data are promising candidates. However, the use of physiological data is challenging in terms of data processing and could be cumbersome for the user, and behavioral data is typically task dependent and can be difficult to generalize.

One particular cognitive state that is relevant for interaction and training purposes is the mental workload. Mental workload was defined by Wickens [76] as "the relation between the (quantitative) demand for resources imposed by a task and the ability to supply those resources by the operator". From an interaction standpoint, it is interesting to monitor the user's mental workload to assess the interaction tasks (e.g. which interaction modality reduces mental workload), and from a training standpoint, the induction of mental workload could be of interest to customize and adapt the training procedures (e.g. increase or decrease the users' mental workload to assess its performance). Our research, has mainly focused on mental workload, and addressed two main aspects. First, the work presented in Section 6.1 describes a study aiming to explore the potential impact of virtual reality on the users' mental workload, while the work presented in Section 6.2 discusses a methodology to design training procedures in which the user's mental workload can be measured and adapted in real-time.

6.1 Assessment of Mental Workload in Virtual Reality

A large amount of studies carried out in VR have considered tasks which required the involvement of motor processes, focusing for instance on the influence of VR training on task performance. However, there have been way less experiments targeting the user's subjective

Figure 26: Real (left) and virtual (right) environments used to explore the impact of virtual reality in mental workload.

mental effort in VR using a cognitive task. While task performance is greatly correlated to mental workload, it greatly depends on the user's motor skills and may not be representative of the user's mental effort if the task is not purely cognitive. Besides, depending on their engagement, some users can show a great performance on a task while involving great cognitive resources, yet, the opposite is also possible. In contrast, assessing the impact of a VR application on the user's mental effort in the long term could be more insightful.

As a first step, we wanted to characterize the potential impact of being immersed in VR to the user's mental workload. In this regard, we conducted a controlled user study [77] to explore the potential differences in the user's mental workload when users were immersed in a virtual environment using a HMD compared to a real environment (see Figure 26). In particular, an auditory N-back letter task was chosen to induce different levels of mental workload [78]. For each sequence of letters, the participants were instructed to react if (and only if) the current letter was the same as the N-th previous one and before the next letter presentation. The value of N determined the number of letters the participants had to remember, thus, the difficulty of the task. Three values of N were chosen: 1, 2, and 3 as they are the most common difficulty levels in experiments relying on the N-back working memory paradigm. We chose the 3-back task over the 0-back task since the latter is mostly used as a control condition, and may not elicit enough difference in mental workload with the 1-back task. In addition, in order to see if a basic task had an impact on the reported mental effort in both environments, participants were asked to perform the N-back task while sitting or while walking. While the performance of the N-back task is a good indicator of the user's mental workload, we also gathered additional objective measurements to explore the potential impact of other indicators, such as physiological (heart rate, skin conductance levels) and behavioral (reaction time, pressure exerted to the confirmation button).

Our main hypotheses concerned the effect of the environment on mental effort. First, the relative cumbersomeness of HMDs, the usual absence of user's virtual body, and other visual artifacts or mismatches with reality could actually require additional information processing from the user, which in turn might increase his/her mental effort. Thus, we hypothesized that being immersed in a VE through an HMD will influence mental effort [H1]. Second, since users are rarely static in VR the effect of basic and natural task such as walking may have an influence on the user's mental effort difference between the virtual and the real environments. Thus, our second hypothesis was that doing a cognitive task while walking will significantly increase the reported mental effort [H2].

Briefly, the analysis of all measurements revealed no significant effect of being immersed in the VE on the users' mental effort. In contrast, natural walking significantly increased the users' mental effort. Among behavioral and physiological measures, the response time was the best indicator of mental effort, followed by the exerted pressure, then, by the heart rate. However, caution should be taken when using physiological sensors in interactive environments, as their viability tend to greatly decrease with the involvement of movements. Furthermore, in this experiment we only considered simple features of physiological data, such as mean hear rate or the skin conductance level, which was insufficient to detect the subtleties of heart rate and skin conductance. These results encouraged us to further explore how mental workload assessment could be considered for the design of VR applications, and in particular training applications.

6.2 Mental Workload Assessment for VR Training

Several VR studies have explored the modulation of the user's mental workload during virtual reality experiences. These studies explored how task difficulty could be adapted based on mental workload indicators, such as task performance [79, 80] or physiological signals [81]. While such VR experiences considered single tasks, real training contexts generally imply complex environments with multiple tasks to perform in parallel. To cope with this limitation, we proposed a new methodological approach to consider mental workload in the design of multitask VR training scenarios [82].

The concept was to assess the mental workload for each potential combination of tasks, number of parallel tasks and their associated level of difficulty using a state machine. We assumed that the training environment is composed of a set of tasks, each having different levels of difficulty in which, at a given moment, a subset of these tasks are enabled. All the potential combinations between tasks and levels of difficulty define the set of possible tasks configurations, and thus the number of potential states in the state machine. Then, the transitions between states are defined based on the possible state changes among different tasks configurations. The final step is to determine experimentally the mental workload that is induced at each potential state. This step is challenging as it requires to assess the mental workload of trainees when performing the training routine for each state. Initially, we restrained the assessment of the mental workload through the use of self-reports, and in particular to the Instantaneous Self Assessment (ISA) scale. The ISA scale enables the user to rate the mental workload using five different ratings (under-utilized, relaxed, comfortable, high, excessive) and

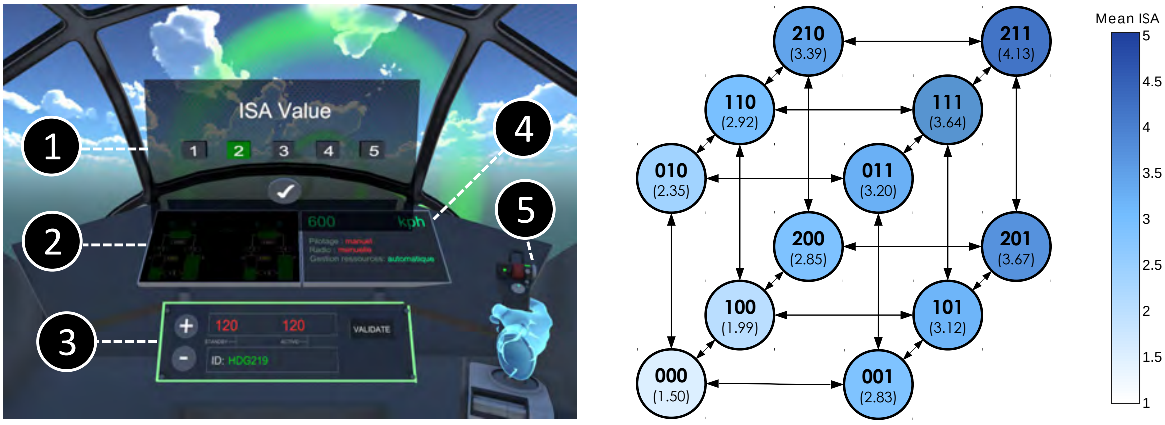


Figure 27: Left, virtual cockpit. (1) Instantaneous Self Assessment (ISA) interface. (2) Resources management task interface (deactivated); when activated, the interface lit up (with a green outline). (3) Communication task interface (activated); when deactivated, the interface lit off (no green outline). (4) Informative panel which gives information about which task is activated or not at the current time. (5) Virtual representation of the joystick used to pilot the aircraft and of the right hand. Right, state machine of the designed VR flight simulator. Each state is labelled using 3 digits. The first digit refers to the level of the piloting task (0-easy, 1-medium, 2-difficult), the second one to the level of the communication task (0-deactivated, 1-activated) and the third one to the level of the resources management task (0-deactivated, 1-activated). The mean ISA value is displayed in brackets.

has been especially used during air traffic control tasks [83]. Once the state machine is build and the mean mental workload is known for each state, the state machine can be used to define training scenarios capable of modulating the mental workload over time.

To illustrate the proposed approach, we designed a VR application based on a flight simulator. In a first experiment ($N = 38$), we collected mental workload data in all the potential tasks configurations (12 different configurations). The chosen application was designed based on a VR flight simulator while the considered tasks were inspired by the second version of the Multi-Attribute Task Battery (MATB-II), a computer-based application designed to induce and evaluate an operator's performance and workload developed by NASA. In particular, three tasks from the MATB-II were selected and adapted to a VR environment: a piloting task, a communication task, and a resources management task (see Figure 27, left). Each task had different levels of difficulty and users had to rate their mental workload using the ISA scale at intervals of 30 seconds.

The mental workload results of the first experiment are summarized in (see Figure 27, right) for each task configuration. The designed tasks configurations were successful into inducing different levels of mental workload to the users. Overall, self-reported mental workload increased as the number of parallel tasks increased and as the level of difficulty of each task increased. Although the real mental workload state of the users during the task cannot be quantified precisely, the coherency between the ISA scores and the theoretical mental workload

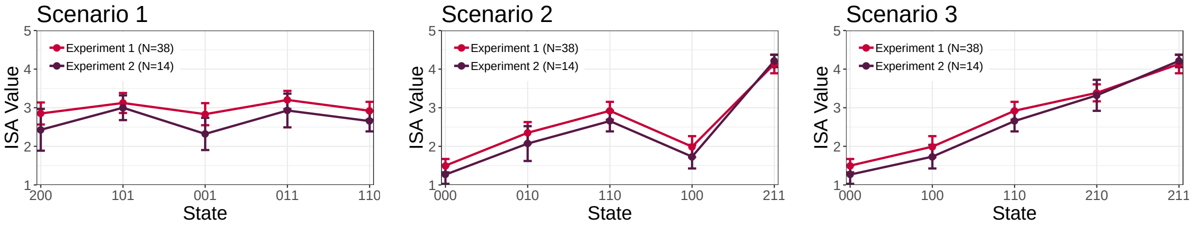


Figure 28: Comparison of the users' subjective mental workload (mean ISA value) for between 3 scenarios between the first and the second experiment. The first scenario was intended to induce a mental workload around $ISA = 3$ all the time. The second one aimed at inducing a low mental workload ($ISA <= 3$) followed by a sudden increase. Finally, the last scenario objective was to progressively increase mental workload.

that is achieved by each task (difficulty and combination) is a promising result. In summary, the results showed that the proposed methodology could be suitable for designing training protocols aiming to the induction of different levels of mental workload. These results motivated a second experiment ($N = 14$) with the purpose to use the created model to design scenarios based on training objectives, and to compare the mental workload results to the expected outcomes. Three different training scenarios were considered,

- **Scenario 1**: Induce a medium mental workload along the entire scenario.
- **Scenario 2**: Induce a low mental workload level first with a sudden increase at the end.
- **Scenario 3**: Induce a progressively increasing mental workload.

The ISA scores and the tasks performances followed the expected results for scenarios 1, 2, and 3, which were designed based on the first experiment measures (see Figure 28). First, even with a different pool of participants, the ISA scores were consistent among studies showing the replicability and potential generalization of the approach. In addition, we also wanted to explore whether participants were able to realize the "theoretical" profile of the mental workload (constant, increasing, sudden increase). However, some users were found to answer in unexpected ways compared to their ISA scores, even just after having completed the scenario. Our hypothesis was that users mainly remembered the first and the last states of the scenarios, which were the same in scenario 2 and 3. Therefore, subjective ratings have to be performed cautiously and delays of rating, considered carefully. The subjective mental workload rating throughout the scenarios was more accurate than the post-experiment self-report questions and the performance measures. In light of these results, we explored how we could exploit the proposed method using physiological data to obtain a direct measure of mental workload.

In particular, we proposed an all-in-one system for the monitoring the user's mental workload while immersed in VR using a head-mounted display [84]. The main features of our work, is that we proposed a custom hear-rate and electrodermal activity system embedded in an HMD (see Figure 29), and a series of machine learning methods to classify in real-time the

Figure 29: Experimental setup. Left, close-up of the Vive Pro Eye HMD with the additional hardware: (1) PPG sensor, (2), electrodes to assess the EDA, (3) electronic card, (4) 3D printed case. Right, overview of the VR setup in which the user is also equipped with the Shimmer3 GSR+ monitoring equipment.

user's mental workload. Due to the complexity of physiological data, machine learning (ML) methods are typically required in order to extract the users' mental workload. A particular challenge remains the training of the ML models as they require the gathering of labeled physiological data and specific induction protocols to ensure that different levels of mental workload are generated. Induction protocols tend to focus mainly on standardized tasks (e.g., N–back task) or relied on task-difficulty metrics to label mental workload data. In contrast, we used our flight simulator training application, capable of generating a wide range of mental workload states, in a large experimental campaign ($N = 75$) to record labeled physiological data. Unlike previous research exploring the recognition of MW in VR [85, 86, 87], the data were labeled using subjective responses.

From the data gathered, the trained models (random forests) were able to classify 4 levels of mental workload with an accuracy up to 65%. Yet, it should be noted that the misclassified levels were mostly contained in the adjacent classes, as shown in Figure 30. As such, the classification accuracy reached 95.5% for the best configuration when considering the classes that were adjacent to the predicted mental workload level. This highlights the good performance of our approach. In addition, we compared the recognition accuracy considering commercial-grade sensors and our all-in-one solution. The normalization method, as well as impact of the HR and EDA features and the combination of sensors were also tested in regards to the classification accuracy. First, normalizing the dataset features greatly improved the classification performance. As for the type of measures, ocular activity features were found to be especially relevant, followed by EDA, cardiac activity, and task performance features. Finally, a proof-of-concept system capable of recognizing in real-time the mental workload of the user was proposed. We envisioned the use of the system to automatically drive the state of the training procedure to match pre-defined mental workload goals.

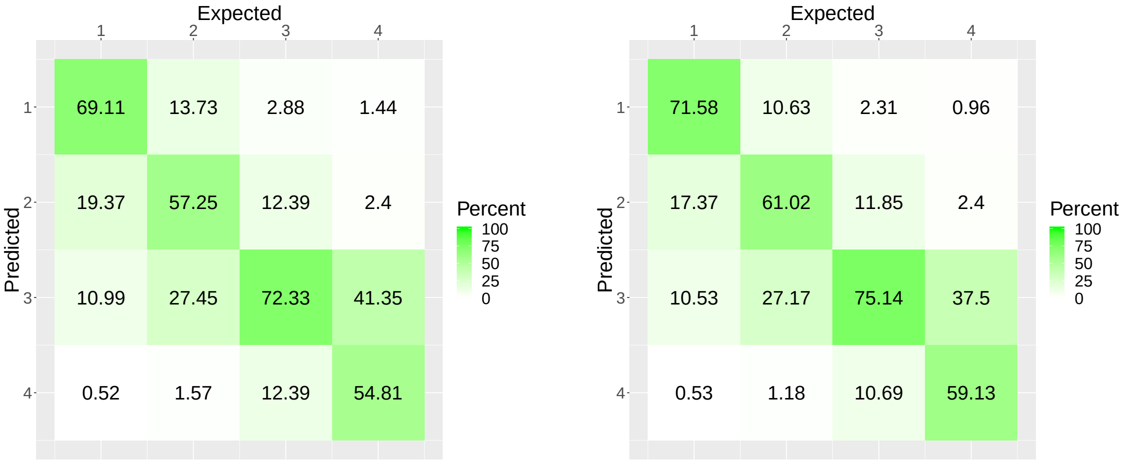


Figure 30: Confusion matrix, expected vs predicted classes, for (left) “commercial-grade” sensors and (right) “custom-grade” sensors.

6.3 Conclusion

This Chapter has presented our research work focusing on the exploitation of mental workload in virtual reality environments and in particular, for virtual training purposes. However, before being able to exploit mental workload, we focused on the potential impact that VR could have in the user’s mental workload (Section 6.1) and on methods to induce and monitor in real-time the user’s mental workload (Section 6.2).

We hypothesized that the degraded cues for understanding the virtual environment could have a potential negative impact on the users’ mental workload. However, the results discussed in Section 6.1 did not showed any significant differences between the real and the virtual conditions. Although we cannot rule out that any difference exists, the difference might be marginal, and thus have a minimal impact on the user. Furthermore, it also showed that the impact was also marginal for a number of physiological features tested (heart-rate and skin conductance levels). Encouraged by this first result, we aimed at proposing a methodology for designing virtual training systems based on mental workload assessment. Two main challenges were addressed, how to design complex training systems and how to monitor in real-time the user’s mental workload. The approaches proposed in Section 6.2 enable to design arbitrary training protocols and task agnostic mental workload recognition algorithm. Yet, the several challenges remains, as the proposed methodology requires a moderate amount of labeled training data, which is potentially task dependent and difficult to generalize among different training systems, and that the proposed classification methods were only efficient in a subject-dependent scenario, requiring to train the system for each user. Nevertheless, the results obtained are promising for the use of mental workload assessment in virtual training systems.

Chapter 7

Assessing the Users' Mental State

In the previous Chapter, we discussed work aimed at determining the user's mental state by considering either physiological measurements (ANS) or estimations gathered from questionnaires and behavioral measurements. Another alternative is to extract the information directly from the CNS, and in particular, through brain imaging techniques. Yet, the use of CNS measures are challenging from a technical and practical point of view, specifically in its VR/AR usage.

From a human-computer interaction perspective, the information gathered from the CNS could serve different purposes in the context of Brain Computer Interfaces (BCIs). From understanding and assessing the user's state to adapt the interaction state (passive BCIs), providing direct input commands (active BCIs), or detecting specific mental states after a specific perceptual stimulation (reactive BCIs). Our research in the field of the brain computer interfaces have proposed a number of proof-of-concepts exploiting already well known neurophysiological processes in order to assess the users experience or enable direct interaction. In particular, Section 7.1 describes two experiments focusing on the assessment of VR experiences, either at detecting system errors or the user's agency over a virtual hand. These two experiments illustrate how recording and analyzing EEG data could be of interest for the automatic assessment of VR experiences. Then, Section 7.2 presents a series of experiments aiming to assess the feasibility and design of SSVEP interfaces (reactive BCI) in the context of AR.

7.1 Assessing VR Applications using EEG measurements

One neurophysiological signature, the error-related potential (ErrP) [88], holds great promise for improving interactive systems, as it appears shortly after an erroneous event happens. In a nutshell, the generation of an ErrP happens when the user commits or witnesses an error and can be observed in the user's EEG signals as a specific EEG pattern. The ErrP mainly consists in a negative potential deflection over the fronto-central scalp areas appearing 50 to 100ms after the error, and is followed by a centro-parietal positive deflection [89]. From

Figure 31: Left, *Feedback error (Fe)* condition. The participant receives a wrong feedback after the completion of the task. Right, *Background anomaly (Be)*. The picture frame in front of the participant flips and gets into an unrealistic position, penetrating the table.

an HCI context, being able to detect when a user perceives an error without requiring an explicit communication can be useful in many scenarios [90], being particularly interesting for the assessment and adaptation of interactive systems. For example, it could be used to automatically assess the robustness of a system (e.g. linking the state of the system with erroneous events) or it could be used directly in the interaction process (e.g. automatically correct application or user errors). However, the study of ErrPs have been scarcely explored in virtual reality.

In [91] we studied the presence and the detection of ErrPs originating from Virtual Reality systems. In particular, our goal was to investigate the presence of error-related potentials in different erroneous situations that could happen in VR. Precisely, we identified 3 types of situations that could potentially elicit error-related potentials: (Te) The loss of tracking of a manipulated object; (Fe) An unexpected or erroneous confirmation feedback; (Be) An unrealistic background anomaly (see Figure 31). To assess whether or not these conditions elicit error-related potentials, we conducted a user study ($N = 15$) to record EEG activity under each one of them. From the literature and previous work, our hypotheses were that (Te) would elicit Event-Related Negativity (ERN) associated with execution errors, that (Fe) would trigger a Feedback Related Negativity whereas (Be) would trigger an ERN for the users who notice the anomaly and that it was possible to classify and detect the elicited ErrPs in single trial.

Taken together, the results obtained in the experiment showed that it was possible to accurately classify "tracking error" vs "correct" conditions in single trial, but that "feedback error" and "background anomaly" conditions did not show a clear ErrP pattern. The first observation was regarding the "tracking" errors as participants could easily notice them. As when they happened, participants were no longer able to achieve the pick-and-place task. Thus, this kind of error jeopardized the task and generated ErrPs (see Figure 32). Besides, the pattern of the ErrP obtained in our study was very similar and well consistent with state-of-the-art literature regarding error signals in 2D paradigms [92]. In contrast, the results of the analysis of the "feedback error" and "background anomaly" did not show clear ErrPs patterns. This could be explained that this errors did not had a direct impact on the task and could

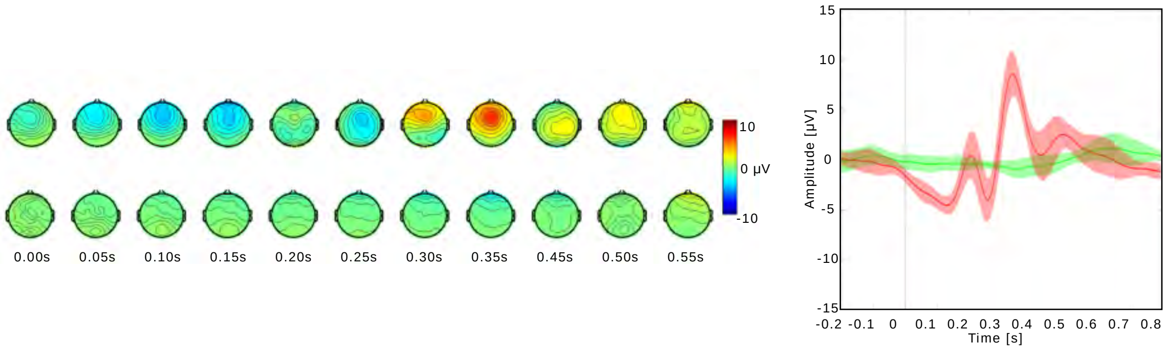


Figure 32: (Left) Topoplots of the grand average for the conditions with tracking errors and without (top and bottom rows respectively). (Right) Grand average for the signal at channel FCz for the conditions with tracking errors and without (green and red lines, respectively).

sometimes been overlooked by the participants. One key aspect for the direct usage of ErrP analysis in VR, is the ability to detect the ErrP in a single trial basis, and not considering the grand average. In this respect, the results were similar as the previously detailed. Single trial detection had a high detection rate for the "tracking" error condition, with an average accuracy of 84.9% while the classification accuracy for the other errors were not significantly above the chance-level.

The second aspect that we explored when assessing the user's experience through CNS monitoring, was whether it was feasible to quantify the sense of agency, one component of the sense of embodiment, towards a virtual hand. When embodying an avatar, experiencing a high Sense of Agency (SoA), i.e. feeling in control when performing an action, is essential to ensure the efficiency of an interaction between a user and any technology-mediated application. The sense of agency as a process can be divided into two components [93]: the feeling of agency (which precedes the action outcome) and the judgment of agency (which results from a comparison between the predicted and actual outcomes). On the one hand, as the pre-motor cortex is being solicited at very early stages of the action, we hypothesized that its activation could specifically reflect the "feeling of agency" component. On the other hand, the fact that the right posterior parietal cortex is activated when there is a mismatch between the predicted and the actual sensory outcome, it suggests that its degree of activation may reflect the "judgment of agency" component. In order to experience a judgment of agency, the priority, consistency and exclusivity principles must be respected [94]. Thus, we proposed three potential modulators of them when immersed in virtual reality and controlling a virtual hand:

- Priority principle (the conscious intention to perform an act must immediately precede the action, which itself must immediately precede the outcome): in order to modulate the SoA based on the priority principle, we proposed to add a visual latency between the user's movement and the feedback (i.e., the virtual hand's movement). For instance, we proposed this latency to be of either 1s, 1.5s or 2s.

- Consistency principle (the sensory outcome must fit the predicted outcome): here, we proposed to manipulate the feedback through the inversion of two finger motions in order to modulate the consistency principle.
- Exclusivity principle (one's thoughts must be the only apparent cause of the outcome): finally, in order to modulate the exclusivity principle, we proposed to make one of the fingers not responsive to users' movements. In other words, this finger would "move by itself". This "mad finger" could be any of the fingers. Here, we proposed to select the thumb, the middle and the little fingers.

In a controlled experiment ($N = 24$), we faced participants with the three potential modulators of the sense of agency. We hypothesized that the perceived SoA, which was measured using questionnaires, would be higher when the principles were not modulated than when they were. Concerning the measure of the SoA, we also considered EEG analysis. Regarding EEG measurements, we hypothesized that a stronger activation of the pre-motor cortex, reflecting a high feeling of agency, will be revealed at the early stages of the trials in "high-agency" conditions (i.e., when the principles were not modulated), while a stronger activation of the right posterior-parietal cortex, reflecting a low judgment of agency, will be revealed after the feedback was perceived in "low agency conditions" (i.e., when one of the principles was modulated).

Concisely, the results revealed that participants rated their SoA significantly lower in the manipulated than in the non-manipulated trials, whatever the principle modulated. It suggested that the three principles were relevant to be considered in VR. Indeed, the manipulations

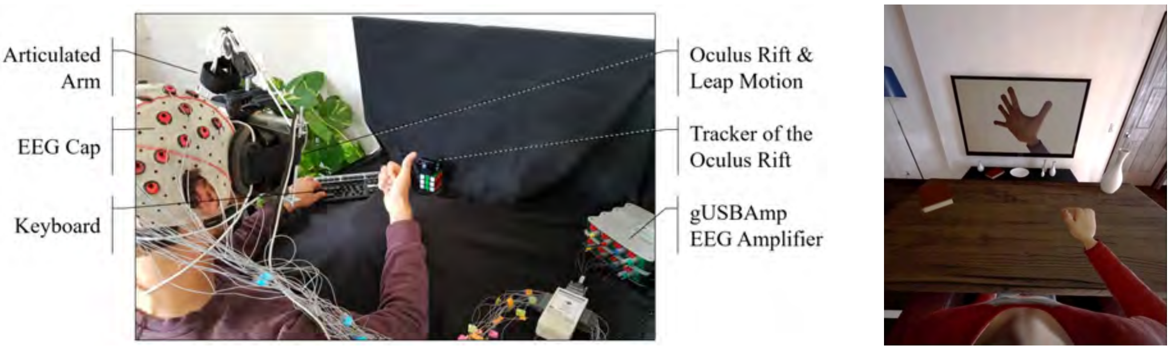


Figure 33: (Left) Experimental set-up. The participant was equipped with an EEG cap and the Oculus Rift DK2 VR head-mounted display. The HMD was supported by an articulated arm (to avoid any pressure on the EEG cap and reduce the risk of muscular fatigue). The user's right hand was tracked by a Leap Motion. (Right) User's first person perspective while immersed in the virtual environment.

of the priority, exclusivity and consistency principles all induced a significant decrease of the SoA. The neurophysiological measures revealed the same patterns as the ones described in the literature [95, 96, 97]. First, a stronger activation in the θ frequency-band was revealed at early stages of the trial when the participants felt in control, potentially reflecting a high feeling of agency. Second, a stronger activation in the same θ frequency-band was reported once the feedback was perceived when participants did not feel in control, potentially reflecting a low judgment of agency. The attribution of these EEG modulations to the feeling and judgment of agency specifically is only a hypothesis that should be further investigated in future works. The next step will consist in applying source reconstruction algorithms in order to determine more precisely the sources of the signal. Indeed, due to the poor spatial resolution of the EEG, we should stay cautious when inferring the precise location of the zones activated. Nevertheless, the strong differences between the high and low conditions as well as the apparent consistency between the literature and the obtained scalp maps is very encouraging.

7.2 Brain-Computer Interfaces for Augmented Reality

EEG data, in addition to assess the user mental state, can also be used as an input mechanism. As such, Brain-Computer Interfaces (BCIs) enable the design of novel interaction schemes based directly on brain activity and mental states of the user [98]. A BCI translates the brain activity into "mental commands" that can be exploited in a wide number of applications ranging from assisting people with disabilities or reeducation therapies, to entertainment and video games. One key feature of BCIs is that they do not require the user to preform any motor action (i.e. hands-free) to interact with the system, which is a promising input method for augmented reality applications where users might require their hands to perform other tasks, and in particular SSVEP as it provides a relatively high throughput compared to other BCI paradigms. During our research, we conducted a number of works exploring alternatives to increase the efficiency of Steady-State Evoked Visual Potentials (SSEVPs) in desktop systems [99, 100]. The SSVEP (Steady-State Visually-Evoked Potential) is a specific brain pattern that occurs when the human visual system is stimulated by a periodic flickering stimulation. The brain responds with an activity at the very same frequency in the early visual cortical area [101]. When facing multiple targets flickering at different frequencies, it becomes possible to determine which target the user is focused by analyzing the user's EEG signals. By associating each target with a particular command, it becomes then possible to create a BCI with multiple mental commands [102].

Yet, combining Brain-Computer Interfaces with Mixed Reality applications raises a lot of scientific and technological challenges. It poses theoretical questions regarding the design of effective interaction paradigms but it also involves feasibility studies regarding the compatibility of these two complex technologies. Through a series of experiments, we explored the potential combination of SSVEP interfaces and Augmented Reality (AR) [103], from the tech-

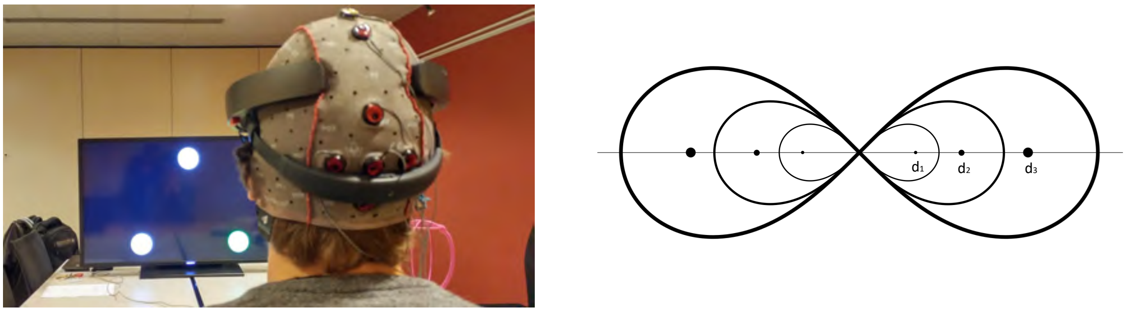


Figure 34: Left, experimental setup of the experiment assessing the compatibility of EEG and AR equipment, in which the SSVEP targets were displayed on an external screen. Right, motions that the SSVEP targets followed during the study exploring the impact of head motions: Lemniscates of Bernoulli with half focal distances $d_1(40°) < d_2(90°) < d_3(140°)$. The time to perform the motion was constant.

nical challenges to its usability in a real application. First, in two studies, we explored the feasibility of using SSVEP interfaces combined with a wearable AR system, and in particular an Optical See-Through (OST) head-mounted display. The main focus of these experiments were the assessment of the hardware compatibility (see Figure 34, left) and the impact of head motions (see Figure 34, right). EEG data is sensible to external stimuli such as electromagnetic interferences and electrical impulses generated by muscular activity, such external noise can compromise the classification accuracy.

The experimental results showed that the combination of the BCI and AR equipment did not provide any significant decrease of classification accuracy. Although this result is dependent of the equipment used, it was needed for the follow up experiments, and also showed that the SSVEP targets could be displayed using an OST system without any drop of classification accuracy. More importantly, the results of the second experiment, showed a clear impact of the head motions on the classification accuracy. This was an expected result from the literature, but provided insights on the tolerance of head motion for AR interfaces. Furthermore, the smallest/slowest head motion condition did not have a significant impact on the classification accuracy.

Encouraged by the results obtained in these two experiments, we furthered explored the potential usages of SSVEP for AR interfaces [103]. Precisely, we explored the design space for a 3 target SSVEP interface for the control of an unmanned vehicle (robot) with three commands: “Forward”, “Turn Left” and “Turn Right”. In order to propose the most coherent association of each command with the AR target, we proposed a 5-dimension design space that described all possible target layouts for the 3 commands. These 5 dimensions, namely Orientation, Frame-of-reference, Anchorage, Size and Explicitness (see Figure 35), provided several configuration that could have an impact on the interaction (e.g. relative placement of the targets) and on the SSVEP classification accuracy (e.g. distance among the targets, size of the targets):

- **Orientation**: This dimension corresponded to the plane containing the 3 targets. It was either "Transversal" or "Frontal". When transversal, the plane of the targets was parallel to the horizon. When frontal, the plane would always face the user.
- **Frame-of-reference**: This dimension defines whether the frame of reference of the targets was "Exo-centered" or "Ego-centered". In case of the "Exo-centered", the targets followed the robot orientation while for the "Ego-centered" targets remained fixed with respect to the user.
- **Anchorage**: This dimension defined whether targets were anchored to the "Robot" or remained fixed at the center of the user's field of view ("User").
- **Size**: This dimension defined whether the size of the targets was "Absolute" or "Adaptive". In the absolute condition the size of the targets were constant (i.e. visually gets smaller as the robot went further away from the user), while from the "Adaptive" the visual size of the targets remained visually constant.
- **Explicitness**: This dimension focused whether the command for each target was "Explicit", in which a red segment will visually connect the target with the corresponding wheel of the robot (front wheel, right wheel and left wheel), or "Implicit" in which the user had to mentally figure it out.

The combination of the different dimensions resulted in a total of 32 potential layout designs, discouraging the formal evaluation of all potential combinations. Thus, we first conducted a user study to determine the layouts that users would perceive as the easiest to understand considering the scenario of robot control. An online evaluation ($N = 42$) enabled to compare the different display strategies by asking the participants to evaluate "how coherent/intuitive was the association between the targets and the commands" while watching pre-recorded videos of a robot performing a series of commands. The final ranking of the strategies was performed following a majority judgment procedure [104] in order to find the most relevant configurations. In a nutshell, the results showed that participants preferred a transversal orientation, meaning that the targets should remain in the same plane as the robot's motion. In addition, the user strongly preferred when the frame-of-reference for the targets coordinates was set to the robot (exo-centered). We hypothesized that these two configurations (transversal orientation and exo-centered frame-of-reference) minimized the mental rotation necessary to map the command with the direction. Although with smaller differences, the "adaptive" size, "robot" anchored and "explicit" were also the most preferred.

However, the highest subjective preference of the users did not guarantee the best objective performance of the SSVEP detection, thus, we performed a final evaluation ($N = 24$) testing four of the most representative configurations from the majority judgment results. In a final experiment, participants had to guide a virtual robot through a pre-defined path using an

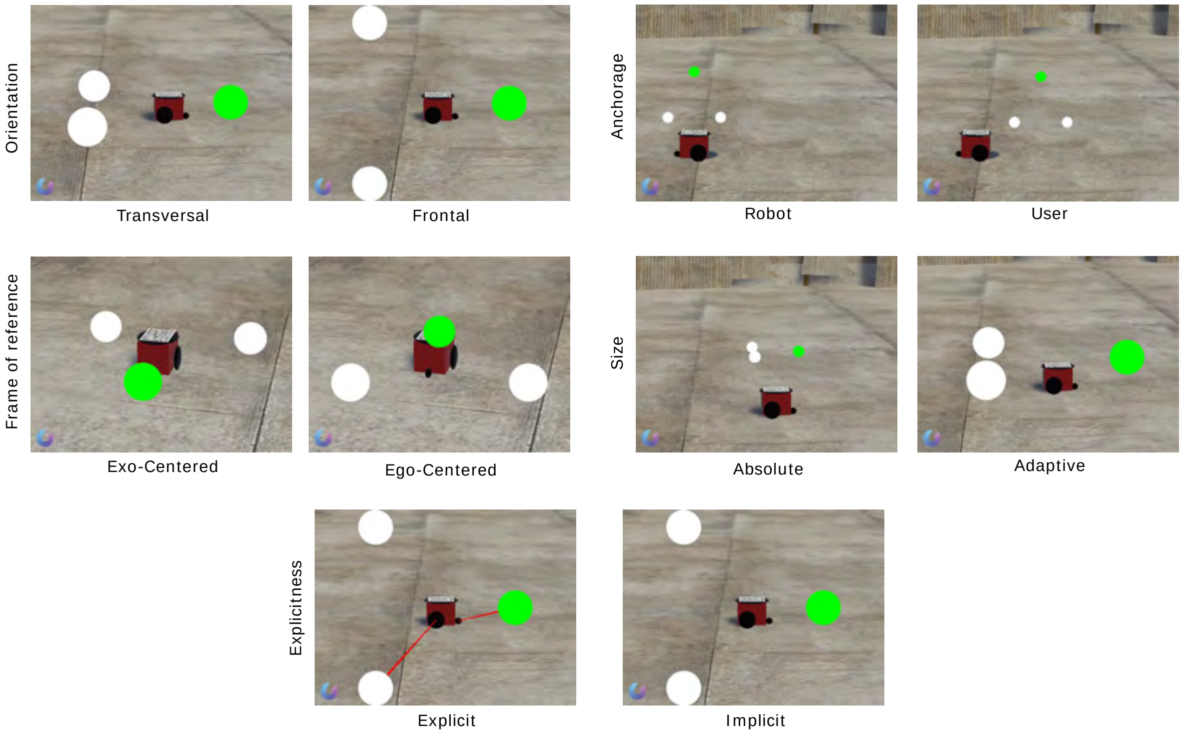


Figure 35: Design space for the placement of the three SSVEP targets in order to control the three robot commands: "Forward", "Turn Left" and "Turn Right". The figure provides an illustration of the 5 dimensions of the design space each having two levels.

OST-HMD. In the experiment, participants had to select the SSVEP targets indicated by the system in order to avoid potential errors by choosing a wrong target. The aim for the experiment was not to assess the usability of the system, but to specifically assess the SSVEP classification accuracy for different layout configurations. The main results indicated that the preferred configuration in the subjective evaluation had a comparable performance as the "default" SSVEP configuration and mainly the size of the target had a significant impact on SSVEP accuracy.

Taken together, these experiments showed the complexity and challenge of adapting BCI interfaces for AR applications. In the case of SSVEP, it involves the placement of flickering targets in a comprehensive way. The design also relates to the final task/application and for instance to the number of possible commands. Designing an AR-BCI application often requires: (1) testing hardware compatibility, depending on the EEG cap, the AR device and the exploited cortical area, (2) testing performance in real conditions (light, movement), (3) designing the user interface related to all the BCI components (command semantics, stimulations and feedbacks) with respect to the final application and (4) validating the effect of the user interfaces on the BCI task performance.

7.3 Conclusion

In this Chapter we have presented research work aiming to exploit an additional input modality to assess and interact in mixed reality systems through direct measurement of the user's brain activity. First, the evaluation of VR systems is highly complex due to all the factors involved in its design (e.g. interaction techniques, visual appearance, simulation, narrative). The ability of objectively assess VR systems analyzing the users' mental state could invigorate the research of novel evaluation methodologies. Moreover, these methods are especially interesting as they can provide a continuous assessment of the user's experience rather than punctual assessments along the experience using questionnaires. While our work have mainly focused on two particular aspects, system errors [91] and agency [105], other mental states such as attention or mental workload could be further explored.

Secondly, we thoroughly explored the technical and design challenges of using BCIs in an augmented reality context [103]. Our analysis highlighted the complexity of introducing BCIs to real case applications, even for simple interactions. In this respect, our analysis also raised a number of challenges that have yet to be solved. From the research of efficient signal filtering methods and artifact removal procedures, to the increase the input throughput. Furthermore, the presented work only focus in a particular mental state, SSVEP. Yet, other mental states such as motor imagery and attention based could constitute additional research paths. In particular, motor-imagery based BCIs are typically used for controlling smart wheelchairs and enable severely impaired patients to regain autonomy. It seems promising to couple such assistive technologies with AR systems in order to augment the environment and provide new hand-free interaction possibilities.

Conclusion

Through this last Part of the manuscript, I have presented work aimed to extract additional information from the user's state during the interaction process. As discussed, this information can serve for a number of different purposes, from purely assessment/evaluation purposes, as a mean to adapt the VR experience, to new input modalities. Nevertheless, a number of challenges are still required in order to fully exploit their potential.

First, while the body of literature on neuroscience and physiology is broad, its direct application on VR/AR is still limited and the potential impacts of VR/AR technology remain scarcely explored. For example, the mere fact of being immersed on a virtual environment could modulate the user's physiological and neural state, such as cybersickness. Thus, a more precise relationship between the user's physiological/neurological state and cognitive/emotional levels is still needed. Second, the actual measurements of the user's physiological and neurological state poses technical and methodological challenges. From the coupling of VR/AR equipment, the generalization of classification and recognition methods, to the support of inter and intra-subject differences. Although, a small number of VR/AR equipment already integrate physiological and brain imaging technology, the quality of the measurements can be far away from actual "medical-grade" systems and the results obtained in controlled experiments outside VR could be challenging to replicate. Third, apart from the offline assessment of the user's state, it remains largely unexplored how this information can be used as an additional interaction modality. As discussed in Section 6, the monitoring of the user's mental workload can be used to adapt VR training scenarios to accommodate the requirements of different training routines (e.g. keep high mental workload). Yet, there is still a large number of unexplored combinations of cognitive states and VR/AR applications. Finally, the user's state can be a direct input parameter than the user can control. On one hand, this new input mechanism could enable interactions that can be achieved without any muscular activity from the user, enabling interactions that are not otherwise possible. However, this new input modality presents a number of caveats that require further research. From the increase of its throughput (traditional systems have much higher throughput), to the potential fatigue that such "unnatural" methods could generate. Humans do not modulate their physiological/neurological signals consciously, which could potentially require intensive training to fully master.

Towards the Perfect Loop?

> *"Why has man changed the shapes and substances of his environment? To change what it affords him. He has made more available what benefits him and less pressing what injures him."*
>
> — James J. Gibson*, The Ecological Approach to Visual Perception*

This manuscript entitled "The Infinite Loop - Towards Adaptive 3D User Interfaces" has provided a broad overview of my research activity on virtual reality and 3D user interfaces during and after my PhD. Using the perception-action loop as a canvas, my research activity has focused on the improvement of user's interaction in virtual environments by (1) studying the user's perception and behavior in virtual environments and (2) proposing novel interaction methods to overcome perceptual limitations and better support the user. This research further highlighted the need to consider additional information to improve user interaction, which motivated the analysis and exploitation of cognitive and mental states. The ultimate goal being to create adaptive 3D user interfaces. Interfaces that are aware of the perception and interaction capabilities of the users. Interfaces that are able to efficiently support the user while performing 3D interaction tasks. However, one might ask which is the optimal level of adaptation, and whether a flawless, "perfect", perception-action loop is achievable.

As evoked by Gibson in the citation that opened this final conclusion, humans have continuously adapted their environment to fit their limitations and their purposes. In real life, the environment cannot adapt to us, we have to adapt to it, either by training or by using tools capable of overcoming our physical and perceptual limitations. In contrast, when interacting in virtual environments, everything is potentially adaptable, and all physical and perceptual limitations could be potentially overcome by the 3D User Interface (3DUI). In a first step, this flexibility, with potentially infinite combinations, has lead to a wide range of research aiming to provide better 3DUIs, in which, some of these examples have been presented in this manuscript. Nevertheless, although we can argue that most of these interfaces have been designed to adapt to the user, they provide mostly a static adaptation and are not tailored to a specific user. The plethora of existing 3DUIs, raises two major challenges in this respect. First, existing 3DUIs are designed to perform precise and specific tasks, and remains unclear how the user will adapt to them in the long term. How will 3DUIs support the user over time? How will the user adapt to the 3DUI over time? Will perceptual biases fade after long exposures? Will motor performance steadily increase after extensive use? Second, due to the high variety of 3DUIs their exhaustive evaluation remains unfeasible. For example, the complete evaluation of all locomotion and selection techniques is virtually impossible, either by the number of techniques or by the potential interactions that should be evaluated. These two challenges are orthogonal, and longitudinal studies will be required to bring light to these questions. With the ultimate goal of finding the 3DUI which will better support the user's interaction in the long run.

Moreover, during the last years, there has been a strong democratization of VR systems that has reinvigorated the use of the virtual reality in the industrial and entertaining field. However, we are far from standardized 3DUIs, which could become as widespread as traditional 2D interfaces such as the WIMP (windows, icons, mouse and pointer) interfaces. This lack of standardization generates heterogeneous interaction systems, and the user is required to learn and adapt for each application, which can be tedious and discouraging for novice users. This clearly contrasts with desktop interfaces in which the interaction process and design principles are centered in the WIMP interfaces. What is more, it remains unclear the potential negative effects this might generate. Can interaction be degraded by the continuous change of interaction methods? Again, these are long term effects hard to assess, as the potential combinations are even higher than when comparing 3DUIs. Furthermore, the transition between real and virtual context might generate a negative effect on the overall interaction process. Once immersed in VR, how much time the user needs to recalibrate himself to achieve an optimal interaction? Is this dependent on the user's experience? Does the 3DUI play an important role? Similarly, longitudinal studies will be required to answer such questions.

In addition to the standardization effort, which could reduce the number of 3DUIs are being used for each application, a user customization approach could also be envisioned. This approach will require for each user to determine their "optimal" 3DUI for his/her needs, knowledge and expertise. The customization could be done at different levels, first, enabling users to choose and configure the interaction methods that they are more comfortable with. Second, by adjusting the interface during its usage by monitoring and assessing the interaction process. While the first step will be manual and might require a number of calibration procedures, the second step should be fully automatic and transparent to the user. The major interest of this approach is that the user not only uses this "customized" 3DUI in a single application, but with all applications.

Finally, another complementary aspect to this standardization/customization relates to the user's representation. Since 2016, my interests in the implications of the user's avatar in the interaction process have been growing. I believe that the avatar could become the cornerstone of the the standardization/customization process. However, the use of the avatar as an interaction tool is still scarcely explored and it mainly focus on enabling realistic manual interactions. During preliminary works, we have started to explore the role of the avatar during the interaction process while performing other interaction tasks, such as manipulation [43] or navigation tasks [46]. This research has highlighted a bi-directional interaction between the user’s performance and the user’s experience, in particular on the sense of embodiment, i.e. the feeling of owning the virtual body. Indeed, the different characteristics of the avatar (e.g. appearance, control scheme) have a strong influence on how users perceive their avatar. The perception of the avatar can improve user experience, modify cognitive performances and even change the way users interact. At the same time, the avatar is also key to drive the interaction process, as it provides feedback regarding the user’s motions, influencing performance, the perception of their actions and even the perception of the environment. However, in order to fully exploit the avatar capabilities, additional research is required to better understand how users perceive their virtual self while interacting, including methodological approaches for their assessment, potentially considering physiological measurements, and research on avatar-compliant 3D user interfaces, which not only provide efficient interaction but allow the user to appropriate their virtual avatar. This could lead not only to a visual customization of avatars, which is mainstream in most VR chat/conferences systems, but also a customization at the interaction level, interaction schemes adapted to each user and maybe reaching the “perfect loop”.

Bibliography

[1] Bruce E. Goldstein and James R. Brockmole. Introduction to Perception. In *Sensation and Perception*, pages 3–19. Tenth edit edition, 2017.

[2] Rebekka S. Renner, Boris M. Velichkovsky, and Jens R. Helmert. The perception of egocentric distances in virtual environments - A review. *ACM Computing Surveys*, 46(2):1–40, 2013.

[3] Holly C. Gagnon, Dun Na, Keith Heiner, Jeanine Stefanucci, Sarah H. Creem-Regehr, and Bobby Bodenheimer. The Role of Viewing Distance and Feedback on Affordance Judgments in Augmented Reality. In *IEEE Conference on Virtual Reality and 3D User Interfaces*, pages 922–929, 2020.

[4] Gerd Bruder, **Ferran Argelaguet**, Anne-Héléne Olivier, and Anatole Lécuyer. CAVE Size Matters: Effects of Screen Distance and Parallax on Distance Estimation in Large Immersive Display Setups. *Presence: Teleoperators and Virtual Environments*, 25(1):1 – 16, 2016.

[5] Etienne Peillard, Yuta Itoh, Jean-Marie Normand, **Ferran Argelaguet**, Guillaume Moreau, and Anatole Lécuyer. Can Retinal Projection Displays Improve Spatial Perception in Augmented Reality? In *IEEE International Symposium on Mixed and Augmented Reality*, 2020.

[6] Etienne Peillard, **Ferran Argelaguet**, Jean-Marie Normand, Anatole Lécuyer, and Guillaume Moreau. Studying Exocentric Distance Perception in Optical See-Through Augmented Reality. In *IEEE International Symposium on Mixed and Augmented Reality*, pages 1–8, 2019.

[7] **Ferran Argelaguet**, Anne-Hélène Olivier, Gerd Bruder, Julien Pettré, and Anatole Lécuyer. Virtual Proxemics: Locomotion in the Presence of Obstacles in Large Immersive Projection Environments. In *IEEE Virtual Reality Conference*, pages 75 – 80, 2015.

[8] Etienne Peillard, Thomas Thebaud, Jean-Marie Normand, **Ferran Argelaguet**, Guillaume Moreau, and Anatole Lécuyer. Virtual Objects Look Farther on the Sides: The

Anisotropy of Distance Perception in Virtual Reality. In *IEEE Conference on Virtual Reality and 3D User Interfaces*, pages 227–236. IEEE, 2019.

[9] Ronald T. Azuma, Yohan Baillot, Reinhold Behringer, Steven Feiner, Simon J. Julier, and Blair MacIntyre. Recent advances in augmented reality. *IEEE Computer Graphics and Applications*, 21(6):34–47, 2001.

[10] Gurjot Singh, Stephen R. Ellis, and J. Edward Swan. The Effect of Focal Distance, Age, and Brightness on Near-Field Augmented Reality Depth Matching. *IEEE Transactions on Visualization and Computer Graphics*, 26(2):1385–139, 2018.

[11] **Ferran Argelaguet** and Carlos Andujar. A Survey of 3D Object Selection Techniques for Virtual Environments. *Computers and Graphics*, 37(3):121–136, 2013.

[12] **Ferran Argelaguet**, Carlos Andujar, and Ramon Trueba. Overcoming Eye-Hand Visibility Mismatch in 3D Pointing Selection. In *ACM Symposium on Virtual Reality Software and Technology*, pages 43–46. ACM, 2008.

[13] **Ferran Argelaguet** and Carlos Andujar. Efficient 3D Pointing Selection in Cluttered Virtual Environments. *IEEE Computer Graphics and Applications*, 29(6):34–43, 2009.

[14] I Scott MacKenzie. Fitts' law as a research and design tool in human-computer interaction. *Human-computer interaction*, 7(1):91–139, 1992.

[15] Regis Kopper, Doug A Bowman, Mara G Silva, and Ryan P McMahan. A human motor behavior model for distal pointing tasks. *International journal of human-computer studies*, 68(10):603–615, 2010.

[16] **Ferran Argelaguet**, André Kunert, Alexander Kulik, and Bernd Froehlich. Improving co-located collaboration with show-through techniques. In *IEEE Symposium on 3D User Interfaces*, pages 55–92, 2010.

[17] **Ferran Argelaguet**, Alexander Kulik, André Kunert, Carlos Andujar, and Bernd Froehlich. See-through Techniques for Referential Awareness in Collaborative Virtual Reality. *International Journal of Human-Computer Studies*, 69(6):387–400, 2011.

[18] Claudio Pacchierotti, Stephen Sinclair, Massimiliano Solazzi, Antonio Frisoli, Vincent Hayward, and Domenico Prattichizzo. Wearable haptic systems for the fingertip and the hand: taxonomy, review, and perspectives. *IEEE Transactions on Haptics*, 10(4):580–600, 2017.

[19] Anatole Lécuyer, Jean-Marie Burkhardt, Sabine Coquillart, and Philippe Coiffet. Boundary of illusion: An experiment of sensory integration with a pseudo-haptic system. In *IEEE International Conference on Virtual Reality*, pages 115–122, 2001.

[20] Louise Moody, Alan Waterworth, John G. Arthur, Avril D. McCarthy, Peter J. Harley, and Rod H. Smallwood. Beyond the visuals: tactile augmentation and sensory enhancement in an arthroscopy simulator. *Virtual Reality*, 13(1):59–68, 2008.

[21] Knut Drewing, Andreas Ramisch, and Florian Bayer. Haptic, visual and visuo-haptic softness judgments for objects with deformable surfaces. In *World Haptics Conference*, pages 640–645. IEEE, 2009.

[22] **Ferran Argelaguet**, David Antonio Gómez Jáuregui, Maud Marchal, and Anatole Lécuyer. Elastic Images: Perceiving Local Elasticity of Images Through a Novel Pseudo-Haptic Deformation Effect. *ACM Transactions on Applied Perception*, 10(3):17:1–17:14, 2013.

[23] Antoine Costes, **Ferran Argelaguet**, Fabien Danieau, Philippe Guillotel, and Anatole Lécuyer. Touchy : A Visual Approach for Simulating Haptic Effects on Touchscreens. *Frontiers in information and communication technologies*, 6:1–11, 2019.

[24] Anatole Lécuyer, Jean-Marie Burkhardt, and Chee-Hian Tan. A study of the modification of the speed and size of the cursor for simulating pseudo-haptic bumps and holes. *ACM Transactions on Applyed Perception*, 5(3):14:1–14:21, 2008.

[25] Hiroyuki Kajimoto, Naoki Kawakami, Susumu Tachi, and Masahiko Inami. Smarttouch: Electric skin to touch the untouchable. *IEEE Computer Graphics and Applications*, 24(1):36–43, 2004.

[26] David Antonio Gómez Jáuregui, **Ferran Argelaguet**, Anne-Hélène Olivier, Maud Marchal, Franck Multon, and Anatole Lécuyer. Toward "Pseudo-Haptic Avatars": Modifying the Visual Animation of Self-Avatar Can Simulate the Perception of Weight Lifting. *IEEE Transactions on Visualization and Computer Graphics*, 20(4):654–661, 2014.

[27] Lionel Dominjon, Anatole Lécuyer, J-M Burkhardt, Paul Richard, and Simon Richir. Influence of control/display ratio on the perception of mass of manipulated objects in virtual environments. In *IEEE Virtual Reality Conference*, pages 19–25, 2005.

[28] Merwan Achibet, Maud Marchal, **Ferran Argelaguet**, and Anatole Lécuyer. The Virtual Mitten: A novel interaction paradigm for visuo-haptic manipulation of objects using grip force. In *IEEE Symposium on 3D User Interfaces*, pages 59–66, 2014.

[29] Betty J. Mohler, Sarah H. Creem-Regehr, William B. Thompson, and Heinrich H. Bülthoff. The effect of viewing a self-avatar on distance judgments in an hmd-based virtual environment. *Presence: Teleoperators and Virtual Environments*, 19(3):230–242, 2010.

[30] Sungchul Jung, Gerd Bruder, Pamela J. Wisniewski, Christian Sandor, and Charles E. Hughes. Over my hand: Using a personalized hand in vr to improve object size estimation, body ownership, and presence. In *Proceedings of the Symposium on Spatial User Interaction*, SUI '18, page 60–68, 2018.

[31] Daniel Roth and Marc Erich Latoschik. Construction of the virtual embodiment questionnaire (veq). *IEEE Transactions on Visualization and Computer Graphics*, 2020.

[32] Konstantina Kilteni, Raphaela Groten, and Mel Slater. The Sense of Embodiment in Virtual Reality. *Presence: Teleoperators and Virtual Environments*, 21(4):373–387, 2012.

[33] Thomas Waltemate, Dominik Gall, Daniel Roth, Mario Botsch, and Marc Erich Latoschik. The impact of avatar personalization and immersion on virtual body ownership, presence, and emotional response. *IEEE Transactions on Visualization and Computer Graphics*, 24(4):1643–1652, 2018.

[34] Henrique Galvan Debarba, Sidney Bovet, Roy Salomon, Olaf Blanke, Bruno Herbelin, and Ronan Boulic. Characterizing first and third person viewpoints and their alternation for embodied interaction in virtual reality. *PLOS ONE*, 12(12):1–19, 12 2017.

[35] Rebecca Fribourg, **Ferran Argelaguet**, Anatole Lécuyer, and Ludovic Hoyet. Avatar and sense of embodiment: Studying the relative preference between appearance, control and point of view. *IEEE Transactions on Visualization and Computer Graphics*, 26(5):2062–2072, 2020.

[36] Richard Skarbez, Solene Neyret, Frederick P. Brooks, Mel Slater, and Mary C. Whitton. A psychophysical experiment regarding components of the plausibility illusion. *IEEE Transactions on Visualization and Computer Graphics*, 23(4):1369–1378, 2017.

[37] Rebecca Fribourg, **Ferran Argelaguet**, Ludovic Hoyet, and Anatole Lécuyer. Studying the Sense of Embodiment in VR Shared Experiences. In *IEEE Conference on Virtual Reality and 3D User Interfaces*, pages 273–280, 2018.

[38] Rebecca Fribourg, Nami Ogawa, Ludovic Hoyet, **Ferran Argelaguet**, Takuji Narumi, Michitaka Hirose, and Anatole Lécuyer. Virtual co-embodiment: Evaluation of the sense of agency while sharing the control of a virtual body among two individuals. *IEEE Transactions on Visualization and Computer Graphics*, pages 1–13, 2020.

[39] Diane Dewez, Rebecca Fribourg, **Ferran Argelaguet**, Ludovic Hoyet, Daniel Mestre, Mel Slater, and Anatole Lécuyer. Influence of Personality Traits and Body Awareness on the Sense of Embodiment in Virtual Reality. In *IEEE International Symposium on Mixed and Augmented Reality*, pages 1–12, 2019.

[40] Gregoire Richard, Thomas Pietrzak, **Ferran Argelaguet**, Anatole Lécuyer, and Géry Casiez. Studying the Role of Haptic Feedback on Virtual Embodiment in a Drawing Task. 2020.

[41] Ludovic Hoyet, **Ferran Argelaguet**, Corentin Nicole, and Anatole Lécuyer. ”Wow! I Have Six Fingers!”: Would You Accept Structural Changes of Your Hand in VR? *Frontiers in Robotics and AI*, 3(27), 2016.

[42] Richard Skarbez, Frederick P. Brooks, and Mary C. Whitton. A survey of presence and related concepts. *ACM Computing Surveys (CSUR)*, 50(6):1–39, 2017.

[43] **Ferran Argelaguet**, Ludovic Hoyet, Michaël Trico, and Anatole Lécuyer. The role of interaction in virtual embodiment: Effects of the virtual hand representation. In *IEEE Virtual Reality Conference*, pages 3–10, 2016.

[44] Doug A. Bowman, Ernst Kruijff, Joseph J. LaViola, and Ivan Poupyrev. *3D User Interfaces: Theory and Practice.* Addison Wesley Longman Publishing Co., Inc., USA, 2004.

[45] Majed Al Zayer, Paul MacNeilage, and Eelke Folmer. Virtual locomotion: a survey. *IEEE Transactions on Visualization and Computer Graphics*, pages 1–12, 2018.

[46] Diane Dewez, Ludovic Hoyet, Anatole Lécuyer, and **Ferran Argelaguet**. Studying the Inter-Relation Between Locomotion Techniques and Embodiment in Virtual Reality. In *IEEE International Symposium on Mixed and Augmented Reality*, 2020.

[47] Diane Dewez, Ludovic Hoyet, Anatole Lécuyer, and **Ferran Argelaguet**. Towards Avatar-Friendly 3D Manipulation Techniques: Bridging the Gap Between Sense of Embodiment and Interaction in Virtual Reality. In *ACM Conference on Human Factors in Computing Systems*, pages 1–14, 2021.

[48] **Ferran Argelaguet** and Carlos Andujar. Visual Feedback Techniques for Virtual Pointing on Stereoscopic Displays. In *ACM Symposium on Virtual Reality Software and Technology, VRST ’09*, pages 163–170, 2009.

[49] Carlos Andujar and **Ferran Argelaguet**. Anisomorphic Ray-Casting Manipulation for Interacting with 2D GUIs. *Computers & Graphics*, 31(1):15–25, 2007.

[50] Carlos Andujar and **Ferran Argelaguet**. Friction Surfaces: Scaled Ray-Casting Manipulation for Interacting with 2D GUIs. In *Eurographics Symposium on Virtual Environments*, pages 101–108, 2006.

[51] Carlos Andujar and **Ferran Argelaguet**. Virtual Pads: Decoupling Motor Space and Visual Space for Flexible Manipulation of 2D Windows Within VEs. In *IEEE Symposium on 3D User Interfaces*, pages 25–32, 2007.

[52] Carlos Andujar, Marta Fairen, and **Ferran Argelaguet**. A Cost-Effective Approach for Developing Application-Control GUIs for Virtual Environments. In *IEEE Symposium on 3D User Interfaces*, pages 45–52, 2006.

[53] Isaac Cho, Jialei Li, and Zachary Wartell. Evaluating dynamic-adjustment of stereo view parameters in a multi-scale virtual environment. pages 91–98, 2014.

[54] James McCrae, Igor Mordatch, Michael Glueck, and Azam Khan. Multiscale 3D navigation. In *Symposium on Interactive 3D Graphics and Games*, pages 7–14, 2009.

[55] Colin Ware, Curil Gobrecht, and Mark Andrew Paton. Dynamic adjustment of stereo display parameters. *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, 28(1):56–65, 1998.

[56] Richard H. Y. So, Andy T. K. Ho, and W. T. Lo. 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, 2001.

[57] **Ferran Argelaguet** and Carlos Andujar. Automatic speed graph generation for predefined camera paths. In *International Symposium on Smart Graphics (SG)*, SG'10, pages 115–126, 2010.

[58] **Ferran Argelaguet**. Adaptive Navigation for Virtual Environments. In *IEEE Symposium on 3D User Interfaces*, pages 91–94, 2014.

[59] **Ferran Argelaguet** and Maignant Morgan. GiAnt: stereoscopic-compliant multi-scale navigation in VEs. In *ACM Conference on Virtual Reality Software and Technology*, pages 269–277, 2016.

[60] Colin Ware. Dynamic stereo displays. In *ACM SIGCHI Conference on Human Factors in Computing Systems (CHI)*, pages 310–316, 1995.

[61] Maud Marchal, Julien Pettré, and Anatole Lécuyer. Joyman: A human-scale joystick for navigating in virtual worlds. In *IEEE Symposium on 3D User Interfaces*, pages 19–26, 2011.

[62] Thierry Pozzo, Alain Berthoz, and Loic Lefort. Head stabilization during various locomotor tasks in humans. *Experimental Brain Research*, 82(1):97–106, 1990.

[63] Hugo Brument, Iana Podkosova, Hannes Kaufmann, Anne-Hélène Olivier, and **Ferran Argelaguet**. Virtual vs. Physical Navigation in VR: Study of Gaze and Body Segments Temporal Reorientation Behaviour. In *IEEE Conference on Virtual Reality and 3D User Interfaces*, pages 680–689, 2019.

[64] Hugo Brument, Anne-Hélène Olivier, Maud Marchal, Gerd Bruder, and **Ferran Argelaguet**. Understanding, Modeling and Simulating Unintended Positional Drift during Steering Navigation in Virtual Environments. *IEEE Transactions on Visualization and Computer Graphics (in press)*, 2021.

[65] Hugo Brument, Anne-Hélène Olivier, Maud Marchal, and **Ferran Argelaguet**. Studying the Influence of Translational and Rotational Motion on the Perception of Rotation Gains in Virtual Environments. In *ACM Symposium on Applied Perception (in review)*, 2021.

[66] Grégoire Courtine and Marco Schieppati. Human walking along a curved path. i. body trajectory, segment orientation and the effect of vision. *European Journal of Neuroscience*, 18(1):177–190, 2003.

[67] Delphine Bernardin, Hideki Kadone, Daniel Bennequin, Thomas Sugar, Mohamed Zaoui, and Alain Berthoz. Gaze anticipation during human locomotion. *Experimental Brain Research*, 223(1):65–78, 2012.

[68] Mark A. Hollands, Aftab E. Patla, and Joan N. Vickers. “Look where you’re going!”: gaze behaviour associated with maintaining and changing the direction of locomotion. *Experimental Brain Research*, 143(2):221–230, 2002.

[69] Stéphane Vieilledent, Yves Kerlirzin, Stéphane Dalbera, and Alain Berthoz. Relationship between velocity and curvature of a human locomotor trajectory. *Neuroscience letters*, 305(1):65–69, 2001.

[70] Halim Hicheur, Stephane Vieilledent, Magnus J. E. Richardson, Tamar Flash, and Alain Berthoz. Velocity and curvature in human locomotion along complex curved paths: a comparison with hand movements. *Experimental brain research*, 162(2):145–154, 2005.

[71] Colas Authié, Pauline Hilt, Steve N’Guyen, Alain Berthoz, and Daniel Bennequin. Differences in gaze anticipation for locomotion with and without vision. *Frontiers in Human Neuroscience*, 9:312, 2015.

[72] Sebastian Freitag, Benjamin Weyers, and Torsten W Kuhlen. Automatic speed adjustment for travel through immersive virtual environments based on viewpoint quality. In *IEEE Symposium on 3D User Interfaces*, pages 67–70, 2016.

[73] Bernd Frohlich, Jan Hochstrate, Alexander Kulik, and Anke Huckauf. On 3d input devices. *IEEE Computer Graphics and Applications*, 26(2):15–19, 2006.

[74] Hugo Brument, Anne-Hélène Olivier, Maud Marchal, and **Ferran Argelaguet**. Does the Control Law Matter? Characterization and Evaluation of Control Laws for Virtual Steering Navigation. In *International Conference on Artificial Reality and Telexistence Eurographics Symposium on Virtual Environments*, pages 1–10, 2020.

[75] Niels Christian Nilsson, Stefania Serafin, and Rolf Nordahl. A comparison of different methods for reducing the unintended positional drift accompanying walking-in-place locomotion. In *IEEE Symposium on 3D User Interfaces*, pages 103–110, 2014.

[76] Christopher D Wickens. Multiple resources and mental workload. *Human factors*, 50(3):449–455, 2008.

[77] Tiffany Luong, Nicolas Martin, **Ferran Argelaguet**, and Anatole Lécuyer. Studying the Mental Effort in Virtual Versus Real Environments. In *IEEE Conference on Virtual Reality and 3D User Interfaces*, pages 809–816, 2019.

[78] Jacky Au, Ellen Sheehan, Nancy Tsai, Greg J Duncan, Martin Buschkuehl, and Susanne M Jaeggi. Improving fluid intelligence with training on working memory: a meta-analysis. *Psychonomic bulletin & review*, 22(2):366–377, 2015.

[79] Florian Grimm, Georgios Naros, and Alireza Gharabaghi. Closed-loop task difficulty adaptation during virtual reality reach-to-grasp training assisted with an exoskeleton for stroke rehabilitation. *Frontiers in neuroscience*, 10:518, 2016.

[80] Thomas D. Parsons and James L. Reinebold. Adaptive virtual environments for neuropsychological assessment in serious games. *IEEE Transactions on Consumer Electronics*, 58(2), 2012.

[81] Arindam Dey, Alex Chatourn, and Mark Billinghurst. Exploration of an EEG-based cognitively adaptive training system in virtual reality. In *IEEE Conference on Virtual Reality and 3D User Interfaces*, pages 220–226, 2019.

[82] Tiffany Luong, **Ferran Argelaguet**, Nicolas Martin, and Anatole Lécuyer. Introducing Mental Workload Assessment for the Design of Virtual Reality Training Scenarios. In *IEEE Conference on Virtual Reality and 3D User Interfaces*, pages 662–671. IEEE, 2020.

[83] Andrew J. Tattersall and Penelope S. Foord. An experimental evaluation of instantaneous self-assessment as a measure of workload. *Ergonomics*, 39(5):740–748, 1996.

[84] Tiffany Luong, Nicolas Martin, Anais Raison, **Ferran Argelaguet**, Jean-Marc Diverrez, and Anatole Lécuyer. Towards Real-Time Recognition of Users' Mental Workload Using Integrated Physiological Sensors Into a VR HMD. In *IEEE International Symposium on Mixed and Augmented Reality*, pages 1–13, 2020.

[85] Christoph Tremmel, Christian Herff, Tetsuya Sato, Krzysztof Rechowicz, Yusuke Yamani, and Dean J Krusienski. Estimating cognitive workload in an interactive virtual reality environment using eeg. *Frontiers in Human Neuroscience*, 13, 2019.

[86] Felix Putze, Christian Herff, Christoph Tremmel, Tanja Schultz, and Dean J Krusienski. Decoding mental workload in virtual environments: a fnirs study using an immersive n-back task. In *International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 3103–3106, 2019.

[87] Jonny Collins, Holger Regenbrecht, Tobias Langlotz, Yekta Said Can, Cem Ersoy, and Russell Butson. Measuring cognitive load and insight: A methodology exemplified in a virtual reality learning context. In *IEEE International Symposium on Mixed and Augmented Reality*, pages 351–362, 2019.

[88] Ricardo Chavarriaga, Aleksander Sobolewski, and José del R Millán. Errare machinale est: the use of error-related potentials in brain-machine interfaces. *Frontiers in neuroscience*, 8:208, 2014.

[89] Michael Falkenstein, Jörg Hoormann, Stefan Christ, and Joachim Hohnsbein. Erp components on reaction errors and their functional significance: a tutorial. *Biological psychology*, 51(2-3):87–107, 2000.

[90] Thorsten O. Zander, Laurens R. Krol, Niels P. Birbaumer, and Klaus Gramann. Neuroadaptive technology enables implicit cursor control based on medial prefrontal cortex activity. *Proceedings of the National Academy of Sciences*, 2016.

[91] Hakim Si-Mohammed, Catarina Lopes-Dias, María Duarte, **Ferran Argelaguet**, Camille Jeunet, Géry Casiez, Gernot Müller-Putz, Anatole Lécuyer, and Reinhold Scherer. Detecting System Errors in Virtual Reality Using EEG Through Error-Related Potentials. In *IEEE Conference on Virtual Reality and 3D User Interfaces*, pages 653–661, 2020. Virtual conference.

[92] Jason Omedes, Iñaki Iturrate, Javier Minguez, and Luis Montesano. Analysis and asynchronous detection of gradually unfolding errors during monitoring tasks. *Journal of Neural Engineering*, 12(5):056001, 2015.

[93] Chlöé Farrer and Chris D. Frith. Experiencing oneself vs another person as being the cause of an action: the neural correlates of the experience of agency. *Neuroimage*, 15(3):596–603, 2002.

[94] Daniel M. Wegner and Thalia Wheatley. Apparent mental causation: Sources of the experience of will. *American psychologist*, 54(7):480, 1999.

[95] Fatta B. Nahab, Prantik Kundu, Cecile Gallea, John Kakareka, Randy Pursley, Tom Pohida, Nathaniel Miletta, Jason Friedman, and Mark Hallett. The neural processes underlying self-agency. *Cerebral Cortex*, 21(1):48–55, 2010.

[96] Suk Yun Kang, Chang-Hwan Im, Miseon Shim, Fatta B Nahab, Jihye Park, Do-Won Kim, John Kakareka, Nathanial Miletta, and Mark Hallett. Brain networks responsible for sense of agency: an eeg study. *PloS one*, 10(8):e0135261, 2015.

[97] Gonçalo Padrao, Mar Gonzalez Franco, Maria V. Sanchez Vives, Mel Slater, and Antoni Rodriguez Fornells. Violating body movement semantics: Neural signatures of self-generated and external-generated errors. *Neuroimage*, 124:147–156, 2016.

[98] Jonathan R. Wolpaw and E. Winter Wolpaw. Brain-computer interfaces: something new under the sun. *Brain-computer interfaces: principles and practice*, pages 3–12, 2012.

[99] Andéol Evain, **Ferran Argelaguet**, Nicolas Roussel, Géry Casiez, and Anatole Lécuyer. Can I Think of Something Else when Using a BCI? Cognitive Demand of an SSVEP-based BCI. In *ACM Conference on Human Factors in Computing Systems*, pages 5120–5125, 2017.

[100] Andéol Évain, **Ferran Argelaguet**, Géry Casiez, Nicolas Roussel, and Anatole Lécuyer. Design and Evaluation of Fusion Approach for Combining Brain and Gaze Inputs for Target Selection. *Frontiers in Neuroscience*, 10(454):14, 2016.

[101] Christoph S. Herrmann. Human EEG responses to 1–100 Hz flicker: resonance phenomena in visual cortex and their potential correlation to cognitive phenomena. *Experimental brain research*, 137(3-4):346–353, 2001.

[102] Andéol Évain, Nicolas Roussel, Géry Casiez, **Ferran Argelaguet**, and Anatole Lécuyer. Brain-Computer Interfaces for Human–Computer Interaction. In Maureen Clerc, Laurent Bougrain, and Fabien Lotte, editors, *Brain-Computer Interfaces 1: Foundations and Methods*. Wiley-ISTE, 2016.

[103] Hakim Si-Mohammed, Jimmy Petit, Camille Jeunet, **Ferran Argelaguet**, Fabien Spindler, Andéol Evain, Nicolas Roussel, Géry Casiez, and Anatole Lécuyer. Towards BCI-based Interfaces for Augmented Reality: Feasibility, Design and Evaluation. *IEEE Transactions on Visualization and Computer Graphics*, 26(3):1608–1621, 2020.

[104] Michel Balinski and Rida Laraki. *Majority judgment: measuring, ranking, and electing*. MIT press, 2011.

[105] Camille Jeunet, Louis Albert, **Ferran Argelaguet**, and Anatole Lécuyer. “Do you feel in control?”: Towards Novel Approaches to Characterise, Manipulate and Measure the Sense of Agency in Virtual Environments. *IEEE Transactions on Visualization and Computer Graphics*, 24(4):1486–1495, 2018.