Doctoral Dissertation
Doctoral Program in Control and Computer Engineering (35th cycle)

Interaction in Virtual Reality Simulations

Davide Calandra
* * * * *

Supervisor
Prof. Fabrizio Lamberti

Doctoral Examination Committee:
Prof. Claudio Pacchierotti, Referee, IRISA, Université de Rennes, Inria, France
Prof. Michele Russo, Referee, Sapienza Università di Roma, Italy
Prof. Sofia Seinfeld, Universitat Politècnica de Catalunya, Spain
Prof. Bill Kapralos, OntarioTech University, Canada
Prof. Guido Marchetto, Politecnico di Torino, Italy

Politecnico di Torino
July 18, 2023

I hereby declare that, the contents and organisation of this dissertation constitute my own original work and does not compromise in any way the rights of third parties, including those relating to the security of personal data.

. . . .

Davide Calandra

Turin, July 18, 2023

Summary

Virtual Reality (VR) has emerged as a powerful technology for creating immersive and engaging simulations that enable users to interact with computer-generated environments in a natural and intuitive way. However, the design and implementation of effective interaction methods in VR remain challenging. The lack of proper haptic feedback, and the need to rely on input devices such as controllers or gestures, for example, can result in awkward or unnatural interactions, reducing the perceived level of realism and the immersion related to the VR experience. At the same time, the employment of poorly designed interaction paradigms may impair usability, reduce the the sense of presence, and even cause unpleasant effects related to the so called cybersickness.

This doctoral thesis addresses the challenges related to interaction in VR simulations, focusing on the design and implementation of effective interaction methods. VR is a powerful technology that enables users to interact with computer-generated environments, but issues such as the lack of proper haptic feedback and reliance on input devices can result in awkward interactions and reduced realism. The thesis explores various topics, including locomotion interfaces, haptic interfaces, and alternative forms of Human-Computer Interaction (HCI) and Human-Human Interaction (HHI) using voice and body gestures. The case studies conducted in the thesis cover aspects like realism, usability, and engagement in VR simulations, validating proposed approaches and methodologies. The thesis contributes to enhancing User eXperience (UX) by designing and evaluating interaction paradigms that help users achieve their simulation objectives.

The doctoral thesis is structured into four chapters. The initial chapter functions as an introductory section, while the subsequent two central chapters focus on exploring specific aspects of the topic studied during the PhD program. Finally, a section dedicated to considerations and remarks is utilized to summarize the primary findings.

Chapter 1 serves as an introduction, while the central chapters focus on specific areas investigated during the PhD program. Each of these central chapters is further divided into sections. The first section provides an overview of the research context, while the second section delves into the literature review. The third

section describes the study methodology, and the fourth section presents the obtained results. A considerations and remarks section summarizes the main findings. Chapter 2 concentrates on locomotion and haptic simulation, two crucial elements of Human-Machine Interaction (HMI) in VR.

Locomotion remains a significant challenge in VR, with numerous devices and techniques proposed for navigating Virtual Environments (VEs), each with its own strengths and weaknesses. Selecting the most suitable locomotion technique for a particular use case is not straightforward. To address this issue, a novel evaluation testbed for locomotion techniques and devices in VR is introduced. The testbed aims to address the long-standing research problem of the lack of a well-defined method for evaluating and comparing locomotion approaches in VR. This tool introduces a research methodology for acquiring both objective and subjective metrics, along with a scoring system that ranks motion methods based on specific criteria. It has started to gain attention and adoption in other literature works on the subject.

Similarly, haptics is a widely studied topic in the field of HMI in VR. Standard interfaces, such as handheld VR controllers, lack sophisticated methods for realistic haptic feedback. However, the number and variety of haptic solutions are continuously expanding. Two investigations on haptics are conducted. The first compares the use of Passive Haptic (PH) interfaces in an interactive VR Training Scenario (VRTS) for forest firefighting with a traditional training approach. The second investigation focuses on the trade-off between UX performance and complexity of haptic solutions based on configurations of Active Haptic (AH) and PH technologies (i.e., VR gloves, standard controllers, and physical props), specifically in the context of simulating an active electromechanical tool (an Electric Screwdriver, ES).

The first study revealed that incorporating the devised VRTS alongside traditional course lessons significantly improved procedural learning. Trainees demonstrated better retention of the concepts related to the use of the simulated tool, as the VR experience enabled them to identify and correct errors before the actual examination. This resulted in better performance on the practice test compared to the control group. However, there was no significant advantage in terms of conceptual learning. Additionally, participants found the overall learning experience more engaging and stimulating due to the hands-on practice session in the VRTS. These findings highlight the numerous benefits that VR technology can offer in simulating physical objects, as long as their essential characteristics, such as shape and weight, are faithfully replicated.

The results of the second study showed that using a physical prop alongside VR gloves positively affects various aspects such as mental demand, frustration, attractiveness, perspicuity, efficiency, input, overall fidelity, comfort, interaction, PH fidelity, fidelity of different elements (hand and finger tracking, ES trigger feedback, screwing action phases), control, efficiency, and overall preference. Moreover, if the physical prop includes a higher-fidelity PH component, like the 3D-printed shell of the ES, there are additional benefits in terms of task control, attractiveness, perspicuity, dependability, simulation, input, presence, satisfaction, interaction, PH fidelity, fidelity of the ES trigger, and overall preference. Thirdly, using only the standard hand controller as a prop negatively impacts the perceived novelty of the setup compared to having no props at all.

Chapter 3 shifts the focus to two important HMI and HHI means: voice and body language. These elements play a fundamental role in supporting novel VR interaction paradigms. The use of speech is initially studied as an input method for hands-free navigation in a training-oriented simulation scenario. Subsequently, the combined use of voice and body language is investigated in two different use cases. The first study explores the effects of avatars' body representation in a multi-user simulation focused on emergency training. The second study evaluates different techniques for conveying emotions in shared VEs within a social VR-oriented scenario.

The first study aims to explore the potential use of voice as an HMI method. Three different speech-based techniques are evaluated to achieve hands-free teleport-based navigation in a large VR environment. The first technique relies solely on speech recognition of navigation intents, the second combines head/gaze tracking with voice input for teleporting, and the third technique combines the functionalities of the other two. Results demonstrate the significant superiority of the two approaches that involve the use of descriptions (speech-only and compound techniques) from various perspectives. The speech-based technique proves to be equally effective as the compound technique, except for a slight decrease in the need for additional references to identify Points Of Interest (POIs). The compound technique, however, leads to a less cohesive interaction scheme and increases the likelihood of confusion, suggesting that combining too many heterogeneous functionalities does not always result in better performance in this particular scenario.

The second study focuses on the HHI aspects in multi-user shared VR simulations. While voice communication is commonly utilized in these contexts, finding the optimal technique for representing VR users within a shared VE is challenging and can vary across different use cases. Two techniques for avatar representation are compared, differing in the level of visibility (head & hands vs. whole-body) and realism (minimal vs. realistic), using a relevant use case in emergency training—a road tunnel fire simulation. Results indicate that utilizing a whole-body approach to depict avatars of other users in multi-user training simulations significantly enhances the shared experience, provided that the use of inverse kinematics (IK) and animations generates credible and realistic outcomes. However, findings regarding the representation of the user's own avatar were mixed. In some cases, the whole-body avatar was perceived as distracting when the estimated posture of the body deviated significantly from the user's actual posture, and users preferred not to see the avatar at all, similar to the minimalist approach. Thus, in certain circumstances, a minimalist approach may still be more effective than a whole-body representation.

After confirming the general preference for the whole-body approach, the investigation explores various possibilities for enabling these avatars to effectively convey emotional content in a multi-user scenario. In a second study, realistic whole-body avatars were paired with two common approaches for managing facial expressions in VR. The first approach leverages facial and eye-tracking hardware to directly map the user's facial and eye movements to the 3D avatar, while the second approach utilizes a software solution to approximate lip movements of the avatar based on the user's voice input. The focus is on conveying emotions in a simulated social VR scenario. Results highlight the importance of accurately portraying the user's eye movements and facial expressions in the investigated scenario, but this is not crucial for all emotions. Apart from sadness, disgust, and to some extent, happiness and fear, there were no conclusive findings for the other two emotions (anger and surprise). This suggests that in certain conditions, the simultaneous use of voice and body movements could reduce the visibility of facial expressions, especially in intense moments. Thus, the less precise lip-sync approximation can be seen as a reasonable trade-off between additional complexity and conveying emotions faithfully.

Finally, Chapter 4 concludes the research and discusses potential future directions for development.

This Ph.D. thesis has been typeset by means of the TEX-system facilities. The typesetting engine was pdfLATEX. The document class was `toptesi`, by Claudio Beccari, with option `tipotesi=scudo`. This class is available in every up-to-date and complete TEX-system installation.