

Research Proposal: Analyzing and Understanding Embodied Interactions in Virtual Reality Systems

Florent Alain Sauveur Robert, Marco Winckler, Lucile Sassatelli, Hui-Yin Wu

▶ **To cite this version:**

Florent Alain Sauveur Robert, Marco Winckler, Lucile Sassatelli, Hui-Yin Wu. Research Proposal: Analyzing and Understanding Embodied Interactions in Virtual Reality Systems. ACM MMSys 2022 - 13th ACM Multimedia Systems Conference, Jun 2022, Athlone, Ireland. 10.1145/3524273.3533928 . hal-03698833

HAL Id: hal-03698833

https://inria.hal.science/hal-03698833

Submitted on 5 Jul 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Research Proposal: Analyzing and Understanding Embodied Interactions in Virtual Reality Systems

Florent Robert
florent.robert@inria.fr
Centre Inria d'Université Côte d'Azur, CNRS, I3S
Sophia-Antipolis, France

Marco Winckler
Université Côte d'Azur, CNRS, Inria, I3S
Sophia-Antipolis, France

Hui-Yin Wu
Université Côte d'Azur, Inria
Sophia-Antipolis, France

Lucile Sassatelli
Université Côte d'Azur, CNRS, I3S
Institut Universitaire de France
Sophia-Antipolis, France

ABSTRACT

Virtual reality (VR) offers opportunities in human-computer interaction research, to embody users in immersive environments and observe how they interact with 3D scenarios under well-controlled environments. VR content has stronger influences on users physical and emotional states as compared to traditional 2D media, however, a fuller understanding of this kind of embodied interaction is currently limited by the extent to which attention and behavior can be observed in a VR environment, and the accuracy at which these observations can be interpreted as, and mapped to, real-world interactions and intentions. This thesis aims at the creation of a system to help designers in the analysis of the entire user experience in VR environment: how they feel, what is their intentions when interacting with a certain object, provide them guidance based on their needs and attention. A controlled environment in which the user is guided will help to establish a better intersubjectivity between designer intention who created the experience and users who lived it and will lead to a more efficient analysis of the user behavior in VR systems for the design of better experiences.

CCS CONCEPTS

• **Human-centered computing** → **Systems and tools for interaction design**; • **Computing methodologies** → **Virtual reality**.

KEYWORDS

Embodied experiences, interactive task modeling, virtual reality, user experience analysis

ACM Reference Format:
Florent Robert, Marco Winckler, Hui-Yin Wu, and Lucile Sassatelli. 2022. Research Proposal: Analyzing and Understanding Embodied Interactions in Virtual Reality Systems. In *13th ACM Multimedia Systems Conference (MMSys '22), June 14–17, 2022, Athlone, Ireland.* ACM, New York, NY, USA, 5 pages. https://doi.org/10.1145/3524273.3533928

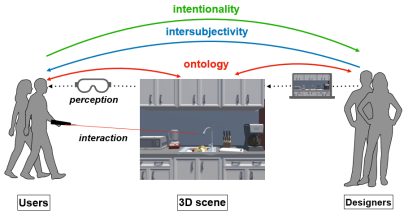


Figure 1: According to Dourish [3], in the creation of interactive embodied experiences, gaps of perception between users and designers are introduced in ontology, intersubjectivity and intentionality, corresponding respectively to the understanding of the environment, the goals, and the users experiences by the designers.

1 INTRODUCTION

Virtual reality usage is growing, increasingly used in various professional fields for education and training. It allows a high level of control of the VR environment and therefore effective monitoring of the users throughout their experience. The designers in charge of the environment must be given the necessary tools for them to create the experience, observe, and analyze to follow up with guidance adapted to the users needs and what they focus their attention on during the exploration. The intentions of the user do not always match what designers expected when they created the experience, and thus communicating the intentions between the two parties is a complicated task, especially with all the possibilities of freely exploring the VR environment and non-linear course of progression offered by VR, increasing the gap in perception.

Inspired by Dourish's work on embodiment theory [3], we identify three important components to reduce this gap in perception of the immersive environment experience between designers who created it and user who experience it:

ACM acknowledges that this contribution was authored or co-authored by an employee, contractor or affiliate of a national government. As such, the Government retains a nonexclusive, royalty-free right to publish or reproduce this article, or to allow others to do so, for Government purposes only.
MMSys '22, June 14–17, 2022, Athlone, Ireland
© 2022 Association for Computing Machinery.
ACM ISBN 978-1-4503-9283-9/22/06…$15.00
https://doi.org/10.1145/3524273.3533928

- **Ontology**: building semantically understandable 3D scenarios in which the user's purposeful interactions (e.g., navigating to a location, picking up an object) can be understood without being predefined,
- **Intersubjectivity**: designing real-time visualization and control systems that can establish shared understanding for user tasks and ongoing experience between multiple actors (e.g., between the designer and user or between multiple users), and
- **Intentionality**: designing computational methods for the analysis and identification of users' actions and their purpose of enacting an effect on the world (e.g., taking a key in order to open a door enables navigation to a previously inaccessible space).

Gaps in perception between user, designer, and system can occur at all of these levels as shown in Figure 1.

This thesis targets the investigation and design of a unified workflow enabling the understanding and analysis of user experience on these different levels of meaning in XR. For this a system will be created for the design of serious games in virtual reality, GUsT-3D (Guided Users Tasks, a tool aimed at creating interactive tasks in a VR environment). This tool allows the definition of a domain specific language (DSL) used to build an ontology to describe a VR environment as an interactive environment. Once the DSL has been defined, the tool allows the designer to create playable game scenarios and provide guidance to the users depending on their needs. The guidance also leads to a better intersubjectivity between the designer who created the experience and the users playing it by controlling the experience and orienting the user in the direction intended by the designer. Finally, multiple metrics about the users are recorded during their experience, used to analyze what the users interacted with during their experiences and how they felt in order to understand their intentionality.

2 RESEARCH QUESTIONS

The analysis of embodied interactions in XR raises research challenges both theoretical and technical that have been only partially explored in specific domains of usage. We identify three main research questions that are yet to be addressed to improve the XR experience and design a workflow enabling the understanding and analysis of users experience:

- **RQ1 : What are the qualitative measures characterizing user perception and interactions in XR ?**
- **RQ2 : How can these measures be quantified such that we can observe and understand them?**
- **RQ3 : How can designers use this information for the creation of more efficient, guided and meaningful experiences?**

A good understanding of the user experience can help in the creation of efficient guided scenarios and for a better control of the experience. This control allows designers to create experiences more aligned with their original intentions, helping them to develop scenarios focused on specific context.

3 RELATED WORK

Here I present the selected findings of my review of the work published in the last five years in IEEEVR and SIGCHI using keywords virtual reality, embodiment, training simulation, and user experience analysis. 75 papers were selected for in-depth review. Around 30 papers were measuring and quantifying user perception and interaction in VR to understand user experience as done by Hawes et al. [4] who investigated VR as a way to reduce anxiety of students, improve their mental health, using questionnaires to measure the evolution of their anxiety. Pfeuffer et al. [8] created a workflow using motion tracking to analyze behaviors in VR while doing different type of tasks (i.e., pointing, grabbing, walking, and typing). Hochreiter et al. [5] designed an exploration scenario in VR environments with and without mismatch between the real and VR environments, showing that users were less confident in mismatched cases while moving around, taking smaller steps and with elevated skin conductance responses. Existing works that measure user perception and interaction in VR uses multiple methods to analyse the experience: questionnaires, motion capture, eye tracking, and physiological sensors such as skin conductance and heart rate.

Designers can use these measures to understand the user embodied experience and guide it. A system providing these analysis would be valuable for research on the creation of training systems such as that of Clifford et al. [2] who designed a system to build a stressful context in VR in order to train firefighters by providing a context close to a real firefighting situation. Questionnaires and heart rate were used to quantify the level of stress felt by the user. A deeper analysis of the feelings and the level of stress felt in real time in the VR environment could help the designer know how to influence user experience which can in turn be used to improve training systems. Hurd et al. [6] and Wu et al. [10] respectively created a VR system designed to train users with amblyopia disability and to train users to perform pin pong spin technique. An analysis of the attention during the training could let the designer know precisely what are the weak points and difficulties of the user to provide them with more personalized training as in Lang et al. [7] where the system learns the driving habits of the user in order to create a personalized training program. As amblyopia training is usually long and repetitive, an analysis of the experience could help detect exactly what motivates the user and use it to make the experience more entertaining. In the paper of Hochreiter et al. [5] mentioned previously studying the impact of mismatching VR environments on arousal, tests are made in a free exploration environment. A well-controlled environment would help the designer make sure no other factors external to the experiment interfere on the user behavior and the subsequent analysis.

There are very few works on the modelization of the lived experience in VR. Bouville et al. [1] built the system #FIVE on the modelization of interactive VR environments by the annotation of elements composing the environment with properties, similar to GUsT-3D ontology part presented in part 4 however without the creation of a DSL. [9] worked on a prototyping workflow for AR applications, immersing the designer in the modelized experience

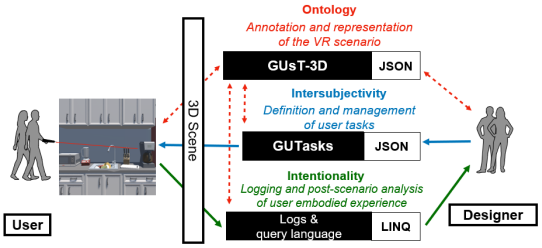


Figure 2: Implementation of the three senses of embodiment in GUsT-3D

with a miniature physical representation of the AR experience. By showing the designer a visualization of how the user will perceive and live the experience, this work helps improve the intersubjectivity between the designer creating the experience and the user living it. Xue et al. [11] created a workflow to analyze user physical and emotional behavior over time using physiological sensors, eye tracking, and questionnaires. Answering relatively similar research questions as ours on the analysis of user lived experience immersed in a VR context. This work, however, analyzes user experience watching 360° videos, where the exploration of the environment is more linear than that in VR game, since the users can only turn on the spot, making the investigation relatively different as the perception gap problem become easier to handle.

My next readings will be focused on SIGCHI and EICS conferences, on the thematic of ontology and intersubjectivity in order to find efficient ways to help the designer transmit their vision of the experience through the VR environment to the user such as Zhao et al. [12], who show efficient ways to attract the attention of the user on certain elements in order to help the user or orient their experience in the direction desired by the designer.

4 METHODOLOGY AND PRELIMINARY RESULTS

We created the GUsT-3D system to seek answers to the research questions. Divided into three components inspired from the three senses of embodiment from Dourish's work [3] as shown in Figure 2. It is composed of an **ontology** component for the creation of an interactive environment understandable by both the user and the designer, an **intersubjectivity** component allowing the designer to control and guide the user experience in the VR environments and an **intentionality** component recording multiple elements about the user and the environment during the scenario and providing visualization about the users logs to the designer to help them with the analysis of the user embodied experience. The GUsT-3D system is developed on Unity using the C# language.

The ontology component intervenes in the design of an interactive scenario, defining the way each object composing the environment can interact with other objects of the environment and with the user. The DSL establishes an important basis of understanding in the experience. An important gap in perception could be created here if the properties or possible interactions are not properly defined and understood. This part allows the modification of a domain specific language (DSL) vocabulary to fit designer's needs. The DSL is composed of three elements : the environment properties (e.g., time and distance units, parameters regarding interactive and affordance properties), user properties (e.g., height of the body, parameters such as field of view and keybinding) and the layers (i.e., type of objects understood by the vocabulary of the DSL). The layers are separated in four categories :

- *interactive layers* describe objects including interaction properties with other objects of the environment and with the user. For example, a $fork$ is a *movable* object the user can move while a *table* is a *support* object on which others objects can be placed.
- *navigation layers* describe objects structuring space where users progress, defining the navigation possibilities and constraints in the VR environment. For example, a *ground* defines a localisation while a *door* is an *entryway* used to create a link between two different localisations in space such as $Kitchen <=> Bedroom$.
- *environment layers* are assigned to objects with special properties composing the environment, like a *camera* with it's own field of view or a source of *light*.
- *object layers* are objects added by the designer using layers from other categories as sub layers to extend the vocabulary depending on their needs. The designer can for example create a $shelf$, composed of two sub layers from the interactive categories, *support* and $container$, meaning that a shelf can have *supporting* and *containing* interactions with objects composing the environment.

Once the vocabulary needed by the designers has been defined, it will be used to annotate objects composing the VR environment with layers to describe it as an interactive environment, similarly to the system #FIVE [1]. This allows us to create an ontology to make a semantically understandable 3D interactive environment for both the user and the designer. This ontology is then used for both remaining parts of the system.

The intersubjectivity component acts in creation of tasks and the guidance of the user, in order to orient their experience, allowing the designer to have a better perception of the user experience and how to help them properly. The ontology of the annotated environment is used for the design of game scenarios. A scenario is a list of tasks, each task with one objective the user will have to achieve. A short scenario could be composed of two tasks :

(1) $retrieve$ a $fork$ placed on a *table* (*support* object) in the kitchen,
(2) $store$ the $fork$ in a $shelf$ ($container$ object) in the kitchen.

The system is composed of a query language to provide help for each task to the user depending on their need and attention. Queries are structured as follows :
[How many/where is]+[an object/a layer]+[a constraint (e.g., on an object, at a localisation)]+[(the object/the layer X)]
It can be used to find precise elements needed by the the user in the scenario above like this :
Where is + *fork* + on a + *table* + in the localisation + *kitchen*
The designer could decide then to provide different guidance to

help the users after certain conditions if they couldn't find the fork by attracting their attention on the right direction using methodologies presented in the paper [12] such as guideline or highlight.

The intentionality component is used by the designer to analyse the user experience, to support designers on the understanding of user reaction and interaction during playtime. Logs are recorded over time during the experience on multiple metrics, regarding the environment (e.g., where they are, what are they interacting with, gaze, motion) and how they felt (e.g., skin conductance, heart rate) to analyze their emotions and intentions. This part provides some statistics and correlation analysis on this data in order to provide meaningful information to the designers by making the link between the observed experience (what they did) and the observed feelings (what they felt). The designer can for example observe that at a precise moment during the experience, the user saw a spider, and at the same moment they observe a spike in skin conductance, rise in heart rate, or longer fixation, potentially meaning that the user was surprised by a spider. This allows designers to fully understand how the user lived the experience. Visualisation methods are created such as graphs or heat-maps to provide efficient ways to communicate relevant information to the designer.

Thus, this system allows creation of an ontology understandable by both the designer and the user with a fully annotated environment with clear properties and interactions possibilities. It create a strong intersubjectivity by allowing the designer to build a guided scenario and orient the user in the VR environment to reduce the possible gap of perception of the experience between them. Finally, the intentionality part provides the designer an in-depth analysis tool with visualisation to understand the embodied experience of the user during the scenario and use these informations as feedback to improve the experience and provide better guidance.

In it's current state, GUsT-3D allows the creation of an ontology with layers and environment properties, annotation of objects composing the VR environment and the creation of scenarios : a list of tasks to achieve in a precise order with goals such as going somewhere, taking something, etc. Guidance of the user is based on the time spent on a task. Records of the user experience includes interactions of the users with the VR environment (where they are, what object they interact with, what they are looking at, etc.).

5 ONGOING ACTIVITIES

To assess the effectiveness of the tool in it's current state, we conducted user tests with several designers with experience in creating 3D environments. The purpose of this study was to improve GUsT-3D by collecting feedback from potential designers who would be using the system. For this, we carried out individual tests in the form of interviews on the creation of a serious game in an indoor environment (a house composed of three rooms). We recruited 6 expert users (3 men 3 women). The average number of years of experience of participants in the development of immersive 3D applications is 3 years (min 1, max 7). Regarding experience with tools available on the market, the 6 participants declared having experience with Unity and some have experience with other software such as Blender and Unreal Engine. The interview was composed of three stages: presentation of the study, demographic questions, then a simulation with and without our tool in order to evaluate several elements of it. Due to the sanitary situation, the interviews were carried out by video conference, therefore the users did not interact directly with the system. First, we provided the participant with an environment consisting of 3 rooms: a kitchen, a bedroom, and a garage. Participants were then asked to rate the perceived ease of use of the tool to complete the following goals:

- **Environment Annotation**: adding and annotating an object to an existing environment in order to add interactive properties to it.
- **Create an interactive scenario**: defining the tasks that the user must carry out in the scenario (i.e., go to the bedroom, take the lamp placed on the desk, go in the kitchen, place the lamp on the table).
- **Analysis of the user experience**: questioning and visualizing the recordings created during the scenario in order to observe and analyze the user experience.

Overall, participants found that the tools made the workflow more efficient and responses were positive. Participants with more experience in 3D development were more critical of the flexibility of the tools. They indicated that depending on the scenario and the environment they were trying to create, they could potentially be restricted by the limitations imposed by the tool. We also observed that the most difficult step, the creation and management of a game scenario presented a steeper learning curve, but with notable efficiency. These results have been presented at the conference JFIG 2021 (Les Journées Françaises de l'Informatique Graphique), and has been accepted for publication in PACM HCI and will be presented at ACM EICS 2022.

6 CONCLUSION AND FUTURE WORK

This doctoral thesis aims at the improvement of VR experience creation, analysis and understanding by providing a workflow inspired by Dourish's work [3]. This workflow would help research in multiple domains where the understanding of the user behavior is essential (e.g., IT, health, neurosciences). The next steps on GUsT-3D will be focused on the extension of the DSL in order to create more complex game scenarios with help provided focused on specific metrics like attention of the user with eye-tracking, for example when crossing a road, to make sure the user saw the traffic light beforehand. Records of physiological sensor data will be added to the recorded logs and analyzed by the system in relation with the other data in order to make a link between what the users did in the environment and how they felt during the scenario to provide a snapshot of the embodied experience felt by the user over the time to the designer in the form of: *at the second 22 of the experience, the user was at the sight of a pink flower*. A user experiment will be conducted using the more advanced version of GUsT-3D to test its efficiency, using physiological sensors (skin conductance, heart rate) and motion records (head motion, eye-tracking) to validate our first hypothesis and to see what metrics are shown to be effective in reducing the gap of perception between a user and a designer.

ACKNOWLEDGMENTS

This work has been partially supported by the French National Research Agency through the ANR CREATTIVE3D project ANR-21-CE33-00001 and the EUR DS4H Investments in the Future projects ANR-17-EURE-0004.
Thesis started in October 2021.

REFERENCES

[1] Rozenn Bouville, Valerie Gouranton, Thomas Boggini, Florian Nouviale, and Bruno Arnaldi. 2015. #FIVE : High-level components for developing collaborative and interactive virtual environments. In *2015 IEEE 8th Workshop on Software Engineering and Architectures for Realtime Interactive Systems (SEARIS)*. IEEE, Arles, France, 33–40. https://doi.org/10.1109/SEARIS.2015.7854099

[2] Rory M.S. Clifford, Sungchul Jung, Simon Hoermann, Mark Billinghurst, and Robert W. Lindeman. 2019. Creating a Stressful Decision Making Environment for Aerial Firefighter Training in Virtual Reality. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. 181–189. https://doi.org/10.1109/VR.2019.8797889 ISSN: 2642-5254.

[3] Paul Dourish. 2004. *Where the action is: the foundations of embodied interaction*. MIT press, USA.

[4] Dan Hawes and Ali Arya. 2021. VR-based Student Priming to Reduce Anxiety and Increase Cognitive Bandwidth. In *2021 IEEE Virtual Reality and 3D User Interfaces (VR)*. 245–254. https://doi.org/10.1109/VR50410.2021.00046 ISSN: 2642-5254.

[5] Jason Hochreiter, Salam Daher, Gerd Bruder, and Greg Welch. 2018. Cognitive and Touch Performance Effects of Mismatched 3D Physical and Visual Perceptions. In *2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. IEEE, Reutlingen, 1–386. https://doi.org/10.1109/VR.2018.8446574

[6] Ocean Hurd, Sri Kurniawan, and Mircea Teodorescu. 2019. Virtual Reality Video Game Paired with Physical Monocular Blurring as Accessible Therapy for Amblyopia. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. 492–499. https://doi.org/10.1109/VR.2019.8797997 ISSN: 2642-5254.

[7] Yining Lang, Liang Wei, Fang Xu, Yibiao Zhao, and Lap-Fai Yu. 2018. Synthesizing Personalized Training Programs for Improving Driving Habits via Virtual Reality. In *2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. IEEE, Reutlingen, 297–304. https://doi.org/10.1109/VR.2018.8448290

[8] Ken Pfeuffer, Matthias J. Geiger, Sarah Prange, Lukas Mecke, Daniel Buschek, and Florian Alt. 2019. Behavioural Biometrics in VR: Identifying People from Body Motion and Relations in Virtual Reality. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–12. https://doi.org/10.1145/3290605.3300340

[9] Maximilian Speicher, Katy Lewis, and Michael Nebeling. 2021. Designers, the Stage Is Yours! Medium-Fidelity Prototyping of Augmented & Virtual Reality Interfaces with 360theater. *Proceedings of the ACM on Human-Computer Interaction* 5 (June 2021), 205:1–205:25. https://doi.org/10.1145/3461727

[10] Erwin Wu, Mitski Piekenbrock, Takuto Nakumura, and Hideki Koike. 2021. SPin-Pong - Virtual Reality Table Tennis Skill Acquisition using Visual, Haptic and Temporal Cues. *IEEE Transactions on Visualization and Computer Graphics* 27, 5 (May 2021), 2566–2576. https://doi.org/10.1109/TVCG.2021.3067761 Conference Name: IEEE Transactions on Visualization and Computer Graphics.

[11] Tong Xue, Abdallah El Ali, Tianyi Zhang, Gangyi Ding, and Pablo Cesar. 2021. CEAP-360VR: A Continuous Physiological and Behavioral Emotion Annotation Dataset for 360 VR Videos. *IEEE Transactions on Multimedia* (2021), 1–1. https://doi.org/10.1109/TMM.2021.3124080

[12] Yuhang Zhao, Edward Cutrell, Christian Holz, Meredith Ringel Morris, Eyal Ofek, and Andrew D. Wilson. 2019. SeeingVR: A Set of Tools to Make Virtual Reality More Accessible to People with Low Vision. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. ACM, Glasgow Scotland Uk, 1–14. https://doi.org/10.1145/3290605.3300341