CONVRGENCE PROCEEDINGS

Editor: Simon Richir
Publisher: Laval Virtual
2022, April 12-14
www.laval-virtual.com

VIRTUAL REALITY INTERNATIONAL CONFERENCE
VRIC conVRgence 2022 PROCEEDINGS

FOREWORD

This document presents the proceedings of the VRIC - ConVRgence conference held on April 12-14, 2022. We would like to thank the authors who submitted their research works, as well as the reviewers for their contributions.

Simon Richir, Arts et Métiers Institute of Technology, Scientific Director of Laval Virtual

LAVAL VIRTUAL EUROPE

VIRTUAL REALITY INTERNATIONAL CONFERENCE
VRIC conVRgence 2022 PROCEEDINGS

TABLE OF CONTENTS

The articles included in these proceedings should be cited as follows.

Official reference to the proceedings

Richir S. (2022), Ed. Proceedings of Virtual Reality International Conference (VRIC), 12-14 April, Laval, France, Publisher: Laval Virtual, www.laval-virtual.com, ISBN 978-2-9566-2517-9

Official reference to a paper included in the proceedings

Mai, X.H., Loup, G. and Didier, J.-Y. (2022). A new method of designing an interactive scenario of serious games based on panoramic videos. In *Proceedings of Virtual Reality International Conference (VRIC)*, S. Richir Ed., 12-14 April, Laval, France, Publisher: Laval Virtual, www.laval-virtual.com, ISBN 978-2-9566-2517-9

ART LONG PAPER

Montage as a narrative vector for virtual reality experiences

Sagot-Duvauroux Rémi[1], Ronfard Rémi[1,2], Garnier François[1]

[1]ENSAD, PSL University, EnsadLab, Spatial Media, Paris, France
[2]Univ. Grenoble Alpes, Inria, CNRS, Grenoble INP, LJK, 38000 Grenoble France

Corresponding author: Sagot-Duvauroux Rémi, remi.sagot@gmail.com
Keywords: Virtual Reality – Montage – Narrative – Presence – Interaction – UX Design

Abstract
The goal of our paper is to demonstrate through artistic experimentation that the concept of montage can be an effective vector for conveying complex narratives in immersive, room-scale virtual reality experiences. We review the particularities of virtual reality mediations regarding narrative, and we propose the new conceptual framework of « spatial montage » inspired by traditional cinematographic montage, taking into account the specific needs of virtual reality. We make the hypothesis that spatial montage needs to be related to body action of the user and can play a significant role in relating the immersive experience with an externally authored narrative. To test those hypotheses, we have built a set of tools implementing the basic capabilities of spatial montage in the Unity 3D game engine for use in an immersive virtual environment. The resulting prototype redefines the traditional concepts of shots and cuts using virtual stages and portals and proposes new interaction paradigms for creating and experiencing montage in virtual reality headsets.

1. Introduction

The art of montage, practised and theorised since the beginning of the 20th century, has consequently determined the sets of cinematographic forms, both from an aesthetic and semiotic point of view. Today, Virtual Reality narrative experiences are more and more emerging and provoking a certain reflexivity on their forms and conventions, implying a reconsideration of the concept of montage.

One of the specificities of virtual reality systems is the sense of presence they can provoke on an "immersant"[1]. This sensation of "being there", combining a place illusion and the plausibility of the experience (Slater, 2009), is often considered by designers as an objective of immersive systems (Bowman & McMahan, 2007). Breaks in presence, revealing the seams of the mediation, are considered as factors that diminish the quality of the experience, encouraging designers to favour continuous time-spaces. Putting to the test the sense of presence within the representation space, the practice of montage in virtual reality is therefore quite rare and has not yet enabled the empirical research of an aesthetics of fragmentation and articulation of immersive environments.

Regarding this practical deficiency, we seek to demonstrate, through artistic experimentations, that a narrative power of montage persists in virtual reality mediation. We hypothesise that by linking it to the subject's body actions and playing on the variation of presence in the virtual environment, montage contributes to the evolution of the distance with the mediation, linking the experience lived virtually to the fluctuation of a reminder of the real world. To feed and test these hypotheses, we are developing the "Spatial Montage" system, a tool for staging, cutting and articulating narratives of virtual environments.

[1]Person immersed in a virtual environment.

2. Virtual realities and montage concept

2.1. Presence and discontinuities

Virtual reality technologies are now reaching a certain maturity in terms of reception quality. The screens are better defined, fields of vision are widening, the calculation of images in real time is accelerating, interactions are becoming more precise, the equipment is becoming lighter and less expensive. Artists, authors, designers, producers, engineers and developers can thus take hold of these technologies, hijack them, distort them, question them, and try to offer the public new forms and new uses.

Virtual environments constitute spaces of representation in which the subject is immersed by his proprioceptive senses and with which he can bodily (inter)act through his motor activity (Fuchs, 2018). This results in a phenomenon of presence in the virtual environment, a sensation of "being there" when interacting with a mediated environment (Schubert, Friedmann & Regenbrecht, 1999). Depending on the nature of the experience, the scenography and her engagement, the person immersed can access different depths of presence (Slater, Usoh & Steed, 1994). The stronger the phenomenon of presence, the more the subject can be led to have the same emotions and reactions as in a real experience (Riva, Mantovani, Capideville, Preziosa, Morganti, Villani, Gaggioli, Botella & Alcañiz, 2007) and engage physiological and behavioural responses (Slater, 2003).

Thus, if immersive system, resulting from algorithms that can be programmed at will, are a way of exploring new psycho-perceptive potentialities and new types of spaces with arbitrary properties (Quéau, 1993), they also redefine the relationship of distance between a work of art and a subject and thus open up new fields of practical and theoretical reflection on narration.

Today, there are several trends in the creation of virtual reality narrative experiences. On the one hand, there is the utopia of the perfect simulation allowing for a real-time, continuous experience, where presence would be a primordial qualitative factor. As the purpose defined by Philippe Fuchs indicates, virtual reality is implicitly evaluated by its capacity to reproduce or imitate certain aspects of reality, particularly with regard to the phenomenon of presence and therefore the subject's conviction of being physically “there” in the virtual environment. This often results in the search for transparency of the mediation. In other words, they seek to make the immersed person forget that he or she is dealing with mediated images in order to create an illusion of non-mediation (Lombard & Ditton, 2006).

This transparency, which media theorists Jay David Bolter and Richard Grusin refer to as immediacy (Bolter & Grusin, 1999), enables the creation of controlled simulations and a number of applications in the military, industry, health and of course video games. The subjective first-person point of view has been used since 1974 (with the game *Maze War* [2]) and is today a very popular genre in the video game industry, which explains quite well the ease with which this industry has seized on virtual reality devices which are by definition perceived in the first person. It should also be noted that the main manufacturers of virtual reality headsets have close ties with the video game industry[3]. Thus, the resulting experiences are often in line with the continuity of video game blockbusters, advocating experiential continuity rather than the perception of a discontinuous space of representation.

On the other hand, there is a growing number of artists, often from the film, digital and performing arts, who are trying to break free from this “presence at all costs” in the virtual environment and from the utopia of real world simulation. There are a number of works that demonstrate attempts to use interactive virtual environments as theatres of discontinuous narratives with heterogeneous fragments where the mediation, although immersive, remains quite visible.

---

[2] Maze War, NSI Games, 1974.

[3] The HTC Vive headset was developed by HTC, in collaboration with the video game development company Valve. The Oculus Rift headset’s Kickstarter fundraising campaign was initially entitled "Oculus Rift : step into the game", <https://www.kickstarter.com/projects/1523379957/oculus-rift-step-intothe-. game> accessed 17 march 2022.

This approach of composite assemblages was particularly striking in the "immersive" selection of the *2021 Tribeca Film Festival*[4] and especially in *A Life in Pieces: The Diary and Letters of Stanley Hayami*[5]. Inside a shack, which constitutes a first frame of representation, 2D mixed with 3D animations and even cut-out videos are presented to our perception. The settings regularly change in texture. For example, we go from environments reconstituted in photogrammetry to scenery drawn with the *Tilt Brush*[6] application. This freedom in the heterogeneity of the representational fragments can be found in *Missing Pictures, EPISODE 1: Abel Ferrara's BIRDS OF PREY*[7] where fragments of storyboard are spatialized in the middle of 3D sets. *The Hangman At Home*[8] also mixes 3D walk-through spaces with small theatres in which traditional animation sketches are cut out and spatialized in depth. Another example is *Paper Birds Pt. 1 & 2*[9] where the environment in which we are immersed is all black, leaving us to perceive only the scenery in which the narrative evolves.

2.2. Interactivity and narrative

But if these mentioned works show a richness in the search for new narrative forms, they often come up against the difficulty of coupling narration and interactivity. In the literature studying interactive narration, Ruth Aylett calls the compromise between interactivity and narrative experience the "narrative paradox" (Aylett, 2000). An environment can be fully interactive, but this makes it difficult to control the temporality and content of a particular narrative. In her research, she develops the concept of "emergent narrative" as the result of a dialectic between an author's creation of a "story landscape" which is only actualised and structured around a process of "storification" depending on the engagement of the immersant (Louchart, Swartjes, Kriegel & Aylett, 2008). William Urrichio observes that virtual reality today "builds on the experiential narratives enabled by some game genres far more effectively than on the narrative structures inherited from cinema and literature" (Uricchio, 2018). The freedom of movement and gaze immediately places the subject in a playful relationship with the mediation. In the context of a video game, a particular contract is established with a player, who is ready to put himself in a playful state of exploration, action/reward loop and embodiment.

How then can we combine what seems to be essential in virtual reality, which is a playful state involving bodily actions, while at the same time freeing ourselves from the doxa of immediacy by proposing fragmented time-spaces, allowing us to imagine a diegesis that goes beyond the limits of the experienced simulation?

In a text entitled "Montage and Discourse in Film", Christian Metz explains how, at the beginning of the 20th century, in order to move from the cinematograph to the cinema, the camera needed to "free itself" from the static frame (Metz, 1967). In order to move away from the frontal static shot associated with theatrical representation, the camera had to be mobilised. This liberation could either emerge through the movement of the camera or through the montage. However, history shows that it was the montage that completely prevailed, particularly with the Russian filmmakers and the practices and theories of "montage roi". Christian Metz explains that the reason montage prevailed was that the montage effects were magical, poetic, while the camera movements referred to the filmic nature of the fragment. In Virtual Reality, camera movements are induced by the fact that our point of view is the camera. We must therefore find what would free this point of view and no longer perceive ourselves as a "casked body" (Chatelet & Di Crosta, 2018).

---

[4] https://tribecafilm.com/festival/immersive
[5] *A Life in Pieces: The Diary and Letters of Stanley Hayami*, Nonny de la Peña and Sharon Yamato, Emblematic Groupe, 2021.
[6] https://www.tiltbrush.com/
[7] *Missing Pictures, EPISODE 1: Abel Ferrara's BIRDS OF PREY,* Clement Deneux, Atlas V, Arte France, BBC, Serendipity Films, 2020.
[8] *The Hangman At Home*, Michelle & Uri Kranot, Late Love Production, Floréal Films & ONF, 2020.
[9] *Paper Birds Pt. 1 & 2*, German Heller, Federico Carlini, 3DAR Studios, Baobab Studios, Oculus, 2021.

2.3. Movements and montage

Many virtual reality narrative works attempt to reconcile the phenomenon of presence, interaction and narration through the participation of the immersant as a diegetic character. Either continuously, as in *Amends*[10], or more often in a more fleeting and detached way, as in *The Book of Distance*[11], where, without embodying a character, interactions with the virtual environment are proposed to us in order to engage ourselves physically and emotionally in the world of the story while keeping a certain distance from it. It seems that it is precisely this distance that allows us to accept extra-dialectical spatio-temporal discontinuities (belonging to a narrative structuring, not perceived by the characters). What if our interactions, rather than giving us a role in the diegesis, gave us a role in its functioning, in its workings, in its rhythm, in its montage?

In this dynamic, we hypothesise that the immersant should be thought of as a battery for the narrative, as a crank that operates the gears of the narrative structure. His/her role would thus be more evident and he/she would have a real corporeal perception of the narrative without necessarily having to be present in the spaces of the diegesis. This body, both spectator and actualiser, would be a sort of evolution of the concept of kino-eye proposed by Dziga Vertov and the Kinoks (Vertov, 2018). A body-machine, present yes, but in a meta way as an instrument for setting a dramatic structure in motion. The phenomenon of presence would then make it possible to inhabit a narrative space at a fair distance from the diegetic entities. However, this intuition comes up against the fact that in order to mentally immerse ourselves in a story, we need to detach ourselves from our consciousness of the space-time in which our body evolves. How can we then combine the essential involvement of bodily action with mental projection into the space-time of a diegesis?

We hypothesise that it is precisely the editing that will allow us to forget the use of this body-machine as an instrument of capture. In the same way that a long sequence-shot makes us perceive the movement of the camera as an aesthetic element at a given moment, a too long space-time in virtual reality reminds us of ourselves as immersed in the process of moving. The montage should therefore allow our body-machine the necessary time and space to activate the world while orchestrating the breaks necessary to distance ourselves from this machination and maintain our mental immersion in the story. The point is therefore to move from the "body-camera" to the " body-montage".

For the Kinoks, cinema enables us to perceive the world through the eyes of the machine, which surpasses human imperfections. Focusing on the essence of the technique, they aim that the understanding of the internal movement of the cinematographic machine can link people and machines in a "creative joy" that they call "kinokism": "the art of organising the essential movements of things in space and time into a rhythmic artistic ensemble, in accordance with the properties of the material and the inner rhythm of each thing". For them, it is the intervals (between two movements) that constitute the material that "drives the action towards the kinetic denouement" (Vertov, 2018). If we follow this theory, the quality of the editing in virtual reality should be highly impacted by the quality of the movement of the person experiencing it. The kinetic power of the montage would then be proportional to the energy deployed by the viewer's movements.

If the movements are to influence the quality of the montage, in what way do the possibilities of rupture could have an influence on our movements?

In a movie theatre, the spectator accepts to be placed in an environment where his body must remain still, his gaze directed towards a screen. He can certainly decide on his distance from the screen, but this will remain fixed for the duration of the film. Through the frame and the montage, his vision is organised by the filmmaker, who imposes a distance with the filmed elements. In virtual reality, the body's distance from things depends in part on the movements of the immersant. So, how can the immersed body move if it knows that it is evolving in a universe affected by spatio-temporal discontinuities?

---

[10] *Amends,* Mikkel Battefeld, The Animation Workshop (DK), 2020.

[11] *The Book of Distance,* RANDALL OKITA, NFB CANADA, 2020.

2.4. Discontinuities and UX Design

Considering spatial breaks in virtual reality comes up with certain cognitive constraints related to the perception of virtual environments involving spatial reasoning (Nnadi, Fischer, Boyce & Nitsche, 2008). There is a difficulty in repositioning oneself in a new virtual environment (Amorim, Trumbore & Chogyen, 2000) and a need to give the user certain visual cues so that he/she can orientate him/herself (Steiner & Voruganti, 2004). It must therefore be taken into account that in a synthetic virtual environment, sensorimotor immersion can indeed open up new relationships between perceptions and body movements, but inconsistencies between our different senses can be a source of physiological discomfort. This partly explains why virtual reality systems base their designs on certain rules of our human sensory-motor schema before exploring borderline cases.

Several studies have been conducted on transitions between virtual spaces (Kohn & Rank, 2016; Oberdörfer, Fischbach & Latoschik, 2018; Husung & Langbehn, 2019). They show the influence of the visible or invisible, predictable or unpredictable, chosen or undergone aspect of the transition on the sensation of presence and the subjective appreciation of the user. They show that the engagement of the body in the mediated space implies a profound reconsideration of the way these transitions are designed if we want to take into account a certain experiential continuity centred on the user. While these researches are mainly based on the impact of transitions on the phenomenon of presence, we seek to study the link between these discontinuities and the distance they may create between the user's experience and the represented story world.

We therefore engage in a study of the narrative potential of a form of montage in virtual reality linked to a body-machine of the immersant, engaging his/herself at a precise distance from a diegesis previously spatially structured by a creator. Based on this theoretical positioning, we are developing an artistic and experimental virtual reality system entitled "Spatial Montage" acting as a reflexive process for our research questions.

3. Experimental system: Spatial Montage

3.1. A research-creation process

We propose a research-creation process inspired by the practice of film editing to consider a new form of UX design for virtual reality narrative works. The aim is to create a tool that enables us to reconstruct the oscillation of the editor between creation and enunciation, between the space of fabrication and that of diegesis. However, if this oscillation is essential in the practice of montage, it is mainly because enunciation standards are strongly rooted in our image culture. But in virtual reality, the standards that emerge are still fragile and year, creators propose new borderline cases. It seems therefore important, in the articulation between each practice and theory, to distinguish the study of creation from that of enunciation. We propose to build a tool for the study of both creative techniques and their influence on the perception of the meaning expressed.

The "Spatial Montage" system offers an editing tool in virtual reality. It is developed with the Unity 3D game engine[12]. It is a 6 dof (degrees of freedom) device using the HTC Vive Pro Eye headset and its controllers[13]. Thanks to a virtual interactive table, it allows the arrangement of visual, sound and proprioceptive elements in space and time for a narrative and interactive experience in virtual reality. It also sets the basis for a study in narratology and design on the quality of fragmented narratives experienced by an immersant.

Link toward a video presentation of the system : https://vimeo.com/660613582

[12] https://unity.com/fr
[13] https://www.vive.com/fr/product/vive-pro-eye/overview/

3.2. Creation and fragmentation

The first objective of the system is to define the fragments enabled for the creator in order to set his/her staging. The functionalities enabling the manipulation of these fragments, their juxtaposition in time and space and their relationships with each other are therefore the keystone of the system.

First, the creator works fully immersed in the virtual environment, reducing the cognitive load of decoding a 2D interface. He/she composes the space directly in the headset with its controllers in proprioceptive immersion, using low-level cognitive functions. By this way, it makes it possible to imagine the representation space while manipulating it with a top level viewpoint above the model. We thus combine the ego, hetero, object and allo-centric representation systems.

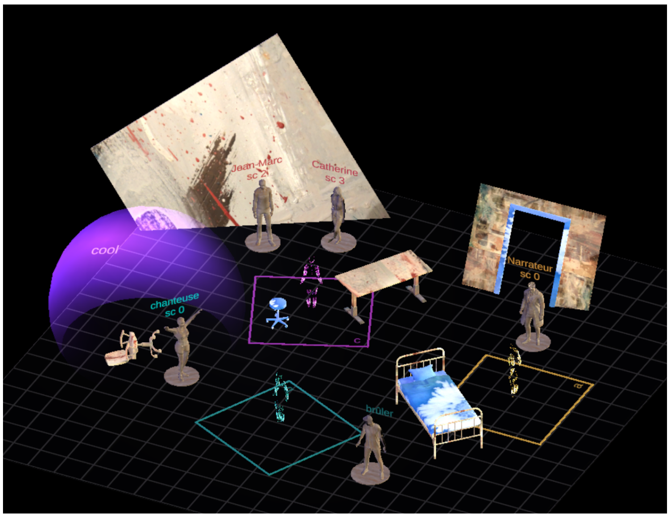


Figure 1: ***The virtual table***

On a virtual table (Figure 1), we can assemble blocks to structure the geometry of a set. We can also manipulate small characters. A word written above each character refers to its text or action. Sound bubbles (coloured transparent spheres), corresponding to sound atmospheres or music, can be placed within the set.

Then, in order to envisage the montage in virtual reality, it is necessary to question the nature of these fragmentary sets to be assembled in time and space. What would be the molecular units to be connected to structure the spatio-temporal unfolding of a virtual reality experience? In cinema, the frame circumscribes the space of the representation and separates it from the spectator's physical environment. Sergei Eisenstein describes the frame as an editing cell (Eisenstein, 1976). According to Igor Babou, in virtual reality, the concept of the frame, in its semiotic status, does not disappear but is shifted (Babou, 1999). The limits of the representation space take on another form, both in their diegetic (what we are given to perceive) and indicial (reference to a larger form) functions. How then can we define these limits and thus establish editing cells?

The perceptual boundaries of a continuous virtual reality experience are based on the subject's freedom of movement in the real environment. The perimeter of the physical space tracked by the device therefore seems to correspond to a frame. In the virtual environment, it defines a scaled, spatialized and oriented zone in which the subject acquires a limited walking and acting area. We therefore decide to take this zone of potential action as our editing cell. The frame in virtual reality no longer delimits a portion of the image as in cinema, but a zone of possible actions in a space.

On our virtual table, we can therefore move and orientate three zones (a, b and c) representing the spaces to be explored when immersed in the set. The sizes of these zones are relative to the area we can walk through in the physical space. In each zone, we can see the location of our relative position thanks to a small ghost that moves and orients itself according to our physical movements. By manipulating these zones, we orientate the virtual spatial geometries according to the limits of the bodily action of the person who will be immersed in our narration.

3.3. From creation to reception

The second objective is to allow an artist to fragment and articulate his/her narrative while being able to experience its reception at any time. So, when we look at an area and press the thumb button, we quickly smoothly change scale, position and orientation and find ourselves in the place of the little ghost (Figure 2). The light changes. The model set becomes our environment. The little characters become human-sized, change texture, activate and speak according to a system of distance cues (which we will discuss later). If we find ourselves in the location of a sound bubble, we hear the soundscape whose level varies according to our proximity to the centre of the sphere.

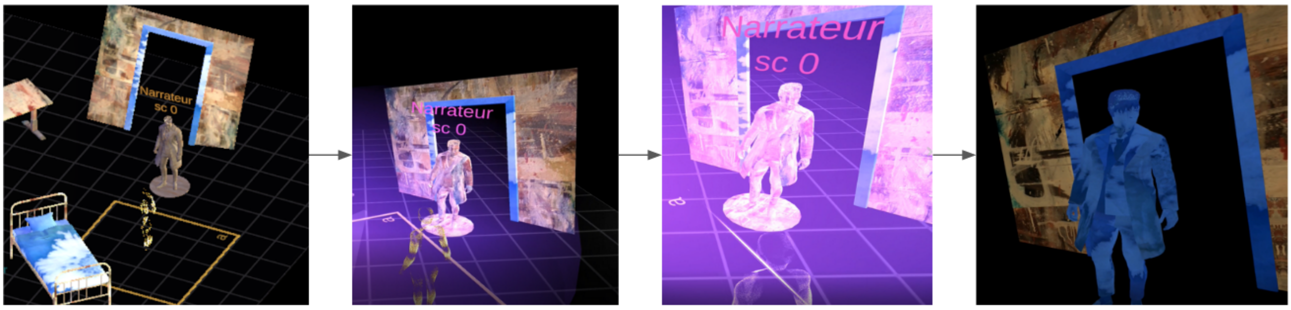


Figure 2: ***Passage from the point of view of the creator to the one of the receiver.***

Thus, we quickly switch our status (from creator to receiver) and our point of view (from third to first person) without experiential discontinuity. By pressing the thumb button again, we can go back to the creator position. This function enables us to experience in real time the evolution of the narrative form in construction. In this way, we attempt to reproduce the rhythmic and playful aspect of film editing.

3.4. Activation of the narrative

The third objective is to test our hypothesis of a montage intimately linked to our bodily actions in order to generate meaning from the coupling between our movements and our narrative perception. Two strategies can be followed. The first is to assume that the immersant will move according to the virtual scenography. In this case, in order to avoid him/her being “misplaced”, the creator's job is to organise the events in space and time in the aim to attract the immersant's attention and to influence his wandering movements. Another strategy, on the contrary, is to make the environment evolve according to the movements of the immersant. According to his/her gaze, the speed of his/her body movements, his/her position in space, the environment evolves, fractures and articulates itself. In this case, it is the immersant who "puppeteers" the virtual environment and more generally the narrative. This device must therefore be able to analyse those two cases and see how different perception/action couplings can provide new narratological paradigms.
So, in order for the actualization of the narrative experience to be favourable for the receiver, taking into account the desires of the creator, the virtual environment must react according to a system of cues defined beforehand. The two variables we can play with are the movements of the immersant and the characters' animations in the virtual environment. These data are therefore monitored in real time in order to serve as variables for the cues that trigger the narrative sequences in space and time.
Each character has an activation sphere. When this sphere collides with the immersant's sphere (also linked to the point of view), the character changes texture and becomes active. Depending on the character's narrative role, the activation sphere can be larger or smaller. Thus, some characters will only come to life in the intimate sphere of the immersant, while others will only be activated from a large distance and will freeze

if they get too close. Depending on the field of vision and the immersant's distance from the elements, the characters come to life, slow down or freeze. In a sensory-motor loop, the environment is puppeteered by the bodily actions of the immersant, which are themselves influenced by the scenography of the unfolding environment.

3.5. Transitions between zones

The final issue is how to move from one zone to another? And how can the creator make him/herself visible or not in this space discontinuity? Initial tests were done by placing a teleporter in each zone. They became invisible when we were immersed in the set. As soon as we arrived in the teleporter's position, we were suddenly teleported (like a cinematographic "cut") to the next zone. The fundamental problem with this mechanism was that, having previously consciously placed the teleporter on the set, we knew where it was and were searching for its location. So we were biased concerning the probability of an unaware immersant being in the right place.

We then tested teleportation not on the basis of spatial but time data. Depending on the length of the characters' dialogues present in the zone, a calculation was made to give a specific time for the zone before being teleported to the next. But as they did not correspond to any particular event or movement, these cuts were often too sudden and meaningless for the immersant. In addition, the relationship between his/her position and orientation before and after the cut became arbitrary.

We also tested the triggering of teleportation once a dialogue ended. Since the characters' dialogues are only activated when you look at them, we were able to determine roughly where the immersant was going to be before the cut, and thus set up the arrival area accordingly. The fact that the characters only speak and move when you look at them requires the viewer to be in a particular position and orientation if they are to activate and follow the action. Thus, when the cut occurs, the immersant becomes aware of the composition of the image he/she is following. In being aware of the transition from one composition to another, he/she seems to shift briefly from a spatial perception to a pictorial perception. In this fraction of a second, he/she loses the perception of his/her body and presence in the virtual environment. A movement of distancing takes place and the presence of a creator appears.

Finally, the last case we find really relevant is that of the portals. Within each zone, we place and orientate a small frame (Figure 3). The image of this frame is visible in the other zones. This association of the frame and its image constitutes a portal between the zones and we call it a "wormhole". Inside the set, it appears to the immersant if he/she approaches it and by passing his/her head through, he/she changes zone (Figure 4). In this case, the choice to move from one area to another is left entirely to the immersant, who must make the necessary movement to pass. He/She is obviously in control of the transition moment, but his/her movements must conform to the placement and orientation of the frame planned by the creator. In order to move from one zone to another, the immersant must consciously adopt the extra diegetic narrator's point of view whose presence is therefore signified during the transition.

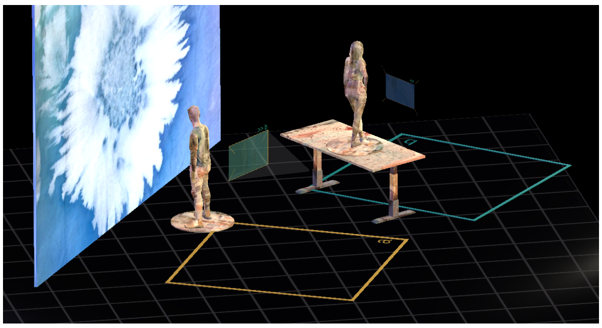


Figure 3 : Wormhole set by the creator in third person view.

Figure 4 : Wormhole experienced in the narrative in first person view.

4. Discussion and conclusion

The Spatial Montage system is complete and effective, and sets the basis of a method for structuring a fragmented narrative space reacting to the bodily actions of an immersant. Preliminary experiments suggest that wormholes are especially relevant for triggering teleportations through bodily movements, allowing active participation in the ongoing story. By requiring consent from the immersant, this also seems to improve the acceptability of the teleportations and to reduce discomfort and disorientation.

The Spatial Montage system is currently being used as part of the "Lines of universe"[14] artistic installation, where participants can create and experience short narratives using a triptych form with three zones and three portals, as shown in Figure 1. The three zones can for example reproduce the Aristotelian notion of an atomic story sequence with a beginning, a middle and an end. But it can also be the basis for looped non-linear stories. This is a useful and generic set-up for experimenting with a variety of film idioms and patterns, and evaluating their effectiveness in the context of virtual reality. Typical examples include navigation between scenes, navigation between characters, temporal ellipses and flashbacks. We are actively investigating those cases as building blocks for assembling more complex narrative structures.

In future work, we plan to augment our empirical, subjective auto-evaluation of the quality of the discontinuities with a more quantitative analysis of the effectiveness of the proposed montage for a naive audience, not involved in the making of the montage. This will require extensive user studies based on an evaluation protocol which still needs to be defined. Future work is also needed for monitoring the immersant's actions during the experience, measuring the progress of the story and connecting atomic story sequences together into long-form narratives and non-linear, branching interactive stories.

Finally, this future research will have to integrate several identified issues: According to which modalities and contracts can the author make the immersant accept a rupture of the experience in favour of the narrative? In what way, precisely by influencing the phenomenon of presence, do discontinuities acquire a narrative role? And how can the spatio-temporal articulations resulting from discontinuities place an immersant at different distances from the diegesis and its author?

5. References

Amorim, M.-A., Trumbore, B., & Chogyen, P. L. (2000). Cognitive Repositioning inside a Desktop VE: The Constraints Introduced by First-versus Third-Person Imagery and Mental Representation Richness. *Presence: Teleoperators and Virtual Environments*, *9*(2), 165–186. https://doi.org/10.1162/105474600566709

Aylett, R. (2000). Emergent narrative, social immersion, and "storification." Proceedings of the 1st International Workshop on Narrative and Interactive Learning Environments, 35-44.

Babou, I. (1999). *L'absence de cadre comme utopie des "réalités virtuelles"*. Champs visuels. Revue interdisciplinaire de recherches sur l'image. L'Harmattan.

Bolter, J. D., & Grusin, R. A. (1999). *Remediation: Understanding new media*. MIT Press. https://doi.org/10.1108/ccij.1999.4.4.208.1

[14] https://remisagduvr.wordpress.com/327-2/

Bowman, D. A. & McMahan, R. P. (2007). Virtual Reality: How Much Immersion Is Enough? *Computer* 40, 36–43. doi:10.1109/MC.2007.257

Chatelet, C. & Di Crosta, M. (2018). Écran incorporé, corps casqué. *Interfaces numériques*, 7 (2), 264-285.

Eisenstein, S. (1976). *Le film: Sa forme, son sens*. C. Bourgois.

Fuchs, P. (2018). *Théorie de la réalité virtuelle: Les véritables usages*. Mines Paristech PSL.

Husung, M., & Langbehn, E. (2019). Of Portals and Orbs: An Evaluation of Scene Transition Techniques for Virtual Reality. *Proceedings of Mensch Und Computer 2019 on - MuC'19*, 245–254. https://doi.org/10.1145/3340764.3340779

Kohn, J., & Rank, S. (2016). You're the Camera!: Physical Movements For Transitioning Between Environments in VR. *Proceedings of the 13th International Conference on Advances in Computer Entertainment Technology - ACE2016*, 1–9. https://doi.org/10.1145/3001773.3001824

Quéau, P. (1993). *Le virtuel: Vertus et vertiges*. Champ Vallon ; Institut national de l'audiovisuel.

Lombard, M., & Ditton, T. (2006). At the Heart of It All: The Concept of Presence. *Journal of Computer-Mediated Communication*, *3*(2), 0–0. https://doi.org/10.1111/j.1083-6101.1997.tb00072.x

Louchart, S., Swartjes, I., Kriegel, M., & Aylett, R. (2008). Purposeful Authoring for Emergent Narrative. In U. Spierling & N. Szilas (Eds.), *Interactive Storytelling* (Vol. 5334, pp. 273–284). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-540-89454-4_35

Metz, C. (1967). "Montage" et discours dans le film: Un problème de sémiologie diachronique du cinéma. WORD, *23*(1–3), 388–395. https://doi.org/10.1080/00437956.1967.11435493

Nnadi, O., Fischer, U., Boyce, M., & Nitsche, M. (2008). Effect of dynamic camera control on spatial reasoning in 3D spaces. *Proceedings of the 2008 ACM SIGGRAPH Symposium on Video Games - Sandbox '08*, 157. https://doi.org/10.1145/1401843.1401873

Oberdörfer, S., Fischbach, M., & Latoschik, M. E. (2018). Effects of VE Transition Techniques on Presence, Illusion of Virtual Body Ownership, Efficiency, and Naturalness. *Proceedings of the Symposium on Spatial User Interaction - SUI '18*, 89–99. https://doi.org/10.1145/3267782.3267787

Riva, G., Mantovani, F., Capideville, C. S., Preziosa, A., Morganti, F., Villani, D., Gaggioli, A., Botella, C., & Alcañiz, M. (2007). Affective Interactions Using Virtual Reality: The Link between Presence and Emotions. *CyberPsychology & Behavior*, *10*(1), 45–56. https://doi.org/10.1089/cpb.2006.9993

Schubert, T., Friedmann, F., & Regenbrecht, H. (1999). Embodied Presence in Virtual Environments. In R. Paton & I. Neilson (Eds.), *Visual Representations and Interpretations* (pp. 269–278). Springer London. https://doi.org/10.1007/978-1-4471-0563-3_30

Slater, M., Usoh, M., & Steed, A. (1994). Depth of Presence in Virtual Environments. *Presence: Teleoperators and Virtual Environments*, *3*(2), 130–144. https://doi.org/10.1162/pres.1994.3.2.130

Slater, M. (2003) A note on presence terminology. *Emotion*, 3, 1–5.

Slater, M. (2009). Place illusion and plausibility can lead to realistic behaviour in immersive virtual environments. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1535), 3549–3557. https://doi.org/10.1098/rstb.2009.0138

Steiner, K.E., & Voruganti, L. (2004). A comparison of guidance cues in desktop virtual environments. *Virtual Reality, 7*, 140-147. https://doi.org/10.1007/s10055-004-0125-1

Uricchio, W. (2018). "VR: Between Hope, Hype, and Humbug." *Los Angeles Review of Books*, June. https://lareviewofbooks.org/article/vr-between-hope-hype-and-humbug/.

Vertov, D (2018). *Le ciné-oeil de la révolution, écrits sur le cinéma*. les Presses du réel.

SHORT PAPERS

A new method of designing an interactive scenario of serious games based on panoramic videos

Xuan Hien Mai[1], Guillaume Loup[1], Jean-Yves Didier[1]

[1]Université Paris-Saclay, Univ Evry, IBISC, 91020, Evry-Courcouronnes, France

Corresponding author: Xuan Hien Mai, maixuanhien.vn93@gmail.com
Keywords: interactive 360° video – serious game – authoring tool – finite state machine

Abstract

The increasingly low prices of virtual reality devices as well as omnidirectional cameras are stimulating the development of serious game applications using interactive 360 videos for the entertainment and training industries. While 360° video production is getting easier and faster, on the other hand, developing interactive content for scenarios of serious games remains time-consuming and manual, and requires a lot of programming knowledge. The lack of an interaction design standard based on a defined architecture can create misunderstandings between game designers and game developers, also making the development process costly and resource-intensive.

As a case study, we present a new design approach dedicated to serious games developed based on interactive panoramic videos which have the special control requirements in both temporal and spatial terms. This design will be introduced as a finite state machine (FSM), suitable for developing multithreaded scenarios.

1. Introduction

Omnidirectional cameras are becoming growingly popular, so producing 360° videos is easy and does not require specialist knowledge or complicated operations. Virtual reality headsets are also becoming cheaper and more accessible to the general public. Moreover, viewing 360° videos on a head mounted display (HMD) device provides a better sense of presence [3].

Traditional 360° videos are enriched by supporting additional interactive elements such as text information, sounds, 2D/3D objects, quizzes and so forth. These types of videos are transforming themselves into interactive 360° videos that not only improve user experiences, but also increase narrative and educational elements. These content-enriched interactions are typically structured to conform to the spatial and temporal dimensions of 360° video. That means they have a scripted fixed behavior. In addition, some other interactions that are based on user behavior, such as those that suggest vision navigation, will behave dynamically. Therefore, unlike editing tools of 2D videos that primarily deal with the temporal dimension, creating interactive 360° videos requires a whole set of tools to handle the combination of spatial and temporal dimensions.

In the production of serious games in general and educational applications in particular, the storyline is extremely important. A number of models have been proposed to help define the main components of this type of application; however, there are additional factors that need to be specifically considered when developing serious games based on the main content of interactive 360° videos.

In this article, we rely on *scenario-based design* [2] approaches to propose a design approach specific to serious games that uses interactive 360° videos and identifies the core components of a system architecture. In the future, we will explore the interactive modules and tools necessary for a useful authoring tool in this context, where an author can create his application while being totally immersed in virtual reality.

2. Related work

2.1. Interactive 360° videos

Using a virtual reality headset to watch 360° videos allows users to be completely immersed in a virtual scene and see the videos in the direction of their choice by turning their head [1]. Regarding the design tools for 360° videos in virtual reality, the usability of the rapid object creation solution was evaluated [9] through tasks such as adding and removing 3D objects in space and time. The panoramic view provides high-fidelity environments [11] and contributes to a sense of presence [6]. Although much of the work consists of simple interactions, an experiment was performed to assess the continuity of the integration of video animations, 3D interactive objects as well as 3D audio [4]. For the design of immersive panoramic videos, the designer's assistance is mainly limited to primitive virtual behavior techniques such as selection [7]. When it comes to game mechanics, interaction techniques are also limited to simple tasks such as managing a score [5].

2.2. Serious game designs

Designing an effective and enjoyable serious game requires not only knowledge of learning domain or pedagogy, but also an understanding of scenario design and game design components. In this context, several models of design have been proposed. Stephen Tang and Martin Hanneghan [10] have proposed a *Game Content Model* which provides definitions of design structure, helping to shed light on important elements of serious game design. Yen-Ru Shi and Ju-Ling Shih [8] investigated 11 important elements of game design, namely purpose, mechanics, fantasy, value, interactivity, freedom, storytelling, sensory, challenge, sociability and mystery. These studies are of great help in the analysis and development of architectural archetypes of serious educational games.

Concerning design methodology, in order to increase the educational value of serious games, it is necessary to add mechanisms that allow the player to interact with the environment. These additional interactions enhance the user's sense of presence as it creates a sense of control and real participation in the scenario [16]. In addition, the content of serious games should be divided into several difficulty levels. The access will be granted according to the results of the previous level [12]. In other words, it is an adaptation that allows the playing environment to change to suit the style or skill level of the player. This further confirms the need to take adaptive mechanics into serious gaming. For this purpose, the scenario of the game can be thought of as a state transition system (finite state machine) and there will be evaluation and adaptation mechanisms based on checking and modifying specific states of the system [13].

2.3. Scenario of serious game based on 360° video

Videos in general and panoramic videos in particular are narrative in themselves. Their content was determined before the production process [14]. In other words, the video itself is part of the serious game storyline. 360° videos can be seen from multiple angles (active) instead of just being observed from the director's point of view (passive) [6]. Therefore, additional interactions are needed to get the player to follow the correct progression of the storyline and watch the video in the right direction at critical times.

Not only the spatial factor, but also the temporal factor must be evaluated when considering the effect of additional interactions on the timeline of the video, and more broadly their effects on the entire timeline of the scenario. The studies by Toro, M. [15] initially confirmed the viability of fixed-time and flexible-time models in conditional branching scenarios. However, they must have certain conditions to be maintained in the scenario.

An example of a scenario is the virtual visit of a museum composed of interactive panoramic videos of each room presented by a guide. The aim is to make a visit more accessible and more adapted to the preferences of each visitor. Thus, it is possible to navigate from room to room quickly and effortlessly. In addition, interactive buttons positioned on the video allow the visitor to get additional information on each work. 3D areas are also positioned on the video to highlight the work evoked by the guide. Finally, a questionnaire can also appear at the end of the video to measure the visitor's understanding.

We examined a few authoring tools that support 360° video and multi thread scenarios to learn about their programming knowledge requirements and operational complexity. Table 1 below lists some typical tools for VR applications, which have been analyzed qualitatively and statistically, including: Amazon Sumerian[1], InstaVR[2], Dataverse[3] and FlowMatic [18]

Table 1 : authoring tools for VR experiences

	360° video support	Authoring tool platform	Multi-thread scenario support	Progamming knowledge	Complexity
Amazon Sumerian	yes	PC	--	high	high
InstaVR	yes	PC	no	no need	medium
Dataverse	yes	PC	no	no need	--
FlowMatic	no	VR	yes	medium	high

There are very few authoring tools built directly on the virtual reality platform but mainly still authoring tools on PC. The simulated viewpoint (on the 2D screen) of these tools therefore deviates from the actual viewpoint in VR, which can lead to deviations in spatial positioning when designing interactions.

In order to be able to design standard tools suitable for a no-code immersive authoring tool, it is first necessary to have a standard method of design that describes in detail the components of an interactive scenario of a serious game based on 360° videos.

3. Method of design towards a no-code immersive authoring tool

Towards the goal of building an intuitive and easy-to-use immersive authoring tool that allows non-programming savvy authors to easily develop their serious game apps based on 360° videos, firstly, we will introduce in this paper a new method of design for interactive scenario that we find appropriate, considering the characteristics of serious games as well as the special control requirements in both temporal and spatial terms for interactive 360° videos.

In the architectural design for serious games based on 360° videos that we come up with, the most basic unit is the scene (Figure 1). Each scene includes a segment of the video with additional interactions (text, sound, 2D/3D objects and so on.). The video segment is spanning between two time values and therefore additional interaction configuration time parameters are bounded by these two values.

Linear video time is irreversible. Therefore, a special object will allow the video to be paused at a specific time. Interactions installed in this object when the video is paused will behave differently from linear interactions. This design guarantees the possibility of controlling the temporal in the scene, allowing the addition of many interesting gameplay such as answering quizzes without any time limit.

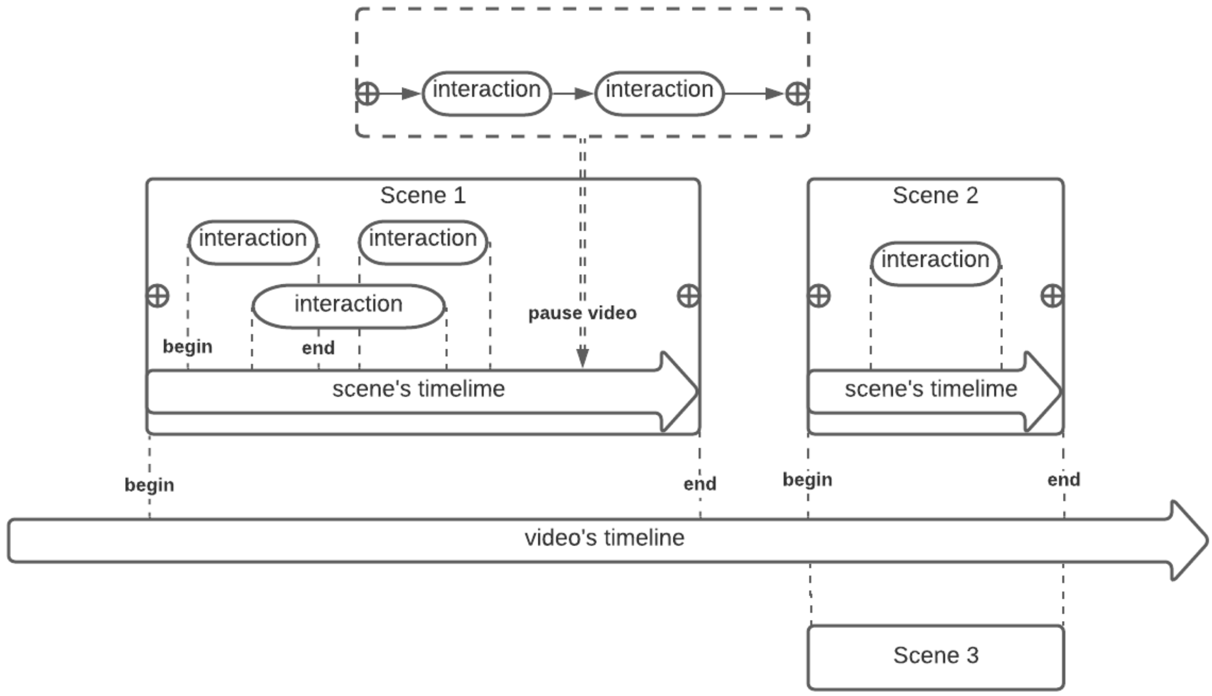


Figure 1 : The elements in the scene based on 360° video

This design is distinguishable by its ability to visually depict the occurrence of interactions on the video's timeline, suggesting to the author about the layout arrangement of objects and interactions in the scene, helping to develop a better scenario structure.

The goal is to develop an immersive authoring tool, so the spatial parameters will be adjusted by the author directly in the virtual environment. A preview mechanism will be provided so that the author can "try it out" from the perspective of the end user.

This is essentially a timeline tree model [17] in the interactive scenario. Scenes 2 & 3 are transitions after scene 1 but on the same video. In the design of the operating system (application system), the two *Data Manager* and *Scene Manager* modules will work independently but cooperatively to ensure the seamlessness of the video content as well as the multi-threading of the scenario. This architecture also helps to save video storage resources as multiple scenes can be edited on the same video. The author does not need to go through the pre-production process to cut and edit the video for each individual scene.

Using scenes as the basic material for scenario building, we have 3 basic components of an interaction scenario: the scenario segment, the transition and the condition.

Scenes will be placed in scenario segments and can be re-used in the same segment or between segments, depending on the scenario design. The *transitions* between the segments of a scenario will be made after evaluation of the results on the basis of the binding *conditions*. Figure 2 will explain the transition state of the finite state machine (FSM) in the system.

All player actions and interactions in the game will be logged in the event log. The analysis of event log data will help determine the state of the system. However, it is not necessary to define the entire state at each evaluation. This design allows the author to build his own evaluation rules at each transition point on the scenario, with a finite number of events analyzed.

The output of each evaluation will be a binary result leading to two different branches of the scenario. The whole scenario will have the structure of a binary network. Therefore, it is necessary to develop algorithms to detect infinite loops in the scenario.

A segment can also reside within another segment, which increases the possibility of designing multi-threaded scenarios. But at the same time this multi-tier architecture also poses an ergonomic design challenge on the development of authoring tools.

The design approach we propose has integrated a FSM, suitable for interactive scenarios where user interactions in the scene will affect scenario progress.

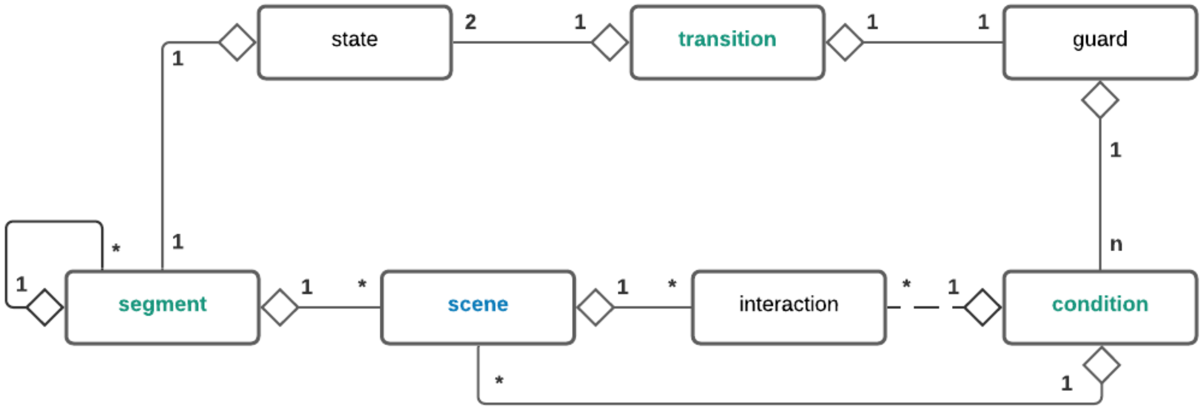


Figure 2 : The transition state of FSM in the system

A loose link between *condition* and *interaction* means that the evaluation conditions of the scene will be analyzed based on the results of user interactions in the scene. It could be for example the results of answering the questions, the time it took to answer the questions, or the total time spent in the scene and the like.

Thus, we have just introduced an overview of the design approach for the interactive scenario of serious games based on panoramic videos. This approach is suitable to help the author develop his game by describing scenarios using storyboards. It should be noted that for this part of the architecture, the multithreading scenario is purely mechanical, which means that the system has a finite state and the state changes are only based on the results of real-time user interaction. In the future, a larger architecture based on adaptive learning models [21][22] will be developed, using the resources of this design approach in combination with artificial intelligence techniques, for example machine learning and deep learning to analyze data profile and user behavior to offer appropriate suggestions.

4. Conclusion

This article proposes a new method of design that enables authors to build serious gaming applications based on interactive 360 videos. The proposed model follows scenario-based design, which is the result of a serious analysis of different game architectures, considering the particular properties of 360 video and additional interactions.

Encapsulation of video segment within the scene ensures the integrity of the story in the video. This design gives the author more control over the video's timeline. Additionally, using video segments instead of the entire video, along with FSM integration, makes it easier for authors to build multi-threaded scenarios with additional interactions in the scene.

Based on this architecture, we are developing an immersive authoring tool that allows users to create applications directly in a virtual reality environment without the need for programming knowledge. Our challenge was to study ergonomics in order to develop a suitable storyboard interface in virtual reality.

In the next stage, based on the scenario classification [19][20], after finalizing the authoring tool, we will design and conduct certain appropriate experiments to test the adaptability of this design to different scenario types.

In the future, we will study the possibility of integrating artificial intelligence as well as machine learning, deep learning to complete the adaptive learning model for this architecture.

5. Acknowledgments

This project was supported by the ANRT (PhD fellowship) and the WideWebVR company (CIFRE #2020/0813).

6. Notes

[1] https://aws.amazon.com/sumerian

[2] https://www.instavr.co

[3] https://dataverse.xyz

7. References

[1] Oliver, A., del Molino, J., & Bibiloni Coll, A. (2020, June). Palma360: A 360-degree Web-based Multi-device Hypervideo. In *ACM International Conference on Interactive Media Experiences* (pp. 165-169). doi:10.1145/3391614.3399394

[2] Carrol, J. M. (1999, January). Five reasons for scenario-based design. In *Proceedings of the 32nd Annual Hawaii International Conference on Systems Sciences. 1999. HICSS-32. Abstracts and CD-ROM of Full Papers* (pp. 11-pp). IEEE.

[3] Chambel, T., Chhaganlal, M. N., & Neng, L. A. (2011, November). Towards immersive interactive video through 360 hypervideo. In *Proceedings of the 8th international conference on advances in computer entertainment technology* (pp. 1-2). doi:10.1145/2071423.2071518

[4] Choi, K., Yoon, Y. J., Song, O. Y., & Choi, S. M. (2018). Interactive and immersive learning using 360 virtual reality contents on mobile platforms. *Mobile Information Systems*, *2018*. doi:10.1155/2018/2306031

[5] Argyriou, L., Economou, D., Bouki, V., & Doumanis, I. (2016, December). Engaging immersive video consumers: Challenges regarding 360-degree gamified video applications. In *2016 15th International Conference on Ubiquitous Computing and Communications and 2016 International Symposium on Cyberspace and Security (IUCC-CSS)* (pp. 145-152). IEEE. doi:10.1109/IUCC-CSS.2016.028

[6] Violante, M. G., Vezzetti, E., & Piazzolla, P. (2019). Interactive virtual technologies in engineering education: Why not 360° videos? *International Journal on Interactive Design and Manufacturing (IJIDeM)*, *13*(2), 729-742. doi:10.1007/s12008-019-00553-y

[7] Horst, R., Diez, S., & Dörner, R. (2019, October). Highlighting Techniques for 360 Video Virtual Reality and Their Immersive Authoring. In *International Symposium on Visual Computing* (pp. 515-526). Springer, Cham. doi:10.1007/978-3-030-33720-9_40

[8] Shi, Y. R., & Shih, J. L. (2015). Game factors and game-based learning design model. *International Journal of Computer Games Technology*, *2015*. doi:10.1155/2015/549684

[9] Adão, T., Pádua, L., Fonseca, M., Agrellos, L., Sousa, J. J., Magalhães, L., & Peres, E. (2018). A rapid prototyping tool to produce 360° video-based immersive experiences enhanced with virtual/multimedia elements. *Procedia computer science*, *138*, 441-453. doi:10.1016/j.procs.2018.10.062

[10] Tang, S., & Hanneghan, M. (2011, December). Game content model: an ontology for documenting serious game design. In *2011 Developments in E-systems Engineering* (pp. 431-436). IEEE. doi:10.1109/DeSE.2011.68

[11] Pan, Y., & Mitchell, K. (2020, March). PoseMMR: a collaborative mixed reality authoring tool for character animation. In *2020 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)* (pp. 758-759). IEEE. doi:10.1109/VRW50115.2020.00230

[12] Barbosa, A. F., Pereira, P. N., Dias, J. A., & Silva, F. G. (2014). A new methodology of design and development of serious games. *International Journal of Computer Games Technology*, *2014*. doi:10.1155/2014/817167

[13] Moreno-Ger, P., Burgos, D., Martínez-Ortiz, I., Sierra, J. L., & Fernández-Manjón, B. (2008). Educational game design for online education. *Computers in Human Behavior*, *24*(6), 2530-2540. doi:10.1016/j.chb.2008.03.012

[14] Feurstein, M. S. (2018). Towards an integration of 360-degree video in higher education. Workflow, challenges and scenarios.

[15] Toro, M., Desainte-Catherine, M., & Baltazar, P. (2015). A model for interactive scores with temporal constraints and conditional branching. *arXiv preprint arXiv:1510.02836*.

[16] Argyriou, L., Economou, D., Bouki, V., & Doumanis, I. (2016, December). Engaging immersive video consumers: Challenges regarding 360-degree gamified video applications. In *2016 15th International Conference on Ubiquitous Computing and Communications and 2016 International Symposium on Cyberspace and Security (IUCC-CSS)* (pp. 145-152). IEEE. doi:10.1109/IUCC-CSS.2016.028
[17] Hirzalla, N., Falchuk, B., & Karmouch, A. (1995). A temporal model for interactive multimedia scenarios. *IEEE MultiMedia*, *2*(3), 24-31. doi:10.1109/93.410508
[18] Zhang, L., & Oney, S. (2020, October). FlowMatic: An Immersive Authoring Tool for Creating Interactive Scenes in Virtual Reality. In *Proceedings of the 33rd Annual ACM Symposium on User Interface Software and Technology* (pp. 342-353). doi:10.1145/3379337.3415824
[19] Rosson, M. B., & Carroll, J. M. (2002). *Usability engineering: scenario-based development of human-computer interaction*. Morgan Kaufmann.
[20] Börjeson, L., Höjer, M., Dreborg, K. H., Ekvall, T., & Finnveden, G. (2006). Scenario types and techniques: towards a user's guide. *Futures*, *38*(7), 723-739. doi:10.1016/j.futures.2005.12.002
[21] Arsovic, B., & Stefanovic, N. (2020). E-learning based on the adaptive learning model: case study in Serbia. *Sādhanā*, *45*(1), 1-13. doi:10.1007/s12046-020-01499-8
[22] Yang, T. C., Hwang, G. J., & Yang, S. J. H. (2013). Development of an adaptive learning system with multiple perspectives based on students' learning styles and cognitive styles. *Journal of Educational Technology & Society*, *16*(4), 185-200.

Cognitive assessment in virtual environments: How to choose the Natural User Interfaces?

MARONNAT Florian[1], DAVESNE Frédéric[1], OTMANE Samir[1]

[1]Université Paris-Saclay, Univ Evry, IBISC, 91020, Évry-Courcouronnes, France

Corresponding author: MARONNAT Florian – florian.maronnat@univ-evry.fr
Keywords: Alzheimer disease – Virtual reality – Cognitive assessment – Natural user interface

Abstract

Facing an aging population with an increasing prevalence of dementia, the challenge lies in improving cognitive assessment and screening. Numerous paper tests are used daily as Mini Mental Status Examination (MMSE) or Montréal Cognitive Assessment (MoCA). Several assessments in virtual environments were developed to enhance ecology and appeared to be as efficient as classical tests. Using the Oculus Quest® device, a novel immersive environment was created and composed of thirteen cognitive tasks. Each scene was conceived using the navigation, selection, and manipulation 3D interaction tasks. This new environment browses several cognitive functions easily despite some technical limits.

1. Introduction

In cognitive assessment, numerous tests exist for a specific field (memory, language) or global cognition. These usual tests normally proceed in natural conditions with a patient and an examinator asking the questions orally. Most of the answers are also orally given, but some need to write, draw, or make movements. Nevertheless, these tests do not appraise the impact of cognitive impairment on daily activities, and Virtual Reality (VR) appears to be an ecological and efficient tool to detect cognitive decline throughout real situations (Parsons, 2015). Depending on the degree of immersion, the virtual environment can be classified as non–immersive, semi-immersive, or fully – immersive (García-Betances et al., 2015a). Mainly developed for video games, VR has become in a few years a new help in medical practice (Li et al., 2017) as in surgery (Schmidt et al., 2021; Winkler-Schwartz et al., 2019) or psychiatric (Freeman et al., 2017). Only five immersive environments have been previously published in cognitive assessment, as presented in the review from Clay and al in 2020 (Clay et al., 2020). In 2020, Maronnat and al (Maronnat et al., 2020) presented an immersive environment inspired from classical cognitive tests: Mini Mental Status Examination (MMSE) (Folstein et al., 1975), Montréal Cognitive Assessment (MoCA) (Nasreddine et al., 2005) or the five words of Dubois (Dubois et al., 2002). Contrary to the other tests developed (Clay et al., 2020), this autonomous assessment browses seven cognitive domains (attention, orientation, abstraction, executive functions, visuospatial abilities, language, and memory) (Maronnat et al., 2020), which is a prominent force in the goal to perform a global evaluation. Human-computer interactions aim to reproduce natural human behaviors in virtual environments and are named Natural User Interfaces (NUI). The term "natural" means that the interface must be intuitive, easy to use, and understand (O'hara et al., 2013). The interface is not a barrier as the communication is natural and direct with the use of own senses. Either in augmented or virtual reality, several senses were exploited (Djelil et al., 2013) as touch (Marton et al., 2012; Song et al., 2008), voice (McGlashan, 1995), body movements (Hasenfratz et al., 2004), lips' movements (Jian et al., 2001), thoughts or cerebral activity (Friedman et al., 2004; Lécuyer et al., 2008) and facial expression (Busso & Narayanan, 2007). NUI can be classified according to their feasible interaction task (Bowman et al., 2005): navigation, selection, manipulation, and application control. This paper will present the immersive environment's NUI choice for each cognitive task.

2. Materials

Before using the virtual environment, choosing an adequate material to upload the environment was necessary. This material should respond to multiple characteristics selected by the authors considering that the material should be easily usable and understandable by older adults with cognitive impairment (Appendix 1) and easy to transport. Thus, they opted for the Oculus Quest (OQ)® (Appendix 2) (Oculus technologies, 2019).

3. Immersive assessment

3.1 Cognitive functions

Cognition can be summarized as the mental processes of knowledge and thinking. It comprises several functions, each participating in a global functioning (Harvey, 2019). Global cognition can be assessed throughout general tests as MMSE or MoCA. The environment (Maronnat et al., 2020) explored seven cognitive functions: Attention, memory, praxis, executive functions, language, abstraction, and orientation. PD Harvey (Harvey, 2019) has defined and classified the different cognitive domains in his state of the art. Attention is the process of attending to and sustaining relevant information while ignoring other nonrelevant information (distraction). Memory refers to information's encoding, storage, and retrieval either in the short term (working memory) or long term. Praxis corresponds to motor skills like drawing, writing, or gestural memory. Executive functions are the domain of reasoning, solving problems, or planning. Language skills are the ability to understand language, access semantic memory, identify objects with a name, and respond to verbal instructions with behavioral acts (example: "close your eyes" in MMSE). Abstraction (Jaegwon, 2017) is the cognitive process of isolating, or "abstracting," a common feature or relationship observed in several things, or the product of such an approach (example: "what is the similarity between an orange and a banana?" in MoCA). Finally, orientation is the ability to orient oneself both in time (example: "what day is it?") or space (measure: "where are we?")

3.2 3D Interactions tasks

The description of the virtual actions was made according to Bowman's classification (Bowman et al., 2005; Ouramdane et al., 2009). This classification proposes four 3D interaction tasks: navigation, selection, manipulation, and application control. This classification aims to traduce actions in the real world into virtual tasks. Navigation includes all the methods that let to know an object's position but also the ability to move inside an environment. Selection refers to the choice of an object to accomplish an action inside the environment. Close from the selection task, manipulation defines the processes leading to changes in the object's properties (position, orientation, color…). Application control corresponds to the commands which change the environment's state and properties. These commands are often included in the application's services.

3.3 Virtual scenes

The cognitive tasks are presented as independent and successive scenes. Interactions used in these scenes are presented in Appendices 4, 5, and 6. Firstly, oral questions given by the examinator were replaced by a verbal modality with instructions delivered through the headphones in every scene. Aiming to get the most realistic immersion possible, navigation was performed by the body's movements and gaze directed. In usual tests, patients use their own hands, so it was essential to making them appear in the environment as virtual hands. Oral answers were difficult to program due to technical limits, and it was decided to opt for item selections in attention, language, memory, and executive functions tasks. Manipulation movements were accomplished with the virtual hands to reproduce natural movements. For example, in the number series present in the MoCA test, it was decided to virtualize it as a list of words to remember. Firstly, the patient hears a list (square, circle, triangle) and needs to remember it. Then it will be asked to replace the figures in the correct order by moving the objects. In the abstraction task, a manipulation substituted an oral answer. For praxis evaluation, drawing and written order were also replaced by a manipulation task. Before beginning the test, was uploaded a welcome task to present the environment and a training task to test the excellent understanding of the functioning (Appendix 3). The selected NUI of these tasks are presented in Appendix 2 and correspond to those present later in the assessment.

4. Discussion

This paper presents a new virtual tool in an immersive environment to assess cognition. One of the first conditions was to provide an autonomous test without the intervention of an exterior examiner. Thus, it was necessary to find a material that led free movements and autonomous functioning, as described in part 2. OQ responded to these characteristics (Appendix 1). Moreover, OQ had already been used for cognitive training (Varela-Aldás et al., 2020) but also for cognitive assessment as for navigation memory (Ijaz et al., 2019) or visual capabilities (Foerster et al., 2016), which comforted the choice of using it. Before beginning the tasks was integrated a training scene. Thus, the user could discover the environment and try selection or manipulation tasks. Contrary to a classical leap motion, the hand's tracking in OQ is performed through two touch controllers, which let virtual hand's appearance and movements for selection and manipulation. The user needs to hold the controllers and push on buttons to accomplish these tasks. However, people with cognitive impairment might not clearly understand touch controller's functioning, and hand tracking with a leap motion might seem more accessible and more intuitive to use; the reason why we integrated a training task.

In classical assessments, patients need to write, speak, or draw and use their voices and hands. Although rapid and significant improvements occurred in VR in a few years leading to more realistic environments, several technical limits remain to reproduce natural actions. Firstly, some limitations deal with the patient oneself in the case of a sensory loss as vision or hearing impairment. Naturally, the sound level can be modulated but needs an exterior intervention. Vision's alterations as macular degeneration can also trouble results. For patients wearing glasses, OG is adapted (Appendix 1). The objective was to create an autonomous system (without an examiner). So, all the oral questions given by the examinator were replaced with vocal instructions delivered through the headphones integrated into the helmet. The choice was made for displayed answers (orientation, memory, language) that should be selected into the environment. Indeed, current advances in speech recognition are still insufficient (González Hautamäki et al., 2019; Shneiderman, 2000) to use in the system and may lead to a false or biased recognition (as an accent or a modulated pronunciation) so to an inaccurate result. Displaying written answers partially resolves the insufficiencies of speech recognition. Indeed it also has limits among people presenting a low degree of education, especially illiteracy (Franzen et al., 2020). Nevertheless, remembering displayed words also explores visual memory. For the abstraction test, oral answers were replaced by a selection and manipulation task respecting the initial evaluation where the patient must find the similarities between two objects (Nasreddine et al., 2005). When writing actions were necessary for usual assessments (clock test and drawing test), simplifications occurred in tests because patients may not understand how to use a connected pencil or manipulate a virtual one. Moreover, these scenarios would need the checking of an exterior examiner contrary to the wish for an autonomous system. So, the drawing task was transformed into a manipulation task to reproduce a scheme and the clock test into a choice task between two clocks. Withal these changes might lead to an unprecise evaluation of executive functions. In the MMSE attention task, patients must realize subtractions by steps of 7. As discussed before, transposing these operations would need speech recognition. A new task was so created with balls. In this task, the patient must select a ball always situated in the same place as the first one on the left. After the selection, the ball disappears, and the test starts again without giving the order once again. Praxis' task is close to the classical one as it uses both selection and manipulation. The patient needs to select a ball and move it to another place with a specific order (p.e yellow ball on the red plate).

5. Conclusion

Whether in immersive or non-immersive environments, VR appears to be an excellent tool to assess cognition (García-Betances et al., 2015b). The thirteen scenes browse multiple cognitive functions by using several interactions such as navigation, manipulation, or selection to reproduce as close as possible natural interactions. Some limits remain as speech recognition due to technical limitations. With continuous improvements, virtual environments and natural user interfaces will become more and more realistic, increasing immersion and ecology (Parsons, 2015). This environment remains untested among an elderly population and needs to be evaluated in natural conditions. Further studies are required to explore new human-computer interactions closer to natural actions.

6. References

Bowman, D. A., Kruijff, E., Jr, J. J. L., & Poupyrev, I. (2005). *3D User Interfaces: Theory and Practice*. Addison-Wesley.

Busso, C., & Narayanan, S. S. (2007). Interrelation Between Speech and Facial Gestures in Emotional Utterances: A Single Subject Study. *IEEE Transactions on Audio, Speech, and Language Processing*, *15*(8), 2331–2347. https://doi.org/10.1109/TASL.2007.905145

Clay, F., Howett, D., FitzGerald, J., Fletcher, P., Chan, D., & Price, A. (2020). Use of Immersive Virtual Reality in the Assessment and Treatment of Alzheimer's Disease: A Systematic Review. *Journal of Alzheimer's Disease: JAD*, *75*(1), 23–43. https://doi.org/10.3233/JAD-191218

Djelil, F., Otmane, S., & Wu, S. (2013, October). Apport des NUIs pour les Applications de Réalité Virtuelle et Augmentée: État de l'Art. *8èmes Journées de l'Association Française de Réalité Virtuelle, Augmentée, Mixte et d'Intéraction 3D*. https://hal.archives-ouvertes.fr/hal-02455134

Dubois, B., Touchon, J., Portet, F., Ousset, P. J., Vellas, B., & Michel, B. (2002). ["The 5 words": A simple and sensitive test for the diagnosis of Alzheimer's disease]. *Presse Medicale (Paris, France: 1983)*, *31*(36), 1696–1699.

Foerster, R. M., Poth, C. H., Behler, C., Botsch, M., & Schneider, W. X. (2016). Using the virtual reality device Oculus Rift for neuropsychological assessment of visual processing capabilities. *Scientific Reports*, *6*, 37016. https://doi.org/10.1038/srep37016

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, *12*(3), 189–198.

Franzen, S., van den Berg, E., Goudsmit, M., Jurgens, C. K., van de Wiel, L., Kalkisim, Y., Uysal-Bozkir, O., Ayhan, Y., Nielsen, T. R., & Papma, J. M. (2020). A Systematic Review of Neuropsychological Tests for the Assessment of Dementia in Non-Western, Low-Educated or Illiterate Populations. *Journal of the International Neuropsychological Society*, *26*(3), 331–351. https://doi.org/10.1017/S1355617719000894

Freeman, D., Reeve, S., Robinson, A., Ehlers, A., Clark, D., Spanlang, B., & Slater, M. (2017). Virtual reality in the assessment, understanding, and treatment of mental health disorders. *Psychological Medicine*, *47*(14), 2393–2400. https://doi.org/10.1017/S003329171700040X

Friedman, D., Leeb, R., Angus, A., Garau, M., Guger, C., Keinrath, C., Steed, A., Pfurtscheller, G., & Slater, M. (2004). Navigating Virtual Reality by Thought: First Steps. *Proceedings of the 7th Annual International Workshop Presence 2004*, 160–167.

García-Betances, R. I., Arredondo Waldmeyer, M. T., Fico, G., & Cabrera-Umpiérrez, M. F. (2015a). A succinct overview of virtual reality technology use in Alzheimer's disease. *Frontiers in Aging Neuroscience*, *7*, 80. https://doi.org/10.3389/fnagi.2015.00080

García-Betances, R. I., Arredondo Waldmeyer, M. T., Fico, G., & Cabrera-Umpiérrez, M. F. (2015b). Corrigendum: A succinct overview of virtual reality technology use in Alzheimer's disease. *Frontiers in Aging Neuroscience*, *7*, 235. https://doi.org/10.3389/fnagi.2015.00235

González Hautamäki, R., Hautamäki, V., & Kinnunen, T. (2019). On the limits of automatic speaker verification: Explaining degraded recognizer scores through acoustic changes resulting from voice disguise. *The Journal of the Acoustical Society of America*, *146*(1), 693. https://doi.org/10.1121/1.5119240

Harvey, P. D. (2019). Domains of cognition and their assessment. *Dialogues in Clinical Neuroscience*, *21*(3), 227–237. https://doi.org/10.31887/DCNS.2019.21.3/pharvey

Hasenfratz, J.-M., Lapierre, M., & Sillion, F. X. (2004). A real-time system for full body interaction with virtual worlds. *Virtual Environments 2004-10th Eurographics Symposiumon Virtual Environments*, 147–156.

Ijaz, K., Ahmadpour, N., Naismith, S. L., & Calvo, R. A. (2019). An Immersive Virtual Reality Platform for Assessing Spatial Navigation Memory in Predementia Screening: Feasibility and Usability Study. *JMIR Mental Health*, *6*(9), e13887. https://doi.org/10.2196/13887

Jaegwon, K. (2017, June 16). *“abstraction”. Encyclopedia Britannica*. https://www.britannica.com/science/abstraction

Jian, Z., Kaynak, M. N., Cheok, A. D., & Chung, K. C. (2001). Real-time lip tracking for virtual lip implementation in virtual environments and computer games. *10th IEEE International Conference on Fuzzy Systems. (Cat. No.01CH37297)*, *3*, 1359–1362 vol.2. https://doi.org/10.1109/FUZZ.2001.1008910

Lécuyer, A., Lotte, F., Reilly, R. B., Leeb, R., Hirose, M., & Slater, M. (2008). Brain-Computer Interfaces, Virtual Reality, and Videogames. *Computer*, *41*(10), 66–72. https://doi.org/10.1109/MC.2008.410

Li, L., Yu, F., Shi, D., Shi, J., Tian, Z., Yang, J., Wang, X., & Jiang, Q. (2017). Application of virtual reality technology in clinical medicine. *American Journal of Translational Research*, *9*(9), 3867–3880.

Maronnat, F., Seguin, M., & Djemal, K. (2020). Cognitive tasks modelization and description in VR environment for Alzheimer’s disease state identification. *2020 Tenth International Conference on Image Processing Theory, Tools and Applications (IPTA)*, 1–7. https://doi.org/10.1109/IPTA50016.2020.9286627

Marton, F., Agus, M., Gobbetti, E., Pintore, G., & Balsa Rodriguez, M. (2012). Natural exploration of 3D massive models on large-scale light field displays using the FOX proximal navigation technique. *Computers & Graphics*, *36*(8), 893–903. https://doi.org/10.1016/j.cag.2012.06.005

McGlashan, S. (1995). Speech interfaces to virtual reality. *In: 2nd International Workshop on Military Applications of Synthetic Environments and VR*.

Nasreddine, Z. S., Phillips, N. A., Bédirian, V., Charbonneau, S., Whitehead, V., Collin, I., Cummings, J. L., & Chertkow, H. (2005). The Montreal Cognitive Assessment, MoCA: A brief screening tool for mild cognitive impairment. *Journal of the American Geriatrics Society*, *53*(4), 695–699. https://doi.org/10.1111/j.1532-5415.2005.53221.x

Oculus technologies. (2019). *Metaquest*. https://www.oculus.com/

O’hara, K., Harper, R., Mentis, H., Sellen, A., & Taylor, A. (2013). On the naturalness of touchless: Putting the “interaction” back into NUI. *ACM Transactions on Computer-Human Interaction*, *20*(1), 5:1-5:25. https://doi.org/10.1145/2442106.2442111

Ouramdane, N., Otmane, S., & Mallem, M. (2009). Interaction 3D en Réalité Virtuelle—Etat de l’art. *Revue Des Sciences et Technologies de l’Information - Série TSI : Technique et Science Informatiques*, *28*(8), 1017–1049. https://doi.org/10.3166/tsi.28.1017-1049

Parsons, T. D. (2015). Virtual Reality for Enhanced Ecological Validity and Experimental Control in the Clinical, Affective and Social Neurosciences. *Frontiers in Human Neuroscience*, *9*, 660. https://doi.org/10.3389/fnhum.2015.00660

Schmidt, M. W., Köppinger, K. F., Fan, C., Kowalewski, K.-F., Schmidt, L. P., Vey, J., Proctor, T., Probst, P., Bintintan, V. V., Müller-Stich, B.-P., & Nickel, F. (2021). Virtual reality simulation in robot-assisted surgery: Meta-analysis of skill transfer and predictability of skill. *BJS Open*, *5*(2). https://doi.org/10.1093/bjsopen/zraa066

Shneiderman, B. (2000). The limits of speech recognition. *Communications of the ACM*, *43*(9), 63.

Song, P., Yu, H., & Winkler, S. (2008). Vision-based 3D finger interactions for mixed reality games with physics simulation. *Proceedings of The 7th ACM SIGGRAPH International Conference on Virtual-Reality Continuum and Its Applications in Industry*, 1–6. https://doi.org/10.1145/1477862.1477871

Varela-Aldás, J., Palacios-Navarro, G., Amariglio, R., & García-Magariño, I. (2020). Head-Mounted Display-Based Application for Cognitive Training. *Sensors (Basel, Switzerland)*, *20*(22), E6552. https://doi.org/10.3390/s20226552

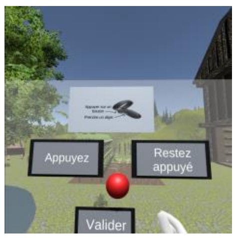
Winkler-Schwartz, A., Yilmaz, R., Mirchi, N., Bissonnette, V., Ledwos, N., Siyar, S., Azarnoush, H., Karlik, B., & Del Maestro, R. (2019). Machine Learning Identification of Surgical and Operative Factors Associated With Surgical Expertise in Virtual Reality Simulation. *JAMA Network Open*, *2*(8), e198363. https://doi.org/10.1001/jamanetworkopen.2019.8363

APPENDICES

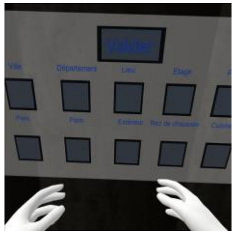
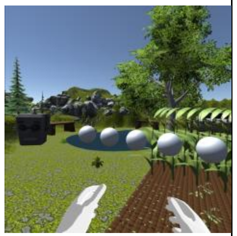
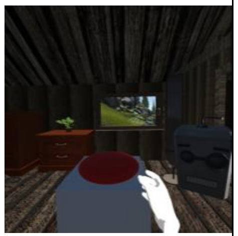
Desired characteristics	**Characteristics of OQ**
Full immersion	Six degrees of freedom
Possibility of appearance of hands	Two touch controllers with four integrated cameras
Sounds compatible with hearing impairment or hearing aids	3D positional audio directly broadcasted from the headset
Compatible with glasses (vision impairment)	OLED screen 1440 × 1600 Resolution Per Eye 72 Hz Refresh Rate Supported **Glasses Compatible**
Own storage	128 GB
Free movements (no cables)	Battery included
Could be proceeded while sitting	Possible
Low weight	Less than 600 grams

Appendix 1: Characteristics of Oculus Quest

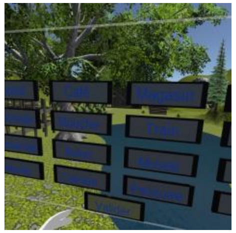
Appendix 2: Views of Oculus Quest®

Task	3D Interaction task				View
	Navigation metaphor	Selection metaphor	Manipulation metaphor	Application control	
Welcome	Body's movements and gaze directed	Virtual hand	None	None	
Training	Body's movements and gaze directed	Virtual hand	None	None	

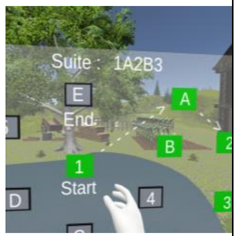
Appendix 3: Characteristics of welcome and training task

Cognitive domain		Orientation	Attention		
Task		Spatial and Temporal	Calculation	List of letters	Number series
Natural action		Oral questions	Oral answer	Oral answer	Oral answer
Virtual action		Selection of items	Selection of items	Touching a virtual button	Selection of items
3D Interaction task	Navigation	Body’ s movements and gaze directed	Body’ s movements and gaze directed	Body’ s movements and gaze directed	Body’ s movements and gaze directed
	Selection	Virtual hand	Virtual hand	Virtual hand	Virtual hand
	Manipulation	None	None	None	Virtual hand
	Application control	None	None	None	None
View		  			

Appendix 4: Characteristics of cognitive tasks (1/3)

Cognitive domain		Language	Praxis		Abstraction	Memory
Task		Naming	Figure	Written order	Associations	Words with immediate and delayed recall
Natural action		Oral answer	Drawing	Writing	Oral answer	Oral answer
Virtual action		Selection of items	Selection and movements of items	Selection and movements of items	Selection and movements of items	Selection of items
3D Interaction task	**Navigation**	Body' s movements and gaze directed	Body' s movements and gaze directed	Body' s movements and gaze directed	Body' s movements and gaze directed	Body' s movements and gaze directed
	Selection	Virtual hand	Virtual hand	Virtual hand	Virtual hand	Virtual hand
	Manipulation	None	Virtual hand	Virtual hand	Virtual hand	None
	Application control	None	None	None	None	None
View						 

Appendix 5: Characteristics of cognitive tasks (2/3)

Cognitive domain		Executive functions	
Task		Logical Sequence	Clock
Natural action		Drawing	Drawing and writing
Virtual action		Selection of items	Selection of items
3D Interaction task	Navigation	Body’ s movements and gaze directed	Body’ s movements and gaze directed
	Selection	Virtual hand	Virtual hand
	Manipulation	None	None
	Application control	None	None
View			

Appendix 6: Characteristics of cognitive tasks (3/3)

POSTERS

Animal Embodiment: Embodying a beaver in immersive virtual environments to create empathy and teach about the impact of global warming in a playful way

Sierra Rativa A.[1], Bakker G. A.,[2], Sierra Rativa, A. G.,[3]

[1]Breda University of Applied Sciences, Breda, the Netherlands
[2]The Future Partners, Rotterdam, the Netherlands
[3]Universidade Federal do Rio Grande do Sul, Porto Alegre, Brazil

Corresponding author: Sierra Rativa Alexandra, profealexandrasierra@gmail.com
Keywords: animal embodiment – virtual reality movie – immersion – empathy –climate change –Justin Beaver Survivor– Justin Beaver VR

Abstract: The beaver is an internationally renowned symbol of engineering. Through virtual reality, users can reflect on how difficult it is for these animals to survive in the face of climate change and with an inability to control the natural conditions around them. An experiment in virtual reality gave rise to the possibility of embodying a virtual animal. The results of this experiment showed a significant effect of the visual body appearance on immersion, but no effects on animal embodiment and empathy. The upcoming investigation will allow an investigation of the effects of virtual hands and a more natural (i.e. not anthropomorphized) animal appearance on animal embodiment, immersion, and empathic learning through a virtual reality movie focusing on climate change and global warming.

1. Introduction

One of the main obstacles to modern education is that it largely remains uninspiring for students and teachers alike, despite major technological innovations. The teaching of climate change awareness through virtual reality presents an unflattering scenario. Markowitz et al. (2018) developed two exciting experiments: In the first of these, they created a combination of a scuba diving trip in a physical world with virtual reality simulation to present the acidification of the ocean, an issue caused by global warming. The results showed that knowledge gain (Ocean Acidification Knowledge) remains stable after weeks of the immersive VR experience. However, there is no significant effect on Environmental Attitude and Presence after using the VR simulation. The second experiment compared two self-avatar representation conditions: a coral avatar and a human scuba diver avatar. The data obtained show that these different avatar appearances or conditions do not improve knowledge gain, evoke no positive attitude toward the environment, and do not create a presence effect on students. A recent study developed by Barnidge et al. (2021) explored whether students learn more about climate change in three conditions: virtual reality (VR), text-with-images (TWI), and 360° video (360°). The results showed that there is no significant effect on learning outcomes in the VR Condition. This study leads us to wonder whether immersive learning with virtual reality simulation could be developed to teach climate change and global warming topics on the basis of virtual worlds.

Animal embodiment is one of the greatest challenges of virtual reality and Human-Computer Interaction Studies. One particular issue is the control of an avatar of certain animals, which have different postures, skeletal arrangements, and shapes compared to human bodies (Krekhov, Cmentowski and Krüger, 2019). Pimentel and Kalyanaraman (2021) described a lack of visuomotor synchrony in virtual reality, for example, when attempting to control a turtle's virtual tail. Likewise, a tiger has a significantly different posture from a human, largely due to the fact that it walks on four legs (Krekhov, Cmentowski and Krüger, 2019; Škola and Liarokapis, 2021). As illustrated in Figure 1, this study of animal embodiment employed the design of a virtual beaver with anthropomorphic characteristics and mixed biological representations, in order to preserve the natural human posture of participants.

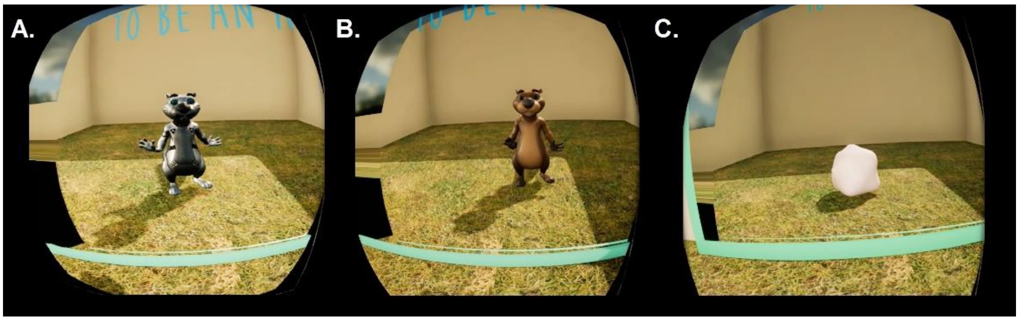


Figure 1. Animal embodiment with three conditions: (A) Virtual beaver, (B) virtual robot beaver, and (C) amorphous figure.

2. Methodology

In our research, we explored the virtual reality value of fostering an immersive and empathic experience where people can 'walk in the shoes' of the natural world. An interesting aspect is that the virtual scenarios and self-representation in virtual reality are entirely artificial, although they simulate the natural world. The experiment is called "Justin Beaver VR", in which users could be a beaver with a natural (virtual beaver) or artificial (robot virtual beaver) body, or could instead be an amorphous figure (Sierra Rativa, Postma and Van Zaanen, 2020b). This experiment was designed in virtual reality, and haptic sensory feedback was included, whereby users could feel vibrations when swimming, eating branches, and receiving the impact of a bullet when being hunted. Previously, the same experiment was created in a desktop version (Sierra Rativa, Postma, and Van Zaanen, 2020a). In total, 90 participants were recruited from the university student population of the Netherlands. The participants were between 18 and 35 years old, and 57.8% male and 42.2% female. Each condition had the same number of participants. Before the experiment, the participants filled out a dispositional empathy questionnaire called the Interpersonal Reactivity Index (IRI) (Davis, 1980). After the virtual simulation experience, participants completed a post-questionnaire on their experience, containing questions on situational empathy Sierra Rativa, Postma and Van Zaanen, 2022)a perceived pain questionnaire (Das et al. 2005), an immersion questionnaire (Jennett et al. 2008), and an avatar embodiment questionnaire (Gonzalez-Franco and Peck, 2018). We referred to this avatar embodiment questionnaire as 'Animal Embodiment', and it was analyzed according to the method proposed by Peck and Gonzalez-Franco (2021).

3. Results

In Figure 2A, we can see that the subscales of Animal Embodiment had a higher mean in the virtual animal compared to the virtual robot beaver and amorphous figure. Moreover, as shown in Figure 2B, we observed a statistically significant effect of character appearance on immersion, $F(2, 89) = 3.284$, $p = 0.042$, $\eta^2 = 0.070$. However, no significant effect of character appearance was discovered for self-reported Situational Empathy $F(2, 89) = 2.057$, $p = 0.134$, $\eta^2 = 0.045$, or Animal Embodiment $F(2, 89) = 1.578$, $p = 0.604$ $\eta^2 = 0.012$.

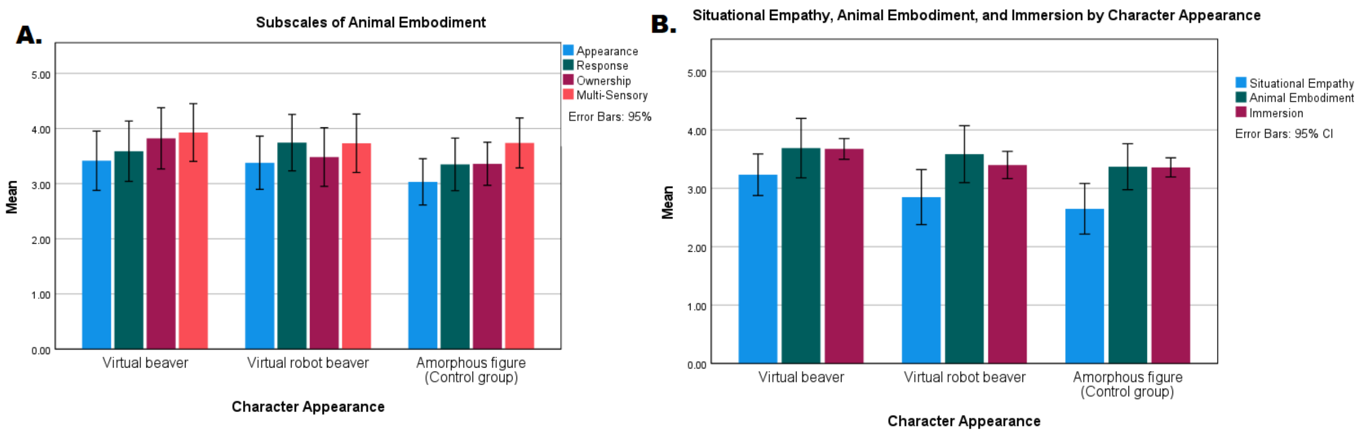


Figure 2. Animal Embodiment: (A) Subscales of Animal Embodiment questionnaire, and (B) Comparison between Situational Empathy, Animal Embodiment, and Immersion, depending on character appearance.

4. Future work on animal embodiment in an Interactive Virtual Movie

The results showed that bodily appearance is not the main factor affecting animal embodiment. A possible aspect relating to this can be explored when users view their embodied avatar in front of a mirror in the virtual reality movie environment with their new virtual animal body, moving their animal appendages, as illustrated in Figure 3. Future research will use two conditions to investigate whether real-life hands – as opposed to the appendages of the virtual animal – can affect animal embodiment. We will use an interactive movie called "Justin Beaver Survivor" for this purpose. Moreover, we will explore whether the storyline of this virtual reality movie can affect the learning outcome and empathy of the users. This new understanding can help improve expectations of the impact of virtual reality on natural and environmental education.

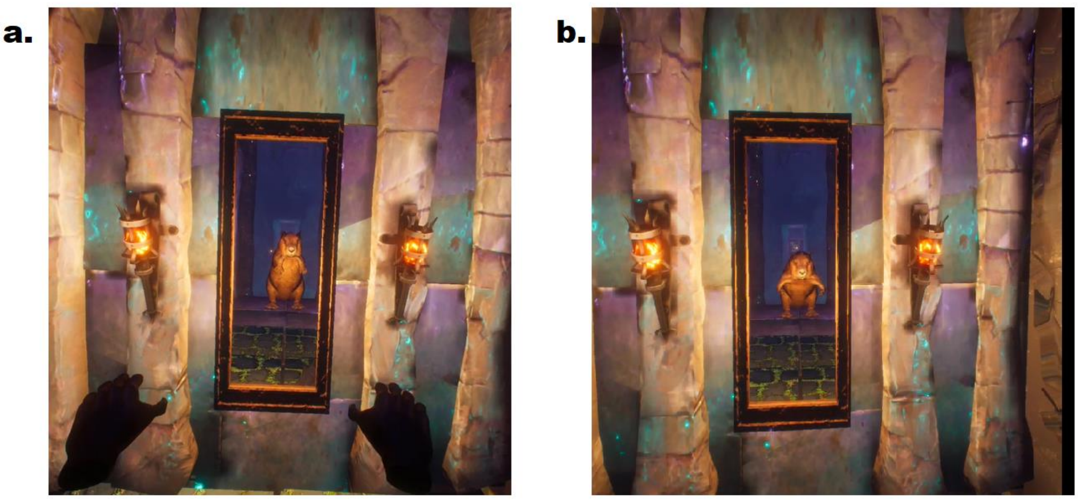


Figure 3. (a) Virtual animal representation where participants can move their animal hands (https://youtu.be/KRa-DQVB0RE), and (b) Virtual animal representation where participants only visualize their virtual animal body in the mirror during the interactive movie (https://youtu.be/u2yNRIHR9gc).

5. Data repository

Our contribution to Open Science and a transparent Community, our data about "Justin Beaver VR", and our experimental video about the "Justin Beaver Survivor" movie is publicly available on the Zenodo platform, restricted only for academic purposes. (https://doi.org/10.5281/zenodo.6390959)

6. Acknowledgments

We would like to especially thank Marie Postma and Menno van Zanne for their contribution to the VR experiment "Justin Beaver VR".

7. References

Barnidge, M., Sherrill, L. A., Kim, B., Cooks, E., Deavours, D., Viehouser, M., ... & Zhang, J. (2021). The Effects of Virtual Reality News on Learning about Climate Change. Mass Communication and Society, 1-24.

Das, D.A., Grimmer, K.A., Sparnon, A.L., McRae, S.E., Thomas, B.H., 2005. The efficacy of playing a virtual reality game in modulating pain for children with acute burn injuries: a randomized controlled trial [isrctn87413556]. BMC pediatrics 5, 1–10.

Davis, M.H., et al., 1980. A Multidimensional Approach to Individual Differences in Empathy. Technical Report. The University of Texas at Austin.

Jennett, C., Cox, A.L., Cairns, P., Dhoparee, S., Epps, A., Tijs, T., Walton, A., 2008. Measuring and defining the experience of immersion in games. International journal of human-computer studies 66, 641–661.

Krekhov, A., Cmentowski, S., & Krüger, J. (2019, August). The illusion of animal body ownership and its potential for virtual reality games. In 2019 IEEE Conference on Games (CoG) (pp. 1-8). IEEE.

Gonzalez-Franco, M., & Peck, T. C. (2018). Avatar embodiment. towards a standardized questionnaire. Frontiers in Robotics and AI, 5, 74.

Markowitz, D. M., Laha, R., Perone, B. P., Pea, R. D., & Bailenson, J. N. (2018). Immersive virtual reality field trips facilitate learning about climate change. Frontiers in psychology, 9, 2364.

Peck, T. C., & Gonzalez-Franco, M. (2021). Avatar embodiment. a standardized questionnaire. Frontiers in Virtual Reality, 1, 44.

Pimentel, D., & Kalyanaraman, S. (2021). How Multi-Sensory Animal Embodiment in Virtual Reality Influences Environmental Threat Processing and Conservation Behaviors

Sierra Rativa, A. S., Postma, M., & Van Zaanen, M. (2020b). Can virtual reality act as an affective machine? The wild animal embodiment experience and the importance of appearance. Proceedings of the MIT LINC, 3, 214-223.

Sierra Rativa, A., Postma, M., & Van Zaanen, M. (2020a). The influence of game character appearance on empathy and immersion: Virtual non-robotic versus robotic animals. Simulation & Gaming, 51(5), 685-711.

Sierra Rativa, A. S., Postma, M., & Van Zaanen, M. (2022). Try Walking in My Paws: Is it Possible to Increase Empathy, Immersion, and Perceived Pain in Virtual Reality Environments by Manipulating Animal Character Appearance?. Available at SSRN: https://ssrn.com/abstract=4014076 or http://dx.doi.org/10.2139/ssrn.4014076

Škola, F., & Liarokapis, F. (2021). Study of full-body virtual embodiment using noninvasive brain stimulation and imaging. International Journal of Human–Computer Interaction, 37(12), 1116-1129.

Contextual Body Postures in UMI3D-Based CVE

Nicolas Pacqueriaud[1]
[1]Inetum, Paris, France

Corresponding author: Nicolas Pacqueriaud, nicolas.pacqueriaud@inetum.com
Keywords: UMI3D – Collaborative Virtual Environment – Embodiment – Body Posture

Abstract

UMI3D is a recent data exchange protocol dedicated to the creation of collaborative virtual environments, accessible via device-specifics applications called UMI3D browsers [1]. User tracking and embodiments are already supported, allowing the animation of avatars thanks to dedicated data sent by the connected browsers [2]. However, the tracking capabilities of the used devices being unequal and generally not including the whole body, it appears that certain gestures or postures cannot be performed with precision. To complete the tracking data, we propose through this poster a system for UMI3D allowing to set up body animations depending on the UMI3D environment.

1. Objectives

Full Body Postures. The data model used within the UMI3D tracking system enables the transmission of information concerning the spatial location of the body parts tracked by a browser. Hardware tracking is supplemented by an interpolation model based on the constraints of the human body and can also sometimes be reinforced with locally played animations. However, these animations are often very basic and related to the browser's navigation capabilities. It appears that with the open-source browsers available today, it is not possible to manipulate the fingers to grab an object, or the lower body to sit on a chair for example. For that reason, it seems useful to improve the ability to animate the whole avatars.

Contextual Postures. Browsers allow users to interact with components present in the virtual environment. It then seems natural to combine these interactions with dedicated animations and postures, to reinforce immersion and non-verbal communication between users. The great diversity of assets used to create a virtual environment will also interfere the design of so-called generic animations, which can create glitches and visual inconsistencies during their application. Therefore, we consider it valuable to facilitate the creation of custom animations, depending on their use and on the content involved.

Browser Adaptative Animations. We have previously discussed the differences in tracking capabilities between existing UMI3D browsers. The main difference between these browsers is the ability to track users' hands. This aspect must therefore be considered to adapt the amount of pre-designed animations to be incorporated into the overall animation of an avatar. Our objective then is to obtain an illusion resulting from the real movements and the artificial additions, which will not disturb the user experience.

2. Constraints

Initially, we think that the stage of creation of postures must necessarily be carried out during the phase of creation of the UMI3D environment. It is indeed at this time that the interactions are designed, and it seems natural to make the links between interactions and animations before their use. In addition, the implementation and use of posture design tools also echo the tools mentioned in [1], legitimizing our opinion. However, on the other hand, it is important to remember that a network latency exists between a browser and a remote UMI3D environment, latency that can strongly impact the user experience. We can also recall that within the UMI3D protocol, the interaction system delegates to the browser the ability of triggering different types of interactions. For these reasons we think that the very use of postures must also be done locally on the browsers, to avoid the impacts of the existing latency as well as to remain synchronized with the triggering of interactions.

3. Data Model and Application

Data Model. First, for the purposes of testing and ease of design, we distinguished the postures dedicated to the hands (Hand Pose) and the postures managing the rest of the body (Body Pose). These data models to describe the postures are however quite similar. A Hand Pose object allows the transmission in one go of the data necessary for the management of each hand, leaving it up to the browser to decide which hand to use, and therefore which data to select. These data contain the information of activation or deactivation of the posture, the type of interaction linked to this posture, allowing us to set up a prioritization of the postures, as well as the spatial information necessary for the positioning of the hands and the phalanges. In the same way, a Body Pose object contains the posture activation or deactivation information, the spatial data necessary for the positioning of the body in the environment, as well as information indicating whether or not it is possible to override certain data by another posture, opening the way to the creation of combinations. In both cases, we relied on the typing of body parts existing within the UMI3D model to link the spatial data that we transmit.

Postures Editors. Since we separated these two data models, we designed two posture editing tools, one dedicated to the two hands, the other to the rest of the body. These tools are based on the use of Unity gizmos. A default posture is created when the tool is initialized, and the different gizmos can be moved to modify the appearance of the posture. Saving, loading, resetting functions and a cache system allow to quickly create, copy and edit a large number of postures. In addition, we have also implemented features allowing to apply symmetries, from one hand to the other in the first case, and of the two halves of the body in the second.

Browser Management. Since one of our objectives is to eliminate the existing latencies when applying postures on the browser, it is necessary for the browser to have access to the list of postures used in the environment. When connecting a browser to a UMI3D environment, all the postures used are sent to it. The browser then stores them like any other entity in the environment. The browser tracking system uses a skeleton animated by the movements of the user, directly or using interpolations, but also using basic animations. When activating a posture, skeleton manipulation processes will take over the animations and inanimate parts of the skeleton, allowing it to move to adopt the desired posture. The tracking data transmission routines will then send the information related to the new posture of the skeleton. This posture will then be perceived by all the other connected browsers.

4. Conclusion and Discussion

Thanks to the use of postures transmitted by UMI3D environments, it is now possible for us to adapt the tracking system to virtual environments and to the realization of tasks, to enhance the realism of the user experience. As we suggested in the previous paragraph, we aim to create compositions of postures in the near future. On the one hand, by mixing body postures and hand postures. On the other hand, by designing so-called partial animations. We plan to leave the possibility of managing only a sub-part of the body. In this case, we can imagine animations and postures only related to the upper or lower body. This could then result in compositions mixing even more closely the data tracked in real time and the data imposed by the postures, such as having the upper body animated by the movements of a user, while having the rest managed by a posture of the server.

5. References

[1] Casarin, J., Bechmann, D., & Keller, M. (2017, November). A unified model for interaction in 3d environment. In Proceedings of the 23rd ACM Symposium on Virtual Reality Software and Technology (pp. 1-7). https://doi.org/10.1145/3139131.3139140

[2] Pacqueriaud, N. & Casarin, J. (2020) UMI3D: Extension of Tracking and Embodiment. Conference: Laval Virtual VRIC ConVRgence 2020 http://dx.doi.org/10.13140/RG.2.2.23728.10247

LAVAL VIRTUAL EUROPE

VRIC 2022 PROCEEDINGS
VIRTUAL REALITY
INTERNATIONAL CONFERENCE

Editor: Simon Richir

Publisher: Laval Virtual
Laval Virtual Center, rue Marie Curie
53810 Changé, France
2022, April 12-14
www.laval-virtual.com