ENABLING EMBODIMENT AND INTERACTION IN OMNIDIRECTIONAL VIDEOS

Fabien Danieau, Thomas Lopez, Nicolas Mollet, Bertrand Leroy, Olivier Dumas and Jean-François Vial

Technicolor, France

ABSTRACT

This paper investigates the role of the embodiment in an immersive video experience. A system allowing to play back omnidirectional videos enhanced with real-time 3D content is presented. It enables the user to be embodied in an avatar and to interact with 3D objects added to the video. A user study was conducted to understand the impact of this embodiment on the user experience. Four different avatars were compared (none, human, ghost and astronaut) as well as three levels of user control on the avatar, and also the possibility to interact with the content. Results show that being embodied has benefits on the user experience but requires the representation to be in line with the story and the avatar to be fully controllable. Indeed a misadapted embodiment may decrease the experience.

Index Terms— head-mounted displays, immersive movies, embodiment, interactive 3D media

1. INTRODUCTION

Although omnidirectional videos or 360° videos exist for decades [1], they recently became popular with the arrival of head-mounted displays (HMD) on the mass market. Now, such audiovisual content may be easily found on most streaming platforms. It opens new possibilities of storytelling and it has been shown to be very immersive for the users since they are at the center of the movie [2].

In order to provide richer experiences, a virtual body, or avatar, representing the user is even displayed in some of those videos (see [3] for example). This type of avatar is not controllable because it has been added and animated in post-production, but in addition to the immersion this feature is meant to increase the feeling of presence: subjective sensation of being present in the content [4]. Indeed numerous studies in the literature have shown that having a self-representation in a virtual environment triggers a sensation of embodiment [5, 6, 7]. The sense of embodiment has been defined by Kilteni et al. as the sense that emerges when a body's properties are processed as if they were the properties of one's own biological body [8]. This concept is divided into three components [9]: the sense of self-location (feeling located inside the body), the sense of agency (feeling of being in control of the

body) and the sense of ownership (feeling that the body is the source of the experienced sensation).

Within this context, the key questions we address here are: To what extent does displaying an avatar improve an omnidirectional video experience? Besides, what kind of avatar should be displayed? Has it to be realistic (a virtual human) or in line with the story (custom design)? Finally, is the control of the avatar mandatory to enhance the experience? In order to answer these questions with a user study, a system allowing a user to experience omnidirectional videos while being embodied in an avatar has been developed.

2. RELATED WORK

The sense of embodiment in virtual reality has been extensively studied, and numerous variations of the rubber hand illusion may be found in the literature [8]. This illusion occurs when a user is seeing a virtual hand touched by a virtual object and, simultaneously, is feeling a tactile stimulus on his actual hand. The user believes then that the virtual hand is a part of his body. The tactile stimulation helps to trigger the illusion, even if the view of a virtual avatar may be enough. For instance, Gonzalez et al. studied the impact of a virtual mirror on the sense of embodiment [10]. Seeing a synchronized reflection of his own body increases the illusion of ownership. Steed et al. conducted an "in the wild" experiment using consumer devices (Samsung Gear VR and Google Cardboard). They studied the influence of the embodiment on the sensation of presence [11]. Participants were immersed in a virtual pub with a singer. Three conditions were tested: having an avatar, being asked by the singer to tap along the music and being looked at by the singer. Only the condition of having an avatar significantly increased the sensation of presence.

The appearance of the avatar also plays a role in the feeling of embodiment. Lugrin et al. examined the influence of three avatars on the perception of the body ownership [12]: a male or a female realistic body, an abstract body made of blocks, and a complex robot. The two last avatars were perceived as acceptable while the realistic avatar led to the uncanny valley effect. Similarly Christou and Michael compared the impact of having a human or an alien representation on the performance in a video game [13]. The appearance of the avatar seems to modify the performance: users are more aggressive when embodied in strong aliens, while both em-

---

bodiments led to the same feeling of ownership and presence. In another study, Guterstam et al. have shown that the illusion of having an invisible body may be enabled [14]. Besides, the illusion of an invisible body also alters the perception of the virtual environment. Participants felt less anxiety in front of a crowd if they believed they were invisible. More specific, the work conducted by Argelaguet et al. was focused on the representation of the user's virtual hand [15]. They studied how the representation of the hand changes the sense of agency and the sense of ownership. Three different virtual hands were tested: an abstract hand (a sphere), an iconic hand (a simplified bone model) and a realistic hand. Participants were asked to grab a virtual cube and to move it to a target location. A "dangerous" element was displayed on the path to the target (a virtual fire or a spinning saw). The authors found out that the sense of agency is correlated to the feeling of control of the virtual hand, while the sense of ownership depends on the realism of the representation of the hand. Interestingly Steptoe et al. have shown that extended avatars can enable a sensation of embodiment [16]. A virtual tail added to the avatar is perceived by the user as a part of his body. This work has also shown that the sensation is not limited to HMDs: a CAVE system with a third person perspective was used. More generally the influence of the representation of the avatar on the user's behavior is named the Proteus effect [17]. It has been shown that the attractiveness or the height of an avatar modulates the user's self-confidence.

These research results strongly suggest that having an avatar provides a sensation of embodiment. The control of the avatar and feedback from the environment also increases this sensation. Adding such features to omnidirectional videos would thus enhance immersive movie experiences.

3. EMBODIMENT AND INTERACTION IN OMNIDIRECTIONAL VIDEOS

We propose a system allowing to play back omnidirectional videos synchronized with real-time and animated 3D objects. This enables the display of a high quality video content which cannot be rendered in real-time (i.e. ray tracing rendering) or which have been captured with a 360° camera. Besides the user's avatar is rendered and animated in order to create a sensation of embodiment. Figure 1 shows an extract of the omnidirectional video augmented with real-time 3D content.

3.1. Architecture of the System

The system relies on a game engine (see Figure 2). Basically the application loads a scene containing four main components. The *Screen*: a sphere surrounding the user on which the video is projected. A video decoder converts each frame as a texture which is applied on the interior surface of the sphere. The *Avatar*: a 3D mesh displayed and animated according to the user's movement. Dedicated trackers are in

Fig. 1. User's point of view. The user is surrounded by the omnidirectional video and sees in real-time the avatar, its reflection in the visor and a floating flashlight which may be pushed.
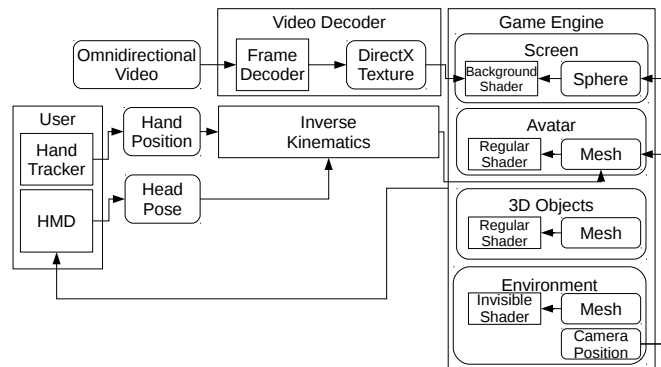

Fig. 2. Overview of the architecture.

charge of capturing his movements. The avatar is animated thanks to an inverse kinematics algorithm. The *3D objects*: regular 3D objects that are present in the scene. They may be animated or controlled by the physics engine. The *Environment*: a 3D animated model of the video sequence. This mesh is not displayed but used by the physics engine to provide the collisions and occlusions between 3D objects and the video content (it may actually contain just the necessary parts of the 3D model of the video). The environment also provides the position of the camera used during the recording. The Avatar's head (user's point of view) and the Screen are set to this position in order to align the video and its corresponding mesh. The whole scene is eventually rendered on a HMD.

3.2. Tracking and Display

In this work we wanted to add the minimal equipment to the mandatory HMD used for watching omnidirectional video. The goal was to improve the user's experience without a complete change of his setup. Hence the architecture has been im-

plemented with Unity engine. The Oculus Rift DK2 served as a HMD and also provided the tracking of the user's head (Oculus Runtime 1.6.0). The tracking of the user's hands was done with a Leap Motion attached to the HMD (SDK Orion 3.1.2). The inverse kinematics was performed thanks to the FinalIK plugin for Unity.

3.3. Omnidirectional Stereoscopic Video

To illustrate our concept we worked with a studio[1] to design the omnidirectional stereoscopic content. They produced a 55-second animation of an astronaut going outside a module of the International Space Station to perform a check-up. Once outside, a meteorite hits the station and the astronaut is projected in a dangerous position. The 3D models and the animations have been created with Blender and the rendering of the frames was performed with the Cycles Rendering Engine. Audio tracks have been added in post-production.

A homemade C++ plugin based on ffmpeg performed the video decoding within Unity. As it was a stereoscopic video (with left and right views), two spherical meshes composed the screen, on for each eye, and the frame was actually split into two textures. This processing was performed by the "Background Shader" which also made sure that the spheres were the first objects rendered (as a skybox).

3.4. Real-time 3D Content

The avatar was the main component of our system to enable the sensation of embodiment. We proposed three different avatars (see Figure 3). Firstly the *realistic human*. This model aimed at directly representing the user. However it may be not suitable to any video content (which is the case with our content). Two variations were available: male and female. Secondly the *ghost*. It was a generic humanoid model with a transparent blueish appearance. This design was meant to suit any video content, but the feeling of embodiment may be weaker than with the previous avatar. Thirdly the *astronaut*. The appearance of this avatar was directly related to the video content and fully integrated into the visual style. This type of avatar requires more efforts in design though.

Interactive 3D objects were also added to reinforce the sensation of embodiment created by the avatar. We proposed to display a "passive" object providing feedback to the user: the astronaut's visor was turned into a mirror so the user could see the avatar representing himself. Technically, a virtual camera was located in the visor, and the captured image was blended to the texture of the visor. Besides we also added an "active" object, a flashlight floating around the user who could push it. The flashlight would then rebound on the walls or on the astronaut, and relight them in real-time.

To co-locate the omnidirectional videos, the avatar and the interactive objects, we relied on the 3D model of the scene.

---

[1]4D Prod - http://4dprodgv.wixsite.com/4dprod/home

They were the actual animated models of the astronaut and the space station used for the rendering of the video. They provided the location of the visor of the astronaut for the mirror effect, and they were used for the collision of the flashlight with the walls of the space station. An "Invisible Shader" was developed for this environment: the meshes were made transparent but set to be rendered before the avatar and 3D objects ("Regular Shader"). They also contained the position of the Blender camera ensuring the alignment between the user's point of view, the video and the 3D model of the scene.

4. USER STUDY

The goal of this study was to understand the influence of the embodiment on the immersive movie experience. Given an omnidirectional video, we wanted to evaluate the impact of the avatar on the experience, as well as the role of its appearance and the importance of the interactivity.

4.1. Experimental Conditions

Three variables were used in this experiment. *The avatar representation*: three representations were compared (see Figure 3), the Ghost, the realistic Human, and the Astronaut as described in section 3.4. *The control of the avatar*: the avatar was not controllable at all (No Movement), not controllable but animated (Animation), or controllable thanks to the tracker (Tracking). The animation was designed by a CGI artist. *The feedback of the environment*: there was either no feedback (Passive) or a feedback from the mirror and the flashlight (Feedback). A control condition (*None*) has also been added were no avatar nor animation was displayed (i.e. regular omnidirectional video experience). This led to $3 \times 3 \times 2 + 1 = 19$ conditions.

Fig. 3. Three avatars: Ghost, Human and Astronaut.

4.2. Measures

Inspired from previous work [7, 12, 15, 16], a questionnaire was designed to evaluate the subjective user experience. The components of the embodiment were evaluated, self-location (Q1: I felt like I really was in the space station, Q2: I felt I was

a character of the story), ownership (Q3: I felt like the virtual body was mine, Q4: I felt disturbed by my virtual representation), agency (Q5: I felt like I fully controlled the virtual body), as well as the general satisfaction (Q6: I enjoyed the movie). These assertions were evaluated on a 5-point Likert scale, from (1) totally disagree to (5) totally agree.

At the end of the experiment an informal interview was conducted to better understand the participants' feelings. They were asked to freely comment the experiment and also to answer open-ended questions like: What was your favorite representation? Why? How would you improve the system?

4.3. Experimental Protocol

Participants were comfortably installed on a swivel chair. Each participant was first introduced to the experiment and personal information was collected: age, gender, expertise in video games and virtual reality from "never used" (1) to "daily use" (5). An informal test of the Oculus Rift and the Leap Motion was proposed to those not familiar with them.

The experiment started by two random conditions to make the participant comfortable with the protocol. Then the 19 conditions were experienced in a random order. The questionnaire was submitted after each conditions (including the two firsts but the results were not recorded). Finally the interview was conducted. The total duration of the experiment was about 25 minutes.

4.4. Results

23 participants took part in this experiment, aged from 21 to 57 ($\bar{x} = 35.83$, $\sigma_x = 10.51$), including 5 females. They had a various expertise in video games ($\bar{x} = 2.35$, $\sigma_x = 1.29$), and none of them were expert in virtual reality ($\bar{x} = 1.74$, $\sigma_x = 0.75$). Non-parametric tests used to analyze those ordinal data: Friedman Anova and pairwise Wilcoxon tests with Holm-Bonferroni correction for post-hoc analysis.

According to Q1, all conditions led to a feeling of being in the space station (see Figure 4). All conditions were statistically similar (F. Anova: $\chi^2 = 67.32$, $df = 18$, $p = 1.3^{-07}$ but Wilcoxon tests did not show significant differences). This suggests that the embodiment do not increase the sensation of self-location already set by the video (None condition).

The feeling of being a character of the story was investigated in Q2 (see Figure 5). The Astronaut-Tracking conditions (Passive and Feedback) were higher than the others (F. Anova: $\chi^2 = 100.20$, $df = 18$, $p = 2.0^{-13}$) and they were significantly higher than nine others conditions (Wilcoxon $p < 0.05$, see stars on Figure 5). This shows that an embodiment in line with the story increases the immersion.

Regarding Q3 (see Figure 6), we observed that only the conditions with the Tracking enabled provided a sensation of ownership (F. Anova: $\chi^2 = 162.02$, $df = 18$, $p = 2.2^{-16}$). Five over the six Tracking conditions were significantly dif-
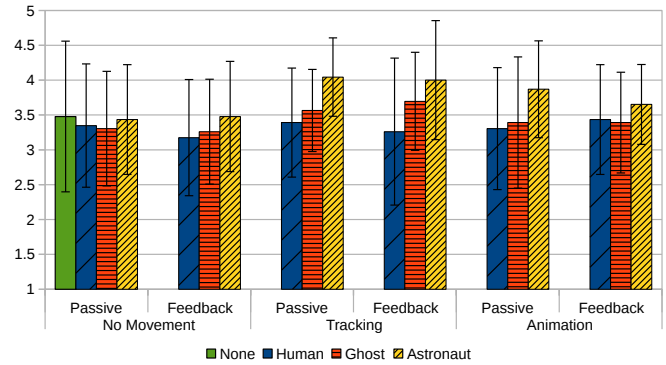

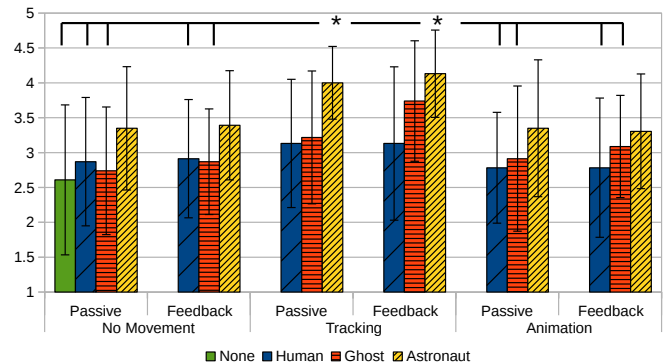**Fig. 4**. Q1 - I felt like I really was in the space station.

Fig. 5. Q2 - I felt I was a character of the story. The stars show the means statistically different.

ferent than the None condition (Wilcoxon: Human-Tracking-Passive $p = 0.001$, Human-Tracking-Feedback $p = 0.003$, Ghost-Tracking-Feedback $p = 0.007$, Astronaut-Tracking-Passive $p = 5.0^{-5}$ and Astronaut-Tracking-Feedback $p = 9.4^{-5}$). The Wilcoxon tests also showed that these conditions were significantly different from the most of others which were similar to None. A static or an animated avatar is thus not different from the absence of body.

In Q4 the comfort of the user with the virtual representation was evaluated (see Figure 7. F. Anova: $\chi^2 = 71.85$, $df = 18$, $p = 2.2^{-08}$). In general the Human embodiment was slightly higher than the neutral score (3) but only Human-Animation-Passive was significantly higher than Astronaut-NoMovement-Passive ($p = 0.03$), Astronaut-NoMovement-Feedback ($p = 0.02$) and Astronaut-Tracking-Feedback ($p = 0.003$). This would suggest that the realistic avatar is more disturbing than the two others embodiments.

Without surprise, the feeling of control was only present when the Tracking was enabled (see Figure 8, F. Anova: $\chi^2 = 226.09$, $df = 18$, $p = 2.2^{-16}$; Wilcoxon against None: Astronaut-Tracking-Passive $p = 4.3^{-4}$, Astronaut-Tracking-Feedback $p = 7.5^{-4}$, Ghost-Tracking-Passive $p = 2.7^{-5}$, Ghost-Tracking-Feedback $p = 3.3^{-4}$, Human-Tracking-Passive $p = 3.3^{-4}$ and Human-Tracking-Feedback $p =$
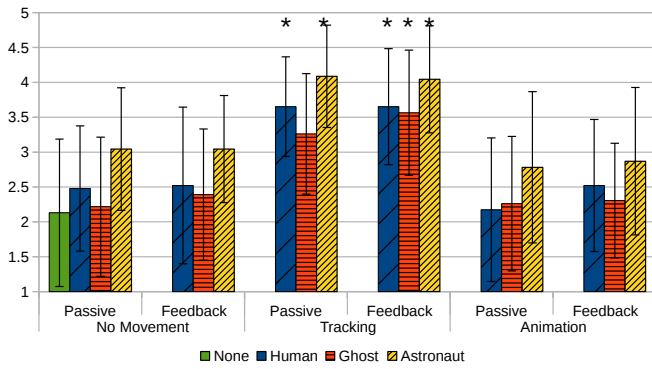
Fig. 6. Q3 - I felt like the virtual body was mine. Stars show the means significantly different from the None condition.
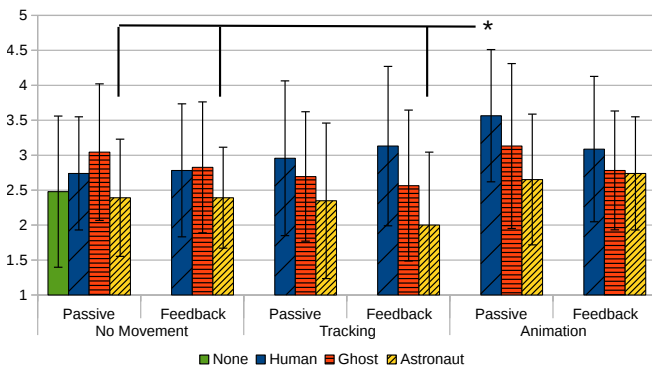

Fig. 7. Q4 - I felt disturbed by my virtual representation. All conditions are similar to the None condition (the lower the better). Only the animated Human is significantly different from three embodiment with the Astronaut.

Fig. 8. Q5 - I felt like I fully controlled the virtual body. The stars show the means significantly different from the None condition.

Fig. 9. Q6 - I enjoyed the movie. The stars shows the means statistically different.

2.8^{-4}). An animated avatar does not provide a feeling of control.

Finally the overall enjoyment of the sequence was evaluated (see Figure 9). All conditions were appreciated in general (F. Anova: $\chi^2 = 82.24, df = 18, p = 3.5^{-10}$). The Astronaut-Tracking (Passive and Feedback) conditions were significantly higher than three others (Wilcoxon $p < 0.05$). This follows the previous results were having an embodiment related to the story and the control of the avatar improves the experience.

4.5. Discussion

We observed that the Astronaut was the preferred avatar. This embodiment strongly related to the story increased the participants' involvement. They felt as characters of the movie while they were actually not. Interestingly the helmet of the space suit limited the field of view ($\sim 170°$ horizontally and vertically) but only four participants reported it was constraining. Depending on the participants, the second preferred condition was either the Ghost or the None. The holographic represen-
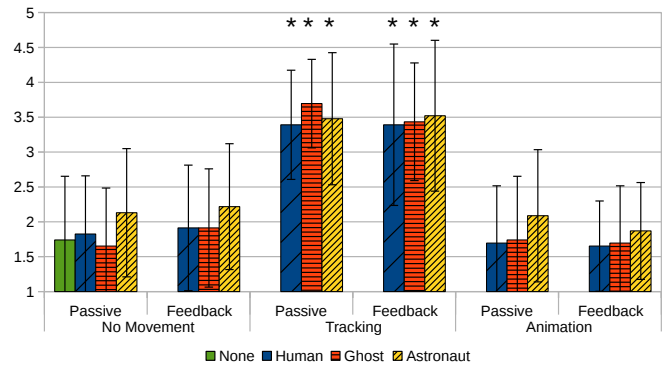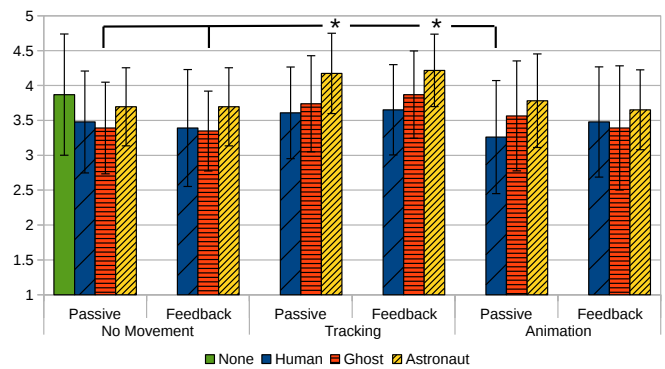
tation of the Ghost was found to be in line with the story by certain participants while disturbing by others. This would mean that designing a generic avatar suitable for any content is not straightforward. A movie provides a strong context in which the avatar has to make sense. Regarding the None condition, most of participants reported that it is better to not have a body than an inappropriate embodiment. An interesting observation is that with this condition the user remains a spectator, while the embodiment involves him in the diegesis. It has to be noted that some participants explicitly preferred to be spectator when watching a movie. This confirm that the embodiment provides a different experience and increases the feeling of self-location. The Human embodiment was the less appreciated. We observed two issues in particular. First the context was clearly a problem. Most of the participants pointed out the conflict of being in space without a space suit. Secondly, a kind of uncanny valley effect was present. The avatar was not an exact representation of the participant, and thus not convincing.

The control of the avatar is also crucial. It clearly appeared from the results that a non-animated avatar provides a strange feeling and that the animated avatar is not acceptable

(feeling of agency not satisfied). This means that the user strongly expects to control his avatar. If this expectation is not met the quality of experience is decreased.

Finally we did not observe a strong difference for the conditions with or without feedback, probably due to the short duration of the video (55"). The mirror effect and interaction with the flashlight were notified and reported as interesting features easing the sensation of embodiment. This feedback has a potential and should be more exploited in future work.

5. CONCLUSION AND PERSPECTIVES

A system enabling the playback of omnidirectional videos enhanced with real-time content was presented in this paper. It allows users to be embodied in avatars and thus to manipulate interactive elements.

A user study was conducted to investigate the role of the embodiment on the immersive video experience. Results showed that a better experience is provided with an avatar related to the narrative of the movie and when the user fully controls it. Also, the avatar has to be properly designed otherwise the quality of experience could be decreased.

In future work we will conduct experiments with realistic content. A content closer to a real-life experience may provide different expectations regarding the embodiment. Besides, the impact of a shared experience on the immersion will also be investigated.

6. REFERENCES

[1] Joshua Gluckman, Shree K Nayar, and Keith J Thoresz, "Real-time omnidirectional and panoramic stereo," in *Proceedings of Image Understanding Workshop*, 1998, vol. 1, pp. 299–303.

[2] Mirjam Vosmeer and Ben Schouten, "Interactive cinema: engagement and interaction," in *International Conference on Interactive Digital Storytelling*. Springer, 2014, vol. 8832, pp. 140–147.

[3] MPC VR, "Suicide Squad VR Experience," www.moving-picture.com/advertising/work/suicide-squad-vr-experience, 2015, [Accessed 04/10/2017].

[4] M.J. Schuemie, P. Van Der Straaten, M. Krijn, C.A.P.G. Van Der Mast, and Mary Ann Liebert, "Research on presence in virtual reality: A survey," *CyberPsychology & Behavior*, vol. 4, no. 2, pp. 183–201, 2001.

[5] Mel Slater, Daniel Pérez Marcos, Henrik Ehrsson, and Maria V Sanchez-Vives, "Inducing illusory ownership of a virtual body," *Frontiers in neuroscience*, vol. 3, pp. 29, 2009.

[6] Maria V Sanchez-Vives, Bernhard Spanlang, Antonio Frisoli, Massimo Bergamasco, and Mel Slater, "Virtual hand illusion induced by visuomotor correlations," *PloS one*, vol. 5, no. 4, pp. e10381, 2010.

[7] Mel Slater, Bernhard Spanlang, Maria V Sanchez-Vives, and Olaf Blanke, "First person experience of body transfer in virtual reality," *PloS one*, vol. 5, no. 5, pp. e10564, 2010.

[8] Konstantina Kilteni, Raphaela Groten, and Mel Slater, "The Sense of Embodiment in Virtual Reality," *Presence: Teleoperators and Virtual Environments*, vol. 21, no. 4, pp. 373–387, 2012.

[9] Matthew R Longo, Friederike Schüür, Marjolein PM Kammers, Manos Tsakiris, and Patrick Haggard, "What is embodiment? a psychometric approach," *Cognition*, vol. 107, no. 3, pp. 978–998, 2008.

[10] Mar Gonzalez-Franco, Daniel Perez-Marcos, Bernhard Spanlang, and Mel Slater, "The contribution of real-time mirror reflections of motor actions on virtual body ownership in an immersive virtual environment," in *IEEE VR*, 2010, pp. 111–114.

[11] Anthony Steed, Sebastian Frlston, Maria Murcia Lopez, Jason Drummond, Ye Pan, and David Swapp, "An "in the wild" experiment on presence and embodiment using consumer virtual reality equipment," *IEEE TVCG*, vol. 22, no. 4, pp. 1406–1414, 2016.

[12] Jean-Luc Lugrin, Johanna Latt, and Marc Erich Latoschik, "Avatar anthropomorphism and illusion of body ownership in VR," in *IEEE VR*, 2015, pp. 229–230.

[13] Chris Christou and Despina Michael, "Aliens versus humans: Do avatars make a difference in how we play the game?," in *6th International Conference on Games and Virtual Worlds for Serious Applications (VS-GAMES)*. IEEE, 2014, pp. 1–7.

[14] Arvid Guterstam, Zakaryah Abdulkarim, and H Henrik Ehrsson, "Illusory ownership of an invisible body reduces autonomic and subjective social anxiety responses," *Scientific reports*, vol. 5, pp. 9831, 2015.

[15] Ferran Argelaguet, Ludovic Hoyet, Michaël Trico, and Anatole Lécuyer, "The Role of Interaction in Virtual Embodiment : Effects of the Virtual Hand Representation," in *IEEE VR*, 2016, pp. 3–10.

[16] William Steptoe, Anthony Steed, and Mel Slater, "Human tails: ownership and control of extended humanoid avatars," *IEEE TVCG*, vol. 19, no. 4, pp. 583–590, 2013.

[17] Nick Yee and Jeremy Bailenson, "The proteus effect: The effect of transformed self-representation on behavior," *Human Communication Research*, vol. 33, no. 3, pp. 271–290, 2007.