# Immersed gaming in Minecraft

Milan Loviska[1], Otto Krause[1], Herman A. Engelbrecht[2], Jason B. Nel[2],
Gregor Schiele[3], Alwyn Burger[3], Stephan Schmeißer[3], Christopher Cichiwskyj [3],
Lilian Calvet[4], Carsten Griwodz[4,5], Pål Halvorsen[4,5],
[1]Territorium-Kunstverein, Austria, [2]University of Stellenbosch, South Africa
[3]University of Duisburg-Essen, Germany, [4]Simula Research Laboratory, Norway, [5]University of Oslo, Norway
milan@loviska.com, info@ottookrause.com, {hebrecht,jason}@ml.sun.ac.za,
{gregor.schiele, alwyn.burger, stephan.schmeisser, christopher.cichiwskyj}@uni-due.de, {lcalvet,griff,paalh}@simula.no

## ABSTRACT

This demonstration will showcase mixed reality technologies that we developed for a series of public art performances in Vienna in October 2015 in a collaboration of performance artists and researchers. The focus of the demonstration is on natural interaction techniques that can be used intuitively to control an avatar in a virtual 3D world. We combine virtual reality devices with optical location tracking, hand gesture recognition and smart devices. Conference attendees will be able to walk around in a Minecraft world by physically moving in the real world and to perform actions on virtual world items using hand gestures. They can also test our initial system for shared avatar control, in which a user in the real world cooperates with a user in the virtual world. Finally, attendees will have the opportunity to give us feedback about their experience with our system.

## CCS Concepts

**•Information systems → Multimedia information systems;** **•Software and its engineering →** *Virtual worlds software;* **•Computing methodologies →** *Tracking;*

## Keywords

Minecraft; position tracking; head tracking; gestures

## 1. INTRODUCTION

Media content should be presented to an audience in the most immersive manner. The means of achieving this differ vastly. Nothing prevents a gamer from becoming totally immersed in the symbolic world of Dwarf Fortress, whereas a simple sharpening algorithm can hinder the biggest fan of a TV series from becoming immersed in a 4K movie on the latest TV set. The pinnacle of immersion is putatively achieved by allowing a user to interact freely in a virtual world by controlling every bodily aspect of their avatar through motions that feel natural and don't have to be trained for the specific activity in the virtual world.

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

*MMSys'16 May 10-13, 2016, Klagenfurt, Austria*

© 2016 Copyright held by the owner/author(s).

ACM ISBN 978-1-4503-4297-1/16/05.

DOI: http://dx.doi.org/10.1145/2910017.2910632

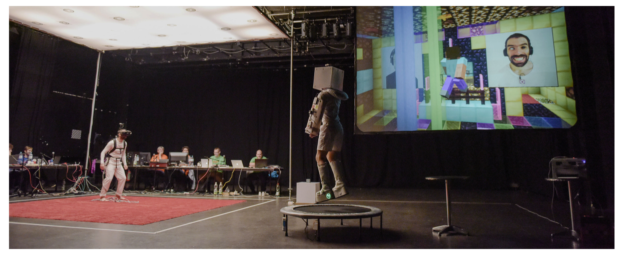

**Figure 1: Collaborative action. Foto: http://eSeL.at**

In recent times, a number of systems have reached the mass market that enable such a direct control over avatars in some manner. A commercial success was triggered by the physical interaction options provided by the input controllers of the Nintendo Wii console. Motion sensors embedded in the controller allow players to perform movements of, e.g., tennis rackets or car wheels. The Microsoft Kinect repeated this success in a different way, by interacting through gestures, without manipulating a physical device. In parallel to these new input devices, there is a re-awakening of the concept of virtual reality (VR) glasses (a 1960s concept [3]). These provide an immersive experience in terms of visual input by giving the user a first-person 3D view into a virtual world, while shutting out the real world. By combining these input and output devices, it is possible that we will achieve an unprecedented level of immersion.

Still, open questions remain. Current systems restrict the user to a relatively small physical area, and moving around in a large virtual world therefore requires different means, e.g., using a classical controller. Interactions require specific gestures that are often totally different from movement for an interaction in real life. As an example, jumping is usually achieved by pressing a button on a controller or by making a special gesture. This requires training, especially if very accurate inputs for complex interactions are needed. It can lead to totally different bodily experiences for the user and break immersion.

Our work aims at offering users in a virtual world a more natural experience with more freedom, e.g., allowing them to move larger distances without resorting to buttons on a device and enabling more complex interactions without lengthy training. In our demo, we invite one attendee at a time to explore a large Minecraft world that was created for a series of art performances in the Third Life Project. They can control their avatars through physical head and body movements, and manipulate the virtual world through hand gestures. Other attendees are not restricted to an observer role, but they are invited to perform real-world actions that exert additional influence onto these avatars. They can actively help or hinder the attendee exploring the world. Figure 1 shows a moment of collaborative dancing from the performance. To get an impression of the user's experience, a video that summarizes the performances in which the system was used can be accessed here: **http://thirdlifeserver.org/media.html**.

In the remainder of this demo paper, we first discuss the context of our work, the Third Life project. Then we present the user interactions of our system and how each one of them is realised. After the artists' experiences during the performances, we elaborate further on the proposed demonstration and conclude the paper with a short summary and outlook.

# 2. THE THIRD LIFE PROJECT

The Third Life Project[1] is an ongoing cooperation between performance artists and computer science researchers on the topic of "third life" – exploring the potential of virtual actions to transgress directly into reality and perform extravirtual actions.

Avatars afford us already the opportunity to live a "second life" in a virtual environment. Thanks to the psychological relativity of human perceptions, a "human mind largely treats virtual people just like physical ones" [1, p.84]. Therefore, avatars can address the issue of dual identities, exercising the gap between online persona and offline self and provide us tools to reconsider our own real life authenticity. The Third Life Project aims to extend the remarkable capacity of this technology and reduce the gap by inventing and implementing strategies and technology for direct engagement with elements of real environments through elements of the virtual ones. Positioning such a project within the context of contemporary digital and live performance art practice not only creates a platform for confronting and sharing ideas between groups that might not have an opportunity to come together otherwise, but through the inclusion of and direct interaction with larger audiences as well grants us a unique opportunity to examine playfully the social effects of representational practices in virtual environments.

The demonstration focuses on a single part of the Third Life Project, the direct influences that a performer's actions in the real world can have onto the virtual world by directly controlling a first-person avatar through a wide variety of movements. Attendees in the demonstration will perform this role. They will experience several aspects of the performance, namely (1) viewing the virtual environment, (2) interacting with virtual objects, (3) natural movement in the virtual environment, and (4) a multi-player control over a single avatar.

## 2.1 Virtual environment

We developed two Minecraft worlds, designed to showcase different actions like jumping, dancing, opening and closing of doors, as well as allowing the avatar to roam freely through different settings like indoor spaces, narrow hallways, staircases, tunnels, and large scale outdoors environments. The worlds help to establish a basic storyline while supporting the implementation of the used technologies in a way that can be intuitively understood by an observing audience. We chose Minecraft because of our previous experience, but applications in other types of virtual world would be possible as well. Both worlds contain two different virtual representations of the WUK performance venue[2], which serve as an entry and exit point from the real world to the virtual one and back. To break the logic of the real world, the first world around WUK is an open space that mashes up greenery with a desert environment, and contains a village, a huge eyeball hanging in the sky above and programming code flying loosely in the air. In the desert, one meets a giant server representing Kubrick's Space Odyssey 2001 monolith with a floating foetus inside. It contains a herd of non-player characters, virtual pigs, which can be released and guided back to the WUK in the course of the presentation. The second world is darker and fantasy-like with mushroom forests, cobwebs, water and lava beams that one can observe while travelling in a mine cart. The long railroad passes along the virtual upside-down version of the Museum of Modern Art in New York City (MoMA)[3] and leads to a discotheque, where the avatars can dance and afterwards get teleported back to WUK.

## 2.2 Viewing the virtual environment

For immersing a user in a virtual environment, it is important to embed the senses of the user into the 3D environment in a manner that is comfortable and natural. VR hardware is currently very popular and affordable as evidenced by the many companies developing VR gear. We chose to use the well-known head-mounted display (HMD) Oculus Rift DK2 (Oculus) for enabling the user to view the virtual environment. The advantage of using the Oculus is that the user is able to change his view of the world by simply moving their head. The Oculus promises low latency, with the goal of minimising simulation sickness for the user [4]. Another motivation for using the Oculus was the existence of an open-source project[4] that had already modified the Minecraft client to support the Oculus.

## 2.3 Interacting with virtual objects

Using the Oculus allows the user to see the virtual environment in a natural manner, but not to interact with virtual objects. It blocks out the real world, which introduces a barrier to the use of traditional input devices (keyboards, mice, etc.). To re-introduce interaction, and to make a natural experience, we want the user to use his hands instead of such devices. We use an off-the-shelf motion controller, the Leap Motion[5] to recognise hand gestures, effectively turning the user into the input device. The Leap Motion contains a pair of infrared cameras and its software recognises the user's hands in the cameras' field of view, which can then be used for gesture recognition.

We mount the Leap Motion on the front of the Oculus (using a mount that can be obtained from Leap Motion specifically for this purpose) and integrate it into the Minecraft client. Gestures are mapped to emulate keyboard or mouse button input, allowing users to interact with virtual objects. The users interact with virtual objects within range by aiming the cursor at the object with their right hand and then

[1] http://www.thirdlifeserver.org

[2] http://www.wuk.at

[3] Minecraft was added to the video games collection of the MoMA in 2013.

[4] https://github.com/mabrowning/minecrift

[5] https://www.leapmotion.com

perform a gesture to interact with the object, e.g., by forming a fist. To select objects and place them in the virtual environment, the users need to navigate and select items in an inventory. They interact with in-game menus (and the inventory) as if the menus are virtual blackboards hanging in front of them. It is noteworthy that the cursor can be moved independently of the users' looking direction.

## 2.4 Moving in the virtual world

The Oculus Rift allows the user to look around in the virtual world, but we wanted to achieve immersion beyond this. It should be possible to translate movement in the real world into movement in the virtual world.

Our approach uses a single camera worn by the user in addition to the Oculus and Leap Motion, which observes a set of pre-installed markers. These markers are CCTags[6] [2], due to their proven resistance to motion blur and partial occlusion. After reconstructing the placement of the markers in space, we can track the position and orientation of the user's torso, and translate this into movement in the virtual world. However, the real-world movement is naturally limited by the demo space. The user movement in most of the demo area is translated directly into movement in the virtual world with a 1:2 ratio, which was determined empirically as a compromise between movement accuracy and faster movement. However, the outer edge of the demo area is turned into a *scrolling area*. As soon as the user enters this region of the demo area, the avatar starts to move continuously in the direction that the user is facing. The user can change the movement direction by turning his body. To stop continuous movement, the user has to step back into the inner demo area.

To provide a haptic feedback for the transition from absolute movement area to scrolling area, we defined the inner area by a soft carpet, distinct from the hard, flat surface of the scrolling area. This allows a user to feel when he enters or leaves the inner area.

In the scrolling area, we use body orientation to determine movement direction, although we contemplated using absolute position for scrolling direction. With respect to movement direction, scrolling is a movement mode that allows the user to cross long distances in the virtual world. The independent movement of the view that is provided by the Oculus provides freedom, similar to looking out of a driving car. Observing a changing situation while scrolling, however, will eventually create the desire to change direction. With the position-based approach, this requires a step back onto the carpet, a move across the carpet, which leads to unintended local movement, and leaving the carpet for another scrolling direction. Instead, we found that relying on torso orientation, achieved better immersion. The user can turn inside the scrolling area, and changes scrolling direction in the virtual world. This is experience in way similar to, e.g., using roller skates and thus feels more natural. In essence, this allows the user to switch between something similar to normal pedestrian movement and floating.

[6]https://github.com/poparteu/CCTag


**Figure 2: Moving in Minecraft, Foto: http://eSeL.at**

## 2.5 Shared avatar control

Extending on the aforementioned interaction techniques, we also experimented with shared avatar concepts. A shared avatar is controlled by multiple users at the same time. This way each user can concentrate on different aspects of the avatar behaviour and can use optimised input devices. This requires the users to coordinate their actions with each other but has the potential to perform more complex avatar behaviours. In our experiment we let two users share one avatar. The first user was fully immersed into the virtual world and was moving and interacting using the techniques discussed before, i.e. an Oculus, a Leap Motion Controller, and our location tracking system. The second user was present in the real world and experienced the virtual world using a traditional 2D screen. This gave him the freedom to move around quickly in the real world and perform activities that would be difficult or potentially dangerous for the user immersed in the virtual world. We chose the second user to control avatar jumping and teleportation. To realise a natural interaction also for this user, these avatar behaviours were initiated by performing analogue activities in the real world. As an example, to make the avatar jump, the second user jumped on a real trampoline with embedded sensors. We also used this to let the avatar dance. To teleport the avatar to different virtual world locations, the second user carried a physical block to different locations on stage, similar to carrying the avatar to different target locations on a miniature map. Again, embedded sensors detected the block movement and reported it to our system. Overall, this can allow the second user to act as a kind of guardian angel for the first one. If the avatar is stuck in a hole in the virtual world, the users can combine their actions to jump the avatar out of the hole. If the avatar cannot get free or needs to get away quickly from a dangerous situation, the second user can initiate a teleport. Depending on the application, this may lead to interesting power dynamics between the users, like the second user playing with the first one by deliberately misusing his avatar behaviours.

Clearly, we only tried initial experiments with the shared avatar idea. Still, the potential for novel interactions in the virtual world as well as between users is very high. It should be examined in more detail.

# 3. PERFORMER'S EXPERIENCE

After describing how a user can perform different interactions, we now discuss how our performers experienced this system on stage. In general, immersion was greatly enhanced by giving the performer the ability to act naturally and experience the virtual world using different senses.

Due to tests and rehearsals, our performer was using our system for several hours per day, often with only a few short breaks. We expected that wearing the Oculus Rift DK2 with its relatively low-resolution display for such extended periods would lead to simulation sickness. This was true for our performer when wearing the Oculus Rift while sitting in front of a computer, even for short periods of use. However, when moving on the carpet in the physical space of the theatre stage with the headset on, no such condition was experienced, regardless of the duration of the interaction. The human body seems to be capable of adapting better if the visual input is accompanied by correspondent motoric activity. Please note that this seems to be true even if the user's movement in the real world is not perfectly reflected in the virtual world. Our system induced noticeable delay and – especially during tests – miscalculated movements.

Another observation is that our simple way of providing the performer some haptic feedback on his location in the real world (using the carpet) worked very well. Being "cut off" from the physical surrounding increases the danger of potentially loosing the orientation and location in the physical space. Walking barefoot on a carpet allowed the performer to be constantly reminded of the borders of his physical area. Besides, the square shape of the carpet relatively matched the blockiness of the Minecraft world and that geometrical correlation smoothed his orientation significantly. The performer could probe with his foot to find the carpet to prepare leaving the scrolling area. More sophisticated optical feedback like virtual shadows could be useful but could also break optical immersion.

An unsolved problem with the Leap Motion gesture recognition is that it requires the hands to be in its cameras' field of view to track gestures. This sometimes confused the performer, especially for abstract gestures. When trying to grab a virtual object, the performer intuitively looked at the object, thus causing no problem. However, we also experimented with hand gestures for more general activities, e.g., sending a signal to the outside world. This caused confusion since after a very short time the performer tried to send signals while looking around in the world. Not seeing the position of his own hands, he failed to notice that they were not visible to the Leap Motion camera, and assumed his signal was not received or ignored. In general, not being able to interact with objects without looking at them does not correspond with our real-life experience. Thus, mounting the Leap Motion on the Oculus Rift seemed to be disrupting immersion. This needs to be explored further.

Finally, we found that interactions that required both performers to cooperate with each other was very engaging for both of them. One performer could teleport the other's avatar from one virtual environment to another. By jumping on a real trampoline he could make the avatar jump repeatedly (dancing in the discotheque) or get him out of places he had fallen into. Especially the jumping case required exact coordination between the performers, because one performer had to walk in a certain direction when the other started jumping. This shows that the real-life multi-user cooperation can enrich immersion by giving the interactions with virtual environments an interpersonal dimension. As a side note, the performer in the real world also felt more immersed because he didn't feel like a mere observer.

# 4. DEMONSTRATION

After conducting several artistic performances, we want to give users the possibility of experiencing our system themselves. The user of the demo can either enter the virtual world by wearing the camera rig with Oculus and Leap Motion, or interfere with another user's actions, by use of the trampoline. After making some progress since the performances, the rig can be used wirelessly in the demo. The user can explore the Minecraft world freely and interact with it. The demo is rather resource-intensive. Whereas we will bring all equipment, there is a considerable demand for space of at least a 5x5$m^2$ area to experience both walking and scrolling operation, and the installation of the CCTags benefits from a ceiling height of at least 5$m$.

# 5. CONCLUSION AND FUTURE WORK

Our results so far are promising. We were able to create a high level of immersion by combining a number of – individually well known – interaction techniques into a single system. Extending sophisticated technology like VR HMDs and gesture tracking with haptic feedback from a passive carpet as well as the idea of a scrolling area, our performers could control their avatar with relative ease and accuracy without extended training sessions.

In future work we plan to continue our collaboration and extend our system with new and enhanced interaction features. First, we plan to optimise the user experience by reducing system latency and giving the user better feedback about the current localisation accuracy. Second we want to extend the system with novel ways of how actions in the virtual world might influence the real world. So far, this is restricted to controlling the lighting on stage. Finally, we also plan to increase audience immersion by giving audience members more abilities to interact with the virtual world themselves.

# 6. ACKNOWLEDGMENT

The Third Life Project was supported by City of Vienna's Department of Cultural Affairs, the Arts Division; the Arts and Culture Division of the Federal Chancellery of Austria; the FP7 project no 609757, FiPS; the H2020 project no 644874, POPART; Stellenbosch University in South Africa. The authors thank the LABO Mixed Realities and the University of Bielefeld for their contributions.

# 7. REFERENCES

[1] J. Blascovich and J. Bailenson. *Infinite Reality: Avatars, Eternal Life, New Worlds, and the Dawn of the Virtual Revolution*. William Morrow, New York, NY, USA, 2011.

[2] L. Calvet, P. Gurdjos, and V. Charvillat. Camera tracking using concentric circle markers: Paradigms and algorithms. In *Image Processing (ICIP), 2012 19th IEEE International Conference on*, pages 1361–1364, Sept 2012.

[3] M. L. Heilig. Stereoscopic-television apparatus for individual use, Oct. 1960. US Patent 2,955,156 A.

[4] E. Regan and K. Price. The frequency of occurrence and severity of side-effect of immersion virtual reality. *Aviation Space and Environmental Medicine*, 65(3):527–530, 1994.