Co-Experiencing Virtual Spaces: A Standalone Multiuser Virtual Reality Framework

Chris Elvis Leisi and Oliver Sahli
Immersive Arts Space, Zurich University of the Arts
Pfingstweidstrasse 96, CH - 8005 Zürich
{chris.leisi,oliver.sahli}@zhdk.ch

Abstract
This paper shows the development of a standalone, co-location multiuser virtual reality tracking system. The assumption is that multiuser VR can greatly increase the immersive aspects and real-world presence and can cause the virtual environment to create a more powerful impression than when experienced by a single user. We developed and released a framework on Github under an CC-license [1] for the accessible game engine Unity [2]. This allows multiple users to join the same virtual environment while also sharing the same real-world space congruently. For this purpose, their tracking areas are aligned, and information is exchanged over a network.

Keywords: Virtual Reality, Multiuser Experience, Tracking Alignment, Immersion, Embodiment, Virtual Environment, Same-Space

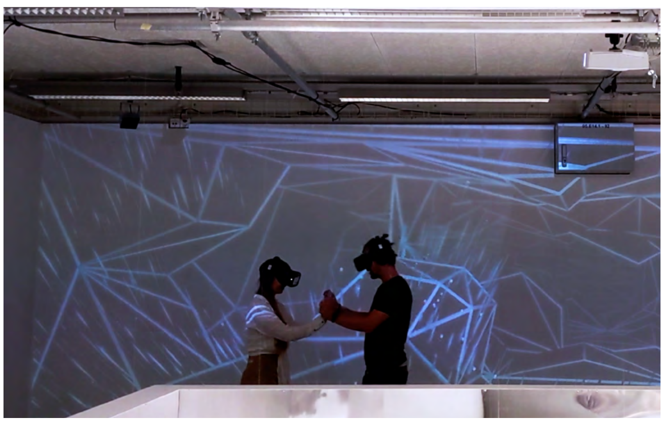

Figure 1: Echolocation at REFRESH#3 Source: Leisi and Sahli

1. Introduction
Conventional co-location multiuser systems [3] require complex tracking systems, calculation machines and usually use backpack computers to power the VR glasses. They are therefore mostly installed in one room and are rarely transported, which makes home use unthinkable. In this paper, we discuss a system that works with standalone VR glasses like the Oculus Quest and a local WLAN network. Such a system can be used in several rooms and even outside without much effort. The most important part of such a system is the linking of the Cartesian tracking areas of the VR glasses. They must all be calibrated to the virtual zero points of the respective Cartesian coordinate systems. After that, the movements and data of the users are transmitted to the other headsets via the local network.

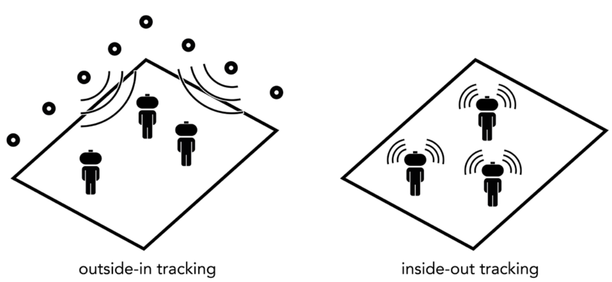

Figure 2: Different tracking systems Source: Leisi and Sahli

2. Technical challenges
2.1. Used tracking systems
Newer VR glasses like the Oculus Rift S and the Oculus Quest/2 [4] rely on inside-out tracking. This is either RGB or BW camera-based tracking, which detects the space via images. In contrast to an outside-in system, the cameras are not installed in the room but are built into the VR glasses. The Oculus Quest, which is used for this system, has four wide-angle black and white cameras installed. These can recognise the environment and thus determine their own rotation and position in it.

2.2. Calibration Process
Using the information obtained with the cameras built into the Oculus Quest, it could be possible to match the orientation of the VR glasses via fixed points in space. However, at the time of publication, access to the camera images is not available for developers. Therefore, another solution for how to align the coordinate systems and tracking of the VR glasses has to be found. Each pair of VR glasses uses its own zero point in space and initial orientation at the outset. For co-location tracking, one could override the zero point with a calibration process that all VR glasses must perform. Since most VR headsets have two controllers that are also recognised in space, they can be used for such a calibration. Thus, if all users place their right controller on a marked point in space, which is ideally the center of the room, and the left controller on a second marked point representing the forward orientation of the virtual environment, each user can generate a more accurate zero point in the real world. It's important that all users place and align the controllers on the same points. To simplify this process, a graphic was developed which can be stuck on the floor as an aid. It shows how and where the controllers have to be positioned. Initial tests

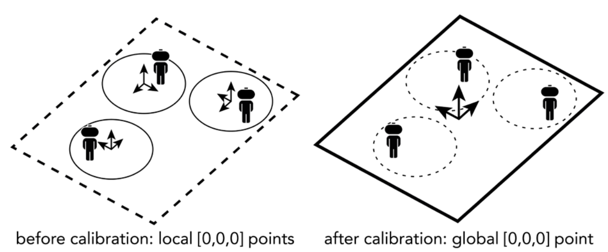

Figure 3: Local and global zero points Source: Leisi and Sahli

show that the offsets of the individual environments are less than a few centimetres. Even if the user is several meters away from the calibration point, the offset is not noticeably larger. Thus, two-point calibration allows the definition of a accurate zero point across a range of virtual environments.

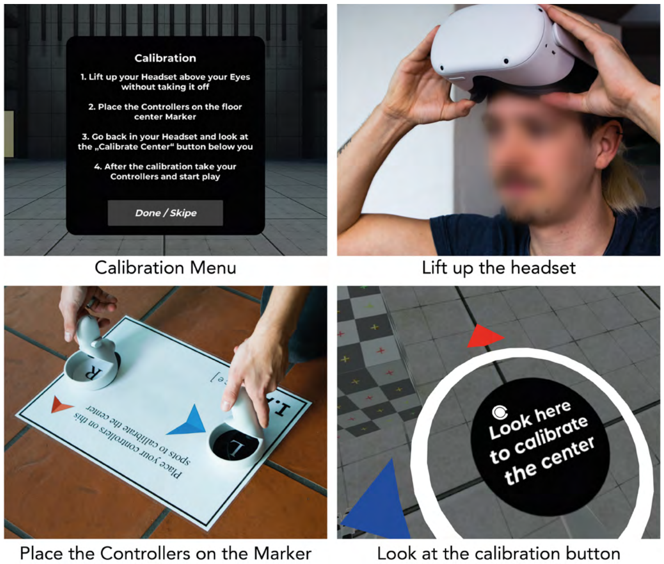

Figure 4: Calibration process Source: Leisi and Sahli

2.3. Virtual mapping of the real room

In the fall of 2020, during the exhibition of the REFRESH#3 [5] conference at the Zurich University of the Arts (ZHdK), we tested the „Echolocation“ project, which visually displays audio input received by the microphone of the VR glasses in a co-location multiuser variant as applicable in an exhibition context. To this end we recreated a gallery space in the ZHdK using 3D software and extended it virtually. To match the real and virtual environment, a fixed zero point is defined on the floor of the virtual environment at which the users can perform the calibration with the controllers in the real environment. The tracking is very stable and even when walking around in other rooms, there are only small displacements of up to 5 centimetres.

2.4. Tracking improvements

Poor lighting conditions or monotonous rooms can lead to position shifts or the suspension of the inside-out tracking of the VR glasses. Under these conditions, glasses can sometimes recover their positioning and only suffer from small displacements but most shifts are larger and the user has to recalibrate the zero point. On the other hand, if the rooms are well illuminated, tracking usually works very well, even over several floors. To improve the tracking in low contrast rooms, we put lines and symbols with white tape all over the floor area, including on some walls. This noticeably improved the tracking of the VR glasses.

2.5. Multiuser Oculus Quest hand tracking

In addition to the tracking alignment, a networking system is also needed to obtain the orientation data of the fellow users and their interactions. For the data exchange we used the open-source system Mirror for Unity [6]. To enhance the immersion and sense of space, the hand tracking of the Oculus Quest has been incorporated.

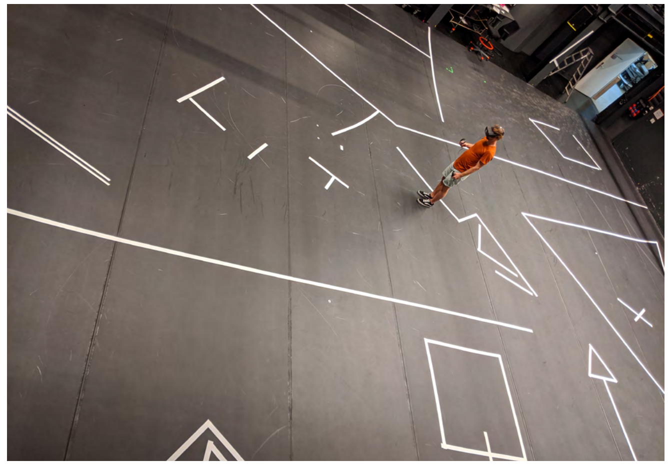

Figure 5: Tracking test in the IASpace Source: Leisi and Sahli

This system can recognise the user‘s hands and transfer their position and finger alignment to the virtual environment. A hand recognised by this tracking consists of 18 bones, one of which is the palm. In order to make the finger data visible on the other devices, the orientations of the individual fingers are sent to the other devices via networking. A quaternion, which describes the rotation of an object in virtual space, contains four floating point numbers. This would add up to 68 floating point numbers per hand, which would have to be transmitted per second at the server synchronisation rate. To reduce this amount of data, we tried to simplify the tracking data by abstracting the hand. The thumb and pinky each have 4 bones, the first of which has very little room to move. Therefore, these are omitted for synchronisation. The first bone of each finger, which is attached directly to the palm of the hand, is the only one that can have large rotational changes in all directions with most people. Therefore, the rotation of these five bones are not simplified and completely transmitted. The last two bones of each finger, for the most part, can only move on one axis. For most people, these two bones always move simultaneously. Therefore, it is possible to map the curvature of the fingers in the system by allocating a floating point number from 0 to 1. The value 1 in this example would be a stretched finger, 0 a completely bent finger. With this method, only 25 floating point numbers per hand need to be exchanged over the network, allowing users to perform more accurate hand gestures.

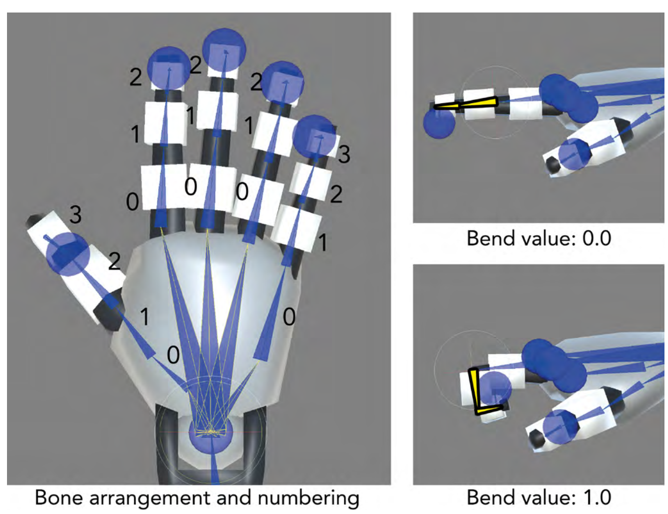

Figure 6: Bone arrangement and approximation Source: Leisi and Sahli

2.6. The view into the virtual environment

The Oculus Quest/2 does have the option of transferring one‘s own image to a cell phone or TV, but this function is passive and only allows the perspective of a headset. Having control over the viewing angle and thus being a part of the virtual environment itself gives more feature for interactions and observations with the virtual environment. Today‘s AR features on Android and iOS devices also have very solid spatial self-location in the virtual environment when starting from a coordinated zero. We used this smartphone tracking functionality via ARCore (Android) or ARKit (iOS) for the respective devices and in order to also be in the same virtual environment relative to VR users, the smartphone must go through a similar zero-point calibration process. Since an image is already attached to the ground for VR device calibration, this can be used via the image recognition functions of the AR libraries. As soon as this image is recognised, the system allows one to determine the position and orientation of the image in space as a zero vector. With this system it is possible to see the virtual environment via the smartphone.

3. Results

The inside-out tracking of the individual Oculus VR glasses works well and shows almost no tracking shifts over time in well-lit and varied rooms. Tracking is stable enough that it works across multiple rooms as well as multiple floors. The calibration process leads to an accurate result. However, it became apparent, especially during the exhibition at REFRESH#3, that particularly those users unfamiliar with VR experience uncertainty when it comes to calibration. The VR glasses have to be lifted to see the calibration mark, but not so far away that the proximity sensor of the VR glasses puts them into power saving mode. In low light conditions or with little discernible differences in the environment, major tracking failures can occur. The users do not receive any information that their tracking no longer matches the other users and run the risk of moving into them or colliding with the environment.

Calibrate the smartphone via image tracking

The VR-headsets and the smartphone share the same tracking space

Figure 7: AR camera calibration and view   Source: Leisi and Sahli

4. Conclusion

We developed a system, which is easy to transport and can be used almost everywhere. It can be extended with other tracking systems like the ARCore tracking with little effort. Also, the calibration process via the two controllers is not tied to the Oculus Quest. All VR glasses with two controllers can be calibrated this way. Here, the different shapes of the controllers would have to be taken into account in order to minimise variation between different devices. Further development and the possible introduction of functions which would increase the mixed reality capability of the Oculus Quest is planned [7]. The calibration process could be simplified and enhanced by using passthrough integration with mixed reality features. Furthermore, we will try to find a solution to detect or even prevent large tracking shifts. Not much is known about the real functions yet, but if access to the 3D tracking data were possible, a system could be developed which would make the calibration process with the controllers superfluous. Instead, the VR glasses might exchange their 3D data with each other directly. With such a system, the virtual environments could recalibrate themselves in the event of a tracking dropout.

5. References

[1] Github Repository, https://github.com/immersive-arts/Co-Experiencing-Virtual-Spaces (28 December 2020, 14:30)
[2] Unity 3D Game Engine, https://unity.com (18 December 2020, 13:00)
[3] Chagué S, Charbonnier C. 2016. Real Virtuality: A Multi-User Immersive Platform Connecting Real and Virtual Worlds, VRIC 2016 Virtual Reality International Conference - Laval Virtual, Laval, France, ACM New York, NY, USA. https://www.artanim.ch/pdf/publications/54.pdf
[4] Oculus Insight Tracking, https://ai.facebook.com/blog/powered-by-ai-oculus-insight/ (14 December 2020, 12:05)
[5] REFRESH#3 conference, https://refresh.zhdk.ch (23 December 2020, 20:00)
[6] Mirror Networking, https://mirror-networking.com (29 November 2020, 12:40)
[7] Oculus MR Function, https://developer.oculus.com/blog/facebook-connect-oculus-quest-2-aaa-gaming-partnerships-and-more/ (29 November 2020, 12:30)