

Digital Twin Construction for Real-World Metaverse: A Case Study of a Collaborative Escape Game

Renta Inoue^{*}, Keigo Hattori¹, Hayato Iwasaki¹, Fumihiko Nakamura¹, Asako Kimura¹ and Fumihisa Shibata¹

¹ College of Information Science and Engineering, Ritsumeikan University, Osaka, Japan

Abstract

We are planning to construct a Mixed Reality (MR) campus that realizes a real-world metaverse by merging physical and digital spaces, where the MR campus is built as a Digital Twin (DT) that mirrors the real-space campus in Virtual Reality (VR) space, aiming to enable collaborative work between users existing in both spaces. In this study, we conducted a basic investigation for the construction of the DT through the production of a collaborative escape room game using an asymmetrical MR environment. In the game, rooms in real space are reproduced as a DT in VR space. We confirmed that the MR user in the real space and the VR user immersed in the VR space with differently sized virtual bodies could see each other and work together interactively to achieve their goals.

Keywords

Multi-user Collaboration, Asymmetric Mixed Reality, Metaverse, Digital Twin

1. Introduction

Various methods of utilizing Digital Twin (DT) in a Mixed Reality (MR) environment have been proposed. Among them, a real-world metaverse proposed by Niantic is a new type of fusion of metaverse and MR [1]. It aims to duplicate the real world into a virtual environment, which can be accessed not only from the real world but also from remote locations.

By combining information in the real space with digital information in the MR space, it is possible to create a sense of presence as if the remote user actually exists in the real space. We expect to see applications in a wide range of fields, including business as well as entertainment.

For example, Zaman *et al.* developed a system that enables remote users to interact with local users by accessing a collaboration space in the real world and demonstrated an improvement in the sense of presence and the ability to perform tasks [2].

Lee *et al.* have shown that sharing a host space with others using 360-degree video and allowing independent viewing ability for each user can lead to improved user presence in collaborative tasks [3].

In addition, a system utilizing asymmetric VR environments has been proposed by Ibayashi *et al.* The proposed system provides both an internal view using VR and a top-down view through a table-type device for architectural design. This system enables communication through ceiling transparency and user gestures [4].

Cho *et al.* proposed an asymmetric VR environment where participants can join from various platforms, including PCs, mobile devices, AR, and VR. It was found that mobile devices and AR lack immersive experience, highlighting the importance of interface design tailored to each platform [5].

We envision an MR campus that reproduces the real-world campus of the university as a DT in a VR space. Users in remote locations can interact with avatars in the VR space. Users in remote locations can also move around as avatars in the VR space constructed as a DT, and at the same time, they can interact with local users by mapping their real-space positions and postures to the VR space. Therefore, in our MR Campus, users in local and remote locations can collaborate as if they were in the same space through the interaction of both spaces by the DT.

In order to realize the concept of the MR Campus, this study created and exhibited cross reality (XR) content using a real stage set (Figure 1 (a)) and its 3D model (Figure 1 (b) and Figure 1 (c)) as a basic study. By constructing a DT of the real space and implementing object and user interactions between the two spaces, we examined the challenges toward realizing the concept.

2. Challenges to realization

One of the challenges in realizing the MR campus is sharing the user's position and posture. Specifically, how to present the position and posture of a user in real space to a user in VR space, and how to present a user in

APMAR'24: The 16th Asia-Pacific Workshop on Mixed and Augmented Reality, Nov. 29-30, 2024, Kyoto, Japan

^{*} Corresponding author.

✉ inoue@rm2c.ise.ritsumei.ac.jp (Renta Inoue);

hattori@rm2c.ise.ritsumei.ac.jp (Keigo Hattori);

iwasaki@rm2c.ise.ritsumei.ac.jp (Hayato Iwasaki);

f-nakamu@rm2c.ise.ritsumei.ac.jp (Fumihiko Nakamura);

asa@is.ritsumei.ac.jp (Asako Kimura);

fshibata@is.ritsumei.ac.jp (Fumihisa Shibata);



© 2024 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).



(a) The real stage set (b) The 3D model of the set (c) The 3D model with virtual objects
 Figure 1: (a) The real stage set, (b) its digital twin and (c) the game scene based on the 3D model.

VR space to a user in real space. In the former case, in particular, there are many issues to be considered, such as how to track and represent not only the position and posture of the head but also the entire body. Other major issues include aligning the two spaces and synchronizing the position and posture of objects.

In this research, we constructed an asymmetrical MR space in which the user in the real world is represented by a life-sized avatar, while the user in the VR space is represented by a scaled-down avatar. We created a collaborative escape room game using this asymmetrical MR space and considered how to share position and posture between the two spaces, how to facilitate interaction between users, and how to align the two spaces. By framing the work as a game, we invited members of the public who are unfamiliar with VR and MR technologies to experience it and discussed the challenges involved.

3. Overview of the game

There are two players in the collaborative escape room game using an asymmetrical MR environment. One player, the human player, takes the role of a phantom thief in the MR space, while the other player, the mouse player, takes the role of a mouse in the VR space. They interact with each other to complete tasks and aim to escape from the room. The human player and the mouse player can see each other, and their movements are reflected in the other's space. The human player can move by walking around in the real space, while the mouse player can move by alternately shaking the controllers.

Additionally, the human player can grab and move the mouse or items using hand-gesture operations (Figure 2). Due to the difference in body size between the human and mouse, tasks in narrow spaces are handled by the mouse player, while tasks such as assisting the mouse's movements or operating levers and other mechanisms are handled by the human player, allowing them to collaborate effectively.

3.1. Flow of the experience

First, both players start in a dark room. The human player can turn on the ceiling light by touching a virtual

switch, allowing them to begin performing various tasks. Next, both players must search for three items required to escape from the room. There are three major challenges, and by clearing each one, they can obtain the item needed for escape. Three items must be acquired within the time limit of 5 minutes.

The first challenge is a maze (located inside the transparent case at the center bottom in Figure 1 (c)). After the human player finds the mouse, they pick it up and carry it to the entrance of the maze. Since the maze is too small for the human player to enter, the mouse player navigates the maze under the guidance of the human player. By pushing out the item inside the maze, the mouse player makes it possible for the human player to grab the item and obtain it.

The second challenge involves pipes and handles (Figure 3). Similar to the maze, only the mouse player can enter the pipes. The two differently colored handles exist as physical objects and some colored pipes are linked to each handle of the same color. By manipulating these handles, the human player rotates the pipes to create a path to the item. The mouse player moves



Figure 2: Hand-gesture of grabbing mouse.



Figure 3: The pipes and handles of the challenge.

forward inside the pipes, collaborating with the human player who manipulates the rotation of the pipes.

The third challenge involves a lift and a lever (Figure 4). The item is placed near the ceiling, out of reach of the human player. By moving the lever up and down, the lift moves correspondingly. The mouse player gets on the lift and the human player manipulates the lever to raise the lift. If the lift moves near the ceiling, they can obtain the item. The books placed on the shelves act as obstacles to the lift's movement. Therefore, players need to remove these books while raising the lift.

After finding the three items, the human player places them inside a virtual attaché case and carries it by grabbing its handle. The game is cleared when both players move with the case to the door. However, if the players exceed the 5-minute time limit, a locking mechanism is triggered on the door, resulting in a failed escape.

4. System configuration

The system configuration is shown in Figure 5. In this setup, both the human player and the mouse player wear the Meta Quest 3 [6] as their head-mounted display (HMD). The Meta Quest 3 was selected due to its support for pass-through, hand tracking, and spatial anchors, and it offers extensive functionality through the Meta XR SDK. The human player utilizes the pass-through feature to overlay virtual objects onto the real-world setup, while the mouse player only sees the VR space. The Unity game engine was used for game development to leverage the Meta XR SDK. The Meta Quest 3 headsets are connected to desktop PCs via Quest Link to run the game on the PCs.

4.1. Synchronization between two players

Photon Unity Networking 2 (PUN2) [7] is used to synchronize actions, position data, and other information between the two players. PUN2 was selected for its easy integration with Unity projects and the extensive API. Data such as coordinates, and rotation information are exchanged via UDP communication through Photon's cloud server. The identification and connection of synchronized information are managed using name servers, regions, and room IDs.

To reduce the bandwidth load, neither the player-specific in-game camera nor the coordinate information from the HMD is synchronized directly. Instead, only the players' avatar information is synchronized using PUN2. The avatar's coordinates are obtained from the head position tracked by the HMD. The movement of the avatar is achieved through inverse kinematics, using four points: the coordinates of the left and right hands, the head, and the floor. For the hand coordinates, the

human player uses the hand positions recognized via the hand tracking feature, while the mouse player's hand positions are obtained from the controllers held in both hands.

Additionally, the status of item acquisition and switch operations is also synchronized using PUN2. The system is designed so that variables on the human player are synchronized and can be used as shared variables.

Furthermore, any virtual object in the MR space can be moved by either player. In PUN2, each object must have an owner, and any movement of an object by anyone other than the owner is not reflected on the other side. Therefore, ownership of the object is primarily held by the human player, and ownership is transferred only when the mouse player touches the object to enable smooth interaction with the object.

4.2. Position alignments

In this game, the human player sees virtual objects like a maze and pipes displayed in the physical room. These objects are aligned using the Meta Quest 3's Spatial Anchor feature, and all virtual objects are positioned relative to a single Spatial Anchor.

The entire room in both spaces is synchronized using the previously mentioned PUN2 in terms of position and orientation. At the start of the game, the entire room in the VR space is aligned with the physical room in the MR space.

Additionally, since the mouse player's starting point is set at a specific position within the room object, the mouse player's position is aligned accordingly. This



Figure 4: The lift and lever of the challenge.

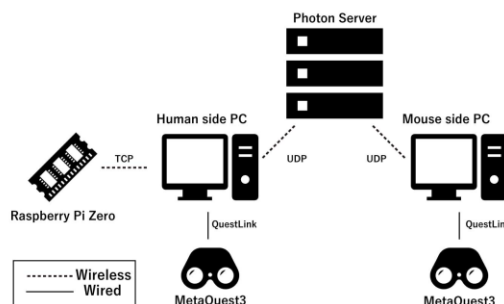


Figure 5: System configuration

process ensures that position alignment is maintained between the physical and virtual spaces, correctly reflecting not only the position of virtual objects but also the direction and posture of both players.

4.3. Operation using physical handles

This section provides a detailed explanation of the challenge involving pipes and handles described in Section 3.1. The virtual pipes that the mouse player passes through have red and blue sections, which can be rotated by turning the corresponding physical handles of the same color.

Initially, the pipes are oriented in different directions and do not connect to each other. The human player can use the physical handles to rotate and connect the pipes, allowing the mouse player to pass through. The handles use a steering controller attachment with a Raspberry Pi Zero featuring a compact size and Wi-Fi capability. An MPU-6050 gyroscope is attached to the axis of each handle, and it rotates together with the handle. The data obtained by the gyroscope is sent to the Raspberry Pi Zero for angle calculation.

The calculated angle data is then transmitted via TCP to the PC for the human player running Unity, where it updates the angles of the virtual pipes in real time.

4.4. Interaction using hand tracking

The human player can interact with virtual objects using hand tracking. In this game, the player performs two types of action.

The first action is a poking gesture. This is implemented using the Poke Interaction feature from the Meta XR Interaction SDK (Interaction SDK). This action is applied to the switch-type object within the space, allowing the player to toggle the switch on and off by pressing it with their fingertip.

The second action is a grabbing gesture. This is implemented using the Grab Interaction feature from the Interaction SDK. This action is used when the human player picks up and moves the mouse and other virtual objects within the space. By making a fist as if they are actually grabbing something, the virtual object follows the movement of the player's hand.

With the grabbing gesture, the human player can move virtual objects to any location within reach, which may occasionally bypass collision detection among objects. There was a concern that if the mouse player were moved into virtual objects by the human player, the maze or pipes could be unintentionally cleared. To prevent this, the system is designed to forcibly release the grab state if the mouse object collides with specific objects, thereby preventing virtual objects from being passed through.

4.5. Mouse player controls

The mouse player is positioned next to the physical set used by the human player. The player moves forward by alternately shaking the controllers. The speed of forward movement is determined by the absolute speed of the controller's movement; the faster the controller is moved, the greater the forward movement. To prevent unintended forward movement due to slight motions of the controller, the player moves forward only when the controller's speed exceeds a certain threshold.

In the development of VR and MR content, for tracking the movement of a controller, it is common practice to detect changes in the controller's position and rotation within Unity rather than directly referencing raw sensor values from the controller. Therefore, in this game, the speed is calculated based on the controller's position relative to the player's head.

The player's direction of movement is determined by the orientation of their head, allowing them to move forward in the direction they are looking. The mouse player encounters a section where they must climb through a vertical pipe. While in this section, forward movement is disabled, and the controls automatically switch to an upward movement, allowing the mouse to climb the pipe. Additionally, when the mouse player is being held by the human player, the mouse player's movement is disabled.

Although the mouse player plays next to the set in this game, the system is designed to allow remote operation from a distant location.

5. Other considerations

5.1. Lighting expression

In this game, there are multiple lighting patterns, such as the player turning on the lights at the start or the lights turning red when the time limit is approaching. These are expressed differently in the VR space and the MR space.

The VR space applies simple lighting expressions with Unity's Directional Light, such as disabling lights and changing colors.

In the MR space, since the real lighting could not be manipulated due to the exhibition, virtual objects are represented in the same way as in the VR space using Directional Light, while physical objects are represented by applying black or red filters to the pass-through video to represent changes in lighting (Figure 6).

5.2. Sound expression

In this game, voice communication between players is possible. This allows collaboration and communication even when the mouse player is in a remote location.

Additionally, some objects emit sound effects. These sound effects can be heard in the appropriate direction and volume in the space not only in VR but also in MR.

5.3. Perspective expression by occlusion

In the MR view of this game, an occlusion feature is implemented using the Meta Quest 3's Depth API. All virtual objects displayed in the MR view, including the mouse, are occluded by the player's hand and other physical objects. In Figure 7, the arm and handles of the physical objects occlude virtual objects such as a wall or pipes.

However, only the mouse is outlined so that the player can locate the mouse when it is occluded. This allows both players to interact smoothly. In Figure 8, the mouse is under the table of the physical object, and the mouse object itself is occluded, but its outline is still visible.

6. Exhibition and findings

6.1. Exhibition results and identified issues

This game was exhibited at the Ibaraki X Ritsumeikan DAY 2024 event (Figure 9) on May 19, 2024. A total of 87 groups, comprising 174 participants aged 12 and older, experienced the game (some participants may have participated multiple times).

Most participants were able to clear the game within the 5-minute time limit and successfully finished the three challenges. However, several issues related to convenience and operability were identified during the exhibition. Two key issues are highlighted below.

Connection Instability: There were various issues with the connection between the Meta Quest 3 and the PC. Problems included the pass-through function failing to activate at the start of the game and position misalignment. These issues are believed to stem from the fact the pass-through feature and the use of spatial data in Quest Link being in the preview stage. It is expected future updates to the Meta Quest 3 will resolve these problems.

Hand Tracking Limitations: Hand tracking only functions within the HMD's field of view. When players faced different directions while operating levers or grabbing the mouse, their hands were not recognized, making it difficult to perform the intended actions comfortably. A fundamental solution cannot be achieved with the Meta Quest 3 alone. To address this issue, additional devices like 360-degree cameras could be used to provide a more comprehensive tracking system. These cameras would allow for hand tracking from all angles, ensuring that the player's hands are always



Figure 6: MR view applied black filter.



Figure 7: Occluded virtual objects.



Figure 8: Outlined mouse occluded by a physical table.



Figure 9: Exhibition setup.

detected, regardless of their position relative to the headset.

6.2. Key findings and future challenges

There are four main findings from this and its exhibition that will help realize the MR campus.

Accurate Positioning and Avatar Synchronization: Accurate positioning ensured that the positional relationship between players was as perceived by the players. This allowed players to complete challenges smoothly by indicating directions and item locations to each other. This is effective for collaborative tasks and important for user interaction.

Avatar Visibility and Voice Communication: Although participants were playing in different spaces, some felt as if they were in the same space. This was likely because they could see each other's avatars and communicate through voice. This enhanced the sense of presence and interaction.

Aligning the two Spaces: In this exhibition, one person played on the MR side and one on the VR side. From a technical standpoint, the VR side supports multiple participants because its positioning is based on the physical space. However, the MR side does not support multiple participants, as each instance has its own coordinate system, and integration between these spaces was not implemented. This requires further study.

Synchronizing Objects in Both Spaces: In this game, synchronization of real-world to virtual objects was achieved by sensing the angle of the handle in the physical space and reflecting it in the VR space. This allows virtual objects to be operated even when the hands are outside the HMD's field of view, as mentioned earlier. Compared to virtual objects like the lever and the lift, this method offers better operability. Since the MR campus envisions interaction in both spaces, synchronization of virtual to real objects should also be achieved. This would eliminate gaps in object alignment between the two spaces, improving communication based on positional relationships. Furthermore, enabling remote users to manipulate physical objects would allow them to take a more active role in collaborative tasks. This could be accomplished by utilizing and developing the system proposed in [8].

7. Conclusion

We developed and showcased content where two players, existing in MR and VR spaces, collaborate to achieve an escape goal. Each space was constructed using a real-world set and a 3D model that replicates it as a digital twin (DT), designed to enable mutual interaction. Several issues were identified during the exhibition, and addressing these challenges is essential to advance the realization of the MR campus concept.

Acknowledgements

This work was partially supported by JSPS KAKENHI Grant Number JP23K21690.

References

- [1] Yuji Higaki: "Building the real-world metaverse," 2022. URL: <https://nianticlabs.com/news/building-the-real-world-metaverse>
- [2] Faisal Zaman, Craig Anslow, Andrew Chalmers, Taehyun Rhee: "MRMAC: Mixed reality multi-user asymmetric collaboration," Proc. International Symposium on Mixed and Augmented Reality (ISMAR 2023): 591-600.
- [3] Gun A. Lee, Theophilus Teo, Seungwon Kim, Mark Billinghurst: "A user study on MR remote collaboration using live 360 video," Proc. International Symposium on Mixed and Augmented Reality (ISMAR 2018): 153-164.
- [4] Hikaru Ibayashi, Yuta Sugiura, Daisuke Sakamoto, Natsuki Miyata, Mitsunori Tada, Takashi Okuma, Takeshi Kurata, Masaaki Mochimaru, Takeo Igarashi: "Dollhouse VR: A multi-view, multi-user collaborative design workspace with VR technology," Proc. SIGGRAPH Asia (2015)
- [5] Yunsik Cho, Myeongseok Park, Jinmo Kim: "XAVE: Cross-platform based asymmetric virtual environment for immersive content," IEEE Access, Vol. 11, pp. 71890-71904, 2023
- [6] Meta Quest 3. URL: <https://www.meta.com/jp/quest/quest-3/>
- [7] Photon Unity Networking 2. URL: <https://doc-api.photonengine.com/en/pun/current/index.html>
- [8] Yumi Fukuda, Ayumu Shikishima, Asako Kimura, Hideyuki Tamura, Fumihisa Shibata: "RV-XoverKit: Mixed reality content creation toolkit to connect real and virtual spaces," Proc. Asia-Pacific Workshop on Mixed and Augmented Reality (2022)