PAPER • OPEN ACCESS

3D telepresence for remote collaboration in extended reality (xR) application

IOP Conf. Series: Materials Science and Engineering **979** (2020) 012005 doi:10.1088/1757-899X/979/1/012005

3D telepresence for remote collaboration in extended reality (xR) application

F E Fadzli [1,2] *, M S Kamson[1,2], A W Ismail [1,2] and M Y F Aladin [1,2]

[1] Mixed and Virtual Environment Research Lab (mivielab), ViCubeLab, Universiti Teknologi Malaysia, 81310 Johor, Malaysia
[2] School of Computing, Faculty of Engineering, Universiti Teknologi Malaysia, 81310 Johor, Malaysia

* fazliaty.edora@gmail.com

Abstract. Extended Reality (xR) encapsulates various computer-altered reality technologies that cover virtual reality (VR), augmented reality (AR), and mixed reality (MR). xR is a technology that merges a virtual element into the real-world with the aims to enhance reality on the virtual world immerse onto real-world space. xR has been improved from time to time as the advanced immersive technologies to extend the reality we experience by either combining the virtual and real worlds or by generating a fully immersive experience. Remote collaboration in xR is a challenge since both parties need to have the same system and to set to parallel to xR environments. In a collaborative interface context, the user can be in a remote collaboration or face-to-face to sense the immersive environment. Human teleportation is transferring a human from a local location to a remote location, where the reconstruction of a human appears in a realistic visual representation. However, creating a fully realistic representation of the human figure need a complex 3D reconstruction method. Therefore, the paper describes the human teleportation in the xR environment using the advanced RGB-D sensor devices. It explains the phases to develop the real human teleportation in xR. The paper also discusses the proposed collaborative xR system that has successfully actualized the user teleportation.

1. Introduction

As technology becomes advance, there is a lot of device far more significant that can replicate and imitate our real-world environment. These technologies called Extended Reality (xR). xR encapsulates various computer-altered reality technologies that cover virtual reality (VR), augmented reality (AR), and mixed reality (MR). MR can bring us into space with computer visual graphics. MR is a technology that merging of real and virtual worlds somewhere along the "Reality-Virtuality continuum," which connects authentic environments to completely virtual ones [1]. MR aims to enhance reality in the virtual world immerse in real-world space. There is much research on Extended Reality (xR), which in every way, makes the human lifestyle more present in the virtual world. In order to create an xR application, accurate positioning and tracking require aimed at aligning the virtual and real world. xR is different from VR in the way of user able to manipulate the object such as in the MR environment while able to compare with real-world space and more similar to AR. MR also make the user aware of the surrounding while not ignoring the surround. Therefore, this paper describes how the MR provides user best realism in virtual and real-world space better than AR and more presence in VR.

The human teleportation in the xR environment using the advanced RGB-D sensor devices [2], often referred to telepresence [3]. This paper explains the phases to develop the real human teleportation in xR. Collaborative xR interface allows users to perform cooperative works seamlessly where the content is a mixture of real and virtual worlds [4]. The paper also discusses the proposed collaborative xR system that has successfully actualized the 3D telepresence in an urban planning prototype.

2. Extended Reality (xR)

What sets MR and VR apart is the ability to be seen and present in a real-world environment. The user would feel separated from the real-world and traditional tool in an immersive VR [5]. MR environments are overlaying of virtual objects in the real world. It can be concluded that MR is allowing the interaction between users on the real-world while maintaining the virtual images through high bandwidth of communication between users and impulsive manipulation on the virtual world [6].

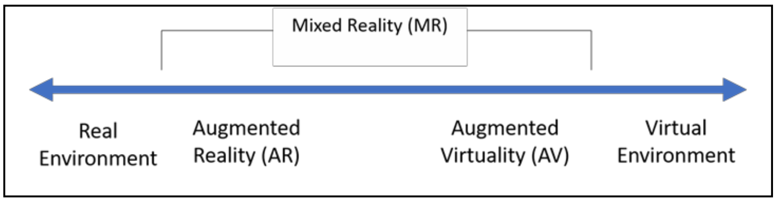


Figure 1. Reality-Virtuality Continuum [1].

From Figure 1, a classification from [1] on Reality-Virtuality Continuum shows that MR consists of AR, AV, and VR. AR is a real-time technology which registered in 3D with display technique. AR is 3D virtual objects which are integrated into a 3D real environment in real-time [7]. AR is referred to as an interface in which 3D computer graphics are overlaid over real objects through head-mounted or handheld displays [8]. The goal of AR is to create the insight that the virtual elements are a part of the real environment. AR is used to project the virtual elements layered on the real object by superimposing a correct view and position of the 3D model onto users' point of view of the real environment [8].

VR is an immersive computer-generated environment that locked the user inside the fabricated environment and cuts the user from the real-world environment. VR is a computer-simulated environment that can convince a human's brain about the existence of the computer-generated environment [9]. VR acts as a medium where it allows users to communicate and exchange with other users while the users immersed and present in the virtual world. The experience in VR is not limited to the fabricated environment only but can also include novels, movies, or any other environment. VR is the combination of display technology with a 3D model in real-time, providing an interactive environment allowing the user to manipulate and, at the same time, feel the immersion of the digital world. In contrast, VR changes the user's perspective of being in a new environment rather than being in a real environment. VR is an immersive and interactive technology using a joystick or mid-air interaction that replaces reality by replacing the real world with the simulated world.

3. Literature Review

3.1. Remote Collaboration Interactions

User collaboration in MR is a challenge since both parties need to have the same system and to set to parallel to MR environments. If the MR collaboration needs a specific space, both users need to replicate both places to have precise and accurate positioning in both users. Lack of co-presence feeling makes collaborative system absence of user experience [4]. A good user collaboration needs a good network architecture to achieve smooth data interchange [5].

A collaborative interface or remote collaboration system is to enable people who are far apart can sense in the same space [10]. The use of a collaborative interface has been utilized in the application to perform an assessment that requires multi-user. The collaborative interface allows its users to work cooperatively towards the shared task [8]. Communication is the most crucial aspect of any cooperative task. In a collaborative interface, visual aid is vital to navigating users around the system. The previous research in collaborative interface produces an issue of lack in social presence. Hence, the adoption of non-verbal communication cues can enhance a collaborative system relating to the issue of social presence [4]. In order to counter this issue, the proposed application include visual aid and avatar to enhance the collaborative interface.

The collaborative interface requires multi-user and network connection to perform the given task. In this research, a remote collaboration takes place between MR and VR users, in which each user use Head Mounted Displays (HMD) device. The use of HMD in MR environments creates the possibility of having a more immersive collaboration experience [9]. The issue may arise through the multi-user interaction as each user have a different platform and input to interact and manipulate in the shared network space [10]. Another remaining issue in the integration is the selection method; users need to select the shared objects at the same time. In order to overcome this issue, the authority method is used to determine which user has given a chance to manipulate the shared object. Using the authority method keeps minimum networking and ensuring a smooth collaborative interface.

3.2. 3D Telepresence

User teleportation can be represented as 3D telepresence [11]. 3D telepresence requires many processes. One of the most complex processes is a 3D reconstruction from input stereo or RGB-Depth camera [12]. Developing a real-world environment with high-fidelity in VR through 3D reconstruction is difficult. The teleportation of the user into the virtual environment requires many algorithms and equipment to compress and extract 3D data. The 3D telepresence gives a critical to the user experience in real-time performance [13]. A slight mistake in placing the camera and sensor would cause a flaw in data and can lead to inaccurate positioning in MR environments. User teleportation in xR environment is a challenge in this research since the process needs hardware that can enable and capture an object in 3D data into the application [12].

The standard camera can produce 2D RGB data, such as image and video. The output of a standard camera is insufficient in enabling and capturing an object in 3D data where the standard camera cannot capture depth data into the application. Therefore, a depth camera is a requirement in this research for stream and capture 3D object into the application [14]. There has been an issue of reconstructing 3D data into the application, in which the previous research leads to creating a low resolution of the 3D model [15]. The user teleportation process requires a loop procedure of 3D reconstruction by post-processing the captured object data and display it into the application. Since the output of user teleportation data is in 3D, the transferred data contains RGB and depth input of the captured object, which leads to a massive amount of data transferred. The remaining issue in user teleportation process is how to stream the post-processing of RGB and depth data over a network because of the vast data transferred required a high bandwidth of internet speed to overcome the issue [16].

Therefore, the proposed solution for the user teleportation issue is by the stream the RGB and depth data from the remote user by extending the connecting wire of the depth camera through the local user. Teleportation is also a common concept involving a distinct movement to a destination [17]. User teleportation is a method to capture 3D objects and teleport into virtual space. Early telepresence researches were focused on capturing dynamics scene using an array of cameras [3]. User teleportation gives the advantage of cutting the cost of travel and time spent to meet each other. In the early systems were developed, hardware and computational were the limitations at that time as the systems were able to capture low-resolution of 3D models and transmitted to the remote viewer. There were numerous researches focusing on 3D telepresence systems, and there were encountered difficulties for real-time capture, transmission, and displays [11]. From the past research, mostly sacrifice realism for faster performance in remote collaboration in telepresence.

IOP Conf. Series: Materials Science and Engineering **979** (2020) 012005 doi:10.1088/1757-899X/979/1/012005

3.3. 3D Reconstruction Method

Depth processing enables objects taken from different viewpoints to be backed up into 3D. Then, the registration and alignment of all points could completely restructure the environment. There are two ways to determine depth, first is through the stereo cameras, and the second is through the structured light sensors. User teleportation needs a stereo camera to provide depth data to capture users in the workspace. A stereo camera is a camera with two or more lenses together with an individual image sensor for each lens. The camera allows us to replicate human binocular vision and therefore delivers the ability to capture 3D images. Nowadays, the stereo camera is equipped with an Infrared (IR) sensor to improve the ability to capture the depth data. Recent research on user teleportation using multi-viewpoint to capture the user in the system [18]. The multi-viewpoint provides the presence of the user in the scene. Structured light is the most precise and fast in-depth estimation [19]. The most stereo camera provided structured light generate from the IR sensor. The stereo camera provides RGB-Depth (RGB-D) data from the lens sensor. RGB-D data is another advanced solution real-time depth estimation. Each sensor in the stereo camera produces an image with estimated depth by match and patch one image to another.

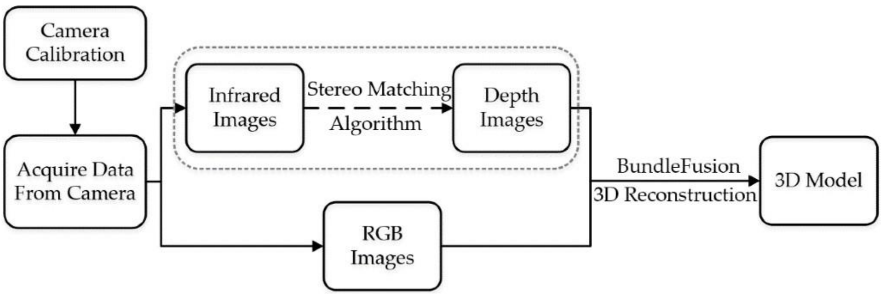


Figure 2. 3D Reconstruction Process using RGB-D camera [20].

The foreground segmentation stage follows the depth estimation. Foreground segmentation extracts the foreground object in an image by grading each image pixel as a foreground or background [3]. The output from depth estimation has created an image with various depths. Foreground segmentation eliminating the background and focuses on the data. This process as shown in Figure 2 plays an important stage in the later process of user teleportation, which is 3D reconstruction.

A modern stereo camera can make exhaustive scans of objects by generating point cloud data containing millions of points [21]. A point cloud is a set coordinate point which represents the surface of the data. One of the methods of 3D reconstruction is by visualizing the 3D model in the form of the point cloud. Point cloud has a problem in which is it lacks the connectivity structure of the underlying mesh and hence requires additional post-processing steps to extract 3D geometry from the model [22]. The output of point cloud may cause an error if not covers the most surface of the captured data as the stereo camera cannot capture the occluded from the camera viewpoint. Point cloud stitching is implemented to reconstruct the captured data by filling the hole

Point cloud data only consists of depth data and no colour data. In order to process a coloured 3D data, colour texturing is applied. Colour texturing approach is by computing the colour of each pixel by blending the RGB images according to its surface normal and the camera viewpoint direction. In colour texturing, depth estimation plays an important role in differentiating the colour between two surfaces. The missing sections of geometry might affect inaccurately projected colour to the surfaces behind it. Past research has implemented stitching and combining the nearest colour texture to overcome the missing colour texture in surface reconstruction [23].

4. Proposed xR Application

4.1. 3D Telepresence Input
The 3D telepresence process and method has been developed to enable user teleportation. This setup for the user teleportation required depth camera, an Intel RealSense D435, it was used to capture the RGB and depth data. The depth camera data undergo point cloud rendering process to display the data in 3D. This process is to map the input data from the depth camera into the output display. From this point cloud rendering, the proposed system captured the user's human body and process it into the application to achieve the user teleportation method.

4.2. xR Apparatuses Setting
In the xR setup, the display technology was required as the user needs to wear HMD to display output in it. The Project NorthStar HMD (PN-HMD) was used for the MR user and Oculus Rift HMD (OR-HMD) for the VR user. For PN-HMD, the gesture input was defined by Leap Motion Sensor for MR user interaction, and for OR-HMD, the Oculus Touch device has been used for VR user interaction. MR user has experienced the application using PN-HMD. The PN-HMD needs to attach along with the Leap Motion device to obtain the gesture input and Intel RealSense T265 on a dedicated mount to acquire the positional tracking. Lastly, the PN-HMD needs a positional tracking input to move around in the scene. Intel RealSense T265 be able to provide the Six Degrees of Freedom (6DOF) positional input.

Figure 3 shows the collaborative interface components overview, which includes all the hardware required to complete this application. Networking was crucial in this research. To integrate into the collaborative interface, interaction for spawning, delete, and manipulation of 3D objects was synchronized in a network. Performing integration and collaboration in the Unity3D game engine, take a plugin to establish a connection between the user. Unity3D engine process, the renderer, and the input handler of the game. In this phase, the Photon Unity Networking (PUN) plugin was used as a networking connection between users as referred to [8]. PUN server act as a host computer that handles all client connection requests maintains and manages client-shared virtual content, and processes data storage and retrieval.

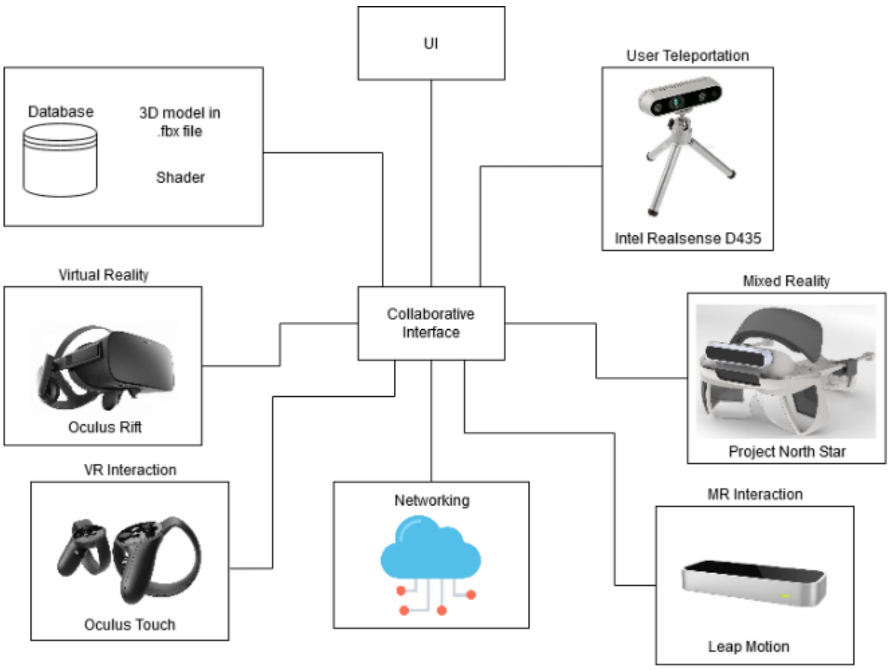


Figure 3. xR apparatuses and interaction in the collaborative interface.

IOP Conf. Series: Materials Science and Engineering **979** (2020) 012005 doi:10.1088/1757-899X/979/1/012005

4.3. Remote Collaboration Interactions

The prototype has been set up a shared environment for the user to experience the same cooperative task in a shared space. The urban planning task in this collaborative interface has several interactions in the shared space. The shared environment consists of a 3D object for urban planning tasks. The application consists of two clients assign as local users as Client 1 and remote users as Client 2. Local user must create room to start the application, and the remote user must enter the room. Both users experienced in a shared space to conduct urban planning activity. The shared room connected to Ethernet for synchronizing action. The data from the user input device when manipulating the object was sent to the server. The server sends the processed data back to the user to update the output in the user view. Figure 4 illustrates the process of collaborative interface application in the MR environment.

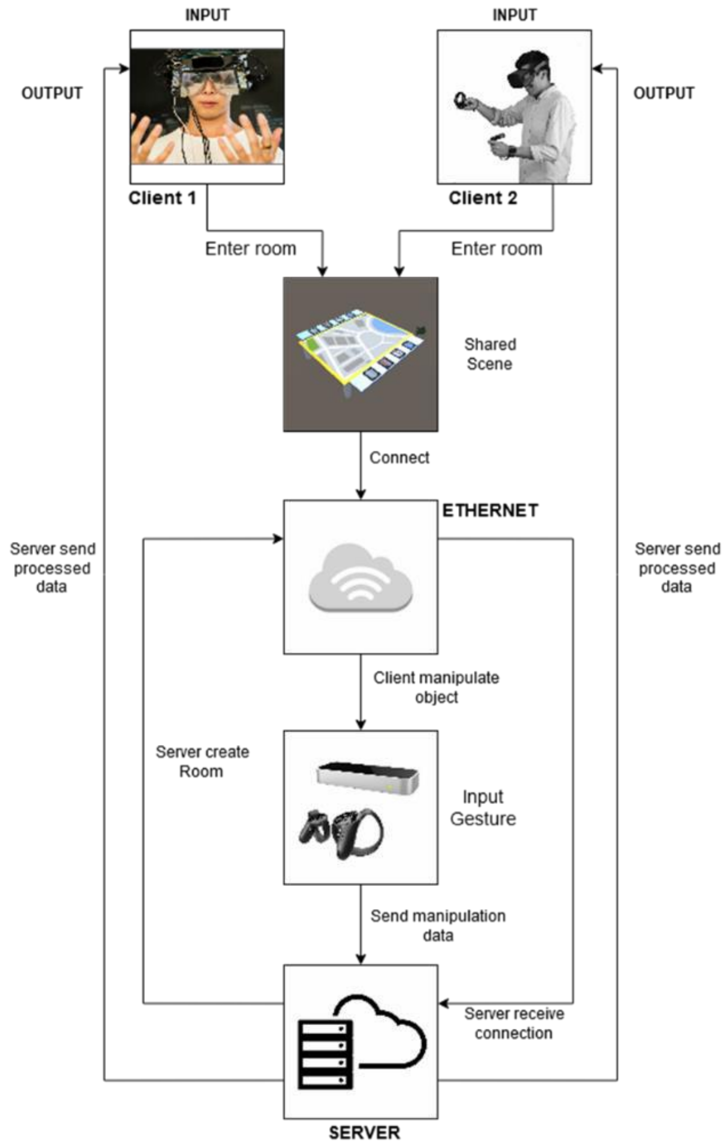


Figure 4. Remote Collaboration Interactions framework.

Both HMD using gesture input for the interaction. There are joystick and Leap Motion devices used in this research. The Leap Motion gesture interaction, gesture consists of drag, tap, pinch and release. Tap gestures are discrete to detect one or more fingers touching the 3D button. The fingers involved in these gestures must not move ominously from the initial touchpoints.

Tap gesture involves in this research is for tapping the 3D button for spawn interaction. A pinch gesture is a continuous gesture that tracks the distance between the first two fingers that touch the 3D object. Pinch gesture is used for grabbing the virtual object in the MR environment. Drag gesture is a motion that brings an object from one location to another. In this research, drag gesture is enabled along with pinch gesture for dragging a virtual object in a collaborative interface. Release gesture is a movement where all fingers are extended. Release gesture is subsequently of pinch gesture. This research uses a release gesture for release user virtual object from hand. Figure 5 shows the hand gesture interaction technique to perform drag, tap, pinch and release. The real hand gesture using Leap Motion has been learnt in [24], their work has been explored in this research.

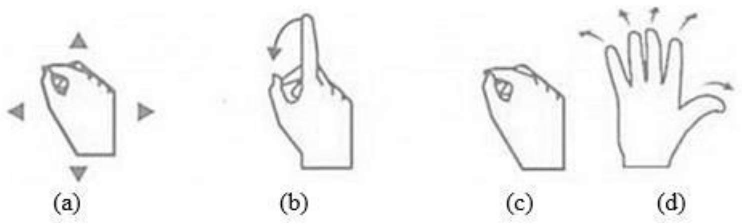


Figure 5. (a) drag (b) tap (c) pinch (d) release.

As both users have different input, the user performs different method to interact with the application, but both shared the same task, which spawns, delete and manipulate the 3D object. Therefore, the collaborative interface has a different executable program to respond to each user. Interaction for object manipulation shares the same process in MR and VR side.

The interaction involves consist of tap the 3D button and grab the virtual 3D object for translate and rotate manipulation. The rigid finger body from the input device is used to detect the collision between the 3D button and finger. If the rigid body is colliding, the object is spawned in shared space. Grabbing the 3D virtual object is performed by getting coordinate of index and thumb finger from input device for pinch gesture. Manipulation of a 3D virtual object can be performed if the object is pinch by the user. The manipulation involves in this research are translate and rotate the object. After performing the manipulation, the user either can repeat the interaction or end the process.

5. Implementation and Results

The user teleportation illustrates that the user teleportation process was achieved by executing the point cloud processing. Then, followed by the experimental setup of the remote user, the RGB-D camera was connected to the local PC. The results of the real-time user teleportation shown in Figure 6. Urban planning has been explored to implement the framework. Urban planning features and requirements have been referred to [25].

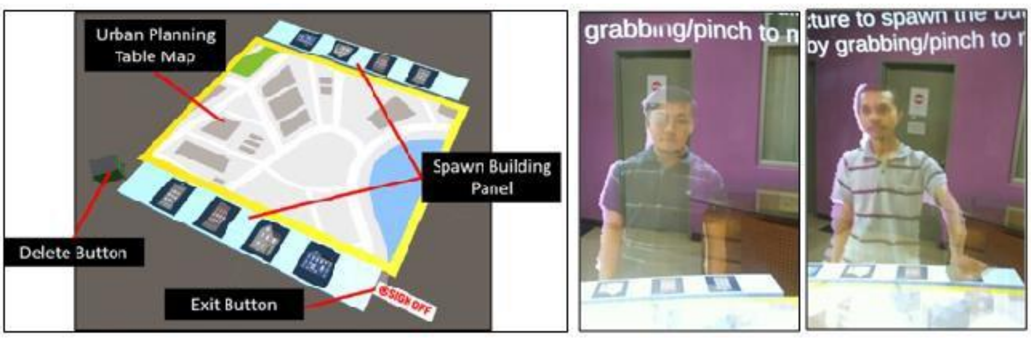
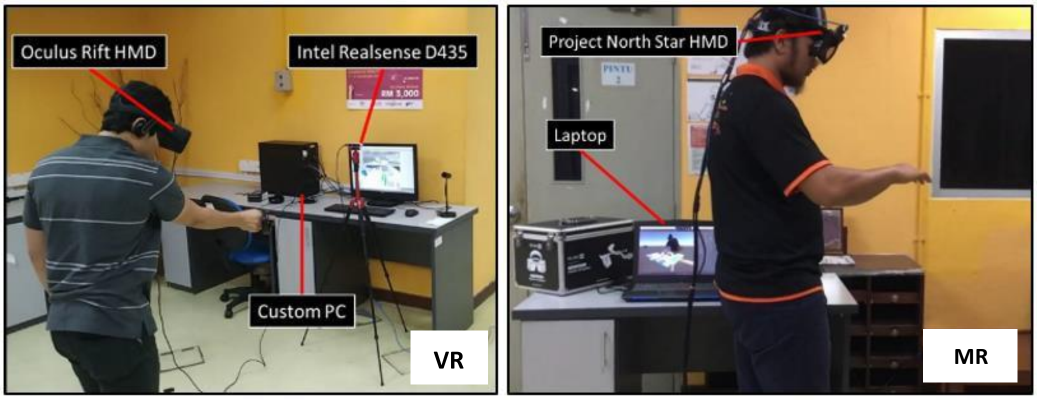


Figure 6. Real-time 3D Telepresence in Urban Planning prototype.

IOP Conf. Series: Materials Science and Engineering **979** (2020) 012005 doi:10.1088/1757-899X/979/1/012005

The collaborative interface has been developed simultaneously with the integration process. In general, both users experienced a synchronous update from each other. All the spawn object spawn at the network space; hence the manipulation also updates on the server, not on individually. The shared space of the collaborative interface is shown in figure 6. In the shared space, the UI only active for each side of the user.

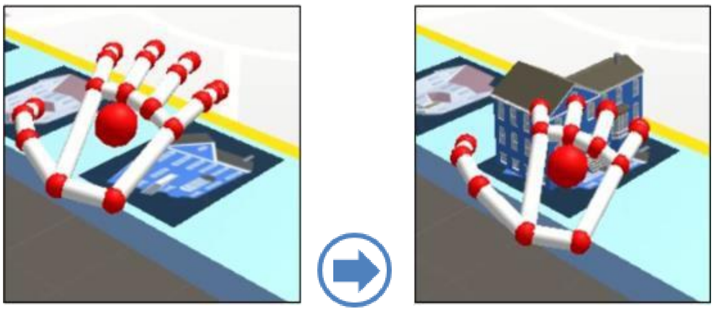
Since the research is working with a remote collaborative interface, there were two setups, which are for the local and remote sites. The experimental setup for this research is shown in Figure 7 (a) and Figure 7 (b).



(a) (b)

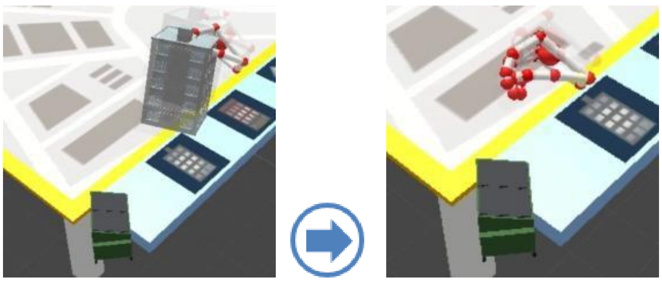
Figure 7. Two users in the two workspace setups in real-spaces, remote site and local site.

The activity involves in the application are spawn, delete, and 3D manipulation of a 3D object. All interaction with the 3D button was user tap the button. Figure 8 shows the spawn object interaction, while Figure 9 shows the delete object interaction. Figure 10 demonstrates both users grabbing a virtual object. Figure 11 shows the collaborative interface output for VR and MR.



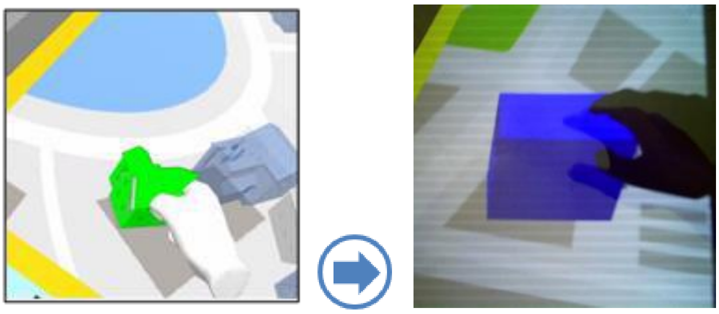
Before After

Figure 8. Spawning object.



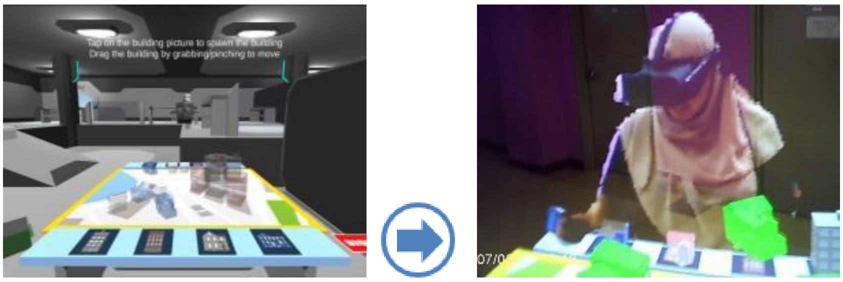
Before After

Figure 9. Deletion object.



Remote site Local site

Figure 10. Translation and Rotation



VR MR

Figure 11. Remote Collaboration Interactions

The xR application creates the synchronous update on 3D manipulation method in the collaborative interface; the 3D object is executed with the function that updates and retrieves the data from the server. All 3D object is involved with Photon Transform View and Photon View script to call the updated positional and rotational data in the server. The PUN component that links to the 3D object to update the manipulation data such as position and orientation coordinates. This proposed application had avoided the network congestion when an authority on the 3D object was given to the current user who has grabbed their desired object. If the 3D object was not assigned to the particular user ownership, the application requested from the other user for the ownership status. From this behaviour and condition, the manipulation of 3D object run in asynchronously. The process of requesting ownership between MR and VR user has started when the MR user is currently having the ownership of the 3D object. As soon as VR user grab the 3D, the ownership of the 3D object is transferred to VR user.

Table 1 shows the list of interaction command and each of the description. The interaction command exist in this research are select, spawn, translate, rotate, release and delete.

Table 1. Interactions and its descriptions.

Interaction	Description
Select	Using tap gesture for tapping the 3D button for activating the button.
Spawn	Instantiate 3D object above the 3D button position.
Translate and Rotate	Using pinch and drag gesture to manipulate the 3D objects by user hand movement.
Release	Using release gesture to discharge current holding 3D object in hand.
Delete	Destroy 3D object by using drag gesture on 3D object into delete 3D button.

6. Implementation and Results

This paper presents the phases to develop the real human teleportation in the xR environment using the advanced RGB-D sensor devices. The paper also discusses the literature reviews of user teleportation, which includes user teleportation method, process, and pre-defining setup. This paper also demonstrates the xR application and explains the development of collaborative interfaces and the experimental setup for the xR. The research acquired a user to wear HMDs for the VR and MR displays. For the collaborative interfaces, the proposed xR system runs the urban planning prototype to execute the cooperative task between two users with the intention of implementing 3D telepresence for remote collaboration. The contents are for urban planning; two users work in a shared environment. Issue in bandwidth network connection that used to connect users from different worlds into a single application has been highlighted. This study has study that distance measurement is one of the factors to affect the shared system. MR user is set as a local user, while the VR user becomes a remote user. The inputs have been captured for both users to interact with the MR environment. Finally, this paper ends with the user interactions and to determine the user interaction for both users, such as deletion, spawning the objects, translation and rotation. The integration of a 3D telepresence to transport the remote user into the shared environment has been successfully demonstrated in this study. The integration is using a network plugin to establish the connection between two users, and the stable internet connection is required to avoid transmission delay. The manipulation in the remote collaboration interface is synchronizing in the network. The proposed xR system has captured the user's human gesture and figure, and process it into the shared environment to complete the user teleportation or being known as 3D telepresence.

Acknowledgement

We would like to express our appreciation to Mixed and Virtual Reality Laboratory (mivielab), ViCubeLab at Universiti Teknologi Malaysia (UTM) for the facilities and technical supports. We also would like to thank Mr Noah Zerkin, who allowed me to have Project North Star HMD and lots of support from the Project North Star community.

References

[1] Milgram P and F Kishino 1994 A taxonomy of mixed reality visual displays. *IEICE TRANSACTIONS on Information and Systems* 77(12) p 1321-1329.
[2] Orts-Escolano S, Rhemann C, Fanello S, Chang W, Kowdle A, Degtyarev Y, Kim D, Davidson

IOP Conf. Series: Materials Science and Engineering **979** (2020) 012005 doi:10.1088/1757-899X/979/1/012005

PL, Khamis S, Dou M and Tankovich V 2016 Holoportation: Virtual 3d teleportation in real-time *In Proc. of the 29th Annual Symp. on User Interface Software and Technology 2016* pp 741-754

[3] Fadzli F E, Ismail A W, Aladin M Y F and Othman N Z S 2020 A Review of Mixed Reality Telepresence. *In IOP Conf. Series: Materials Science and Engineering* Vol 864, No 1, p 012081

[4] Irlitti A, Piumsomboon T, Jackson D, Thomas BH 2019 Conveying spatial awareness cues in xR collaborations *IEEE transactions on visualization and computer graphics* Vol 25(11) pp 3178-89

[5] Billinghurst M and H Kato 1999 Collaborative mixed reality *Proc. of the 1st Int. Symp. on Mixed Reality* p 261-284.

[6] Ismail AW and Sunar M S 2009 Survey on Collaborative AR for Multi-user in Urban Studies and Planning *Springer Berlin Heidelberg*

[7] Aladin M Y F and Ismail A W 2019 A review on multimodal interaction in Mixed Reality Environment *In IOP Conf. Series: Materials Science and Engineering* Vol 551, No 1, p 012049

[8] Nor' a M N A and Ismail AW 2019 Integrating Virtual Reality and Augmented Reality in a Collaborative User Interface *International Journal of Innovative Computing* 9(2).

[9] Craig A, Sherman W and Will J 2009 Developing Virtual Reality Applications: Foundations of Effective Design

[10] Nor'a M N A and Ismail A W 2019 A review on collaborative learning environment across virtual and augmented reality technology *In IOP Conf. Series: Materials Science and Engineering* Vol 551 No 1 p 012050

[11] Zhao M 2018 Towards high-quality 3D telepresence with commodity RGBD camera (Doctoral dissertation).

[12] Fuchs H, State A and Bazin J C 2014 Immersive 3D Telepresence Vol 47 pp 46-52

[13] Córdova-Esparza D M, Terven J R, Jiménez-Hernández H, Herrera-Navarro A, Vázquez-Cervantes A and García-Huerta J M 2019 Low-bandwidth 3D visual telepresence system *Multimedia Tools and Applications* Vol 78(15) pp 21273-90

[14] Manolova A, Neshov N, Tonchev K, Koleva P and Poulkov V 2019 Challenges for real time long distance holoportation to enable human bond communication. *In 2019 42nd Int.Conf.on Telecommunications and Signal Processing (TSP)* pp 529-534