




# Design of Metaverse for Two-Way Video Conferencing Platform Based on Virtual Reality

Dongeon Yoon<sup>1</sup> and Amsuk Oh<sup>2\*</sup> , Member, KIICE

<sup>1</sup>Department of Computer Media Engineering, TongMyong University, Busan 48520, Korea

<sup>2</sup>Department of Digital Media Engineering, TongMyong University, Busan 48520, Korea

## Abstract

As non-face-to-face activities have become commonplace, online video conferencing platforms have become popular collaboration tools. However, existing video conferencing platforms have a structure in which one side unilaterally exchanges information, potentially increase the fatigue of meeting participants. In this study, we designed a video conferencing platform utilizing virtual reality (VR), a metaverse technology, to enable various interactions. A virtual conferencing space and realistic VR video conferencing content authoring tool support system were designed using Meta's Oculus Quest 2 hardware, the Unity engine, and 3D Max software. With the Photon software development kit, voice recognition was designed to perform automatic text translation with the Watson application programming interface, allowing the online video conferencing participants to communicate smoothly even if using different languages. It is expected that the proposed video conferencing platform will enable conference participants to interact and improve their work efficiency.

**Index Terms:** Interaction, Metaverse, Video Conferencing, Virtual Reality, Voice Translation

## I. INTRODUCTION

Virtual reality (VR) is one of the technologies for creating a metaverse environment, thereby allowing various activities in the real world to be performed in a virtual world [1]. Recently, as the quality of VR hardware and software has improved, VR has increasingly been used in realistic games for consumers, product visualizations for social areas and companies, VR collaborations, and business conference areas [2]. Existing online video conferencing platforms are also used for social and business meetings, but they often experience problems such as difficulties in interaction and leakages of personal information [3-4]. This is because the structure, in which the speaker attending the online meeting unilaterally transmits information to the listener, tends to be repeated. Therefore, there is a need to improve the fatigue experienced

during the use of existing video conferencing platforms, as well the lack of adequate functionality.

In this paper, we propose the design of a metaverse VR-based video conferencing platform for improving participation performance and emotional commitment in an online virtual space. Video conferencing design systems are generally divided into virtual conferencing space design systems, realistic VR video conferencing content authoring tool support systems, and real-time voice chat systems. Each system operates organically as connected to the others, thereby combining smooth communication (an advantage of offline meetings) and unlimited space (an advantage of online meetings).

The remainder of this paper is organized as follows. The first section introduces studies related to the proposed platform, such as spatial and horizontal workrooms. The second section describes the design process of the three systems


Received 12 April 2022, Revised 11 July 2022, Accepted 26 July 2022

\*Corresponding Author Amsuk Oh (E-mail: [asoh@tu.ac.kr](mailto:asoh@tu.ac.kr), Tel: +82-51-629-1211)

Department of Digital Media Engineering, Tongmyong University, Busan, 48520 Korea

Open Access <https://doi.org/10.56977/jicce.2022.20.3.189>

print ISSN: 2234-8255 online ISSN: 2234-8883

 This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Copyright © The Korea Institute of Information and Communication Engineering

necessary to complete the platform. Finally, the paper is concluded by drawing conclusions and implications for platform design.

## II. RELATED WORKS

Since 2020, the metaverse has begun to grow significantly, proving its use not only in games and business-to-consumer entertainment, but also in business-to-business areas such as telecommuting and corporate events [5]. Bloomberg Intelligence predicted that the size of the metaverse market will grow from \$478.7 billion in 2020 to \$783.3 billion in 2024 [6]. Global companies such as Meta, Microsoft, and NDI-VIA have also begun to participate in competition to develop virtual video conferencing platforms using metaverse technology, and various platforms are being studied [7].

### A. Spatial

“Spatial” is a platform for helping team members interact by supporting meetings as avatars in virtual spaces [8]. Spatial users can share traditional document files in a virtual space, share their computer screens with colleagues, and talk about work. Post-it notes can also be pasted into virtual spaces.

Figure 1 shows a portion of the Spatial virtual space.



Fig. 1. Virtual space in Spatial

### B. Horizon Workrooms

Horizon Workrooms is a VR-based collaboration platform developed by Meta [9]. It is optimized for the company’s Oculus Quest 2, and its advantage is in that it feels like meeting and working in the real world. It supports a virtual keyboard for inter-user chats and can expand multiple screens (e.g., only those sharable in existing online video conferences) [10-11].

## III. DESIGN OF METAVERSE VIRTUAL REALITY (VR)-BASED VIDEO CONFERENCING PLATFORM

This study aims to design a metaverse platform able to increase the productivity of video conferencing users based on metaverse technologies such as software related to VR functions and hardware for assisting various interactions in the virtual world. First, the virtual conference space design system provides a virtual conference space without restrictions on the size of the space when users access the video conferencing. In this context, a realistic VR video conferencing content authoring tool support system is a VR-based system for helping users use a meta-bus by supporting a series of authoring tools such as laser pointers and avatars, and for executing the applications necessary for meetings and presentations. A real-time voice chat system analyzes a person’s voice pattern using a voice recognition software development kit (SDK) and provides a real-time voice translation function.

Figure 2 shows the configuration diagram of the metaverse VR-based bidirectional video conferencing platform.

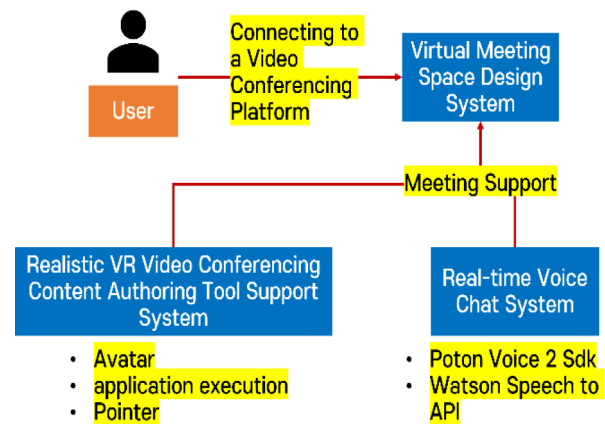


Fig. 2. Composition diagram of a video conferencing platform based on Metaverse virtual reality (VR)

In this study, the design environment was divided into hardware for manipulating the content and software for developing it. A head-mounted display (HMD), i.e., an image display device worn on the head, used Oculus Quest 2 and a VR controller for interaction equipment. It used Unity and 3D Max software to produce avatar and video conferencing platforms. The software development was performed in a PC environment with a Ryzen 7-5800H CPU, RTX 3060 GPU, and 32 GB of RAM.

Table 1 describes the design environment of the metaverse VR-based video conferencing platform.

**Table 1.** Design environment of Metaverse virtual reality (VR)-based video conferencing platform

Classification	Type	Name
Hardware	VR head-mounted display (HMD)	Oculus Quest 2
	Interaction Tools	VR Controller
	PC Environment	Ryzen 7-5800H (CPU)
		RTX 3060 (GPU) 32 GB (RAM) 1 TB (STORAGE)
Software	Video Conferencing Platform Feature Development Tools	Unity (2020.330f1 ver.)
	Metaverse Management Server	Photon Server
	Avatar Design Tools	3D Max, N.CAD

**A. Virtual Meeting Space Design System**

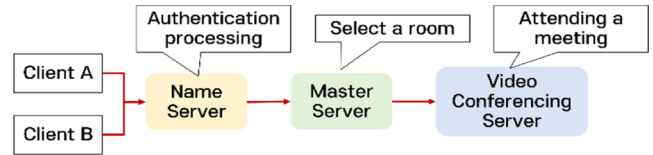
The virtual conference space design system supports remote procedure call (RPC) communication with other users through the hardware and Photon server required for program execution and operation. The main body, VR HMD, and controller hardware use Facebook’s Oculus Quest 2 to develop the platform’s functionality with the Unity game engine. The Photon server allows for RPC communication with other users and produces avatars using voxel design techniques.

**1) Remote Procedure Call (RPC) Communication with the Photon Server**

The Photon server is a Unity package for helping different objects synchronize over a network. When a client connects to the network, it first connects to a name server. The name server performs authentication processing on the application and server, and moves to the master server. The master server checks the information of all rooms before entering the meeting and selects the meeting room the user desires to enter. When the user enters the selected meeting room, they are moved to the video conference server designed to participate in the meeting. On the video conference server, the user can participate in the conference by using an authoring tool supported by the platform.

Figure 3 shows the operation structure of the Photon server.

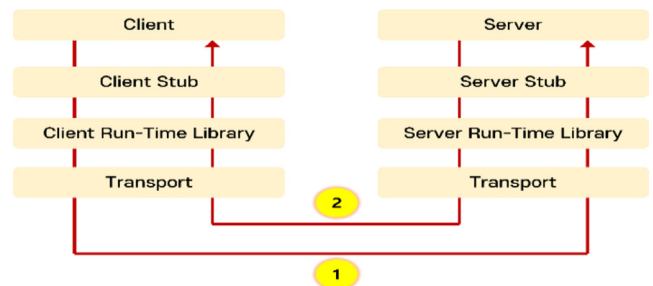
The RPC is a communication method that allows a procedure to be called to a remote server without directly implementing a connection part for the client-server communication. An interface definition language is used to define the call conventions between clients and servers and to convert the parameters used for procedure calls through a stub. Because of the parameter conversion via the stub, this approach can be used to call a local function on a remote server when calling a procedure on the remote server.



**Fig. 3.** Photon Server structure

The client’s RPC communication consists of executing remote procedures, outputting parameters, and generating return values. The client passes parameters to the client stub. The parameters from the client stub are converted. The client invokes the RPC client runtime library function and requests it based on the parameters converted by the server. The server RPC runtime library function accepts the request and invokes a server stub procedure. The server stub searches the network buffer for parameters and converts them from the network transport format to the format required by the server. The server stub calls the actual procedure from the server. The server completes the communication by returning the value to the client. The process for returning a value from the server to the client is similar to that for the previous process.

Figure 4 shows the RPC operation process.



**Fig. 4.** Remote procedure call (RPC) operation process

**2) Voxel Design**

Voxels are representations of two-dimensional pixels as three-dimensional cubes. When designing a voxel, several cubes are used to express objects. Designing in this way consumes less data and the 3D model can feel like pixel art; this makes it a design style suitable for VR design, where performance is the key. In addition, it has the advantage of being able to implement detailed textures and freely deform compared to polygons. We designed a virtual video conferencing space and optimized the operational performance of the data by utilizing a low-poly design.

Figure 5 represents one of the voxel designs.

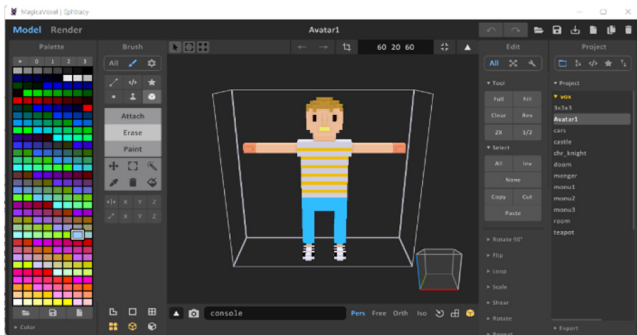


Fig. 5. One of the voxel designs

**B. Realistic VR Video Conferencing Content Authoring Tool Support System**

A person who accesses the metaverse video conferencing platform while wearing an HMD and controller is connected to the virtual conferencing space provided by the platform. The virtual conference space can be built by directly modeling video conference participants. When a user customizes the virtual conference space and starts communicating with others, the support for the real-time voice chat system and realistic VR video conferencing content authoring tool support system begins. A realistic VR video conferencing content authoring tool support system provides functions such as brainstorming, presentation, team planning, product review, collaboration whiteboards, avatars, and interactions.

Figure 6 shows the process of using the virtual conferencing platform.

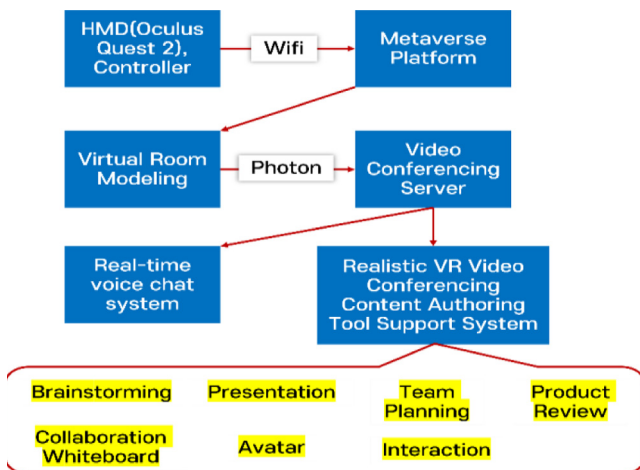


Fig. 6. Process of using the virtual conferencing platform

**C. Real-Time Voice Chat System**

Voice recognition is a technology for recognizing a speaker’s

voice and converting it into text. Automatic translation is a technology for converting voice recognition content into a language used by another party. The real-time voice chat system recognizes another person's voice during video conferencing and automatically translates it into other languages through natural language processing (NLP). NLP is a branch of artificial intelligence that deals with human-machine interactions, and can provide convenience and high immersion in virtual conferencing spaces by interpreting the soft languages used in everyday communication and converting them into other languages.

After the speech recognition process within Unity is performed, the real-time speech translation system allows the voices of other participants to be translated into the language used by the participants themselves for verification. When a videoconferencing participant has a conversation through the Photon Voice SDK in a metaverse virtual conference room, the Watson Speech-to-Text application program interface (API) recognizes the voice and converts it into textual information. The translated text performs automatic translation via the Watson Language Translator and applies a singleton pattern to Unity for memory leakage prevention and smooth data sharing between the different classes. The translated data are transmitted through the server to other clients, completing the system.

Figure 7 shows the real-time voice translation process.

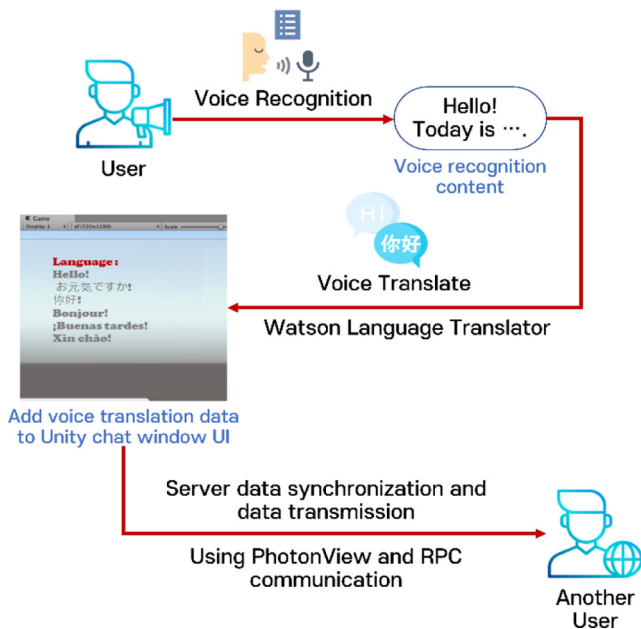


Fig. 7. Real-time voice translation process

**IV. IMPLEMENTATION OF METAVERSE VR-BASED VIDEO CONFERENCING PLATFORM**

The video conferencing platform designed in this study

was tested using the Unity Tool. The test environment consisted of the user A avatar, user B avatar, screen where the other party's voice chat contents were displayed, and screen for document display and object manipulation.

Figure 8 shows the test environment.

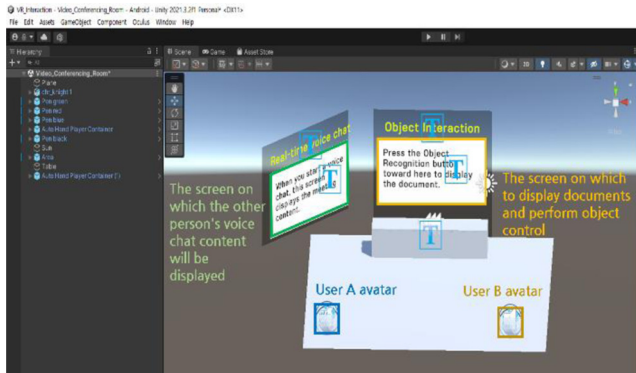


Fig. 8. Test environment for Metaverse platforms

The Oculus Quest 2 controller was used to manipulate objects, load documents, and/or activate voice chat functions. When the object recognition button on the controller was pressed, a raycast was output, which served to read information about the conflicting object. If the raycast did not collide with the object, it appeared as a red line; if it did, it appeared as a sky-blue line. Depending on whether the object with which the raycast conflicted was a UI object or GameObject, the response button was clicked to perform various interactions.

Figure 9 shows an example of object manipulation using the Oculus Quest2 controller.

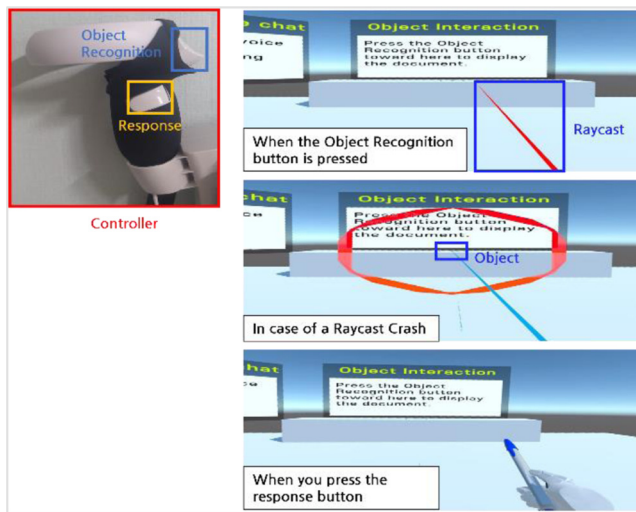


Fig. 9. Example of object manipulation process

When user A outputs a presentation using the controller and proceeds with the presentation, the screen viewed by user B is as shown in Fig. 10. User B can check the information, presentation materials, and memos from converting user A's voice into text.



Fig. 10. Interaction with User A as seen by User B

## V. CONCLUSIONS

In this study, we designed and implemented of a metaverse VR-based video conferencing platform by dividing it into a virtual conferencing space design system, realistic VR video conferencing content authoring tool support system, and real-time voice chat system. The virtual conference space was designed by dividing it into the HMD, controller hardware, Unity engine, and Photon server software. Avatars to be active in the VR adopted a voxel design, and explained the RPC communication process to other users. The RPC could call procedures faster by helping clients and servers connect automatically. In the realistic VR video conference content authoring tool support system, various interactions, as well as the data sharing and collaboration used in existing online video conferences, were supported to improve the immersion. In the real-time voice translation system, voice recognition was performed using the Photon Voice SDK supported by Unity, and automatic translation was performed using the Watson Speech-to-Text API.

The proposed platform will provide experiences similar to real offline conferencing, thereby enhancing interaction and work efficiency among videoconferencing participants. Smooth communication is also possible when conducting business with conference participants of various nationalities. In addition, it is expected that the space can be used in various con-

cepts, such as rest and travel, and not just as a working space.

## ACKNOWLEDGMENTS

This research was supported by BB21plus and funded by the Busan Metropolitan City and Busan Institute for Talent and Lifelong Education (BIT).

## REFERENCES

- [ 1 ] Computer Tech Reviews, What is Virtual Reality? – Definition, Types, Uses, And More [Internet]. Available: <https://www.computer-techreviews.com/definition/virtual-reality/>.
- [ 2 ] Software Policy & Research Institute, A Survey on the Virtual Augmented Reality (VR/AR) Industry in 2020 [Internet]. Available: <https://spri.kr/posts/view/23200?code=research>.
- [ 3 ] K. W. Chang, “A case study on project based learning in distance education: Focusing on learners’ learning experiences,” *Journal of Engineering Education Research*, vol. 36, no. 3, pp. 775-804, Sep. 2020. DOI: 10.17232/KSET.36.3.775.
- [ 4 ] H. S. Choi and Y. H. Cho, “A study on problem analysis for safe online video platform use,” *Journal of the Korea Society of Computer and Information*, vol. 25, no. 10, pp. 269-277, Oct. 2020. DOI: 10.9708/jksci.2020.25.10.269.
- [ 5 ] K. J. Kim, “The evolution of reality and the virtual world from metaverse cases,” *Broadcasting and Media Magazine*, vol. 26, no. 3, pp. 10-19, Jul. 2021.
- [ 6 ] Straight News, [Metaverse innovation wave?] Metaverse changes the industry [Internet]. Available: <https://www.straightnews.co.kr/news/articleView.html?idxno=207979>.
- [ 7 ] H. W. Nam, “XR technology and metaverse platform status,” *Broadcasting and Media Magazine*, vol. 26, no. 3, pp. 30-40, Jul. 2021.
- [ 8 ] J. S. Lee and J. Y. Kim, “A study on spatial characteristics of immersion and reality in cases of VR and AR technology and contents,” *Journal of the Korean Institute of Interior Design*, vol. 28, no. 3, pp. 13-24, Jun. 2019. DOI: 10.14774/JKIID.2019.28.3.013.
- [ 9 ] Money Today, Zoom is also on board... “Metaverse Conference Hall” competition [Internet]. Available: <https://news.mt.co.kr/mtview.php?no=2021091715490173746>.
- [10] Editorial department, “For business use metaverse horizon workrooms,” *Excellence Marketing for Customer*, vol. 55, no. 9, pp. 63-70, Sep. 2021.
- [11] S. Y. Kim, J. Y. Kim, J. Y. Jun, R. I. Ryu and, Y. R. Koo, “Metaverse platform service design proposal based on the value consumption experience factors of the MZ generation -Focused on the use of the kano model-,” *Journal of the Korean Society Design Culture*, vol. 28, no. 1, pp. 17-33, Mar. 2022. DOI: 10.18208/ksdc.2022.28.1.17.



**Dongeon Yoon**

Received his bachelor's degree from the Department of Digital Media Engineering, University of Tongmyong, KOREA, in 2020. He is attending as currently a master's student in Computer Media Engineering, Tongmyong University, Busan, Korea. His research interests include Big Data, Artificial intelligence, and the Metaverse.



**Amsuk Oh**

Received his Ph.D. degree from the Department of Computer Engineering, University of Busan, KOREA, in 1997. He is currently serving as Professor at the Department of Digital Media Engineering, College of Technology, Tongmyong University, Busan, Korea. His research interests include Databases, Big Data, the Internet of Things, Healthcare Systems, and Medical Information Systems.