IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received 16 June 2023, accepted 5 July 2023, date of publication 11 July 2023, date of current version 19 July 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3294390

APPLIED RESEARCH

XAVE: Cross-Platform Based Asymmetric Virtual Environment for Immersive Content

YUNSIK CHO[1], MYEONGSEOK PARK[2], AND JINMO KIM[1,2,3]
[1]Department of Information and Computer Engineering, Graduate School, Hansung University, Seoul 02876, South Korea
[2]Department of Computer Engineering, Graduate School, Hansung University, Seoul 02876, South Korea
[3]Division of Computer Engineering, Hansung University, Seoul 02876, South Korea

Corresponding author: Jinmo Kim (jinmo.kim@hansung.ac.kr)

This research was financially supported by Hansung University for Jinmo Kim.

ABSTRACT This study proposes XAVE–an asymmetric virtual environment in which users on various devices and platforms (personal computer (PC), mobile devices, virtual reality (VR) and augmented reality (AR) users, and actor-based avatars using motion capture) participate together by expanding the range of users considered in the existing asymmetric virtual environment studies. XAVE aims to provide a novel experiential environment that includes non-immersive and immersive participation methods for immersive applications and contents. All users interact with virtual environments, objects, and users easily and conveniently through a common and popular interface that considers the characteristics of the devices and experiential methods. PC and mobile users interact through keyboard, mouse, and touchscreen inputs based on graphical user interfaces (GUIs). VR users interact using controllers while wearing head-mounted displays (HMDs), and AR users interact through the image target and GUI on mobile devices. Avatar users participating in virtual environments using motion capture were considered to construct asymmetric virtual environments that can be used in various metaverse and immersive content fields. Furthermore, an application where all users can participate and experience together was created and used to statistically analyze and examine user satisfaction with the interface and presence in the virtual environment through surveys.

INDEX TERMS Asymmetric virtual environment, non-immersive and immersive interaction, virtual reality, augmented reality, motion capture, immersive content.

I. INTRODUCTION

Metaverse refers to an extended virtual world; this portmanteau blends "meta," which relates to a virtual or transcendent state, and "universe," which relates to the world and environment. The metaverse has been evolving toward providing virtual environments where both reality and unreality are connected throughout society, including economy and culture [1]. These virtual environments are used in various industrial fields, such as games, marketing, gamification, digital humans, and non-face-to-face meetings. Users participating in virtual spaces interact through interfaces, ranging from keyboard, mouse, and touching mobile screens to a variety of device that conveys three-dimensional space information, such as headsets and motion controllers.

The associate editor coordinating the review of this manuscript and approving it for publication was Tai-Hoon Kim.

Furthermore, technology for extended reality (XR), including virtual reality (VR) and augmented reality (AR), is considered important in constructing environments that provide immersion and presence for users to feel as if they are in a real space.

Human-computer interaction (HCI), an academic discipline that studies interactions between humans and computers, is widely studied in various industrial fields, including engineering. Typically, users interact using window, icon, menu, and pointer (WIMP) interfaces based on input methods using keyboard and mouse. Studies on intuitive and user-friendly graphical user interface (GUI) technology are underway, as are studies on technology using post-WIMP methods such as voice recognition and gestures that allow users to feel immersed in virtual environments in an intuitive method through their physical senses such as vision, hearing, and touch [2]. Furthermore, studies have been conducted in designing and analyzing user interfaces that use the user's gaze, hands, and legs to provide a satisfactory presence and experience in interactions between users and virtual environments using immersive technologies such as VR and AR [3], [4], [5], [6], [7], [8]. These studies have been evolving toward feedback physical responses through haptics [3] and the provision of redirected walking to represent free walking in a limited space [9], which are applied in multiuser experience environments beyond the presence of only one user. Studies have been conducted broadly, including a study on collaborative virtual environments where multiple users collaborate and interact [10] and studies on asymmetric virtual environments where non-immersive users, such as personal computer (PC), mobile, and AR users, participate with VR users [11], [12], [13], [14]. To create an immersive content where users of various platforms can participate and immerse themselves deeper in extended virtual worlds, research is required on interactions and asymmetric virtual environments that can encompass various devices, including PCs and mobile devices, and various technologies, including VR and AR.

Existing studies on asymmetric virtual environments are focusing on immersive users (VR or glass-based AR) and constitute an experience environment in which non-immersive users such as non-HMDs participate. However, this study proposes the Cross-platform based Asymmetric Virtual Environment (XAVE), comprising five types of users, ranging from non-immersive PC and mobile device users to VR (PC and mobile) and AR users, and avatars operated through actors using motion capture (Figure 1). Actions are performed easily and conveniently, considering the characteristics of the experiential environment and participation method of each user based on common and popular interfaces. The key objective is to propose the novel experiential environment, XAVE, which provides satisfying experiences and a presence to all users through a process where non-immersive, immersive and AR users coexist and collaborate. And it is to define the structure of the experience environment in which all users can organically collaborate. For this, the proposed user participation environment in asymmetric virtual environments is as follows.

1 **Non-immersive users**
- PC users: interaction using keyboard and mouse based on GUI
- Mobile users: interaction using GUI-based touchscreens

2 **Immersive users**
- PC-based VR users: interaction using VR HMDs and controllers
- Mobile-based VR users: interaction using removable mobile HMDs and controllers

3 **AR users**
- Mobile-based handheld AR users: AR composition and interaction using an image target and mobile device

4 **Actor-based avatars**
- Motion capture avatar: the avatar is controlled using motion capture and projector screen

The proposed XAVE aims to build an experiential environment where all users can easily and conveniently interact with the virtual environment, objects, and other users based on conventional interface structures and a more immersive virtual space while coexisting with non-immersive, immersive, AR, and motion-capture avatars. Accordingly, we conducted a comparative analysis systematically through surveys to determine whether XAVE provides a presence to all users, satisfactory interfaces, and interaction experiences. Previous studies on asymmetric virtual environments [14], [15], [16], [17], [18] considered specific experiential environments and limited user participation. Therefore, it is difficult to build various user participations, expand experiential environments, and produce various application contents based on previous studies. The aim of this study is to build an asymmetric virtual environment considering various types of user participation and define a template structure for function implementation along with interactions according to user characteristics and platforms to produce immersive contents. In this study, a template structure that includes function implementation and interaction according to user characteristics and platforms were defined to produce immersive contents and build an asymmetric virtual environment that considers user participation. In addition, we compared and analyzed the experience, presence, and satisfaction provided by the virtual environment that each user experienced differently within the immersive content. Through this process, we intended to conduct applied research for researchers who want to apply an asymmetric virtual environment to produce various immersive contents. The key objectives of this study were as follows:

1 A template will be defined considering the characteristics and platform of each user to build an expanded asymmetric virtual environment and create various immersive contents.

2 An appropriate asymmetric virtual environment will be possible to customize based on the user experience by analyzing the presence and satisfaction of each user through a survey experiment with participants.

In addition, by utilizing a system such as motion capture, we would like to consider an experiential environment in which the real-world users can participate in an asymmetric virtual environment together.

II. RELATED WORK

Various studies have been conducted on interface design in consideration of satisfaction, experience, and immersion, considering the ease and convenience of users interacting with virtual environments. Shneiderman [19] proposed a direct manipulation of the user interface with a structure similar to that of users performing tasks, while manipulating objects manually in real-world environments. User-oriented interface technologies have been evolving with studies on interactions with virtual environments based on various platforms and devices. Nunamaker et al. [20] designed a kiosk

FIGURE 1. Structure of XAVE, a cross-platform based asymmetric virtual environment: (a) PC user, (b) mobile user, (c) PC-based VR user, (d) mobile-based VR user, (e) mobile-based AR user, and (f) actor-based avatar using motion capture.

that uses intelligent agents to detect changes in arousal, behavior, and cognitive effort and conducted user evaluations. Another study proposed a hand-free interaction for rehabilitating stroke patients based on leap motion device [21], and motions performed in real life, such as stretching, compressing, and squeezing, were applied to a deformation interface of shape modeling [22]. These studies have been evolving toward user interface application studies by combining with immersion technologies to provide immersive interactions with which users can directly interact with virtual environments and realistically control objects.

Various interface application studies have been conducted on providing highly immersive interactions with which users can easily and effectively control movements and express actions using various senses, such as vision, hearing, and touch through their gaze, gestures, and walks in virtual environments. For example, a study was conducted on a gaze-controlled multimedia user interface, which uses eyes to control the interface [23]. Some application studies were conducted on hand-based interactions and hand interfaces [3], [24]. Studies have also analyzed the efficiency of interaction and ease of manipulation by combining gaze and hands to find reasonable alternatives [5], [6], [7]. To implement realistic virtual environments by narrowing the gap between virtual and real worlds, the following technologies were studied on haptic systems and interfaces that feedback physical responses: a wearable cutaneous haptic interface based on multi-DOF (degree-of-freedom) fingers and finger-tracking modules (FTMs) [25]; a visuo-haptic interface that can demonstrate programmable tactile responses with visual information [26]; and an electric turntable haptic platform [27]. Moreover, application studies related to movements have been actively conducted, such as a redirected walking method for free walking interactions in a limited experiential space [9] and an interaction method representing natural walking by recognizing in-place walking [8]. Various methods have been proposed, such as interacting with virtual objects by blowing wind in immersive virtual reality [28], and manipulating objects with freehand gestures through ray-casting in handheld AR [29]. Furthermore, to explore the unit and size of information, a voice user interface has also been designed [30]. Nevertheless, considering popular and practical utilization while increasing user immersion, studies have been conducted on how to skillfully manipulate objects with hands using handheld VR controllers in virtual environments [31], as well as designing a hybrid paper interface based on touch-based inputs in handheld AR [32]. In recent years, interface studies in immersive virtual environments have attracted attention, such as a study on establishing a classification system for immersive interactions in AR [33], which is relatively insufficient compared to VR. Considering the characteristics of platforms and devices is important for providing interactions with which users can immerse themselves in virtual environments, as well as interfaces that are easy and convenient to use. Therefore, considering these points, designing appropriate interfaces for the platforms, devices, and experience methods in virtual environments where various types of users can participate is desirable. This study further intends to define user-specific templates and interactions required for constructing and developing an asymmetric virtual environment that considers popular and realistic utilization and considers various user participation.

Collaborative virtual environments (CVEs) are used for the collaboration and interaction of many users separated by significant distances. They provide highly immersive collaboration environments by combining both non-immersive experience methods (such as in PCs and mobile devices) and immersive technologies (such as VR and AR) to transmit 3D visual data and facilitate non-verbal (such as gaze processing and actions) and verbal communications. Madathil and Greenstein [34] divided collaboration methods into three categories: face-to-face, Web access-based video conference, and VR with HMD. Furthermore, a comparative analysis was conducted through a survey of user groups to provide immersive environments in distributed collaboration [35]. These have been used to study asymmetric virtual environments where various types of users participate together, such as VR users wearing HMDs, non-immersive users of PCs and mobile devices, and AR users. Studies have been conducted through various perspectives and approaches, including studies on asymmetric virtual environments, where HMD users of VR participate with PC or tabletop interface users [36], [37] as well as providing differentiated experiences and an improved presence in virtual environments by providing a different active role to each user [11], [13]. Jansen et al. [38] proposed a method for AR users to perform interesting interactions with non-HMD users. Cho et al. [14] studied interfaces and interactions considering the characteristics of users in an asymmetric virtual environment where VR and AR users participated together. Based on these studies, new experiential environments where users of various platforms and devices can participate together have been proposed, including potential applications to various industrial fields [39], [40]. Studies on user collaboration in asymmetric virtual environments are still actively ongoing. Lyu et al. [18] proposed a WebTransceiVR, a toolkit for asymmetric collaboration where multiple non-VR users can share the virtual space of VR users. Furthermore, Reski et al. [17] presented a system through case studies that allows multiple users to synchronously explore the same data based on a collaboration scenario that combines immersive and non-immersive interfaces in asymmetric virtual environments. In addition, Li et al. [41] conducted a comparative experiment on social presence for users in a hybrid VR and AR environment (HVAR) and a shared VR environment (SVR). Our study aims to propose interfaces and interactions that provide all users with satisfactory experiences based on a new experiential environment that encompasses devices, such as PCs and mobile devices, and immersive technologies, such as VR and AR, to build the virtual experience environment of an immersive content with an asymmetric virtual environment, where various platforms and devices are considered and non-immersive and immersive users participate together.

III. XAVE: CROSS-PLATFORM BASED ASYMMETRIC VIRTUAL ENVIRONMENT

This study proposes XAVE, an asymmetric virtual environment where non-immersive users on PCs and mobile devices, immersive users using VR and AR users can participate and experience together. It considers the characteristics of the platforms, devices, and experience methods of user, aiming to provide an experiential environment where the users can perform communications and social interactions based on various participation methods. As shown in Figure 1, active users form a virtual space through their respective interfaces. Furthermore, we provide an immersive content-based experiential environment that allows users to participate through the actions of actors using motion capture. This study aims to define user-specific interactions as an experiential environment that can be applied in various industrial fields.

The interfaces of the users participating in the proposed XAVE facilitate interactions with virtual environments, objects, and users in an intuitive structure, considering the point of view (POV) and input methods based on the characteristics of the platforms and devices. Furthermore, referring to the four factors (usefulness, ease of use, ease of learning, and satisfaction) of the Usefulness, Satisfaction, and Ease of use (USE) Questionnaire used by Lund [42], we focused on maximizing the use of conventional input methods, GUI structures, and manipulation processes. Using these factors, we focused on interface designs that can improve satisfaction in interactions by allowing users to learn and use interfaces easily and perform desired actions quickly and accurately.

To implement XAVE, PC and mobile device-based VR users used an Oculus Quest 2 HMD, touch controller, mobile HMD, and Xbox controller, and AR users detected the real world through image targets. The environment was constructed by integrating the Oculus Integration software development kit (SDK) [43], Google VR SDK [44], and Vuforia Engine [45] in the Unity 3D engine [46]. Photon Unity Networking (PUN) [47] was used to synchronize actions and communication between users. In addition, motion capture devices (16 OptiTrack Prime 17W cameras [48]) and motive software were used to implement the participation of actor-based avatars in the asymmetric virtual environment through linkage with the Unity 3D engine. The purpose of this study is to provide a method that anyone can easily use by utilizing popular engines and development tools during the development process, while considering the characteristics of the platform and traditional input processing methods. To achieve this, we aim to organize interfaces and interactions for each user.

Table 1 is an overview of templates for user-specific implementation in the proposed XAVE. Devices for each user, input method, plugin required for development, functions for actions and components required to implement them, and constraints in the current template were specified. Basically, it was intended to be easily utilized and applied as immersive

TABLE 1. An overview of the proposed XAVE user template (depends on Unity 3D engine).

User	Device	Input	Plugin	Function (Component)	Constraint
PC	keyboard, mouse, and monitor	defined key, and mouse button	none	**movement** (CharacterController), **change POVs** (Camera), and **action** (Input, EventTrigger, and Rigidbody)	none
Mobile	mobile	click or drag screen, and GUI (button, image, etc.)	none	**movement** (CharacterController, and EventSystems's Handler), **change POVs** (Camera), and **action** (Input, EventTrigger, and Rigidbody)	android
PC-based VR	VR HMD, and controller	controller key and position tracking	OpenXR or toolkit (Oculus integration, etc.)	**movement** (Transform), **camera** (XRRig or OVRCameraRig), and **action** (OVRInput (or XR Controller), RayCast, and Rigidbody)	Oculus
Mobile-based VR	mobile, and mobile HMD	gaze pointer, and controller key	GoogleVR SDK	**movement** (CharacterController), and **action** (Input, EventTrigger, RayCast, and Rigidbody)	Xbox controller, and android
Mobile-based handheld AR	mobile	click or drag screen, and GUI (button, image, etc.)	Vuforia engine	**movement** (CharacterController, and EventSystems's Handler), **camera** (ARCamera), and **action** (Input, EventTrigger, and Rigidbody)	image target, and android

content by configuring an asymmetric virtual environment in a general and popular way without a specific system or additional device. Also, the constraints and types of device are only concrete examples of the currently presented asymmetric virtual environment, and they can be applied in sufficiently different ways and device, and the interaction structure is the same, so it can be seen as a structure with extensibility.

A. PC USER

A PC user on a monitor and desktop computer interacts with the virtual environment through the GUI and provides inputs using a keyboard and mouse, representing the most typical interactive system. The interface was implemented based on the operating method of 2D/3D content, such as games.

Figure 2 shows the PC user experience method and interface structure. First, first- and third-person POVs are provided such that the POV can be changed. The user can move easily and conveniently in the virtual space by moving the character (camera) using the keyboard and controlling the camera rotation using the mouse. The user uses the GUI to perform an appropriate action for an event or situation.

The following is a summary of the PC user input methods and interaction structure based on it.

The presented PC user interaction uses the most common input method of keyboard and mouse. However, of course, the use of additional input devices such as game pads is also possible within the provided structure. In this case, if the directional keys and defined keys are mapped to the joystick and buttons of the gamepad, it is possible to implement the same interaction structure.

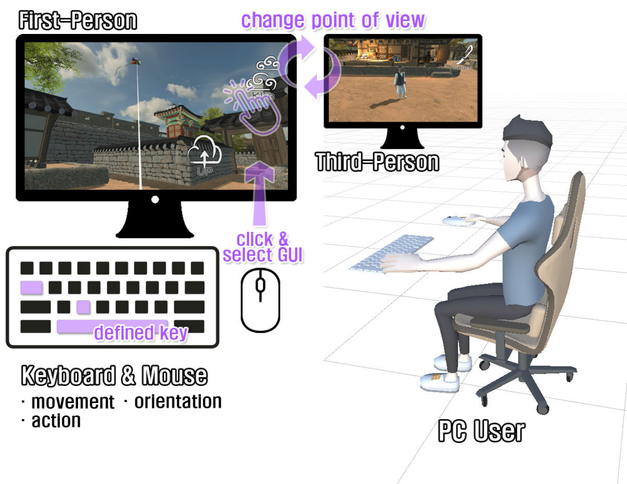


FIGURE 2. PC user interface and interaction structure in XAVE.

B. MOBILE USER

A mobile user performs interactions and actions through the GUI by touching the screen in specific methods, such as taps, long presses, drags, and swipes. Performing actions through the GUI by touch has an advantage in that direct interaction with the virtual environment and objects can be performed in an easy to learn and familiar method. Mobile users in the proposed XAVE also use touchscreen inputs and a GUI and are provided with both first- and third-person POVs, as with PC users, allowing them to freely and conveniently explore the virtual space according to the situation.

Figure 3 shows the mobile user experiential method and interface structure, which facilitates the character (camera) movements using a GUI-based virtual joypad and camera rotation control function through a screen touch and drag.

The following process summarizes interactions defined based on the mobile user's touch-based interface.

procedure PC user interaction
  0. definition
    (a) classify and define the keys and mouse buttons used for interaction.
    (b) design a GUI that corresponds to the action.
  1. keyboard keystroke input
    directional (W/A/S/D) keys: for moving the character (camera).
    defined keys:
      (a) for changing the camera's POV.
      (b) for the character's jumping.
      (c) for performing an action for an activity.
  2. mouse input
    screen click and drag: for rotating the camera.
  3. GUI selection
    button-click: for switching the behavioral state for an activity (experience, game, education, etc.).
end procedure

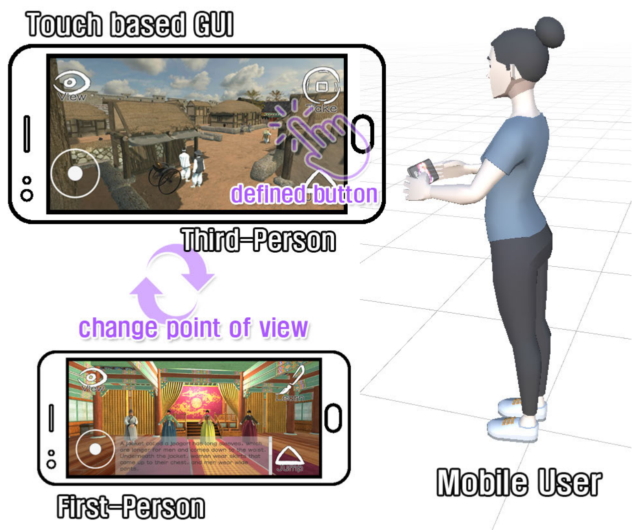


FIGURE 3. Mobile user touch-based interface structure in XAVE.

The first-person POV of VR users and third-person POV of AR users are provided together to partially improve the limited experiential environment of PC and mobile users constructed in non-immersive systems. Therefore, immersion, overall scene exploration, and flow identification can be provided comprehensively. The goal is to increase the immersion or interest while facilitating easy and efficient interactions in structures familiar to PC and mobile users in the asymmetric virtual environment through the multi-POV function.

C. PC-BASED VR USER

In a typical immersive experiential VR environment, a user wears an HMD to receive 3D visual information and interact with the virtual environment through a controller or hand, thereby creating a highly immersive experience. In the proposed asymmetric virtual environment, the VR user–as an independent character in the virtual space–can also perform direct interactions, i.e., perform interactions through controller inputs based on a first-person POV.

procedure Mobile user touch-based GUI
  definition
    (a) differentiate actions related to GUI and screen touch.
    (b) define the GUI and input corresponding to the classified actions.
  1. virtual joypad: for moving the character (camera).
  2. screen click and drag: for rotating the camera.
  3. defined button-clicks:
    (a) for changing the POV.
    (b) for jumping.
    (c) for switching the behavioral state for an activity (experience, game, education, etc.).
end procedure

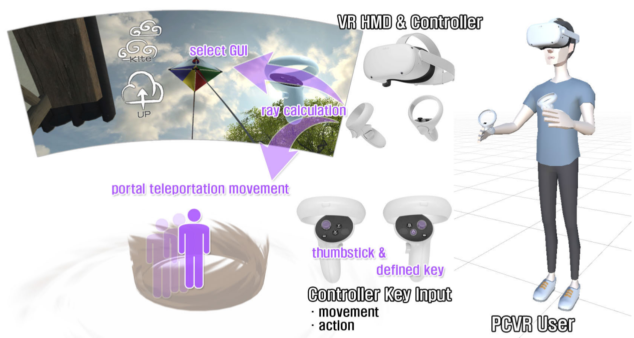


FIGURE 4. Interaction process of PC-based VR user in XAVE.

Figure 4 shows the interaction process of the PC-based VR user. The user performs actions through the key inputs of the touch controller while wearing an Oculus Quest 2 HMD. The function for performing actions based on key inputs is provided, with the character (camera) movements facilitated by two methods–the controller's thumbstick inputs and controller-to-ray-based portal teleportation method–and the information presented through the GUI.

The following summarizes the interaction process using Oculus Quest 2 and a touch controller.

The PC-based VR users presented in this study include interfaces and interactions that consider not only the Oculus Quest 2 HMD used in this study, but also the HTC VIVE HMD, Oculus Rift S HMD, and even console-based PlayStation VR HMD. Therefore, in order to consider versatility and expandability, a controller-based method is proposed, although the hand tracking function of the Oculus Quest 2 HMD can be used for an immersive experience environment.

D. MOBILE-BASED VR USER

Owing to the advancement of high-end portable smart devices, anyone can easily buy and use them. Furthermore,

procedure PC-based VR user interaction
    0. definition
        (a) classify and define the controller keys used for interaction.
        (b) classify and define interactions through raycast.
    1. controller-to-ray calculation
        collision with ground: for calculating the character's teleportation movement (portal).
        collision with a button or image: for menu selection.
    2. controller key input
        thumbstick keys: for moving the character (camera).
        defined keys:
            (a) for switching the behavioral state for an activity (experience, education, game, etc.).
            (b) for portal teleporting through controller-to-ray.
    present the user's action status through the GUI.
end procedure

VR content, which use 360 VR videos based on mobile device or have simple interactive content structures, have been developed for various fields. People can experience them by purchasing mobile HMDs easily and at low costs. They have the advantage that they are easier to use than PC-based VR content, which have constraints owing to the dependency on device and high costs. Considering these advantages, this study builds an environment where mobile-based VR users can participate with PC-based VR users.

Mobile users also exist in the virtual space in a first-person POV. However, using touchscreen-based interfaces are impossible in experiential environments where the mobile device is attached to the HMD. Therefore, mobile VR often implements interactions using users' gaze. In this study, actions were performed by selecting the GUI or directly selecting the virtual environment, object, and user based on the input through a gaze pointer. In addition, a controller compatible with the mobile device was used to facilitate manipulation and control through key inputs.

Figure 5 shows the interaction process of a mobile VR user, and various actions performing functions defined with the character (camera) movements are provided through the gaze pointer or controller inputs.

The following summarizes the interaction process, where a user wearing a mobile HMD uses the gaze pointer and controller to interact.

The VR user exists in a virtual space when the visual information of the real world is completely blocked as the user wears the HMD. Therefore, the main objective is to design an interface and interaction structure that facilitates easy and fast decision-making and actions through the controller and delivers the current situation accurately using the GUI.

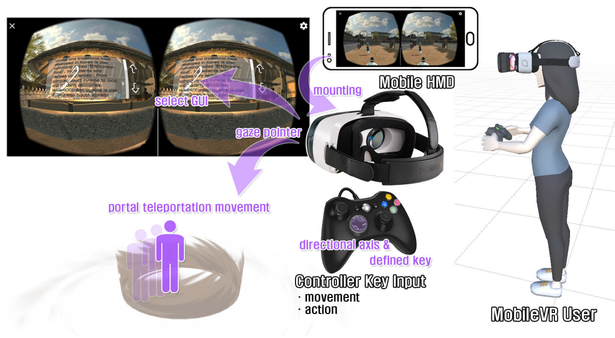


FIGURE 5. Input and interaction process of a mobile-based VR user in XAVE.

procedure Mobile-based VR user interaction
    0. definition
        (a) differentiate actions related to gaze pointer and controller keys.
        (b) define the gaze pointer and input corresponding to the classified actions.
    1. gaze pointer detection
        collision with ground: for calculating the character's movement.
        collision with a button or image: for menu selection.
    2. controller key input
        directional axis key: for moving the character (camera).
        defined keys:
            (a) for switching the behavioral state for an activity (experience, education, game, etc.).
            (b) for portal teleporting through the gaze pointer.
    present the user's action status through the GUI.
end procedure

E. MOBILE-BASED HANDHELD AR USER

Experiential environments using AR technology involve an eyeglasses-type wearable device or a mixed-type HMD AR, a glasses-type device, and a handheld method that uses a mobile or tablet device. VR HMDs are continually lowering in price and can be easily purchased and used by ordinary users. However, eyeglasses-type devices for AR have many constraints for ordinary users, such as high costs and developer-oriented supply methods. Therefore, for this study, we built a handheld AR experience environment using mobile devices in a method that users can access easily, with a reduced device burden. For this purpose, the Vuforia engine and image target were used to create an augmented virtual scene.

Figure 6 shows the interface and interaction structure of the handheld AR user using mobile device. It is based on a touch-based interface because the AR user interacts based on a mobile device, as with the mobile user. Therefore, when the image target is detected through the mobile device's camera, the virtual environment is augmented atop the real world, and interactions are performed in the same method as the mobile user's touch-based GUI. However, AR is in the third-person POV because the overall situation in the scene can be determined by viewing the virtual space from above. The goal is to control the virtual scene while freely moving the mobile device and communicating easily with the virtual environment, objects, and other users.

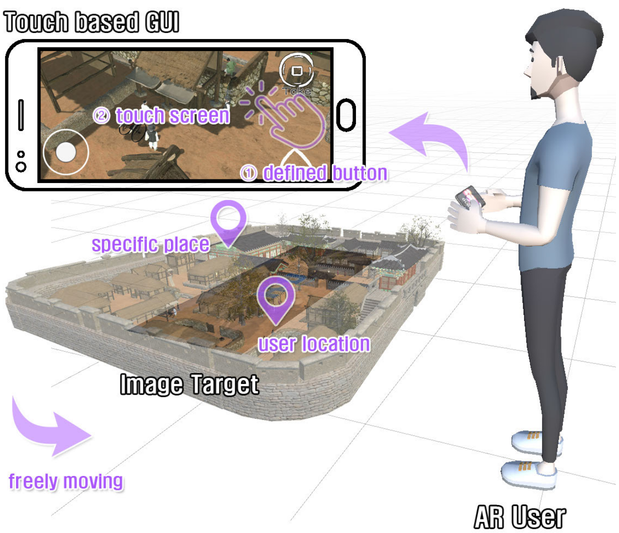


FIGURE 6. **Interface and interaction process of a handheld AR user using mobile device in XAVE.**

The following summarizes the AR interaction process. It basically uses the same touch-based GUI of mobile user. However, in order to fully utilize the advantage of judging and controlling the situation while looking at the scene as a whole, a function to display a specific place and user location information in the experience environment is added. In addition, to provide convenience of movement, it provides the function of directly selecting a destination on the screen and moving through navigation.

In this study, AR user utilizes marker-based AR method using image target as a universal and easy approach. Similarly, although it is possible to apply a highly immersive AR environment using real-world space such as ground plane or area target, the interface and interaction structure do not differ significantly from basic mobile input methods. Therefore, it is expected that the presented mobile-based handheld AR user interaction can be further developed into more advanced and immersive experiences.

F. ACTOR-BASED AVATAR USING MOTION CAPTURE

The proposed asymmetric virtual environment has an experience environment, where actor-based avatars using motion capture can participate as guides, non-player characters (NPCs), and event elements (parades, performances) in addition to the users proactively participating and performing actions in the virtual space. Motion capture refers to digitally recording the human body's movements by attaching sensors

procedure Mobile-based Handheld AR User interaction
    0. definition
        (a) differentiate actions related to GUI and screen touch.
        (b) define the GUI and input corresponding to the classified actions.
    1. touch screen
        collision with ground: for moving the character through navigation.
    2. GUI selection
        defined keys:
            for marking the specific place or user location information as a marker
    secure a blocked view by inspecting obstacles between the camera and the character.
    same as the touch-based GUI of mobile user.
end procedure

on the body or using infrared light and can be classified into optical and inertial types. The optical method uses two or more cameras to capture images of the sensors (markers) and calculate the 3D information; this method has the advantage of representing the result close to real-time. The inertial method calculates the movement through a suit, on which inertial sensors—a combination of acceleration, gyro, and geomagnetic sensors—are attached to body joints and major parts. It can represent movements at less expensive costs than the optical motion capture method. This study used optical motion capture equipment for sophisticated movements, recording and tracking the actor's motions to apply them to the defined virtual avatar.

Figure 7 shows the setup of an actor-based avatar using motion capture and participating in the virtual environment. The actor wearing a marker-attached full-body suite performs motions in a 7.8 $\times$ 7.8 m space, and the motion information captured from 16 cameras can be seen in the Motive software. Real-time skeleton data streaming is set up for the pre-defined humanoid avatar in the Unity 3D engine, and the data captured by the motion capture equipment is streamed in real-time to the Unity 3D engine through the Motive software.

The following summarizes the process of the proposed actor-based avatar. The avatar in the virtual space operating based on the actor's motion cannot be directly observed from the avatar's POV but can be seen indirectly through a projector or monitor. Providing an environment in which the actor wearing the marker-attached full-body suit interacts is limited in terms of the virtual environment or controlling the GUI while performing various actions; therefore, this study defines the role of the actor-based avatar as a limited participant.

G. COLLABORATION

This study defines a collaborative structure in which all users of various participation methods (device, platform, etc.) considered as non-immersive, immersive, and AR users and actor-based avatars using motion capture can perform natural

FIGURE 7. Environment set up and data streaming for the construction of actor-based avatar using motion capture.

procedure Process of Actor-based Avatar using Motion Capture
1. The actor wearing a marker-attached full-body suite performs the action in real space.
2. The human body's movements accumulated through the motion capture camera is calculated as a skeleton of the Motive software.
3. Data streaming from Motive software to Unity 3D engine: synchronize humanoid avatar in Unity 3D engine and skeleton information of Motive software.
4. Run actor-based avatars through immeresive content.
5. Check the virtual scenes and actions from the avatar's POV through the projector screen.
end procedure

experiences in the same virtual space. The most important goal in the collaborative structure is that each user with a different participation method participates in the immersive content through a consistent process. And it is to define a synchronization structure that can effectively perform the presented experience environment. Figure 8 is the proposed collaborative processing structure of XAVE, and shows the network process for efficiently processing data in an immersive content that performs experiences on various topics such as experiences, education, and games with users and actor.

1) USERS

Each user has one controllable character and a camera for the scene. And it synchronizes the character animation information along with the transform (position, rotation) for the camera and character. Also, the user changes the properties and statuses of objects or other users through remote procedure call (RPC) in the process of experiencing them in the virtual environment.

At this time, regardless of the participation method (device, platform, interaction, etc.), it has a flexible structure in which new participation methods can be added with a consistent structure of information delivery and synchronization method.

2) ACTOR

We define a synchronization process to accurately and realistically express the movement of an actor wearing a marker-attached full-body suite and skeleton data calculated from motion capture cameras in a virtual space in which users participate. Specifically, it consists of the process of transmitting and broadcasting the transform (position, rotation) information of each bone constituting the joint to the server.

3) IMMERSIVE CONTENT

Objects and statuses composing content are defined as a list. And it broadcasts to other users by checking the situation changed by users in real time. It also defines a rigidbody synchronization structure to synchronize the physical movement of objects placed in the content.

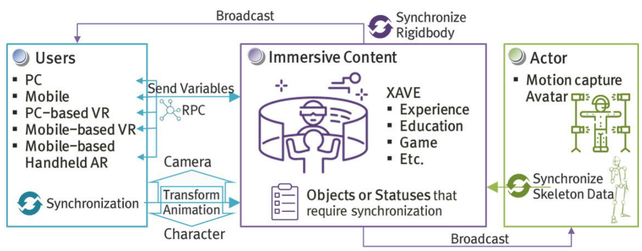


FIGURE 8. Collaboration structure for network-based synchronization of proposed XAVE.

This study defines a collaborative structure for an experience environment in which various user interactions, platforms, and input methods are synchronized in a generalized structure and flow, allowing for participation of an actor-based avatar (real-world participant) using motion capture. There is no technological difference compared to the network synchronization structure for characters in typical games (MMORPG, etc.) because the basic assumption is that all users have the same role. However, the key purpose is to define a structure for creating an effective collaborative experience environment that considers various users and avatar participation in an asymmetric virtual environment.

IV. IMMERSIVE CONTENT CREATION

Specific immersive content was created to construct the asymmetric virtual environment, where five types of users–PC, mobile, PC- and mobile-based VR, and mobile-based AR users–and actor-based avatars using motion capture participate together. All users have the same role, and each user is free to experience various events provided in the virtual environment based on the defined interface and interaction method. Actor-based avatars are NPCs or event elements in the environment and have the role of triggering interest or increasing immersion.

This study uses the graphic resources (Korea Village Palace, Interior, People, Soldier [49]) provided by the asset store in the theme of a virtual folk village to construct the backgrounds and define various activities that can be performed in the folk village. Here, the important goal is to provide a new direction for applying an immersive content creation field by fusing various industrial fields through the proposed asymmetric virtual environment. Therefore, we incorporated various experience factors, education, and games into the content and examined the interaction process of different users in the given experience environment. The activities (experience, education, and games) provided by the immersive content in this study are as follows.

1 **Experience:** This refers to the experiential contents of folk culture, comprising activities (kite flying and riding) that allow users to experience this culture.
2 **Education:** This refers to the educational content of folk culture, which delivers educational content on the theme of traditional clothes and food through the GUI (texts).
3 **Game:** This refers to a folk game content, which provides a game in which all users participating in the virtual environment participate together. A tagger user must find all other hiding users within the game event, which occurs for a limited amount of time.

Figure 9 shows the process of proceeding with the created immersive content, where each user performs the activity provided according to the defined interface. This process was used to conduct surveys with all users to investigate the interface satisfaction and presence. Then, a comprehensively comparative analysis of the experiences based on the experiential method was conducted.

V. EXPERIMENTAL RESULTS AND ANALYSIS

The construction of the experiential environment and immersive content in the proposed asymmetric virtual environment XAVE were implemented using the Oculus Integration SDK, GoogleVR SDK (GoogleVRForUnity_1.200.1), and Vuforia 9.8.11 Engine based on Unity 2019.4.30 (64-bit). Each PC-based VR user used an Oculus Quest 2 HMD and touch controller while the regular, VR, and AR mobile users used Samsung Galaxy models with Android version 12 or higher. Among these models, mobile-based VR users used a Baofeng Mojing III HMD and Xbox controller. The actor-based avatar using motion capture used 16 units of OptiTrack Prime 17W cameras. Finally, the PC for the integrated development environment and experiments was equipped with an Intel Core i7-10875H, 16 GB of RAM, and a Geforce RTX 2060 GPU. Figure 1 shows all users participating in XAVE. A user experiences a space (1.5 × 1.5 m) sufficiently large for sitting or and comfortably. For the AR user, a space of sufficient size (2 × 2 m) was constructed for moving freely on the image target.

A. EXPERIENCE ENVIRONMENT COMPARISON

The proposed XAVE experience environment is different from the existing asymmetric virtual environment studies. Table 2 shows this, and this study intends to confirm it through comparative analysis with ShareVR [11], RoleVR [13], X-Person [14], QuarantivityVR [50], Virtual City Tours [39], and Collaborative Learning [51], which are well-known asymmetric virtual environment studies.

The detailed comparative analysis is divided into platforms and device, roles, interactions, manipulation, and POVs (first and third person) by users in the participation environment. Compared with general asymmetric virtual environment studies in which VR and non-HMD or VR and AR users participate, the proposed XAVE shows that it is possible to participate in various platforms by expanding the user's scope. In particular, while existing studies have practically difficult problems to apply with device configuration specialized for the experience environment, this study has the advantage of being easy to apply to various immersive or metaverse contents by using common and popular device while considering various platforms. In addition, there was a difference in providing a flexible POV in consideration of the immersion and satisfaction of non-immersive users.

The XAVE experience environment in this study takes into account expanded platforms, user participation, and limited freedom in POV compared to previous studies. The important point is to systematically analyze each user experience through surveys in the development process in order to assist effective user or platform selection to apply the asymmetric virtual environment to various immersive or metaverse contents.

B. SURVEY EVALUATION

Surveys were conducted to compare and analyze user experiences, including the interface satisfaction and presence, after experiencing the immersive content created, considering each participation method in the experiential environment of the XAVE. The survey consisted of 23 participants (10 males and 13 females) between the ages of 19 and 40. To eliminate the difficulty or confusion of manipulation when experiencing all the various participation modes, such as PC, mobile, VR, and AR, the participants were recruited from a group of people who had experienced or developed interactive content, such as games. The purpose of the survey was to comparatively analyze the experiences of the users in various modes while participating in the asymmetric virtual environment with the content based on the proposed input methods, interfaces, and interactions. Four or five participants formed a group and experienced content together in their respective participation methods. When an experience was completed, the participation method was changed to proceed with the experience. Therefore, each user experienced content five times, and the experiencing sequences were evenly distributed to reduce the difference in the survey results caused by sequential order.

FIGURE 9. Results of creating immersive content with a virtual folk village background based in the XAVE: (a) experience (kite flying and riding), (b) education (traditional clothes and food), (c) game (hide-and-seek), (d) PC, mobile, PC- and mobile-based VR, mobile-based AR users, and actor-based Avatar.

TABLE 2. Comparison results of experiential environment with existing asymmetric virtual environment studies (2D: interact with 2D GUI, 3D: interact with 3D objects.

Task	ShareVR	RoleVR	X-Person	QuarantivityVR	Virtual City Tours	Collaborative Learning	XAVE
Platform	VR, non-HMD	VR, non-HMD	VR, AR	VR, non-HMD	VR, AR	VR, non-HMD	PC, mobile, VR, AR, motion capture
Device	HTC Vive, 7 inch display mounted on controller, TV, BenQ W1080ST projectors	Oculus Rift, walking simulator (self-designed), leap motion	Oculus Rift S, mobile (image target)	HTC Vive, Microsoft Kinect, web cam	Oculus Quest, Hololens 2	HTC Vive, Microsoft Surface 2 Pro	PC, mobile, Oculus Quest 2, mobile (image target), OptiTrack Prime 17W
Role Diff.	different	different	different	same	different	different	same
Interaction	physical, voice	voice	in content	gesture, voice	in content	voice	in content
Manipulation	3D	3D	2D, 3D	3D	3D	3D	2D, 3D
POVs	fixed	fixed	fixed	fixed	fixed	fixed	limited flexible

In the first survey, the satisfaction with each interface was comparatively analyzed between the PC, mobile, PC- and mobile-based VR, and mobile-based AR users. The objective of interacting through the proposed interfaces is easy learning and convenient use and efficiency based on conventional input methods and GUIs. To examine the experiences on this, we recorded the scores on a 7-point scale based on the four factors and 30 questions of the USE Questionnaire from Lund [42].

Table 3 and Figure 10 list the statistical data for the four factors (usefulness, ease of use, ease of learning, and satisfaction) based on the survey results. PC and mobile users could easily learn and use the system in the most popular and familiar interaction method, demonstrating the advantage of improving the efficiency of actions. Note that the PC-based VR users show relatively high scores in satisfaction, which indicates that once they learned and adapted to the interface, it helps immersion in the virtual environment and may trigger more interest. Furthermore, because the two VR participation modes (mobile-based VR and PC-based VR) were provided simultaneously, users who felt immersed or satisfied in the PC-based VR were confirmed to have a relative influence on the interaction in mobile-based VR. Finally, many responses indicated that AR users did not have a better experience compared to mobile users because none of the additional elements showed distinctive characteristics of AR other than providing an experiential environment by synthesizing virtual scenes over the real world. The result of calculating the statistical significance through one-way analysis of variance (ANOVA) showed no significant differences in terms of the participation methods and all questionnaire factors. Comparative analysis was performed according to the equipment (platform, PC, or mobile), immersive technology (PC- vs. mobile-based VR), and the same interface, but no significant differences were found in the overall satisfaction.

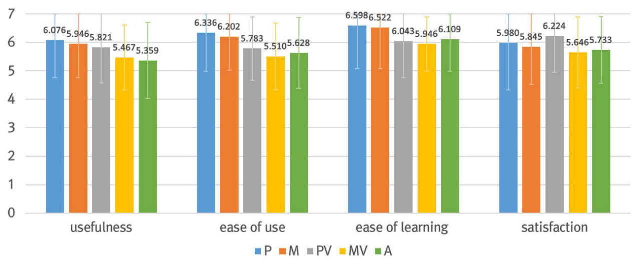


FIGURE 10. **Analysis results of interface satisfaction for each user type in the asymmetric virtual environment XAVE (P: PC, M: Mobile, PV: PC-based VR, MV: mobile-based VR, A: mobile-based handheld AR).**

TABLE 3. **Pairwise comparison analysis results of interface satisfaction for each user type in the asymmetric virtual environment XAVE (P: PC, M: Mobile, PV: PC-based VR, MV: mobile-based VR, A: mobile-based handheld AR); * indicates statistical significance.**

Pairwise Comparison	
usefulness	$F(4,110) = 1.482, p = 0.213$
ease of use	$F(4,110) = 2.314, p = 0.062$
ease of learning	$F(4,110) = 1.928, p = 0.111$
satisfaction (overall)	$F(4,110) = 0.777, p = 0.542$
satisfaction (P vs PV)	$F(1,44) = 0.435, p = 0.513$
satisfaction (M vs MV vs A)	$F(2,66) = 0.146, p = 0.864$
satisfaction (PV vs MV)	$F(1,44) = 2.807, p = 0.101$
satisfaction (M vs A)	$F(1,44) = 0.095, p = 0.759$

The second survey was conducted to analyze the presence based on immersion in the virtual environment. This study assumed that users would feel a different presence depending on the participation method or that the presence felt by non-immersive users would improve while interacting with immersive users through the asymmetric virtual environment, which considers both the non-immersive- and immersive-type participation methods. Therefore, participants evaluated the respective participation methods on a 7-point scale for the 19 questions of the presence questionnaire used by Witmer et al. [52].

Table 4 and Figure 11 list the comparative analysis results of the specific factors based on the recorded scores. In the comprehensive analysis of presence, PC-based users showed a higher overall satisfaction than mobile-based users. The difference in visual information (large screen size or 3D visual information) seemed to play an important role. PC- or mobile-based users showed higher values with the same device because various POVs (first and third person) were provided, which helped users predict and observe situations at various angles and adapt to the virtual environment quickly. As expected, the survey results confirmed that VR users could focus more on actions that benefited in terms of immersion in the experiential environment. Again, mobile-based VR showed relatively low satisfaction because of the experience in PC-based VR. Finally, AR users demonstrated a limitation in providing a differentiated experience compared to the interactions of mobile users in third person POV, as was the case in the survey results on interface satisfaction. However, the difference in visual information was found to act as an element for focusing on or hindering the action (quality of interface), causing a significant difference. Furthermore, in adapting to the virtual environment quickly and performing actions skillfully (self-evaluation of performance), significant differences were observed between the participation methods of the users.

TABLE 4. **Pairwise comparison analysis results of presence for each user type in the asymmetric virtual environment XAVE (P: PC, M: Mobile, PV: PC-based VR, MV: mobile-based VR, A: mobile-based handheld AR); * indicates statistical significance.**

Pairwise Comparison	
total	$F(4,110) = 1.147, p = 0.338$
realism	$F(4,110) = 0.850, p = 0.497$
possibility to act	$F(4,110) = 1.222, p = 0.306$
quality of interface	$F(4,110) = 2.599, p < 0.05$*
possibility to examine	$F(4,110) = 1.080, p = 0.370$
self-evaluation of performance	$F(4,110) = 3.007, p < 0.05$*
total (P vs PV)	$F(1,44) = 0.050, p = 0.823$
total (M vs MV vs A)	$F(2,66) = 0.970, p = 0.384$
total (PV vs MV)	$F(1,44) = 1.832, p = 0.183$
total (M vs A)	$F(1,44) = 1.403, p = 0.243$

Various situations occurring in the experience process demonstrated that non-immersive participation users adapted to the virtual space relatively quickly. However, in adapting to the space, non-immersive-type (PC or mobile) users formed a social presence with VR users, thereby guiding or experiencing them together. Although they could perform independent activities, except for game activities, new experiences could be triggered, such as achieving the goal together and self-assigning unintended roles, considering the characteristics of the participation method. Furthermore, actor-based avatars using motion capture could be used sufficiently as an element of interest to users in the form of events, such as a parade in a theme or amusement park.

C. PERFORMANCE

Each user participates in the proposed asymmetric virtual environment of XAVE using different devices and platforms. Therefore, differences in technical performance for each user may cause problems such as delay and disconnection in the

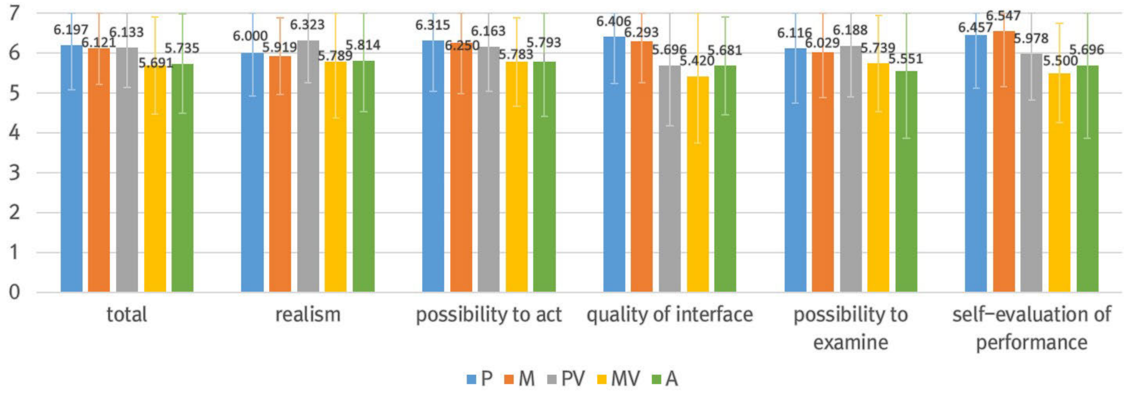


FIGURE 11. Analysis results of presence for each user type in the asymmetric virtual environment XAVE (P: PC, M: Mobile, PV: PC-based VR, MV: mobile-based VR, A: mobile-based handheld AR).

experience process. The technical performance as real-time content was confirmed by measuring the number of frames per second (FPS) for each user in the immersive content produced as the last experiment. Table 5 shows this. FPS was measured while the user freely experienced immersive content for 1 minute. Since there is a change in FPS depending on the scene being rendered in the current camera, the minimum, maximum and average values were recorded. It is a technical environment where all users can experience a natural experience without delay or interruption on average. When it is based on a mobile platform, performance is inevitably relatively low due to hardware differences compared to a PC. In addition, since VR is rendered from the binocular viewpoint, there are performance differences even within the same platform. Nevertheless, it is not unreasonable to experience our immersive content.

The resource data of currently produced content is large, with triangles drawing the most in the rendered scene at 3.1M and vertices at 2.9M. Therefore, it is expected that technical performance can be further improved by constructing a scene composed of lighter resources or adding optimization work such as occlusion culling and light map.

TABLE 5. Comparison result of performance measurement for each user in the proposed asymmetric virtual environment XAVE (P: PC, M: Mobile, PV: PC-based VR, MV: mobile-based VR, A: mobile-based handheld AR). *FPS (frames per second).

	Min	Max	Mean
P	99	595	341
M	22	133	85
PV	87	157	101
MV	20	99	59
A	25	67	50

VI. LIMITATIONS AND DISCUSSION

In the case of user-specific templates in the proposed XAVE, dependence on existing commercial engines, SDKs, and plugins is high. This was because using engines, such as Unity 3D and Unreal, can effectively produce 3D content for various purposes with high immersion, including interactive and immersive contents, compared to building an experiential environment in a native development environment. In addition, it has the advantage of constructing an asymmetric virtual environment in which users of various platforms participate because of its compatibility with external development tools. Therefore, this paper presents a template with a focus on usability. However, these limitations can be overcome if the template is implemented in the form of a package (or library) that can be applied immediately to any representative commercial engine, or an approach to design a generalized structured experiential environment that can reduce dependence on existing tools and engines.

The AR users in this study were unable to utilize many of the advantages and characteristics of AR. As observed from the immersive content created for the experiment, information from the real world was not particularly utilized; therefore, there was no significant difference in experience compared to mobile users. The results confirmed that it is impossible to achieve high expectations in terms of presence and satisfaction. Therefore, from the point of view of AR, if devices such as the Microsoft HoloLens 2 are used or visually improved in a manner that can meaningfully utilize information in the real world, it is expected that the value of their use for users having a new experience will increase.

In the survey experiment, the subjective experiences of the participants were recorded together with the statistical data described above. For example, there was an opinion that interest in AR was higher when there was no experience with immersive technologies, such as VR and AR. Although curiosity about the virtual environment drawn in the real world was high, the advantages were not utilized because of the lack of clear differences with mobile users during the experience process. Next, in the case of VR, mobile-based VR was provided in terms of usability, but the dominant opinion was that the overall satisfaction was the lowest.

The main reason for this was that it could not provide a better experience than PC-based VR, although many constraints existed regarding the interaction. From an economic point of view, mobile-based VR was also considered; however, there was a mistake in the approach that the participants could not directly check this part because there was no cost to experience. Finally, the use of a controller rather than teleportation in the movement of VR users affected VR sickness. This was predicted in a previous study [53]; therefore, not only controller-based movement but also teleportation movement was provided. The problem described in this section was confirmed experimentally.

Ultimately, this study aims to present an experiential environment and interaction that can be used in various industries through the proposed asymmetric virtual environment. However, in content production, experiments in various industrial fields have not been carried out. Therefore, we plan to conduct applied study on the production of immersive convergence contents in various industries such as education, construction, and manufacturing through the proposed XAVE in the future.

VII. CONCLUSION

In this study, we proposed XAVE, an asymmetric virtual environment, where various types of users (PC and mobile-based non-immersive users, VR-based immersive users, AR user, and actor-based avatars using motion capture) can participate and experience together. It allows all users to learn easily and act efficiently through common and popular interfaces, considering the characteristics of respective experience methods. In particular, the experiential environment of non-immersive users was constructed based on interactions through keyboard and mouse input-based GUI selection and key inputs for PC users and interactions through GUI-based touchscreen inputs for mobile users. For the immersive participation methods, VR users wore an HMD and entered inputs through a controller to perform interactions; AR users recognized the image target from a mobile camera and used a GUI and touchscreen inputs to perform interactions. Another participation method included using motion capture to record and track the actor's motions, which were linked to an avatar in the virtual environment, as an event element. Consequently, a novel experiential environment that is useable in a variety of metaverse content fields was constructed. Based on this, immersive content was created, and surveys were conducted with the participants to examine whether all users can have significant experiences by performing interactions according to the respective participation methods and experiential environments. The results show that the users were provided with satisfying, immersive, and interesting experiences in the virtual environment through appropriate interfaces and interaction methods according to their respective participation methods. In the future, we plan to expand the range of user participation, including AR users using eyeglasses-type devices, such as Microsoft HoloLens, which are not currently considered in the XAVE. Furthermore, we will create new, immersive or metaverse content to confirm that the proposed XAVE is applicable in various industries, such as manufacturing and construction.

REFERENCES

[1] T. Sweeney, "Foundational principles & technologies for the metaverse," in *Proc. ACM SIGGRAPH Talks*, New York, NY, USA, Jul. 2019, p. 1, Art. no. 38.

[2] A. van Dam, "Post-WIMP user interfaces," *Commun. ACM*, vol. 40, no. 2, pp. 63–67, Feb. 1997.

[3] M. Kim, J. Kim, K. Jeong, and C. Kim, "Grasping VR: Presence of pseudo-haptic interface based portable hand grip system in immersive virtual reality," *Int. J. Hum.-Comput. Interact.*, vol. 36, no. 7, pp. 685–698, Apr. 2020.

[4] M. Fechter, B. Schleich, and S. Wartzack, "Comparative evaluation of WIMP and immersive natural finger interaction: A user study on CAD assembly modeling," *Virtual Reality*, vol. 26, no. 1, pp. 143–158, Mar. 2022.

[5] J. Wagner, W. Stuerzlinger, and L. Nedel, "Comparing and combining virtual hand and virtual ray pointer interactions for data manipulation in immersive analytics," *IEEE Trans. Vis. Comput. Graphics*, vol. 27, no. 5, pp. 2513–2523, May 2021.

[6] D. Yu, X. Lu, R. Shi, H.-N. Liang, T. Dingler, E. Velloso, and J. Goncalves, "Gaze-supported 3D object manipulation in virtual reality," in *Proc. CHI Conf. Human Factors Comput. Syst.*, New York, NY, USA, May 2021, pp. 1–13.

[7] A. K. Mutasim, A. U. Batmaz, and W. Stuerzlinger, "Pinch, click, or dwell: Comparing different selection techniques for eye-gaze-based pointing in virtual reality," in *Proc. ACM Symp. Eye Tracking Res. Appl.*, New York, NY, USA, May 2021, pp. 1–7.

[8] J. Lee, K. Jeong, and J. Kim, "MAVE: Maze-based immersive virtual environment for new presence and experience," *Comput. Animation Virtual Worlds*, vol. 28, nos. 3–4, p. e1756, May 2017.

[9] N. L. Williams, A. Bera, and D. Manocha, "Redirected walking in static and dynamic scenes using visibility polygons," *IEEE Trans. Vis. Comput. Graphics*, vol. 27, no. 11, pp. 4267–4277, Nov. 2021.

[10] C. Carlsson and O. Hagsand, "DIVE a multi-user virtual reality system," in *Proc. IEEE Virtual Reality Annu. Int. Symp.*, Sep. 1993, pp. 394–400.

[11] J. Gugenheimer, E. Stemasov, J. Frommel, and E. Rukzio, "ShareVR: Enabling co-located experiences for virtual reality between HMD and non-HMD users," in *Proc. CHI Conf. Human Factors Comput. Syst.*, New York, NY, USA, May 2017, pp. 4021–4033.

[12] J. G. Grandi, H. G. Debarba, and A. Maciel, "Characterizing asymmetric collaborative interactions in virtual and augmented realities," in *Proc. IEEE Conf. Virtual Reality 3D User Interfaces (VR)*, Mar. 2019, pp. 127–135.

[13] J. Lee, M. Kim, and J. Kim, "RoleVR: Multi-experience in immersive virtual reality between co-located HMD and non-HMD users," *Multimedia Tools Appl.*, vol. 79, nos. 1–2, pp. 979–1005, Jan. 2020.

[14] Y. Cho, J. Kang, J. Jeon, J. Park, M. Kim, and J. Kim, "X-person asymmetric interaction in virtual and augmented realities," *Comput. Animation Virtual Worlds*, vol. 32, p. e1985, no. 5, Sep. 2021.

[15] K. Jeong, J. Kim, M. Kim, J. Lee, and C. Kim, "Asymmetric interface: User interface of asymmetric virtual reality for new presence and experience," *Symmetry*, vol. 12, no. 1, p. 53, Dec. 2019.

[16] Y. Cho, S. Hong, M. Kim, and J. Kim, "DAVE: Deep learning-based asymmetric virtual environment for immersive experiential metaverse content," *Electronics*, vol. 11, no. 16, p. 2604, Aug. 2022.

[17] N. Reski, A. Alissandrakis, and A. Kerren, "An empirical evaluation of asymmetric synchronous collaboration combining immersive and non-immersive interfaces within the context of immersive analytics," *Frontiers Virtual Reality*, vol. 2, Jan. 2022, Art. no. 743445.

[18] H. Lyu, C. Vachha, Q. Chen, O. Pyrinis, A. Liou, B. T. Kumaravel, and B. Hartmann, "WebTransceiVR: Asymmetrical communication between multiple VR and non-VR users online," in *Proc. CHI Conf. Human Factors Comput. Syst. Extended Abstr.*, New York, NY, USA, Apr. 2022, pp. 1–7.

[19] B. Shneiderman, "Direct manipulation: A step beyond programming languages (abstract only)," *ACM SIGSOC Bull.*, vol. 13, nos. 2–3, p. 143, May 1981.

[20] J. F. Nunamaker, D. C. Derrick, A. C. Elkins, J. K. Burgoon, and M. W. Patton, "Embodied conversational agent-based kiosk for automated interviewing," *J. Manag. Inf. Syst.*, vol. 28, no. 1, pp. 17–48, Jul. 2011.

[21] M. Khademi, H. M. Hondori, A. McKenzie, L. Dodakian, C. V. Lopes, and S. C. Cramer, "Free-hand interaction with leap motion controller for stroke rehabilitation," in *Proc. CHI Extended Abstr. Human Factors Comput. Syst.*, New York, NY, USA, Apr. 2014, pp. 1663–1668.

[22] J. Cui, A. Kuijper, and A. Sourin, "Exploration of natural free-hand interaction for shape modeling using leap motion controller," in *Proc. Int. Conf. Cyberworlds (CW)*, Chongqing, China, Sep. 2016, pp. 41–48.

[23] N. Sidorakis, G. A. Koulieris, and K. Mania, "Binocular eye-tracking for the control of a 3D immersive multimedia user interface," in *Proc. IEEE 1st Workshop Everyday Virtual Reality (WEVR)*, Mar. 2015, pp. 15–18.

[24] T. Kang, M. Chae, E. Seo, M. Kim, and J. Kim, "DeepHandsVR: Hand interface using deep learning in immersive virtual reality," *Electronics*, vol. 9, no. 11, p. 1863, Nov. 2020.

[25] Y. Lee, M. Kim, Y. Lee, J. Kwon, Y. Park, and D. Lee, "Wearable finger tracking and cutaneous haptic interface with soft sensors for multi-fingered virtual manipulation," *IEEE/ASME Trans. Mechatronics*, vol. 24, no. 1, pp. 67–77, Feb. 2019.

[26] S. Yun, S. Park, B. Park, S. Ryu, S. M. Jeong, and K. Kyung, "A soft and transparent visuo-haptic interface pursuing wearable devices," *IEEE Trans. Ind. Electron.*, vol. 67, no. 1, pp. 717–724, Jan. 2020.

[27] H.-Y. Huang, C.-W. Ning, P.-Y. Wang, J.-H. Cheng, and L.-P. Cheng, "Haptic-go-round: A surrounding platform for encounter-type haptics in virtual reality experiences," in *Proc. CHI Conf. Human Factors Comput. Syst.*, New York, NY, USA, Apr. 2020, pp. 1–10.

[28] M. Seo, K. Kang, and H. Kang, "VR blowing: A physically plausible interaction method for blowing air in virtual reality," *IEEE Trans. Vis. Comput. Graphics*, early access, Jan. 20, 2023, doi: 10.1109/TVCG.2023.3238478.

[29] N. A. Halim, A. W. Ismail, and N. M. Suaib, "Summoning method using freehand interaction for occluded object selection in handheld augmented reality," *Adv. Res. Appl. Sci. Eng. Technol.*, vol. 30, no. 1, pp. 228–242, Mar. 2023.

[30] M. C. Cha, H. C. Kim, and Y. G. Ji, "The unit and size of information supporting auditory feedback for voice user interface," *Int. J. Hum.-Comput. Interact.*, pp. 1–10, Feb. 2023.

[31] D. Han, R. Lee, K. Kim, and H. Kang, "VR-HandNet: A visually and physically plausible hand manipulation system in virtual reality," *IEEE Trans. Vis. Comput. Graphics*, early access, Mar. 13, 2023, doi: 10.1109/TVCG.2023.3255991.

[32] M. D. Dogan, A. F. Siu, J. Healey, C. Wigington, C. Xiao, and T. Sun, "StandARone: Infrared-watermarked documents as portable containers of AR interaction and personalization," in *Proc. Extended Abstr. CHI Conf. Human Factors Comput. Syst.*, New York, NY, USA, Apr. 2023, pp. 1–7.

[33] J. Hertel, S. Karaosmanoglu, S. Schmidt, J. Bräker, M. Semmann, and F. Steinicke, "A taxonomy of interaction techniques for immersive augmented reality based on an iterative literature review," in *Proc. IEEE Int. Symp. Mixed Augmented Reality (ISMAR)*, Oct. 2021, pp. 431–440.

[34] K. C. Madathil and J. S. Greenstein, "An investigation of the efficacy of collaborative virtual reality systems for moderated remote usability testing," *Appl. Ergonom.*, vol. 65, pp. 501–514, Nov. 2017.

[35] K.-D. Le, M. Fjeld, A. Alavi, and A. Kunz, "Immersive environment for distributed creative collaboration," in *Proc. 23rd ACM Symp. Virtual Reality Softw. Technol.*, New York, NY, USA, Nov. 2017, p. 16.

[36] T. Duval and C. Fleury, "An asymmetric 2D pointer/3D ray for 3D interaction within collaborative virtual environments," in *Proc. 14th Int. Conf. 3D Web Technol.*, New York, NY, USA, Jun. 2009, pp. 33–41.

[37] H. Ibayashi, Y. Sugiura, D. Sakamoto, N. Miyata, M. Tada, T. Okuma, T. Kurata, M. Mochimaru, and T. Igarashi, "Dollhouse VR: A multi-view, multi-user collaborative design workspace with VR technology," in *Proc. SIGGRAPH Asia Posters*, New York, NY, USA, Nov. 2015, p. 8.

[38] P. Jansen, F. Fischbach, J. Gugenheimer, E. Stemasov, J. Frommel, and E. Rukzio, "ShARe: Enabling co-located asymmetric multi-user interaction for augmented reality head-mounted displays," in *Proc. 33rd Annu. ACM Symp. User Interface Softw. Technol.*, New York, NY, USA, Oct. 2020, pp. 459–471.

[39] N. Feld and B. Weyers, "Mixed reality in asymmetric collaborative environments: A research prototype for virtual city tours," in *Proc. IEEE Conf. Virtual Reality 3D User Interfaces Abstr. Workshops (VRW)*, Mar. 2021, pp. 250–256.

[40] S. Karaosmanoglu, K. Rogers, D. Wolf, E. Rukzio, F. Steinicke, and L. E. Nacke, "Feels like team spirit: Biometric and strategic interdependence in asymmetric multiplayer VR games," in *Proc. CHI Conf. Human Factors Comput. Syst.*, New York, NY, USA, May 2021, pp. 1–15.

[41] Y. Li, E. Ch'ng, S. Cobb, and S. See, "Presence and communication in hybrid virtual and augmented reality environments," *PRESENCE, Virtual Augmented Reality*, vol. 28, pp. 29–52, Dec. 2021.

[42] A. Lund, "Measuring usability with the use questionnaire," *Usability Interface*, vol. 8, no. 2, pp. 3–6, 2001.

[43] *Oculus*, Facebook Technol., LLC, Menlo Park, CA, USA, 2021.

[44] *Google VR SDK for Unity*, Google Developers, Google, Mountain View, CA, USA, 2019.

[45] *VUFORIA Engine*, PTC, Boston, MA, USA, 2021.

[46] *Unity Engine*, Unity Technologies, San Francisco, CA, USA, 2019.

[47] *Photon Unity Networking*, Photon, Exit Games GmbH, Hamburg, Germany, 2021.

[48] *Optitrack*, NaturalPoint, DBA OptiTrack, Corvallis, OR, USA, 2022.

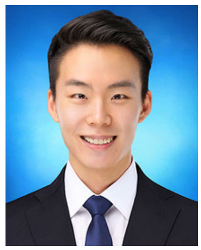
[49] *Village Palace, People, Soldier, Interior*, TRIPLE, Incheon, South Korea, 2021.

[50] A. Yassien, M. A. Soliman, and S. Abdennadher, "QuarantivityVR: Supporting self-embodiment for non-HMD users in asymmetric social VR games," *I-COM*, vol. 21, no. 1, pp. 55–70, Apr. 2022.

[51] T. Drey, P. Albus, S. Der Kinderen, M. Milo, T. Segschneider, L. Chanzab, M. Rietzler, T. Seufert, and E. Rukzio, "Towards collaborative learning in virtual reality: A comparison of co-located symmetric and asymmetric pair-learning," in *Proc. CHI Conf. Human Factors Comput. Syst.* New York, NY, USA: Association for Computing Machinery, Apr. 2022, pp. 1–19.

[52] B. G. Witmer, C. J. Jerome, and M. J. Singer, "The factor structure of the presence questionnaire," *Presence*, vol. 14, no. 3, pp. 298–312, Jun. 2005.

[53] J. Lee, M. Kim, and J. Kim, "A study on immersion and VR sickness in walking interaction for immersive virtual reality applications," *Symmetry*, vol. 9, no. 5, p. 78, May 2017.



YUNSIK CHO received the B.E. degree in computer engineering from Hansung University, South Korea, in 2021, and the M.S. degree from the Department of Computer Engineering, Hansung University, in 2022, where he is currently pursuing the Ph.D. degree with the Department of Information and Computer Engineering. His research interests include virtual reality, augmented reality, HCI, and computer graphics application.

MYEONGSEOK PARK received the B.E. degree in computer engineering from Hansung University, South Korea, in 2022, where he is currently pursuing the master's degree with the Department of Computer Engineering. His research interests include virtual reality, augmented reality, HCI, and computer graphics application.

JINMO KIM received the B.E. degree in multimedia engineering from Dongguk University, in 2006, and the M.E. and Ph.D. degrees in multimedia from Dongguk University, in 2008 and 2012, respectively. He is currently an Associate Professor with the Division of Computer Engineering, Hansung University, South Korea. He was an Assistant Professor with the Department of Software, Catholic University of Pusan, South Korea, from March 2014 to February 2019. His primary research interests include virtual reality, augmented reality, computer graphics, metaverse, game engineering, and related applications.