RESEARCH

Open Access

An empirical study of virtual museum based on dual-mode mixed visualization: the Sanxingdui bronzes

Feng Tian[1], Keying Li[1], Xin Huang[1], Xiaoyue Zhang[1], Ning Wang[2], Ying Song[1], Qiaolian Zhu[3*] and Yuzhi Li[1*†]

Abstract

The design and implementation of virtual museums have evolved over decades, broadening the scope of traditional museum exhibitions and serving as an effective complement to physical museums. However, virtual visual effects have consistently fallen short of replicating the authentic exhibition experience and cannot substitute for physical museums. In this paper, we report a dual-mode mixed visualization system combining computational photography and binocular stereo vision techniques, and introduce an innovative virtual museum for Sanxingdui bronzes. This approach enables visitors to engage with the artifacts in a multifaceted manner-allowing for navigation, interaction, auditory guidance, and the examination of textural details as if handling the artifacts directly. Further analysis of participant responses in terms of visual cognition and aesthetic appreciation across different settings demonstrates that this dual-mode mixed visualization delivers not only a spatiotemporal experience consistent with the real world but also surreal visual effects. This immersive experience allows audiences to experience Sanxingdui art intimately, offering both high-resolution visuals and cost-effectiveness.

Keywords Virtual reality, Mixed visualization, Virtual museum, Binocular vision

Introduction

The Sanxingdui site, with 3000 to 5000 years of history, stands as the most extensive, enduring, and culturally rich ancient city, state, and repository of Shu culture discovered to date in China's southwestern region. Designated as one of the greatest archaeological discoveries of the 20th century, Sanxingdui, together with the Jinsha site, was nominated for UNESCO World Heritage status in 2021. The artifacts of Sanxingdui reveal traces of ancient Chinese civilizations—familiar yet unique, overflowing with boundless imagination. They stand as tangible evidence of the diversity, unity, openness and inclusiveness of Chinese culture.

In physical museums, artifacts are confined to fixed spaces, preventing visitors from closely interacting with or appreciating intricate textures [1]. In Addition, since artifact series, such as the great Yu *ding* and great Ke *ding*, are often dispersed across various physical museums, a virtual museum could unite and present valuable artifacts of the same category. As many artifacts face permanent restrictions on international exhibition, virtual museums offer a new avenue for education and dissemination, facilitating a greater understanding of cultural heritage and exchanges between different civilizations.

Moreover, given China's vast territory, schools in remote areas lack access to provincial museums,

---

†Yuzhi Li: Primary corresponding author.

*Correspondence:
Qiaolian Zhu
gjky@shu.edu.cn
Yuzhi Li
shadowmcv@shu.edu.cn
[1] Shanghai Film Academy, Shanghai University, 149 Yanchang Road, Shanghai 200072, China
[2] Shanghai Academy of Fine Arts, Shanghai University, 149 Yanchang Road, Shanghai 200072, China
[3] School of Foreign Languages, Shanghai University, 99 Shangda Road, Shanghai 200444, China

Tian *et al. Heritage Science*   (2024) 12:146

Page 2 of 14

necessitating affordable virtual alternatives. Even visiting famed institutions like the Louvre, British Museum or Hermitage Museum is prohibitively expensive, highlighting the need for comprehensive virtual museums.

This study, selecting Sanxingdui artifacts as samples, has crafted a free, lifelike, mobile, visually enhanced, physically independent, engaging, and educational virtual museum. We blur the line between real and virtual items, allowing exploration of hyper-resolution 3D texture details.

Hyper-resolution refers to the ability to display texture details of 3D models at higher resolutions than what would typically appear blurred when zooming in on traditional 3D texture maps in virtual environments. By utilizing binocular stereo vision techniques, we can provide sharper and clearer texture details upon close inspection.

Through multidimensional comparison of seven exhibition formats—Physical Museum, Heritage Artifacts Photograph, Video, Mobile Application, Mobile Augmented Reality (AR), Mixed Reality (MR), and Binocular Stereo Vision (BVS)—plus additional experiments on visual cognition and aesthetic experience, we ultimately developed a dual-mode hybrid visualization virtual museum. This extends major museums from urban to rural communities both locally and globally, spreading Sanxingdui's rich cultural heritage to wider audience.

Literature review
Virtual museums
Traditional museums inherently face multiple challenges, including spatial limitations, inventory constraints, low management efficiency, restricted displays, low artifact sharing rates, and high exhibition costs [2]. Tsichritzis et al. [3] was among the first to propose the concept of virtual museums to overcome the shortcomings of physical exhibitions and provide remote visitors with a vivid experience. Mafkereseb et al. [4] conducted an early comprehensive review of AR/VR/MR technologies.

Most museums now have digital websites, offering high-quality images, text, and 360-degree venue tours [5, 6]. Google Arts & Culture, a collaboration between Google and museums worldwide utilizing Street View technology, stands out for capturing ultra-high resolution images of historic paintings for global online appreciation.

In the past decade, 3D scanning and visualization have been gradually adopted for desktop virtual museum applications [7–9]. These applications employ conventional computer screens with keyboards and mice for viewing and interacting with virtual museums.

Notable examples include Kiourt's innovative dynamic web-based framework [10], Barbieri's touchscreen user interaction system for 3D artifacts [11], Jonauskaite's exploration of interactive discovery and aesthetic evaluation [12]. Desktop applications are also utilized in other museum contexts [13, 14], such as science museums. While common, desktop applications lack a vivid, immersive experience.

Beyond desktop applications, three-dimensional models are often employed in AR formats, including mobile AR and head-mounted AR (e.g., Microsoft HoloLens). Wu et al. conducted a study on the educational effects of mobile AR for on-site cultural heritage learning and off-site environments, and found that AR exhibitions improved learning outcomes (including motivation and the effectiveness of learning local history) [15]. Ch'ng et al. investigated how social interaction functions within museum AR contexts [16]. There is extensive research on head-mounted AR, and although MR devices lack a clear definition, head-mounted AR is often categorized as an MR form. O'Dwyer et al. used volumetric video technology to create digital guides, enhancing museum visits with HoloLens AR in an engaging way [17]. Chen et al. conducted a significant comparative study on the learning effects and motivation of head-mounted AR, finding superior performance with AR glasses learning strategies in science museums [18]. Hammady et al. developed an alternative guide system using head-mounted AR to enhance customer experiences and reduce the need for human guides in museums [19, 20]. Aok et al. introduced a helmet-based AR guide system for exhibitions, adding digital commentary and demonstrations to enrich museum visits, enabling visitors to gain deeper insights into exhibits and enjoy a pleasurable experience [21].

In parallel to head-mounted AR, head-mounted VR frequently finds application in virtual museum research. Verhulst et al., by comparing user experiences of museum VR and AR, found that VR scored higher in enjoyment, cognition, and emotional engagement [22]. Wu et al. designed a head-mounted VR virtual museum, enabling visitors to interact with exhibits, access multimedia information, and even ring a bell [23]. Rahimi et al. demonstrated that the use of head-mounted VR technology offers a novel museum experience, exerting a greater impact on learning and enjoyment for audiences [24]. Kim et al. proposed a multisensory digital cultural heritage platform based on head-mounted VR (HMD), proving that novel multisensory experiences can enhance immersion and preference [25].

Mixed Reality (MR) represents the integration of VR, AR and the real environment, thereby creating a blend of real and virtual worlds [4], it combines the advantages

Tian *et al. Heritage Science*   (2024) 12:146

Page 3 of 14

of both VR and AR while redefining the reality-virtual continuum in a unique spatiotemporal environment, reshaping the physical, social, and symbolic spaces [26].

In the MuseumEye project [18, 19], MR technology was utilized within a museum context. Margetis et al. [27] proposed combining augmented reality, virtual reality, and mixed reality technologies to provide a unified X-Reality experience in realistic virtual museums, enabling visitors to partake in the interaction and seamless fusion of the physical and virtual realms.

However, the MR discussed above essentially equates to head-mounted AR technology. AR glasses, like the Microsoft HoloLens, offer a limited field of view—only 30 degrees horizontally and 17.5 degrees vertically—making it easy to lose sight of virtual artifacts. The HoloLens device is relatively heavy (579 g), imposing a significant burden on users [28]. Moreover, semi-transparent AR imagery proves difficult to discern outdoors, limiting applications while the near-$3000 costs hamper widespread adoption.

Despite numerous studies on virtual museums, the vast majority of audiences still opt for physical visits to museums. Why haven't virtual museums partially replaced the functions of traditional physical museums? Wang [29] posited that virtual museums, as a specialized manifestation of physical museums, can never substitute traditional physical museums, regardless of their level of advancement. Instead, they can only serve as a beneficial and necessary complement. Currently, major museums frequently present immersive virtual projections of cultural artifacts, typically utilizing panoramic videos, some desktop platforms on the Internet also offer 3D virtual exhibitions. However, these displays lack interactivity. Practical observations indicate that large museums seldom employ VR headsets to showcase cultural artifacts. Two-dimensional displays eliminate the spatial-temporal behavior associated with traditional viewing, head-mounted VR restricts audience movement, and the virtual visual effects of head-mounted AR fall far short of expectations. Mobile AR represents a commendable compromise but lacks a sense of depth.

Binocular stereovision
Binocular stereovision is an essential component of computer vision [30]. The distance between the two pupils of the human eye is approximately 65 mms. Consequently, when our eyes observe the same target within a certain distance range, the angles differ. When the same target image is projected onto the two retinas, subtle differences arise, known as binocular disparity [31]. Binocular stereovision can effectively simulate the human eye and obtain depth information of objects. The main process involves using two imaging sensor devices to capture the same object from different positions, thereby obtaining two images of the object to be measured. Subsequently, by analyzing the differences between the two images, the position deviations of the feature points in the two images are obtained. Finally, the three-dimensional geometric information of the object is calculated using the disparity principle [32]. The head-mounted display (HMD) places two displays inside the helmet, corresponding to the user's left and right eyes, respectively. When the user wears the VR headset, the left and right eyes perceive different images, creating a stereoscopic visual effect. In this study, we employ this principle to capture images with binocular disparity to simulate human stereoscopic vision, rendering flat images three-dimensional.

Presence in virtual environments
Lombar et. al [33] summarized six related concepts of presence, one of which emphasizes perceptual and psychological immersion. Immersion in a virtual reality environment is a complex psychological phenomenon characterized by the individual's perception of close interaction with the virtual environment and constantly changing stimuli within it. A deeply immersive virtual reality environment often elicits a stronger sense of presence. When users experience a sense of "being there," they become immersed [34]. Factors influencing immersion include, but are not limited to, the degree of isolation from the real world, the intensity of self-presence in the virtual environment, the naturalness and control mode of the interaction process, and the user's perception and experience of their own movement [35]. Individuals' sense of presence is strongest in real space, as our awareness of the surrounding environment inevitably depends on the data obtained through our sensory systems: vision, sound, touch, force, taste, and smell [36].

To achieve the strongest immersion and sense of presence, the virtual environment needs to maintain sensory consistency, a consistent flow of time, consistent proportions of three-dimensional space, and consistent spatial motion, faithfully replicating the four-dimensional space-time of reality. Some scholars have sought to enhance participants' sense of engagement and immersion in virtual reality by studying different motion patterns and interaction modes [37–40]; others have focused on exploring how to enhance participants' immersion in virtual environments by adding haptic feedback [41–44].

Materials and methods
Data acquisition, processing, and modeling
This study employed experimental methods to gauge participants' responses to the interactivity, visual

Tian *et al. Heritage Science*   (2024) 12:146

Page 4 of 14

effects, enjoyment, and aesthetics (dependent variables) of selected environments (independent variables). To control the variables, we implemented seven practical application scenarios: Physical Museum, Heritage Artifacts Photograph, Video Animation, Mobile Application, Mobile Augmented Reality (Mobile AR), Mixed Reality (MR), and Binocular Stereo Vision (BVS).

Within our platform, digital materials included flat artifact photographs, binocular stereo images, 3D models, and vocal narrations.

We first compiled textual records of Sanxingdui artifacts for vocal narration production. Next, diverging from prior laser scanning techniques in [23], following the method of [45], we employed photogrammetry to capture point cloud data. It is a cost-effective technique for obtaining dense 3D geometric data of physical objects from overlapping stereoscopic photographs [46], which typically uses an ordinary digital (static) camera, and then processes the data using 3D software to achieve detailed 3D reconstruction of the scanned object [47]. This involves photographing artifacts, calculating camera positions, and generating point clouds. The advantages include (1) cost-effectiveness without expensive laser scanners; (2) safety from artifact laser exposure; and (3) precision in creating accurate 3D models. We processed the data through filtering, alignment and surface reconstruction in Maya.

Additionally, as depicted in Fig. 1, we obtained binocular stereo photographs through dual-camera positioning, presented on stereoscopic display devices, such as VR headsets. Limited to 100,000 facets, 3D models cannot match stereo image precision. Texture resolution reached 8K full angle (1.5K viewing) in Unity, requiring lighting versus 16K (8K each eye) resolution frontally for binocular stereo photographs. Thus stereo images provide far superior frontal resolution, precision, and color accuracy versus 3D models. As Figs. 2 and 3 show, both formats provide insights into spatial structures, architectures and texture colors.

Architecture design

As Fig. 4 illustrates, the system architecture encompasses mobile augmented reality (AR) and mixed reality (MR) displaying 3D models. Mobile AR comprises AR Foundation, user interface, mobile interactions, etc., while MR leverages level of detail (LOD), XR interactions, and more. Binocular stereo vision consists of stereo cameras, a photo interaction system, a virtual museum environment and additional modules.

Virtual artifacts are organized into two Levels of Detail (LOD) based on the complexity of the artifact models. Low-level LOD models contain between 10,000 to 30,000 triangles, while high-level LOD models range from 50,000 to 100,000 triangles. When viewers engage with or control a specific virtual artifact, it is presented in high LOD mode to render the artifact in greater detail; conversely, low LOD mode is utilized for a less detailed presentation.

The system supports bi-manual and gesture recognition, allowing viewers to use their own hands to drive virtual hands for grasping and rotating artifacts, thus triggering voice narrations and textual prompts alongside the artifact manipulation. Similar to the electronic guides found in museums, the ultimate virtual museum equips each artifact with pop-up subtitles and corresponding narrative voices. When viewers activate a specific virtual artifact, cues such as lighting and sound effects are triggered, accompanied by textual displays and voice narrations.

Experiment design

As Fig. 5 shows, the seven experimental environments comprised: the physical museum (C1) hosted the "Bronze Light" special exhibition of Sanxingdui. Artifact photographs (**C2**) showcased photographs taken

Fig. 1 Stereoscopic work portrait photography (special effects team admitted with work credentials to the "Bronze Light" exhibition in Shanghai)

Tian *et al. Heritage Science*　　(2024) 12:146

Page 5 of 14

Fig. 2 Handcrafted 3D artifact models (material sourced from the "Bronze Light" exhibition in Shanghai, first row: physical artifact photographs; second row: 3D reconstructed white models; third row: 3D textured models; fourth row: stereoscopic photographs)
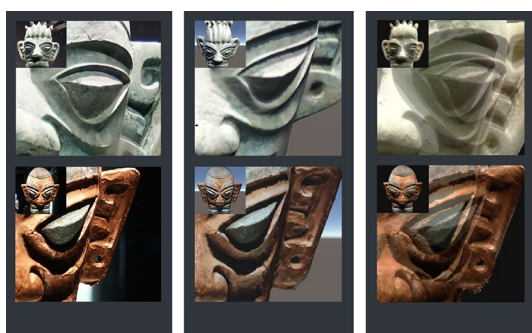

Fig. 3 Contrast of detail textures and colors between models and photographs (first column: real scene photographs constituting stereovision; second column: 3D model visual effects in unity; third column: stereoscopic disparity constituting stereovision)

on-site to the audience. Video (C3) involved playing three-dimensional animation videos on smartphones. The mobile application (C4) was a specially developed interactive Sanxingdui exhibition software that supports the rotation and scaling of three-dimensional models. Mobile AR (C5), developed alongside C4, is an augmented reality exhibition application that supports mobile positioning and full-angle planar browsing, as

shown in Fig. 6. MR (C6) utilized the reality-perspective function of virtual reality headsets for mixed reality exhibitions, while BVS (C7) replaced three-dimensional models with binocular stereo photographs, with the photograph orientation synchronizing with changes in the viewer's perspective, as illustrated in Fig. 7.

Furthermore, C2 to C7 maintain consistency with the artifact displays in C1, including the latest "Bronze Figure with a Gold Mask," "Bronze Standing Figure with a Skirt," and "Bronze Standing Figure Holding a Bird" among other renowned artifacts. C4 to C7 all support interactive triggering of voice narrations, with the narration words being entirely consistent with the voice narrations provide by QR codes in the real Sanxingdui Museum. C6 supports gesture operations. Both C5 and C6 enable mobile viewing experiences.

Our dual-mode hybrid approach combines C6 and C7, equipped with three Pico 3 VR headsets, two Pico 4 Android headsets and two Android phones. We implemented the system in Unity 2021.3.15.

This study was reviewed and approved by the Ethics Committee of Shanghai University (ECSHU 2024-007).

Given the study's comparison with physical museum exhibitions in the Shanghai University Museum, a two-phase experiment was designed to assess the audience's
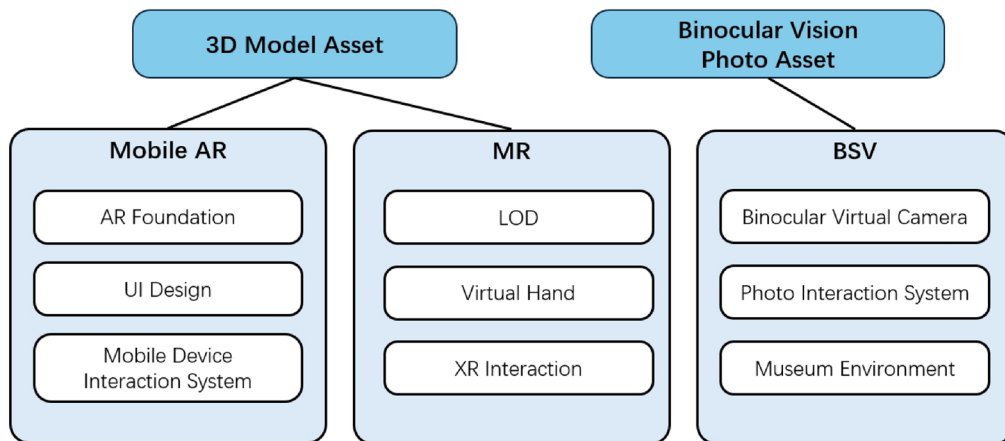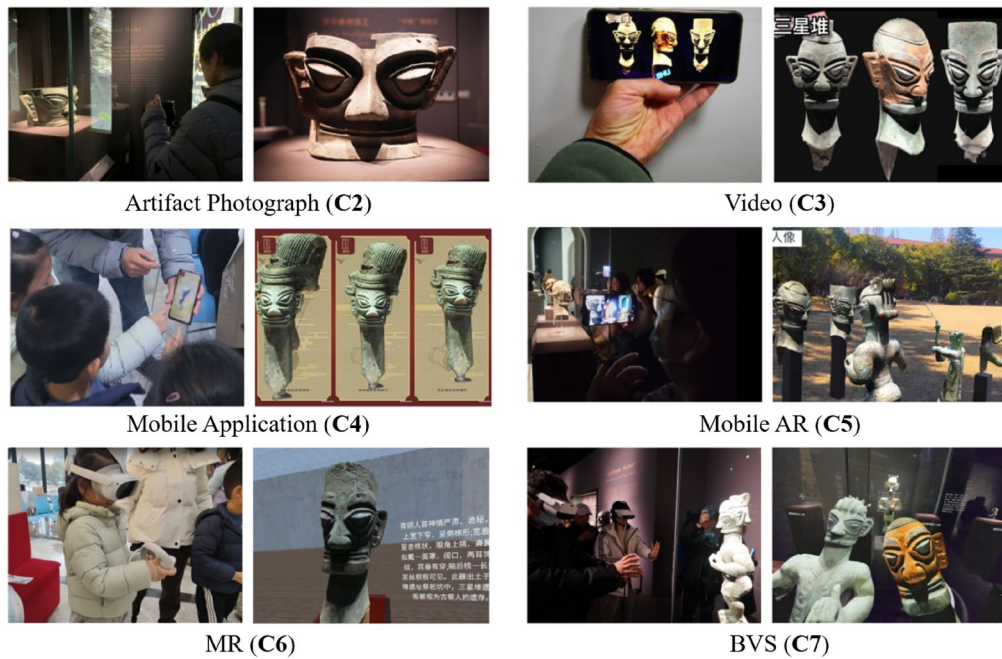
Fig. 4 The framework of our system

Fig. 5 The presentation effects of museums in various modes

appreciation and enjoyment under various digital visual conditions. Conducted in January 2024, the first group consisted of a questionnaire experience group, and the second group comprised a semi-open experience group.

The first group of experiment subjects included 19 faculty and students from our institution (7 males and 12 females) aged between 21 and 31 years old (M = 23.76, SD = 1.97). All participants had not visited the physical Sanxingdui artifacts before the experiment.

After introducing and clarifying the purpose of the study, consent forms were provided to participants, who were informed to first visit the "Bronze Light" special exhibition and then return to the on-campus experimental site to participate in six additional virtual exhibition tasks. The order of the six exhibition experiences was at the participants' discretion, with onsite guidance on how to interact within the AR/ MR spaces. For C2, participants were instructed to view each photograph and label information of the
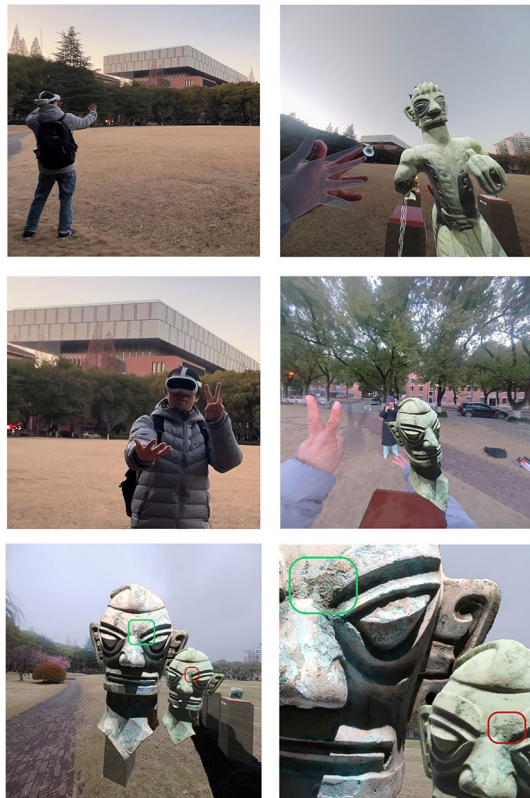
Fig. 6 The experiment in mobile AR

Fig. 7 The experiment in MR and BVS. first row and second row: the effect in MR; third row: detail comparison of MR (red) and BVS (green), the resolution of BVS is higher than that of MR

Sanxingdui artifacts. For C3, they were told to watch the complete Sanxingdui artifact animation videos. For C4, participants were prompted to swipe on the mobile platform. For C5, they were informed they could walk around freely with the mobile platform to observe

virtual artifacts. For C6 and C7, participants were told they could walk freely, grasp artifacts, and admire them. Upon completing the experiment, participants were asked to fill out a questionnaire and rank the seven types of interactions and visual effects.

The questionnaire was inspired by two seminal pieces of literature focused on virtual reality guidance [22, 48], encompassing five dimensions:

(1) Satisfaction: What is the audience's level of satisfaction with the Sanxingdui artifacts across the seven modes?
(2) Perceived Effectiveness of Visual Cognition: Can digital technologies assist audiences in comprehending the texture, color, and structure of Sanxingdui artifacts?
(3) Cognitive Load: How do audiences perceive the difficulty of manipulating digital devices?
(4) Flow Experience: What is the level of concentration exhibited by audiences while appreciating artifacts across the seven modes?
(5) Interaction: Do audiences feel a sense of interaction and participation?

The questionnaire also explored the audience's aesthetic pleasure and empathy towards the Sanxingdui artifacts in the seven modes, covering seven aspects:

(1) Perception: Are the Sanxingdui artifacts significant, relevant, and intriguing to you?
(2) Resting State: Do you feel as though the statues and masks are gazing at you? Do you find yourself staring at the Sanxingdui artifacts for extended periods?
(3) Imagination: Do you sense the mystic atmosphere of ancient Shu and the grandeur of sacrificial scenes emanating from the artifacts?
(4) Association: Do you perceive the ancient pursuit of beauty and creative wisdom through the artifacts?
(5) Understanding: Do you believe you draw inspiration from the designs and decorations of the Sanxingdui artifacts, thereby deepening your appreciation of beauty?
(6) Philosophy: Have you experienced the harmonious concept of "unity between heaven and humanity, the coexistence of all beings"?
(7) Empathy: Would you be willing to travel back to ancient Shu to participate in sacrificial rituals, the sun-shooting challenge, and mask collection activities?

The second group of experimental subjects consisted of 22 members of the public, mostly families visiting

Tian *et al. Heritage Science*   (2024) 12:146

Page 8 of 14

the "Bronze Light" special exhibition. After introducing and clarifying the purpose of the study, consent was verbally obtained from participants who were already visiting the "Bronze Light" exhibition. The exhibition tasks were conducted within the museum galleries and lounges. Similar to the first group, public viewers randomly experienced the other six types of virtual exhibitions. At the end of Experiment One, instead of completing a questionnaire, participants discussed and ranked the seven types of interactions and visual effects, followed by detailed and lively semi-structured interviews, which facilitated a collective discussion to clarify the audience's familiarity with interaction, the naturalness of interaction, cognitive recognition of visual effects, and ranking of aesthetic pleasure.

Result

The survey results indicated that for all participants, it was their inaugural visit to view the Sanxingdui artifacts in person. The audience was unanimously captivated by the visual effects of the artifacts on display. Overall, the participants described MR and BVS as most closely replicating the visual effects and aesthetics.
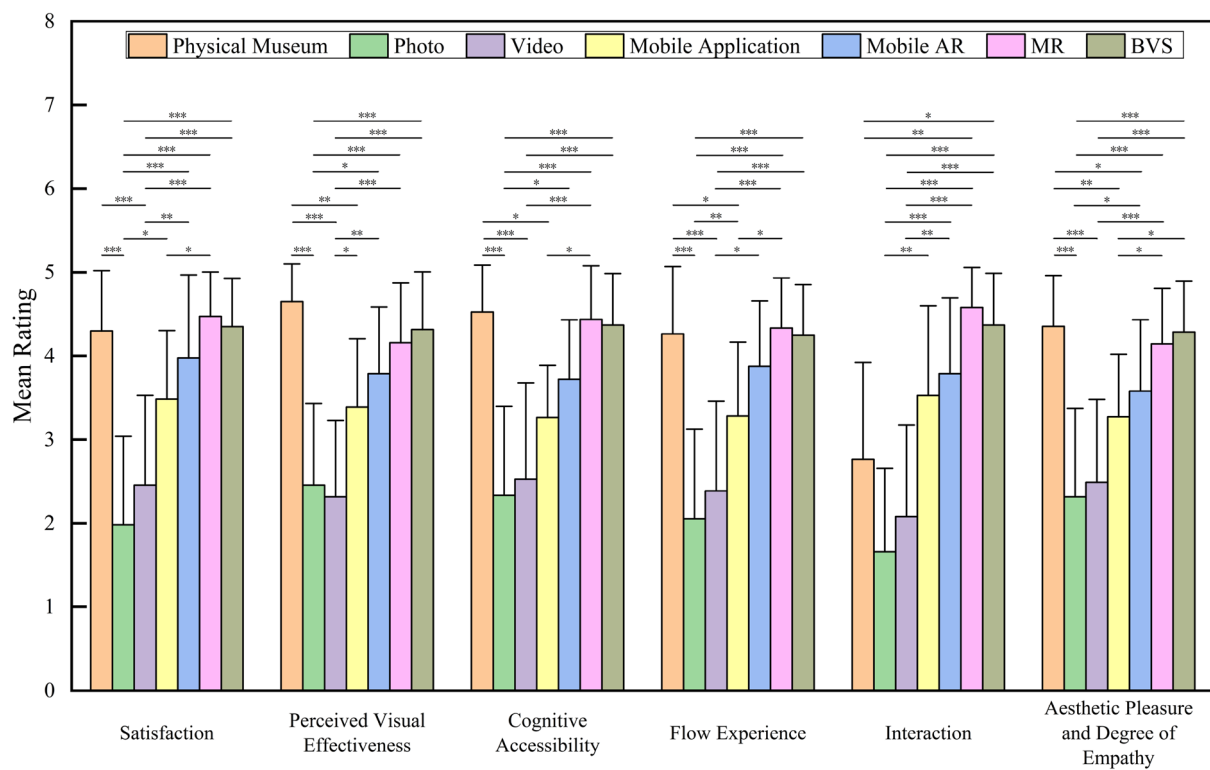

Fig. 8 The results of the first group experiment, the asterisk denotes significant differences observed after conducting the Friedman test (*0.01 < p < 0.05, **0.001 < p <0.01, ***p < 0.001)

Table 1 The mean scores and standard deviations of results for the first experimental group

	C1		C2		C3		C4		C5		C6		C7	
	Mean	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**
Satisfaction	4.30	0.72	1.98	1.06	2.46	1.07	3.48	0.82	3.97	1.00	4.47	0.53	4.35	0.57
Perceived visual effectiveness	4.65	0.45	2.46	0.98	2.32	0.91	3.39	0.82	3.79	0.80	4.16	0.71	4.32	0.69
Cognitive accessibility	4.53	0.56	2.33	1.06	2.53	1.15	3.26	0.62	3.72	0.71	4.44	0.64	4.37	0.62
Flow experience	4.26	0.81	2.05	1.07	2.39	1.07	3.28	0.88	3.88	0.78	4.33	0.60	4.25	0.61
Interaction	2.76	1.16	1.66	1.00	2.08	1.10	3.53	1.07	3.79	0.90	4.58	0.48	4.37	0.62
Aesthetic pleasure and degree of empathy	4.35	0.61	2.32	1.05	2.49	0.99	3.27	0.75	3.58	0.86	4.14	0.66	4.29	0.61

Cognition and interaction

As Table 1 and Fig. 8 show, phase one employed five-point scales where 1 is complete dissatisfaction and 5 is complete satisfaction. Satisfaction order is: C2 < C3 < C4 < C5 < C1 < C7 < C6. The effectiveness of visual cognition order is: C3 < C2 < C4 < C5 < C6 < C7 < C1. Accessibility of cognition order is: C2 < C3 < C4 < C5 < C7 < C6 < C1. Flow experience order is: C2 < C3 < C4 < C5 < C7 < C1 < C6. Interaction order is C2 < C3 < C1 < C4 < C5 < C7 < C1.

Through detailed and passionate semi-structured interviews in the second group experiment, a consensus was reached. The familiarity with interaction is ranked as: C6 < C5 < C4 (C1, C2, C3, C7 had no digital interaction); naturalness of interaction is ranked as: C4 < C5 < C6.

Preschool children (younger than 7 years old) in the second group particularly enjoyed the interactive applications, mobile augmented reality, and mixed reality visual presentations. Through inquiries, children were able to clearly identify that the images in C4 and C5 were identical to the cultural artifacts they had just observed at the museum. Although preschool children have relatively smaller head circumferences and interpupillary distances, observations revealed that in comparison to C4 and C5, they spent more time engaging with C6 and C7, with no instances of unsuitable usage observed for the latter two conditions.

One parent in the second group described, "Preschool children are not particularly interested in the historical and cultural background of the artifacts, but they are universally attracted to visual content that is participatory, interactive and provides feedback."

Another parent from the second group commented, "I found artifact photograph to be quite ordinary, and something that most individuals would undertake themselves. I would share the photographs on social media to garner more attention. Videos permitted me to view the cultural artifacts from multiple angles, rendering it more vivid than photos. I am also curious about the process through which these models were scanned. If I could obtain these animations, I would share them on social media as well, as I believe they would receive more positive engagement. Mobile Application allowed me to freely choose the viewing angle, which I found more appealing to the audience than the video. I believe this application could be released on the app store for a broader audience to download and appreciate. Mobile application is similar to digital collectibles, but I would not purchase them."

The third parent in the second group suggested, "To attract audiences, I believe more experiential content could be added to the interactive applications, mobile augmented reality, and mixed reality in terms of interactivity."

The fourth parent suggested, "To attract audiences, I think more experiential content could be incorporated into the interactivity of Mobile Application, AR, and MR, such as enabling me to wear a golden mask and immerse myself in the Sanxingdui sacrificial rituals."

The fifth visitor in the second group shared, "In mixed reality, bringing a figure closer to view reveals the triangular facets composing the virtual model, yet I cannot imagine how the realistic detail effects in binocular stereo vision are created."

The sixth visitor in the second group recounted, "Compared to mixed reality and binocular stereo vision, I am more familiar with interactive applications and mobile augmented reality, finding them simpler and easier to use, though lacking in stereoscopic effects. Under mixed reality conditions, it felt as though the glass barrier was broken, allowing me to get closer to the artifacts and even pick up virtual masks. Binocular stereo vision allowed me to discern the texture details of artifacts more clearly, a resolution of flat details I could only achieve on-site by zooming in with my smartphone camera, providing both hyper-resolved details and stereoscopic vision, surpassing the visual detail effects I get from statically viewing real masks upfront."

Aesthetic

The order of aesthetic perception for the first group of subjects is: C2 < C3 < C4 < C5 < C6 < C7 < C1; aesthetic resting state order is: C2 < C3 < C4 < C5 < C6 < C1 < C7; aesthetic imagination order is: C2 < C3 < C4 < C5 < C6 < C7 < C1; aesthetic association order is: C2 < C3 < C4 < C5 < C6 < C7 < C1; aesthetic understanding order is: C2 < C3 < C4 < C5 < C6 < C7 < C1; aesthetic philosophy order is: C3 < C2 < C4 < C5 < C1 < C6 < C7; and aesthetic empathy order is: C2 < C3 < C4 < C5 < C1 < C6 < C7.

The consensus among subjects in the second group on cognitive recognition of visual effects and aesthetic pleasure follows the order: C2< C3< C4< C5< C6< C7< C1.

One visitor from the second group expressed, "Although the shapes and colors of artifacts can be reproduced, Photos, Videos and Mobile Application are relatively common formats, and AR also lacks spatial aesthetics, failing to match the actual size of the cultural artifact one-to-one.

In MR, I was able to freely appreciate the figure from all distances and angles–near, far, front, and back–mirroring the viewing experience of a real museum, invoking new perceptions of spatial and distance beauty. Coupled with the high-resolution presentation of BVS, it perfectly captured the static and dynamic beauty of the sculpture."

A parent from the second group described, "The bronze figures in the real museum stood silently, where I saw the mysterious bronze masks and the splendid golden masks. Employing MR to present virtual artifacts within the real museum momentarily blurred the lines between real and virtual figures for me. BVS offered a hyper-resolution mask effect; the mask gazed at me, and I gazed back, creating a 'face-to-face' cultural experience where time seemed to stand still, bridging a 3000-year-old mutual gaze as a form of temporal beauty."

Another audience member recounted, "The style of the bronze figure's decorations was both diversely rich and harmoniously unified, with mask lines that were taut and appropriately curved, conveying a profound and dignified effect throughout. MR gathered the bronze figures in a compact virtual space, showcasing their multifaceted beauty, regular beauty, and the beauty of strangeness more effectively than the real museum could. The solemn expressions of the bronze figures synchronized with my facial expressions, as if I transformed into one of them, ascending to divinity from 3000 years ago, achieving a shared realm of beauty."

Discussion
Pleasures of the experience
In terms of visual effect restitution, C4 and C5 were still based on flat displays, with three-dimensional models rendered and projected on two-dimensional screens. C6 reconstructed three-dimensional artifacts, presenting them to the audience in a one-to-one, face-to-face manner, at low cost and with high resolution, maximizing the audience's acquisition of an almost authentic viewing experience. C7 enhanced the resolution of the reconstructed models, achieving ultra-high-definition presentation. The dual-mode hybrid visualization combining C6 and C7 provided a visual effect that is both virtual and surpasses reality.

Regarding interactivity, C4 enabled touch operations, and C5 facilitated mobile browsing, both offering voice-guided tours and detail magnification of artifacts. C6 allowed audiences complete control over visual space selection, freely choosing the spatial location, viewpoint, and experience distance for observation and learning, with the added functionality of handheld artifact appreciation. [23] and others employ handheld controls for interaction, while the dual-mode hybrid visualization utilizes natural gesture interactions, significantly enhancing usability.

Cognitive and interactive experience
The results of the questionnaire from Sect. "Cognition and interaction" suggest that the Physical Museum (C1), MR (C6), and BVS (C7) modes outperformed the other modes in terms of satisfaction, visual cognition effectiveness, cognitive accessibility, flow experience,

and interaction. Particularly, MR and BVS scored the highest in satisfaction, interaction, and flow experience, indicating that these two modes provided the most engaging and immersive experiences for the participants. The consensus was that MR and Mobile AR (C5) offered the highest familiarity and naturalness of interaction among the digital modes, as they allowed for more intuitive and direct interactions with the virtual artifacts.

These highlight the advantages of employing advanced technologies in virtual museum settings. By offering more natural and immersive interactions, as well as enhanced visual fidelity, these modes can potentially provide a more engaging and enriching experience for visitors, closely replicating or even surpassing the experience of visiting a physical museum.

However, it's worth noting that the more traditional modes, such as Photographs (C2) and Video (C3), scored relatively lower in most aspects. This suggests that while these modes can serve as supplementary tools, they may not be sufficient in delivering a comprehensive and engaging virtual museum experience on their own.

Aesthetic experience
As aesthetic experience is the core focus, we further analyzed differences across modes via Friedman testing, which is a non-parametric statistical test used to detect differences among groups when the dependent variable being measured is ordinal. The test statistic for the Friedman test is denoted as X^2, which measures the degree of difference between the groups, a larger X^2 value indicates a greater difference among the groups. The p-value represents the probability of obtaining a test statistic result, a small p-value (typically < 0.05) suggests that the observed differences among groups are unlikely to have occurred by chance.

The results indicate significant differences between C1 and C2 ($X^2 = 4.553$, $p < 0.05$), C3 ($X^2 = 4.263$, $p < 0.05$), C4 ($X^2 = 2.711$, $p < 0.05$), C5 ($X^2 = 2.263$, $p < 0.05$), but no significant differences with C6 ($X^2 = 0.526$, $p > 0.05$), C7 ($X^2 = 0.421$, $p > 0.05$). Additionally, significant differences were found between C4 and C6($X^2 = -2.184$, $p < 0.05$), C7 ($X^2 = -2.289$, $p < 0.05$). It can be concluded that the aesthetic pleasure of the dual-mode hybrid visualization combined with C6 and C7 is not significantly different from that of the physical museum C1, with their scores being essentially equivalent on average. The mean differences between C2, C3, C4, and C1, C6, C7 are significant, indicating a notable disparity. Although C5 does not show significant differences with C6 ($X^2 = -1.737$, $p > 0.05$), C7 ($X^2 = -1.842$, $p > 0.05$), its average scores are significantly lower than C1, C6, C7, suggesting that C5 is recognized by the majority of the audience for its simplicity and visual effects, second only

Tian *et al. Heritage Science*   (2024) 12:146

Page 11 of 14

to C6 and C7. In summary, in terms of aesthetic pleasure, the performance of the dual-mode hybrid visualization combining C6 and C7 is most closely aligned with that of the physical museum. Among all digital presentation forms, the dual-mode hybrid visualization is the superior solution, supporting the construction of an ultimate virtual museum.

Hardware system
The literature referenced in related research presents numerous exemplary cases in the realm of virtual museums, yet some deficiencies remain when compared to our system.

Firstly, traditional PC-based VR systems can exhibit mobility constraints. The VIVE Pro requires a cable to connect the VR headset to a PC, which limits mobility [24]. The issue of motion sickness remains unresolved during operation, necessitating an assistant to manage the cable to facilitate use. Recently, mobile VR has gained popularity; however, the boxing training systems [22, 23] employ full virtual vision, obscuring the real environment and significantly restricting free movement, increasing the risk of falls and collisions for the audience while moving.

Secondly, traditional head-mounted AR systems may have deficiencies in clarity, field of view, or other aspects. The horizontal field of view of AR glasses is only 30 degrees, with a vertical field of view of 17.5 degrees. Early HoloLens devices were relatively heavy (579 g). Moreover, the virtual images seen through AR glasses are semi-transparent and difficult to discern outdoors.

Thirdly, the cost is another factor. The Apple Vision Pro, based on OLED screens, offers high-definition imaging. Though the headset is also connected by a cable to a portable power and processing unit, its portability is similar to other mobile VRs and significantly better than PC-based VR. However, the cost of the Apple Vision Pro is exceedingly high, roughly equivalent to the cost of 20 standard mobile phones.

In contrast, our system exhibits the following advantages. Firstly, our device offers a 105-degree field of view, providing a broader range of observation unaffected by environmental lighting. The devices we use weigh only 295 g. Secondly, our system is cost-effective, with simple equipment; the VR hardware used in our study costs approximately 1/20th of the Apple Vision Pro. Only high-performance, low-cost devices are practical for widespread implementation in schools. Lastly, the dual-mode hybrid visualization employs perspective techniques to eliminate motion sickness and is unaffected by environmental lighting.

Content presentation
Presentation Volume. [23] utilized photographic techniques to reconstruct a set of the chime bells of Marquis Yi of Zeng, while [7] rebuilt more than a dozen three-dimensional artifact models, and the study in [10] employed fewer than ten three-dimensional artifact models. These academic endeavors fall significantly short of the requirements for a virtual museum. An ultimate virtual museum necessitates the display of thousands of artifacts simultaneously; such extensive artifact scanning is typically undertaken only by large museums. However, these institutions generally exhibit a portion of their artifact data in Web-3D format, not offering VR/MR presentations within the scope of ordinary schools. The majority of participants in the second group expressed a desire to experience world-renowned artifacts beyond those of Sanxingdui. Our system can integrate resources from dozens of museums, showcasing renowned artifacts such as bronzes, ceramics, jades, and stone tools. The resources in our system will be more exemplary, abundant, and unique than those of any single large museum's virtual exhibits.

Presentation Quality. [23] reconstructed a set of the chime bells of Marquis Yi of Zeng using photographic techniques, which resulted in models that significantly diverged from the actual objects in visual effects. Photographic techniques are incapable of modeling under conditions of numerous artifact aggregations, and even laser scanning is unsuitable for modeling under such conditions. It was impossible for [23] to independently disassemble the chime bells, leading to significant discrepancies in textures, lighting, shadows, and colors from the actual objects. As illustrated in Fig. 2, our study modeled individual artifacts, maintaining the utmost consistency in textures, lighting, shadows, and colors with the actual objects.

Challenges
Participants noted that the museum experience augmented by virtual reality still harbors room for enhancement. The virtual reality museum experience, while offering an immersive environment, still lacks opportunities for deeper user engagement and active participation in cultural learning activities.

To address this, we are contemplating the development of a comprehensive narrative scheme for our system, where users embark on a journey as priests traversing 3,000 years, forging a novel, experiential exhibition method. In mixed reality, additions such as a "Rain Requesting" experience (Sanxingdui sacrificial rituals), a "Sun Shooting" task (inspired by the myth

Tian *et al. Heritage Science*   (2024) 12:146

Page 12 of 14

Fig. 9 Principals (Deans) visiting the ultimate virtual museum (2023 World Chinese Conference)

of Houyi shooting the suns), and a "Building Blocks" activity (assembling portraits and masks) would allow participants to don golden masks, immersing themselves in the joy of a face-to-face encounter with ancient Shu culture. This represents a method to ignite visitors' interest in cultural heritage (Fig. 9).

Conclusion and future work

In this study, we introduce a dual-mode hybrid visualization method combining mixed reality and binocular stereo vision experiences. We unveil a portable and innovative ultimate virtual museum system, establishing a digital museum that blurs the lines between virtual and real, fact and fantasy, and extensively explore the presentation effects and aesthetic experiences across seven museum formats.

Key findings are revealed by our research. Firstly, the dual-mode hybrid visualization engendered a surreal aesthetic experience, significantly differing in aesthetic impact from other visual formats, underscoring the practicality of this system in the realm of virtual museums. The mixed reality mode preserved the appreciative format consistent with real museums, while binocular stereo vision offered high-precision stereoscopic visual effects, distinct from images, video animations, and mobile formats.

Secondly, the dual-mode hybrid visualization provided a spatiotemporal experience congruent with the real world, enabling the dissemination of Sanxingdui culture globally, allowing audiences a more portable, relaxed, and free manner to enjoy traditional arts, participate in art interaction, and explore art treasures (see Fig. 9).

Author contributions
FT designed the entire project and wrote the article; KL designed the experiment and contributed to the VR software development; XH contributed to the AR software development; XZ contributed to the experiment procedure; NW guided the visual design; YS built the 3D models; QZ, as the co-corresponding author, provided writing guidance; YL , as the primary corresponding author, supervised the project and revised the paper. All authors read and approved the final manuscript.

Availability of data and materials
The data used in this article are not available.

Declarations

Tian *et al. Heritage Science*   (2024) 12:146

Page 13 of 14

References

1. Wang D. Exploring a narrative-based framework for historical exhibits combining JanusVR with photometric stereo. Neural Comput Appl. 2018;29(5):1425–32.
2. Xian F. Virtual museum. Decoration. 2007. https://doi.org/10.16272/j.cnki.cn11-1392/j.2007.09.040.
3. Tsichritzis D, Gibbs SJ et al. Virtual museums and virtual realities. In: ICHIM, 1991; p. 17–25.
4. Bekele MK, Pierdicca R, Frontoni E, Malinverni ES, Gain J. A survey of augmented, virtual, and mixed reality for cultural heritage. J Comput Cult Herit. 2018;11(2):1–36.
5. Zhang Q, Li H. Exploration of the application of 360-degree panoramic display in virtual museums. J Natl Mus China. 2011;9:154–6.
6. Mah OBP, Yan Y, Tan JSY, Tan Y-X, Tay GQY, Chiam DJ, Wang Y-C, Dean K, Feng C-C. Generating a virtual tour for the preservation of the (in) tangible cultural heritage of Tampines Chinese temple in Singapore. J Cult Herit. 2019;39:202–11.
7. Carvajal DAL, Morita MM, Bilmes GM. Virtual museums: captured reality and 3d modeling. J Cult Herit. 2020;45:234–9.
8. Xu J, Du Q, Xue J. Digital virtual museum. In: 2020 International Conference on Virtual Reality and Visualization (ICVRV), IEEE. 2020; p. 344–5.
9. Obradović M, Mišić S, Vasiljević I, Ivetić D, Obradović R. The methodology of virtualizing sculptures and drawings: a case study of the virtual depot of the gallery of Matica Srpska. Electronics. 2023. https://doi.org/10.3390/electronics12194157.
10. Kiourt C, Koutsoudis A, Pavlidis G. Dynamus: a fully dynamic 3d virtual museum framework. J Cult Herit. 2016;22:984–91.
11. Barbieri L, Bruno F, Muzzupappa M. Virtual museum system evaluation through user studies. J Cult Herit. 2017;26:101–8.
12. Jonauskaite D, Dael N, Baboulaz L, Chevre L, Cierny I, Ducimetiere N, Fekete A, Gabioud P, Leder H, Vetterli M, et al. Interactive digital engagement with visual artworks and cultural artefacts enhances user aesthetic experiences in the laboratory and museum. Int J Hum-Comput Interact. 2022. https://doi.org/10.1080/10447318.2022.2143767.
13. Ma J, Ma K-L, Frazier J. Decoding a complex visualization in a science museum-an empirical study. IEEE Trans Vis Comput Gr. 2019;26(1):472–81.
14. Ma J, Liao I, Ma K-L, Frazier J. Living liquid: design and evaluation of an exploratory visualization tool for museum visitors. IEEE Trans Vis Comput Gr. 2012;18(12):2799–808.
15. Chiang K-C, Weng C, Rathinasabapathi A, Chen H, Su J-H. Augmented reality supported learning for cultural heritage of Taiwan in on-site and off-site environments: The case of a Daxi old street. ACM J Comput Cult Herit. 2023. https://doi.org/10.1145/3589231.
16. Ch'ng E, Cai S, Feng P, Cheng D. Social augmented reality: communicating via cultural heritage. ACM J Comput Cult Herit. 2023;16(2):1–26.
17. O'dwyer N, Zerman E, Young GW, Smolic A, Dunne S, Shenton H. Volumetric video in augmented reality applications for museological narratives: a user study for the long room in the library of trinity college Dublin. J Comput Cult Herit. 2021;14(2):1–20.
18. Chen H-R, Lin W-S, Hsu T-Y, Lin T-C, Chen N-S. Applying smart glasses in situated exploration for learning English in a national science museum. IEEE Trans Learn Technol. 2023. https://doi.org/10.1109/TLT.2023.3276702.
19. Hammady R, Ma M, Al-Kalha Z, Strathearn C. A framework for constructing and evaluating the role of MR as a holographic virtual guide in museums. Virtual Real. 2021;25(4):895–918.
20. Hammady R, Ma M, Strathern C, Mohamad M. Design and development of a spatial mixed reality touring guide to the Egyptian museum. Multimed Tools Appl. 2020;79:3465–94.
21. Aoki S, Itabashi N, Adha R, Sameshima S, Kinoshita Y, Kotoku T. Mixed reality guided museum tour: digital enhancement of museum experience. In: 2023 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW), IEEE. 2023; p. 809–10.
22. Verhulst I, Woods A, Whittaker L, Bennett J, Dalton P. Do VR and AR versions of an immersive cultural experience engender different user experiences? Comput Hum Behav. 2021;125: 106951.

23. Wu L, Yu R, Su W, Ye S. Design and implementation of a metaverse platform for traditional culture: the chime bells of marquis Yi of Zeng. Herit Sci. 2022;10(1):193.
24. Rahimi FB, Boyd JE, Levy RM, Eiserman J. New media and space: an empirical study of learning and enjoyment through museum hybrid space. IEEE Trans Vis Comput Gr. 2020;28(8):3013–21.
25. Kim K, Kwon O, Yu J. Evaluation of an HMD-based multisensory virtual museum experience for enhancing sense of presence. IEEE Access. 2023. https://doi.org/10.1109/ACCESS.2023.3311135.
26. Trunfio M, Jung T, Campana S. Mixed reality experiences in museums: exploring the impact of functional elements of the devices on visitors' immersive experiences and post-experience behaviours. Inf Manag. 2022;59(8): 103698.
27. Margetis G, Apostolakis KC, Ntoa S, Papagiannakis G, Stephanidis C. X-reality museums: unifying the virtual and real world towards realistic virtual museums. Appl Sci. 2020;11(1):338.
28. Yuqing Z, Mingliang C, Haoyang Z, Yong Z. VR technology and application in martial arts. In: 2021 IEEE 7th International Conference on Virtual Reality (ICVR), IEEE. 2021; p. 240–5.
29. Wang Y. Discuss the relationship between virtual museums and physical museums: and discuss the development of digital museums. The Silk Road, 2011; p. 84–6.
30. Liu Y, Li C, Gong J. An object reconstruction method based on binocular stereo vision. In: Intelligent Robotics and Applications: 10th International Conference, ICIRA 2017, Wuhan, China, August 16–18, 2017, Proceedings, Part II 10, Springer. 2017; p. 486–95.
31. Lv C, Wang X, Zhang Q. The head detection method based on binocular stereo vision. In: 2016 9th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI), IEEE. 2016; p. 477–83.
32. Lin X, Wang J, Lin C. Research on 3d reconstruction in binocular stereo vision based on feature point matching method. In: 2020 IEEE 3rd International Conference on Information Systems and Computer Aided Education (ICISCAE), IEEE. 2020; p. 551–6.
33. Lombard M, Ditton T. At the heart of it all: The concept of presence. J Comput-Med Commun. 1997;3(2):321.
34. Palmer MT. Interpersonal communication and virtual reality: Mediating interpersonal relationships. Communication in the age of virtual reality, 1995; p. 277–99.
35. Witmer BG, Singer MJ. Measuring presence in virtual environments: a presence questionnaire. Presence. 1998;7(3):225–40.
36. Slater M, Sanchez-Vives MV. Enhancing our lives with immersive virtual reality. Front Robot AI. 2016;3:74.
37. Kim W, Xiong S. User-defined walking-in-place gestures for VR locomotion. Int J Hum-Comput Stud. 2021;152: 102648.
38. Buttussi F, Chittaro L. Locomotion in place in virtual reality: a comparative evaluation of joystick, teleport, and leaning. IEEE Trans Vis Comput Gr. 2019;27(1):125–36.
39. Schäfer A, Reis G, Stricker D. Controlling teleportation-based locomotion in virtual reality with hand gestures: a comparative evaluation of two-handed and one-handed techniques. Electronics. 2021;10(6):715.
40. Cao L, Zhang H, Peng C, Hansberger JT. Real-time multimodal interaction in virtual reality—a case study with a large virtual interface. Multimed Tools Appl. 2023;82(16):25427–48.
41. Kim W, Xiong S. Pseudo-haptic button for improving user experience of mid-air interaction in VR. Int J Hum-Comput Stud. 2022;168: 102907.
42. Tinguy X, Pacchierotti C, Lécuyer A, Marchal M. Capacitive sensing for improving contact rendering with tangible objects in VR. IEEE Trans Vis Comput Gr. 2020;27(4):2481–7.
43. Tinguy X, Pacchierotti C, Marchal M, Lécuyer A. Toward universal tangible objects: optimizing haptic pinching sensations in 3d interaction. In: 2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR), 2019; p. 321–30. https://doi.org/10.1109/VR.2019.8798205.
44. Feick M, Degraen D, Hupperich F, Krüger A. Metareality: enhancing tactile experiences using actuated 3D-printed metamaterials in virtual reality. Front Virtual Real. 2023;4:1172381.
45. Obradović M, Vasiljević I, Durić I, Kićanović J, Stojaković V, Obradović R. Virtual reality models based on photogrammetric surveys-a case study of the iconostasis of the Serbian orthodox cathedral church of Saint Nicholas in Sremski Karlovci (Serbia). Appl Sci. 2020. https://doi.org/10.3390/app10082743.

46. Remondino F, Del Pizzo S, Kersten TP, Troisi S. Low-cost and open-source solutions for automated image orientation—a critical overview. In: Progress in Cultural Heritage Preservation: 4th International Conference, EuroMed 2012, Limassol, Cyprus, October 29–November 3, 2012. Proceedings 4, Springer. 2012; p. 40–54.
47. Gajic D, Mihic S, Dragan D, Petrovic V, Anisic Z. Simulation of photogrammetry-based 3D data acquisition. Int J Simul Model. 2019;18(1):59–71.
48. Cai S, Wang X, Chiang F-K. A case study of augmented reality simulation system application in a chemistry course. Comput Hum Behav. 2014;37:31–40.

Publisher's Note