

# Practical Research on 3D Scene Construction in Digital Presentation of Cultural Heritage from the Perspective of Immersive Interaction

Jiayu Zhou

Edinburgh College of Art, The University of Edinburgh, Edinburgh, EH8 9YL, United Kingdom

---

**Abstract:** After the introduction of digital technology for the protection and inheritance of culture, immersive experience gradually became a trend in digital presentation. Construction of three-dimensional spaces is related to the sense of reality of the display and communication effectiveness, and it is the foundation for immersive experiences. From the perspective of immersive interaction, this paper investigates the practical logic and technical path for building three-dimensional scenes in the digitalisation of cultural heritage. The problem is not just a lack of technical precision; there is also a conflict between the requirement for determinacy in digital modelling and the openness and contextuality of cultural heritage meaning. The three forms of this tension are: the continuous development of historical awareness versus the fixed nature of a 3D environment; deviation from the user's own exploration path by designers in the construction of a story; and the use of advanced technology for high-quality experience creation that excludes most people. The digital exhibition hall of Bao Gong Temple, the VR museum of Haihun Culture, the virtual exploration of Jinci Temple, and the intelligent Yingxian Wooden Pagoda case study all show that the construction of three-dimensional scenes is changing from "precise replication" to "meaning reconstruction". Therefore, this paper puts forward the following three optimisation ideas: data collection, model generation and scene interaction.

**Keywords:** Immersive Interaction; Cultural Heritage; Digital Presentation; Three-Dimensional Scene Construction; Virtual Reality.

---

## 1. Introduction

Digitalisation of cultural heritage is one of the most active directions for practising in the field of cultural inheritance. New technologies such as virtual reality, augmented reality and digital twins have been iteratively improved and are now changing how cultural heritage is presented. Previously, the main focus of digitalisation efforts was on two-dimensional image recording and three-dimensional model archiving of cultural relics, and the main purpose was "preservation". Immersive Interaction Technology has altered this model by converting the passive "viewing" of digital presentations into an interactive "experience", allowing audiences to enter virtual environments and interact with cultural heritage. The realisation of this change depends to some extent on the quality of three-dimensional scene construction [1].

At the level of policy, corresponding measures are also being taken simultaneously. In 2023, the Ministry of Culture and Tourism released a series of industry standards for "Digital Protection of Intangible Cultural Heritage—Digital Resource Collection and Cataloging" (WH/T 99.1–2023 to WH/T 99.11–2023). In 2026, the National Cultural Heritage Administration issued the "Management Measures for State-owned Cultural Heritage Resources" (Document No. 17 of the Cultural Affairs Department in 2026) and established an all-encompassing life-cycle management system. In July of the same year, the World Internet Conference released the "Case Collection of Cultural Heritage Digitalisation at the World Internet Conference (2026)", which included 45 cases from almost 200 applications around the world. These actions suggest that the digitalisation of cultural heritage has moved from a local technical experiment to an institutional system.

Although there has been a rapid development at the policy level, more attention from the academic community is needed

for the construction of three-dimensional scenes. The existing achievements mainly focus on optimizing technical links or introducing case studies, and lack a systematic study from the perspective of immersive interaction. Problems such as the function position of three-dimensional scene construction, the applicable conditions and limitations of different technical routes, and the sources of contradictions in practice have not been fully resolved. The answers to the above questions will guide the development of technology and help choose one or more paths for the digitalisation of cultural heritage. Based on immersive interaction and drawing representative practical cases from China in recent years, this paper presents an all-around study of three-dimensional scene construction.

## 2. Core Functions and Basic Logic of 3D Scene Construction

Before presenting specific contents, we first need to understand what functions are required for building 3D scenes in the digital presentation of cultural heritage. First of all, this problem should be at the level of immersive interactive experience. Not only technical parameters are restricted; otherwise, one cannot determine how immersive a space feels to the user based on how high-resolution a computer or TV is or how fast it operates. We need to know whether users feel that the virtual environment is their "current place" at the same time in perception, cognition and emotion [2]. Only when the virtual environment can be established in terms of the three dimensions of spatial form, behavioural affordance and meaning-carrying capacity will users be willing to let go of their concerns about physical space and instead immerse themselves in the digital world. In other words, the construction of a 3D scene provides the underlying spatial support for the realisation of these three dimensions.

In terms of spatial form, 3D scene construction recreates

the digital model of buildings, materials, light and shadow, etc., to create a virtual environment that feels more real to the user. If the space logic deviates too far from people's daily lives, it will be difficult for them to experience the desired state; for instance, scale distortion or violations of light and shadow common sense will instantly drive them away. The Behavioural affordance dimension is more direct; that is to say, what can be touched, which paths are available for travel, and what information can be accessed—these are actual divisions in user exploration [3]. If the interaction points are integrated smoothly, more people will be eager to use and engage with them. For the transmission of meaning, 3D scene construction should be used to convey cultural narratives through spatial organisation, visual symbols and dynamic effects, and should not be empty.

Note that 3D Scene construction does not always need to be based on reality. Cultural heritage is often damaged or missing; at the same time, much of the historical information is also unavailable. Therefore, the construction of scenes must choose between evidence-based restoration and narrative-driven reconstruction. The former (restoration) is closer to the archaeological materials and documentary records, whereas the latter (reconstruction) may incorporate some narrative elements based on the basic facts. The two are not in a fixed order, and the selection will be based on whether the heritage is to be preserved or exhibited. Discussing "reconstruction degree" or "expressiveness" in isolation from a specific context is not feasible in terms of methodology.

### **3. Technology-Driven Practice Path: From Data Collection to Scene Presentation**

After clarifying the functional positioning of 3D scene construction, the next step is to explore how it is realised at the technical level. Generally speaking, the three stages in the technical process of 3D scene construction are data collection, model reconstruction and scene rendering. The technical routes of the various steps are not the same, and these choices will ultimately impact how users experience them; thus, they are worth exploring individually.

The first is to collect data in this stage. The goal is to turn the Space of cultural heritage into machine-readable digital data. At present, the general technical means are 3D laser scanning, oblique photogrammetry, close-range photogrammetry and handheld fine scanning. These ways have different degrees of accuracy, operational convenience and application ranges [4]. The strength of 3D laser scanning is that it can acquire high-precision geometric data at the millimeter level, and thus is suitable for recording immovable cultural relics such as old buildings and grottoes; oblique photogrammetry uses multiple-angle aerial photographs to obtain surface texture information and is more suitable for large-scale modelling of sites and landscapes. In practice, these ways are not mutually exclusive; rather, they are frequently used in combination according to the demands of a specific project. During the data collection stage of the Smart Yingxian Wooden Pagoda project, station-based scanning, handheld fine scanning and unmanned aerial vehicle oblique photogrammetry were used, and a total of more than 15 million polygons of the basic material were obtained. It will be relatively accurate and also relatively fast.

The first set of collected data is a discrete point cloud. The aim of model reconstruction is to transform the above point

clouds into an easy-to-use 3D model. Point cloud data is discrete and contains noise, so algorithms need to be used to process it to generate continuous mesh surfaces, and then texture maps can be added to complete the work. The strengths of 3D laser scanning include high-precision geometric data collection at the millimeter level, making it highly suitable for documenting fixed cultural relics such as old buildings and grottoes; oblique photogrammetry is based on multiple angles of aerial photography to acquire surface texture and is more suitable for large-scale modelling of sites and landscapes. The digital art exhibition of the Song Dynasty architecture and murals of Gaoping Kaihua Temple uses 3D modelling and immersive light-and-shadow interpretation to present the murals of the Northern Song Dynasty. The technical problems are to extract the colours of the murals accurately and to improve the quality of model surface bonding.

The model is ready, and now scene rendering will be used to add dynamic lighting and interactive logic. Unreal Engine and Unity3D are real-time rendering engines that provide spaces for displaying changes in time and spatial textures for three-dimensional scenes. The Digital Cultural Education Exhibition Hall of Baogong Temple is based on Unreal Engine 5. Dynamically render light to add different visual layers on the surface of the architectural complex in various light conditions and improve the sense of space in the virtual environment. Data collection to rendering, all stages involve technical decisions, and thus, any deficiencies in a particular step will be exacerbated in the end result [5].

### **4. Experience-Oriented Practice Shift: Scenario Reconstruction in Immersive Interaction**

At the technical level, it refers to "how to build the scenario"; at the level of experience, it is to ask "for whom" and "what kind of experience" to create. The two problems above have led to the construction of three-dimensional scenarios that are now "people-centred" rather than "object-centred".

The Exploration Journey of the virtual world of Jinci Temple provides a case study for this change. All of the above 3D laser scanning, white-light scanning, aerial photography, panoramic photography and tilt photography techniques were used to collect spatial and colour data of the Jinci Temple garden, ancient buildings, painted sculptures, murals and inscriptions, and these were accurately and precisely recreated in a virtual scene at a 1:1 scale. Based on the technical indicators, it is without a doubt a "precise reproduction" method. However, the highlights are not limited to replication; free-interaction areas have been added to the single-person linear narrative, and under the guidance of the cartoon guide "Xiao You", the audience can complete the experience. Therefore, the three-dimensional scenario has taken on a dual nature as both a subject of observation and a way of conveying stories or exploring ideas. By 2019, the cumulative number of visits had exceeded 20,000 and thus proved that interactive scene forms were popular with the public [6].

The VR Museum of Haihun Culture promotes the concept of "scenario as narrative" further. The Pico VR headset and Unity3D engine will be used for the development of this project, and a dual-scenario design of "modern museum + ancient life scenarios" will be adopted. The old and new are

not opposed, but rather in conversation through comparison. Optical-inertial hybrid motion capture is used in the construction of the three-dimensional scenario to reproduce the dance movements of the Haihun culture, and dynamic figures are added to static scenes to move away from the traditional method of static processing.

The Smart Yingxian Wooden Pagoda Project will also be interactive. The project creates "wooden pagoda twins" through spatial computing AIGC technology, has written nearly 60,000 lines of Unity code, and rendered 4.2GB of models for the first to fifth floors and five main exploration scenarios, along with seven interactive designs and 11 large-scale animation effects. In the 2.0 edition of 2025, an embodied intelligent six-legged robot dog is released for non-contact restoration of fine details on millimeter-scale cultural relics. When the scenic area raised the access to the pagoda, a three-dimensional scene was created at the original site. When the space is small, digital scenarios can still offer people immersive and exploratory cultural experiences; this is precisely the special feature of immersive interaction for cultural heritage [7].

## **5. Structural Challenges in 3D Scene Construction and Its Generation Mechanism**

3D Scene Construction faces the following three problems in actual implementation. These problems are not technical deficiencies but rather inherent contradictions in the digitalisation process. Due to the young technology, they are not disclosed.

The first problem is that of time: historical consciousness is evolving, but the 3D space is static. The scene has been generated, and now the architectural style, colour and space have been fixed. Archaeological cognition is also developing continuously, and new discoveries may contradict the previous conclusions at any time. The Yungang Grottoes digitalisation project began in 2003, and more than 20 years have passed since the collection of 80% of the caves; this cycle is much longer than the update cycle for individual academic cognition. Therefore, when a 3D scene is displayed, it may precisely lock onto a transitional academic judgment rather than a final one. This temporal displacement is almost unavoidable in long-term projects, and the timeliness of the results of digitalisation itself is a problem that needs to be solved.

The second problem is a space issue: although the designer wishes to create a whole story, the user is unable to move freely in space. A constructor generally desires to build a stable story world, but in an immersive environment, we cannot predict what users will do. The organised clues may be ignored, and the unobserved details will be noticed. There will always be a difference between the story path drawn by the designer and the actual exploration path of the user [8]. It is not a design defect, but an inherent contradiction between the preset Space and the active subject. Who will be using this scene and in what way is not clear at the time of design, so we cannot precisely know how users will experience it in virtual reality.

The third problem is that the technical level required for high-quality experience and mass dissemination is too high. At present, the above projects require high-performance computing equipment, professional display terminals and customised development environments. The equipment has

relatively high requirements and a small number of people are allowed to use it. The small scale of the limited-time, digitally immersive special exhibition "Exploring the Kingdom of Haihunhou" (only 300 square meters and 20 visitors per session) also points to a problem. The digital cultural education exhibition hall at Baogong Temple has received praise from the professional community, but the personalised solution based on the UE5 engine is not easily reproducible due to issues with venue and technology investment. Find a feasible path that can maintain quality and reduce the threshold [9].

## **6. From Precise Replication to Meaning Construction: Optimization of the Path for 3D Scene Construction**

Considering the problems listed above, an increase in technology is not sufficient; we must also change how we do things. The direction of optimisation is clear: from "precise reproduction of objects" to "construction of meaning through human experience". The three levels of this change are as follows.

First, collect data and then build a model. At present, there is a lack of cooperation among the departments. Technicians handle the technical problems, and the judgments of historians and archaeologists are often required only at the end. A reasonable way is to include people from all walks of life in the collection stage, have the judgments of history and archaeology guide the focus of data collection, and incorporate user experience feedback during the model reconstruction process [10]. The "Beyond the Murals: The Tang Dynasty" project of Shaanxi History Museum is a typical case. The murals of the blue dragon and white tiger in Crown Prince Zhanghuai's tomb have been lost. The team first systematically sorted out the patterns of the four beasts elements in the murals of the same period of Tang tombs, and then, combined with high-precision 3D technology, performed restoration by adding information based on evidence rather than random creation. Research and development of academic institutions and digital technology in the project are happening concurrently, not sequentially. At least this way can provide a feasible path to address the contradiction of "reproducing history and making the scene visually appealing".

Next, I will present the middle part of the interaction and experience design. The two are "guidance" and "openness"; they should be in harmony. If the narrative path is too rigid, users will not be able to explore off-road paths in their experience; if there are no hints at all, it may be difficult for them to learn about the culture deeply. The compromise solution is to design a gradient of interaction, and then, according to one's own interests, decide whether to go to a deeper level. The Jinci Temple left a free-interaction space along the main road, and the Yingxian Wooden Pagoda also set up multiple scenes and interaction methods according to the same idea.

Finally, it is the technical adaptation and promotion of dissemination. After the development of mature cloud rendering technology, high-quality 3D scenes will no longer require high-performance local hardware. Through Web pages and mobile terminals, experience is gradually becoming available. Another problem that is often neglected should be addressed: although academic research has been

advancing steadily, 3D scenes are static. Scenes should have the ability to update dynamically [11]. This is not only a technical problem but also a change in the organisational model to a continuous operation mode for the project instead of a one-time development model.

## 7. Conclusion

Three-Dimensional Scene Construction can transform static displays of cultural heritage into interactive spaces. The typical cases of projects such as Baogong Temple, Haihun, Jinci Temple and Yingxian Wooden Pagoda show that the value of this technology does not lie in the accuracy of the model, but in providing the public with an experience of historical scenes. When visitors interact with the murals in the virtual environment and walk around the suspended wooden pagoda, three-dimensional scene construction will be fully realised.

However, practice has also shown three problems. The more than 20-year digital transformation of the Yungang Grottoes is a typical case: although a three-dimensional environment can be constructed, the development of archaeological knowledge continues steadily. The designer has created a storyline, but it is not entirely in line with what the users actually feel. Investment in equipment for high-quality experiences directly limits the number of participants. The three levels are in conflict; thus, the technical choice needs to meet all of them.

The "historical research + digital reconstruction" model of "Beyond the Murals: The Tang Dynasty" at the Shaanxi History Museum, as well as the gradient interactive design of Jinci Temple and Yingxian Wooden Pagoda, jointly point in the same direction: from precise reproduction to the construction of meaning. Therefore, many disciplines need to be involved in the data collection phase, the user experience feedback must be incorporated into the model-building process, and space for users' free choice should be left in the interactive Design.

The above discussion is based on a few representative cases. Therefore, the deficiencies of this study need to be presented. This paper does not discuss the application of small and medium-sized projects, and there are no quantitative comparisons of the cost-effectiveness of different technical routes. How three-dimensional scene construction can achieve low-cost iteration after the update of academic cognition and how to overcome equipment limitations are still to be explored. With the progress of the national cultural digitalisation plan, three-dimensional scene construction will be used in more complex ways. Cultural heritage in the digital space does not need to be a perfect reproduction of the past; rather, it should strive to answer how to provide the public

with an open and participatory cultural space.

## References

- [1] Rodriguez-Garcia, B., Guillen-Sanz, H., Checa, D., & Bustillo, A. (2024). A systematic review of virtual 3D reconstructions of cultural heritage in immersive virtual reality. *Multimedia Tools and Applications*, 83(42), 89743-89793.
- [2] Ghani, I., Rafi, A., & Woods, P. (2020). The effect of immersion towards place presence in virtual heritage environments. *Personal and Ubiquitous Computing*, 24(6), 861-872.
- [3] Lee, J., Jung, T., Tom Dieck, M. C., García-Milon, A., & Kim, C. S. (2024). Affordance, digital media literacy, and emotions in virtual cultural heritage tourism experiences. *Journal of Vacation Marketing*, 31(4), 1056-1073.
- [4] Li, Y., Zhao, L., Chen, Y., Zhang, N., Fan, H., & Zhang, Z. (2023). 3D LiDAR and multi-technology collaboration for preservation of built heritage in China: A review. *International Journal of Applied Earth Observation and Geoinformation*, 116, 103156.
- [5] Luther, W., Baloian, N., Biella, D., & Sacher, D. (2023). Digital twins and enabling technologies in museums and cultural heritage: An overview. *Sensors*, 23(3), 1583. <https://doi.org/10.3390/s23031583>Mendeley
- [6] Liu, Z., Chen, D., Zhang, C., & Yao, J. (2024). Design of a virtual reality serious game for experiencing the colors of Dunhuang frescoes. *Heritage Science*, 12(1), 1-31. <https://doi.org/10.1186/s40494-024-01477-x>Nature
- [7] Hou, H., Lai, J. H., Wu, H., & Wang, T. (2024). Digital twin application in heritage facilities management: Systematic literature review and future development directions. *Engineering, Construction and Architectural Management*, 31(8), 3193-3221.
- [8] Petousi, D., Katifori, A., Servi, K., Roussou, M., & Ioannidis, Y. (2022). Interactive digital storytelling in cultural heritage: The transformative role of agency. In M. Vosmeer & L. Holloway-Attaway (Eds.), *Interactive storytelling* (pp. 48-67). Springer International Publishing. [https://doi.org/10.1007/978-3-031-22298-6\\_3](https://doi.org/10.1007/978-3-031-22298-6_3)
- [9] Anastasovitis, E., Georgiou, G., Matinopoulou, E., Nikolopoulos, S., Kompatsiaris, I. Y., & Roumeliotis, M. (2024). Enhanced inclusion through advanced immersion in cultural heritage: A holistic framework in virtual museology. *Electronics*, 13(7). <https://doi.org/10.3390/electronics13071396>
- [10] Nikolakopoulou, V., & Koutsabasis, P. (2025). The 'making' of participatory and co-design for digital experiences in cultural heritage: A review. *CoDesign*, 21(4), 876-902.
- [11] Ma, Y., Tong, Q., He, X., & Su, B. (2025). Exploring virtual restoration of architectural heritage through a systematic review. *npj Heritage Science*, 13(1), 167. <https://doi.org/10.1038/s40494-025-02162-3>.