Digital Modeling and Display Method of Tourist Scenic Spots Based on 3D Reconstruction and Neural Radiance Field

Haiyang Shu[1,2], Yandan Peng[1,2], Junjie Li[3], Shiyu Shu[3], Shanglin Li[1,2,*]
[1] *School of Computer Science and Artificial Intelligence, Xiangnan University, Chenzhou, Hunan, China*
[2] *Hunan Engineering Research Center of Advanced Embedded Computing and Intelligent Medical Systems, Chenzhou, Hunan, China*
[3] *School of Clinical Medicine, Xiangnan University, Chenzhou, Hunan, China*
**Corresponding Author*

Abstract: With the growing popularity of tourism in China, the digital preservation and display of culturally significant tourist sites have become increasingly important. This study focuses on the 'Half Quilt' scenic spot in Shazhou Village, Rucheng County, Chenzhou City, Hunan Province, as a case study. By integrating digital modeling and simulation techniques, such as 3D modeling, 3D reconstruction, and neural radiance fields, we develop an immersive panoramic and 3D walkthrough of the site. The constructed 3D model is tested in the Unreal Engine 5 and can be deployed on web, mobile, and PC platforms, with additional support for VR and AR technologies. This enables users to experience virtual tours of the scenic spot, enhancing both its visibility and cultural significance.

Keywords: Tourist Attractions; Digital Exhibition; 3D Modeling; 3D Reconstruction; Panorama; Protecting Culture

1. Introduction

1.1 Research on Digitalization of Tourism Resources in Scenic Spots

Scholarly investigations and practical implementations of digital technologies within scenic heritage domains have garnered escalating interest internationally, rendering digital exhibition and conservation a prominent research frontier [1]. Liu and Deng [2] advocated deploying information technologies for systematic digital documentation, condition monitoring, restoration, reconstruction, and creative reproduction, thereby establishing sustainable frameworks for preserving both physical configurations and intangible cultural values through reusable digital repositories. Nguyen et al. [3] implemented VR and AR solutions to construct immersive virtual museum environments, offering innovative methodological insights that extend to adjacent domains including interactive 3D entertainment and digital tourism platforms. Three-dimensional reconstruction methodologies generate precise geometric models and associated datasets for architectural heritage, furnishing essential visual references that facilitate subsequent conservation interventions and structural upkeep.

Digital preservation initiatives for cultural landscapes emerged earlier within technologically advanced nations. Advancements in 3D reconstruction have opened unprecedented avenues for safeguarding digitized cultural patrimony. In 2003, Allen et al. [4] performed comprehensive laser scanning of Cathedrale St-Pierre, producing detailed three-dimensional digital documentation of the structure. In China, the adoption of 3D technologies for heritage applications began comparatively later. Following pioneering collaborative efforts between academic institutions and the Palace Museum in 2004, laser scanning methodologies have progressively permeated tourism-related research and practice as technological capabilities have matured [1]. Cao et al. [5] integrated landscape gene theory with semiotic analysis and GIS-based spatial processing to systematically identify, extract, and encode the distinctive landscape characteristics of traditional settlements in western Hunan, enabling structured data organization and visual interpretation. Contemporary scenic data have achieved marked improvements in visual fidelity and geometric precision. The documentation and safeguarding of three-dimensional and panoramic datasets, alongside the development of specialized heritage databases, continue to broaden the scope of practical applications. Recent innovations in neural radiance field architectures, incorporating frequency-domain and tensorial decomposition strategies, have substantially enhanced reconstruction fidelity for architectural heritage, particularly in scenarios constrained by sparse photographic coverage [6]. Despite these advances, the current state of digitalization research for tourist destinations remains underdeveloped. Operational methodologies require refinement, and field implementation demands more rigorous validation to bridge the gap between theoretical frameworks and practical deployment.

1.2 Digital Dissemination of Tourist Resources in Scenic Spots

Concurrently, the digital transformation of scenic resources engenders innovative pedagogical modalities and cultivates participatory channels through which digital heritage assets can contribute to cultural dissemination and cross-cultural dialogue. Within the contemporary digital ecosystem, leveraging advanced technologies to diversify preservation strategies and expand communication pathways for cultural heritage constitutes an imperative objective for the sustainable development of scenic cultural industries. By harnessing Internet-based infrastructures alongside virtual reality, immersive panoramic imaging, and intuitive human-computer interaction paradigms, cultural scenic content transcends traditional spatiotemporal limitations, delivering visually compelling experiences directly to end-users.

The digitization and dissemination of historical heritage has become a key node in the preservation and value adding of cultural heritage, such as Fascia et al. [7] developed a replicable prototype platform for immersive virtual Tours (VT) of art collections and the building complexes in which they are located, and the source code to implement the application is available on GitHub, which is reproducible.

Historical and cultural scenic spots embody historical, aesthetic, and architectural attributes that constitute vital carriers of human civilization. Three-dimensional reconstruction, panoramic visualization, and virtual reality (VR) offer immersive, visually rich modalities that serve as critical instruments for heritage conservation and tourism presentation. The evolution of digital technologies fosters novel syntheses between human perceptual experience and analytical reasoning within virtual environments, necessitating that landscape culture dissemination align with digital transformation trends. Employing 3D reconstruction to construct digital scene replicas enables accurate virtual mapping of physical sites, supports material restoration efforts, and facilitates cultural value propagation through emerging media channels.

2. Research Method

2.1 Panoramic Technique

Panoramic imaging leverages computational algorithms to composite multiple photographic captures into seamless spherical environments, enabling comprehensive interactive exploration of spatially accurate scene reconstructions. The underlying mechanism relies on panoramic capture apparatus—whether single-sensor or multi-sensor configurations—to record visual data, which subsequent stitching software merges into continuous spherical projections delivering full 360-degree horizontal and vertical coverage. When coupled with specialized playback systems and display hardware, these immersive media reconstruct dynamic real-world settings, transporting viewers across temporal and spatial boundaries.

During the digital documentation of Shazhou Village, expansive environments such as exhibition halls were captured using DJI Mavic 2 Pro unmanned aerial vehicles to acquire high-resolution panoramic imagery and associated metadata. Such aerial acquisition approaches align with emerging protocols in heritage documentation, wherein drone-based photography integrated with neural reconstruction techniques has demonstrated efficacy in preserving historic architectural fabric [8]. For spatially constrained interiors, including the original "Half Quilt" residence and the former Ministry of Health site, smartphone-based panoramic capture via Google Camera provided appropriately scaled imagery and dimensional data for enclosed environments. These datasets constitute critical inputs for the digital processing and geometric modeling of revolutionary heritage resources, furnishing adequate visual and spatial information to enhance both the accuracy and photorealism of reconstructed digital assets. The renderings of the indoor panorama are shown in Figure 1.

Figure 1. Renderings of Indoor Panorama

The digital exhibition of the "Half Quilt" narrative integrates three-dimensional modeling and reconstruction techniques with panoramic visualization and real-time spatial navigation. The operational pipeline proceeds as follows: terrestrial LiDAR scanning data and aerial photographic surveys are aggregated to generate high-precision three-dimensional point clouds, digital elevation models (DEM), and digital orthophoto maps (DOM); dedicated modeling software then constructs and optimizes geometric representations of architectural elements, public squares, and vegetation within the scenic area; finally, immersive navigation tools render these digital constructs for interactive user exploration. To safeguard sensitive cultural assets, the system incorporates contextual digital backdrops, anthropomorphic character models, and guided interactive presentations that restrict unauthorized visual access to protected zones. The panorama of the Folk Square is shown in Figure 2.

Figure 2. Panorama of Folk Square

2.2 NeRF(Neural Radiance Fields)

Neural Radiance Fields constitute a deep learning-based paradigm within computer graphics and vision for implicitly encoding three-dimensional scenes or objects. The methodology represents spatial configurations as continuous volumetric functions, estimating chromatic properties and optical density along light trajectories, thereby synthesizing novel perspectives through training on multi-view photographic collections.

The operational principle involves training neural architectures to implicitly encode illumination, spectral, and surface material properties within three-dimensional environments. At inference time, the trained model synthesizes photorealistic imagery from arbitrary camera positions and lighting scenarios, substantially enhancing both visual fidelity and perceptual immersion. For expansive architectural scenes, the neural radiance field formulates the environment as a continuous volumetric field, estimating per-point color and density values along incident light paths. The mathematical formulation is presented below, while Figure 3 illustrates the NeRF-based rendering of the Zhu Ancestral Temple.

The formula for pixel color is as follows:

$$C(r) = \int_{t_n}^{t_f} T(t)\sigma(r(t))c(r(t), d)dt,\ where\ T(t) = exp\,(-(s))ds) \quad (1)$$

In the actual rendering process, we can divide the ray into N equal intervals, sample a random point in each interval, and perform some kind of weighted sum of the colors of the sampled points:

$$t_i \sim U[t_n + \frac{i-1}{N}(t_f - t_n), t_n + \frac{i}{N}(t_f - t_n)) \quad (2)$$

$$\hat{C}(r) = \sum_{i=1}^{N} T_i\,(1 - exp\,(-\sigma_i\delta_i))c_i,\ where T_i = exp\,(-\sum_{j=1}^{i-1}\sigma_j\,\delta_j) \quad (3)$$

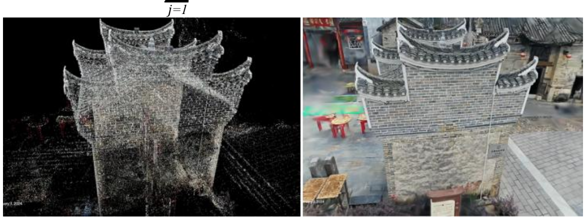

Figure 3. Rendering of Zhu Ancestral Temple

2.3 Multi-view 3D Reconstruction

Intricate architectural structures, sculptural elements, and artifact collections within Shazhou Village undergo systematic digitization via multi-view three-dimensional reconstruction pipelines. This approach faithfully captures and restores geometric intricacies and surface characteristics of complex scenic features, enabling digital preservation and public engagement with distinctive cultural assets while supporting scholarly analysis, conservation activities, and intergenerational transmission.

They are described as follows:

Structure from Motion represents an autocalibration methodology enabling simultaneous camera trajectory estimation and geometric correspondence establishment.

Specifically, i projections of points in world space to image space are observed in j images taken by cameras from different locations. The camera orientation Rj, position vector cj, and focal length fj are defined, then the image coordinate Xij obtained by the camera projection of a point Xi in the world coordinates can be defined as:

$$x_{ij} = f_j \prod (R_j (x_i - c_j)) \tag{4}$$

Where $\prod$ is the projection function. Theoretically, provided that there are adequate observed points, the positions of these points in the world space and the original camera position can be computed by diagonal geometry. Nevertheless, the calculation of inverse projection for a considerable number of images will give rise to the accumulation of errors. To generate high-quality camera parameters and sparse point clouds with reduced errors, the following optimization process has to be accomplished:

$$\arg \min_{x_i, R_j, c_j, f_j} \sum_{i \sim j} \| x_{ij} - f_j \prod (R_j (x_i - c_j)) \|^2 \tag{5}$$

Where i~j means that Xi is visible in image j. This optimization process is called Bundle Adjustment (BA) and is usually solved by the nonlinear Levenberg-Marquardt algorithm. While Structure from Motion paired with Multi-view Stereo continues to dominate heritage digitization workflows, empirical evaluations demonstrate that neural radiance fields and conventional photogrammetry offer complementary strengths across diverse imaging scenarios, notably in negotiating trade-offs between geometric exactitude and perceptual realism [9].

Generating Dense point clouds and meshes After recovering the camera pose using SfM, a sparse 3D point cloud is output simultaneously. Sparse point clouds are not enough to generate mesh models. General dense matching methods, such as binocular Stereo vision and Multi-view stereo (MVS), are needed to generate dense point clouds, and then mesh them to generate mesh models. The 3D reconstruction results are shown in Figure 4, including the point cloud, reconstruction model, and rendered result.

(a) A Rendering of the 3D Reconstructed Point Cloud **(b) 3D Reconstruction Model** **(c) Render the Rendering**

Figure 4. 3D Reconstruction Rendering

2.4 3D Gaussian Splatting

To address inherent limitations of conventional approaches regarding noise artifacts and rendering performance while accelerating visualization throughput, the pipeline incorporates 3D Gaussian Splatting. Contemporary comparative analyses involving specular and low-texture heritage objects demonstrate that Gaussian Splatting exceeds traditional Multi-view Stereo in both rendering velocity and noise attenuation without compromising geometric accuracy [10].Post-reconstruction models achieve superior visual quality within application environments through rapid real-time rasterization, while training convergence occurs substantially faster relative to previous methodologies. The relevant formula is as follows:

The properties of a 3D Gaussian volume are its center (position) μ, opacity α, a 3D covariance matrix (representing the degree of scaling) ∑, and color c. c is represented by spherical harmonics to present a view-dependent appearance. All attributes are learnable and optimized by backpropagation.

The scaling transformation S and the rotation transformation R are combined to obtain Σ:

$$\sum = RSS^T R^T \tag{6}$$

The 2D graphics obtained after projecting a 3D Gaussian volume onto a plane is known as Splatting rendering method.

The 3D model is projected into the 2D image space for rendering. Given the view transformation W and the 3D covariance matrix Σ, the projected 2D covariance matrix

Σ' can be calculated using the following formula:

$$\sum{}' = JW \sum W^T J^T \tag{7}$$

Where J is a matrix that transforms a 3D vector into a 2D vector. The center position and color of the projected 2D Gaussian can be directly obtained from the parameters of the 3D Gaussian. The opacity of the projected 2D Gaussian needs to be adjusted according to the opacity and covariance matrix of the 3D Gaussian. Specifically, the opacity α' of the projected 2D Gaussian can be calculated using the following formula:

$$\alpha' = 1 - \exp\left(-\frac{\alpha}{\sqrt{\det(\Sigma)}}\right) \tag{8}$$

2.5 Digital Man Technology

MetaHuman Creator enables fine-grained control of digital human characters by integrating advanced graphics rendering, artificial intelligence, and machine learning techniques. Users can manipulate facial features, skin texture, muscle movement and tooth shape through an intuitive interface to create highly personalized character appearance, and further enrich the diversity of characters by using a rich preset library. The character data package generated by this technology is complete and has strong compatibility, and can be directly imported into mainstream 3D creation tools such as Unreal engine or Maya, which is convenient for subsequent animation production and scene integration.

MetaHuman Creator streamlines the construction of high-fidelity virtual humans. The pipeline initiates with authentic facial capture data serving as foundational reference, which integrated algorithms and template libraries rapidly scaffold into base character architectures; fine-grained editing instruments subsequently enable individualized refinement. After the construction is completed, the character data package can be seamlessly connected to the animation production process, and realistic action simulation and behavior interaction can be realized by using Unreal Engine or Maya. Within metaverse deployments, non-player characters synthesized via MetaHuman Creator substantially enhance both immersion and interactivity within virtual environments while diversifying applicable usage contexts. In the Shazhou Village virtual environment, these synthetic agents replicate human locomotion, facial expressions, and behavioral tendencies, delivering participants increasingly authentic and fluid engagement experiences. The face data acquisition process and the digital human renderings are shown in Figure 5.

Figure 5. Face Data Acquisition and Digital Human Renderings

3. Technical Framework

To achieve photorealistic scene restoration, the development framework assimilates contemporary computer graphics and immersive technology (VR/AR) paradigms, integrating state-of-the-art capabilities encompassing real-time rendering engines, three-dimensional modeling, neural radiance fields, and panoramic visualization into a cohesive, highly realistic interactive system. Cross-platform deployment targets encompass personal computers, web browsers, and mobile terminals.

Unreal Engine 5's Nanite virtualized geometry system enables real-time rasterization of high-polygon models. Its advanced streaming format and intelligent level-of-detail management substantially enhance both rendering throughput and micro-detail preservation. The Niagara particle simulation framework generates visually compelling atmospheric and environmental effects with high physical plausibility. To maintain environmental authenticity, the Ultra Dynamic Sky system produces adaptable meteorological phenomena featuring naturally evolving cloud formations. Dynamic interactive elements—including real-time deformation effects such as footprints accumulating within snow-covered terrain—substantially elevate both physical realism and user immersion. To preserve real-time performance across these effects, the Lumen global illumination and reflections system dynamically responds to scene and lighting modifications. Through synergistic hardware-accelerated ray tracing and software-based optimization, the architecture delivers premium reflective surfaces, expansive viewing frustums, and computationally efficient global illumination without requiring dedicated ray tracing hardware modifications. The convergence of Nanite's virtualized micro-polygon geometry with Lumen's fully dynamic global illumination and reflection capabilities achieves rendering quality approaching physical-world perceptual fidelity. Furthermore, DirectX 12 integration leverages UE5's low-level rendering capabilities to maximize execution efficiency, enhance three-dimensional graphics and spatial audio fidelity, and optimize overall application performance and temporal smoothness.

MetaHuman Creator's principal advantage lies in its extensive customization depth coupled with streamlined integration. Practitioners can exercise precise control over character attributes—including facial expressions, epidermal pigmentation, coiffure, and apparel coordination—through intuitive direct manipulation interfaces. Moreover, by incorporating authentic facial capture data with advanced rendering methodologies, the platform produces convincingly realistic virtual human representations that substantially enrich visual authenticity within digital environments. To maximize anthropomorphic fidelity, UE5's native rendering subsystem imbues virtual human models with convincingly lifelike dermal textures and responsive facial features The schematic of the technical framework is shown in Figure 6.

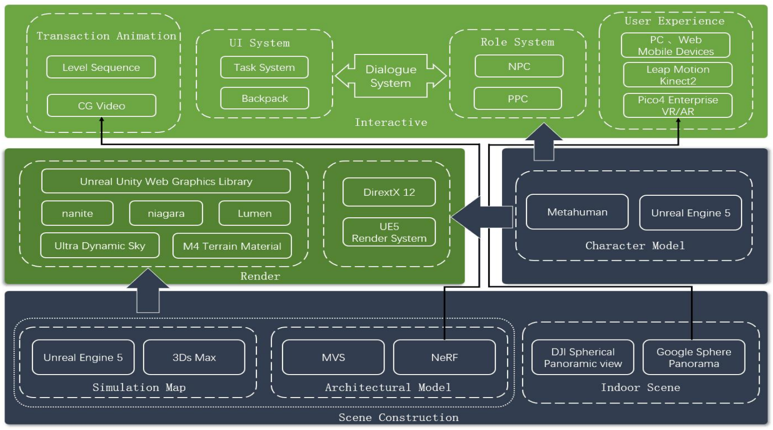

Figure 6. Schematic of the Technical Framework

4. Experimental Result

Utilizing comprehensive photographic datasets captured throughout Shazhou Village, the aforementioned technological pipeline reconstructs the settlement's distinctive architecture and historical artifacts at true-to-life scale, yielding a complete digital facsimile of the village. The modeling result of the Exhibition Hall is shown in Figure 7, and the tourist experience of Shazhou Village is shown in Figure 8.

Figure 7. Exhibition Hall Modeling Diagram

Within web-based deployments, models exhibit rapid initialization times, permitting users to fluidly navigate panoramic imagery and three-dimensional geometry through standard browsers while executing fundamental manipulations including orbital rotation and zoom scaling. Mobile implementations incorporate device-specific performance optimization, ensuring satisfactory experiences across smartphones and tablets through crisp image quality and responsive interaction latency. PC deployments leverage superior computational resources to present enhanced visual detail, while mouse and keyboard interfaces facilitate sophisticated interactions including detailed architectural interior examination. Within the UE5 authoring environment, gameplay mechanics are implemented through visual blueprint scripting. This node-based programming paradigm ensures code comprehensibility while facilitating convenient iterative modification. The game involves 26 player blueprints, 12 vehicle systems, and 20 NPC blueprints. The average number of nodes in each blueprint is more than 50, and the total number of variables involved is up to 30. By completing the main task of the game, the masses, especially teenagers, can deepen their understanding of the story of "Half a quilt", so that historical relics and stories can live forever in the digital world.

Figure 8. Tourist Experience Map of Shazhou Village

The integration of three-dimensional reconstruction, neural radiance field methodologies, panoramic visualization, and virtual human synthesis constructs an immersive gamified representation of Shazhou Village, delivering substantially enriched user experiences. An example of the main interface UI is shown in Figure 9.

Figure 9. Example of the Main Interface UI

5. Conclusion

This investigation demonstrates that digital exhibition methodologies substantially amplify both the communicative reach and cultural significance of the "Half Quilt" heritage site. Through systematic application of panoramic imaging, three-dimensional reconstruction, neural radiance fields, and virtual human agents, these technologies are cohesively woven into the scenic presentation design, vividly resurrecting historical settings and crafting deeply immersive heritage experiences for visitors. Concurrently, the finalized model deploys across web browsers, mobile platforms, and personal computers, augmented by VR and AR modalities to furnish unrestricted virtual access. This multi-platform strategy diversifies presentation formats through enriched multimedia interaction and immersive reality technologies, enabling tourists to comprehend and appreciate cultural distinctiveness with greater depth and comprehensiveness

In conclusion, the Shazhou Village digital twin achieves notable efficacy in exhibiting the "Half Quilt" heritage narrative, offering a replicable methodological reference and practical exemplar for digital presentation initiatives at comparable tourist destinations. Prospectively, digital exhibition technologies are anticipated to assume increasingly pivotal functions within tourism promotion and heritage interpretation strategies.

Acknowledgments

This work was supported in part by Hunan Provincial Natural Science Foundation of China under Grant 2023JJ30565，2022JJ40415, in part by the Scientific Research Fund of Hunan Provincial Education Department under Grant 2022B0818, in part by the Hunan Provincial Social Science Achievement Evaluation Committee Project under Grant XSP24YBZ034, and Chenzhou Social Science Planning Project under Grant CZSSKL2023123.All supports are gratefully acknowledged.

References

[1] Zhan, L., Huang, J., Wang, C., et al. (2022). Three-dimensional digital presentation of red tourism resources based on landscape gene theory: taking the former residence of Mao Zedong and Yang Kaihui in Qingshuitang as an example. Tourism Journal, 37, 54-64.
[2] .Liu, P., & Deng, Y. (2017). Digitized protection: A new way to protect historical and cultural towns and villages. Journal of Peking University (Philosophy and Social Sciences), 54(6), 104-110.
[3] Nguyen, S. V., Le, S. T., Tran, M. K., et al. (2022). Reconstruction of 3D digital heritage objects for VR and AR applications. Journal of Information and Telecommunication, 6(3), 254-269.
[4] Allen, P. K., Troccoli, A., Smith, B., et al. (2003). New methods for digital modeling of historic sites. IEEE Computer Graphics and Applications, 23(6), 32-41.
[5] Cao, S., He, J., & Deng, Y. (2017). Map Features of Traditional Landscape Gene Symbols in Southern China—A Case Study of the Great Western Hunan. Economic

Geography, 37(5), 191-198.
[6] Li, C., Sui, D., Li, J., Li, Y., Li, Y., & Guo, M. (2025). Frequency and tensorial neural radiance fields for digital preservation of ancient buildings. npj Heritage Science, 13, 158.
[7] Fascia, R., Barbieri, F., Gaspari, F., et al. (2024). From 3D Survey to Digital Reality of a Complex Architecture: a Digital Workflow for Cultural Heritage Promotion. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 48, 205-212.
[8] Vasiljević, I., Kuzmanović, N., Draganić, A., Silađi, M., Obradović, M., & Obradović, R. (2025). A Case Study of the Katolička Porta in Novi Sad. Electronics, 14(19), 3785.
[9] Croce, V., Billi, D., Caroti, G., Piemonte, A., De Luca, L., & Véron, P. (2024). Comparative Assessment of Neural Radiance Fields and Photogrammetry in Digital Heritage: Impact of Varying Image Conditions on 3D Reconstruction. Remote Sensing, 16(2), 301.
[10] Clini, P., Perdicca, R., Nespeca, R., Angeloni, R., & Coppetta, L. (2025). Comparing MVS and Gaussian Splatting for the 3D Reconstruction of Reflective and Texture-less Cultural Heritage Artifacts//Digital Heritage 2025. Eurographics.