Journal of Trends in Arts and Humanities
ISSN: 3007-7370
DOI: https://doi.org/10.61784/jtah3047

DIGITAL TWIN RECONSTRUCTION OF MUSICAL INSTRUMENTS IN DUNHUANG MURALS: AN ENGINEERING WORKFLOW GROUNDED IN PHYSICAL REPLICAS

ZiLe Liu*, JiaYang Du
Lanzhou University of Arts and Science, Lanzhou 730070, Gansu, China.
Corresponding Author: Zile Liu, Email: 53636042@qq.com

Abstract: The Mogao Caves murals at Dunhuang preserve a dense iconography of music and dance from medieval times, yet the two-dimensional carrier and age-related degradation hinder quantitative understanding of instrument morphology. This paper proposes and validates a replica-based, 3D-capture-centered, visualization-oriented workflow for digital twin reconstruction. Using historically reasoned physical replicas as targets, we perform multi-sensor acquisition (structured light, photogrammetry), accurate registration, and high-fidelity meshing; we then conduct semantic retopology with lightweighting in parallel to physically based texture repainting. Representative samples—including the gourd lute (huluqin), large gourd lute, petal-edged ruan, guqin, bent-neck pipa, stick pipa, and round-bodied large ruan—are used to establish a traceable data lineage and metadata schema, demonstrating portability and dual effectiveness for exhibition and teaching. Results indicate that the method preserves mural-style cues while markedly improving geometric consistency and interactive legibility. The workflow provides an engineering-grade, reusable path for digital conservation, scholarly comparison, and public outreach of Dunhuang instruments.
Keywords: Dunhuang murals; Ancient musical instruments; 3D scanning; Photogrammetry; Semantic retopology; Unreal engine

1 INTRODUCTION

Instrument imagery in the Dunhuang murals is a primary source for studying medieval music and craft. Since the 1990s, physical reconstructions pioneered by Zheng Ruzhong and collaborators have helped clarify form and fabrication. In 2018, follow-on work by the Dunhuang Academy with industry partners refined shapes and materials, completing four major categories, 97 types, and 245 replicas [1-3]; in 2021, the "Hearing Dunhuang" initiative increased public visibility of these outcomes. Building on this corpus of replicas, we present a digital workflow—"replica as foundation, 3D capture as core, visualization as use"—that aims at engineering repeatability and scholarly verifiability [4-5].
In our pipeline, physical replicas provide a stable geometric baseline. We capture geometry via structured light and AI-assisted photogrammetry, perform semantic retopology and PBR repainting, and implement multi-endpoint interaction in Unreal Engine 5 (UE5) [6]: metric overlays, exploded views, deterministic camera paths, and side-by-side comparison. The present study focuses on geometric and visual recovery and does not address acoustic simulation. Our goal is a high-fidelity, cross-sample-portable procedure that serves exhibition and pedagogy while meeting reproducibility requirements for digital-humanities research (Figure 1).

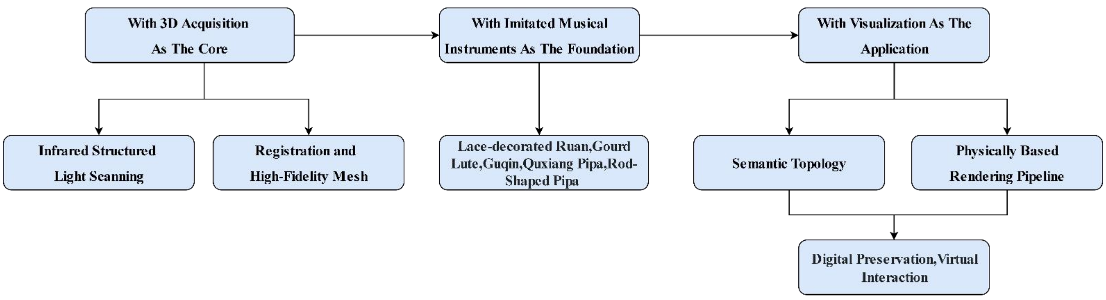

Figure 1 Digital Restoration Flow Chart

2 RELATED WORK AND RESEARCH GAP

Heritage digitization has evolved from manual polygonal modeling to measurement-based scanning. Manual modeling excels at stylistic synthesis but struggles with authentic tool marks and consistent high-frequency detail; scanning samples reality directly yet is sensitive to illumination, reflections, and occlusion. For Dunhuang instruments, the literature emphasizes iconographic scholarship and physical reconstruction. What remains under-specified is a reusable digital workflow that (i) is grounded in physical replicas, (ii) preserves mural stylistics without over-stylization, (iii) encodes traceable metadata, and (iv) packages a cross-device interactive template with measurable evaluation criteria. This paper addresses these gaps by balancing "authentic geometry—mural style—interactive legibility" with explicit thresholds and audit trails.

3 METHODS AND WORKFLOW

3.1 Acquisition Protocol: Multi-Sensor Complementarity

For small- to mid-scale instruments, structured light serves as the primary modality. We begin with far-range mode plus coded markers to capture the global silhouette; markers densely cover the underside to stabilize registration. We then switch to near-range mode and apply feature-based stitching to acquire carved details, reinforcing alignment between detail and global datasets. To mitigate occlusion and shadow, concentric rings of near/mid/far photographs at alternating pitch angles are captured for photogrammetric fusion. All data are imported into Revo Scan 5 [7], where far-range (global) and near-range (detail) sets are registered to form complete, high-accuracy point clouds (Figure 2).

Figure 2 Scanning is Performed Using Methods with Different Distances

3.2 Registration and Mesh Reconstruction

Point spacing during fusion is set to 0.2–0.3 mm to preserve carvings and curvature. Automated detection removes isolated and overlapping points; residual non-target clusters (e.g., matting powder flakes, background noise) are manually lassoed out. For meshing, the quality parameter is set to the maximum (8.0) to increase subdivision and detail capture. Holes induced by occlusion or specular glare are closed manually to ensure surface integrity, enabling stable downstream analysis and archiving.

3.3 Topology Optimization and Detail Transfer

Meshes are exported to Maya for retopology under the principles of clean edge flow and structural completeness, prioritizing evenly distributed quads and planned loops to support later sculpting and materials. Base meshes are then refined in ZBrush to restore carvings and surface texture using appropriate subdivision and brushes. The low-poly returns to Blender for UV unwrapping; shells are segmented by structural logic, with controlled stretch and optimized density. Low- and high-poly pairs are baked in Marmoset Toolbag 4 (normal, displacement, AO), with tuned sampling and cage settings to transfer detail while balancing render cost and fidelity.

3.4 Physically Based Texture Repainting

Baked maps feed into Substance Painter for PBR repainting. After close inspection of roughness distribution, metallic parts (e.g., tuning pegs), and characteristic surface patterns (carving grain, wear) [8], we assign region-specific roughness and metallic values and replicate distinctive textures to match physical behavior. Historic patina cues—non-uniform weathering, edge fading, and micro-polishing at hand-contact zones—are layered to encode "readable craft history" without exaggeration.

3.5 UE5 Interactive Scene and Scripting

To ensure reuse and verification, we formalize the mapping "mural instrument geometry—semantics—interaction" via an isomorphic tri-layer: Data (D), Logic (L), and View (V).

D-layer: A data-asset descriptor standardizes geometric parameters (e.g., soundbox axes, neck length, string spacing), disassembly levels, measurement baselines, and material channels. Traceable metadata (cave, date, panel, version, author, timestamp) support asset versioning, experimental replication, and auditability.

L-layer: A finite-state machine plus an event bus models the verifiable sequence "entry → search → focus → explode → measure → compare → back." Interactive objects are parametric Actors/Components; behaviors are instantiated through data-driven Blueprints to minimize hard-coded coupling and enable horizontal scale-out.

V-layer: Nanite hosts dense geometry and Lumen provides consistent global illumination. CineCamera and Level Sequence guarantee deterministic camera paths and reproducible viewpoints. Textures and meshes stream asynchronously with a lightweight loading widget to suppress perceived stutter. Cross-device input is normalized via Action/Axis Mapping; UMG/Canvas ensures resolution-independent UI.

Performance and interaction telemetry (frame time, draw calls, dwell time) plus contract-based assertions (geometric thresholds, material-channel completeness) close the loop "asset—script—render—evaluation," elevating scenes from engineering artifacts to measurable, reproducible research objects.

4 CASE IMPLEMENTATIONS AND RESULTS

We conducted reproducible experiments on six representative categories under a unified chain—acquisition, reconstruction, semanticization, rendering, and evaluation—with gates for geometric consistency (curvature fields, outline overlap, normal stability), material plausibility (complete PBR channels, continuous specular flow), interactive verifiability (auditable state transitions, parameter traceability), and process auditability (versioning and metadata completeness). Below we summarize working conditions, challenges, responses, and validation.

4.1 Petal-Edged Ruan (Mogao Caves 220/217)

hin petal rims with high specularity and sharp edges risk under-sampling, normal flips, and boundary oscillation. We applied light matting to reduce local specular components and used grazing-angle sidelighting. Near-range, high-density layered scans were rigidly registered; semantic retopology abstracted rims as ring-shaped edge loops, stabilizing exploded-view mechanics and specular-flow continuity. Results show congruent morphology (body diameter, petal width/density, short-neck ratio) with the murals; PBR layering renders lacquer–wood core–wear transitions naturally, and normal-flip rates remain within thresholds (Figure 3).

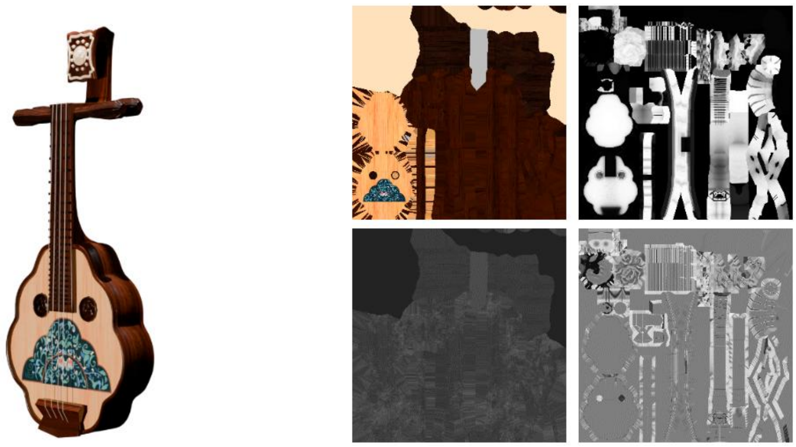

Figure 3 Rendered Images of the Petal-Edged Ruan and Four Texture Maps

4.2 Gourd Lute (Mogao Caves 322/420/262/423)

Weakly asymmetric double-curved soundboxes, compounded by hand-made replica deviations, can produce "bulging" artifacts under global smoothing. We used centroid–principal-axis constrained registration, reconstructed "gourd neck/belly" in sub-domains, and placed constraints at curvature maxima to suppress over-smoothing; a partitioned error field guided global merging. The process yields a traceable chain "mural parameters → physical deviation → digital fit." Exploded views reveal the assembly logic of tuning pegs—nut—soundboard, supporting cross-version auditing and parameter back-tracking (Figure 4).

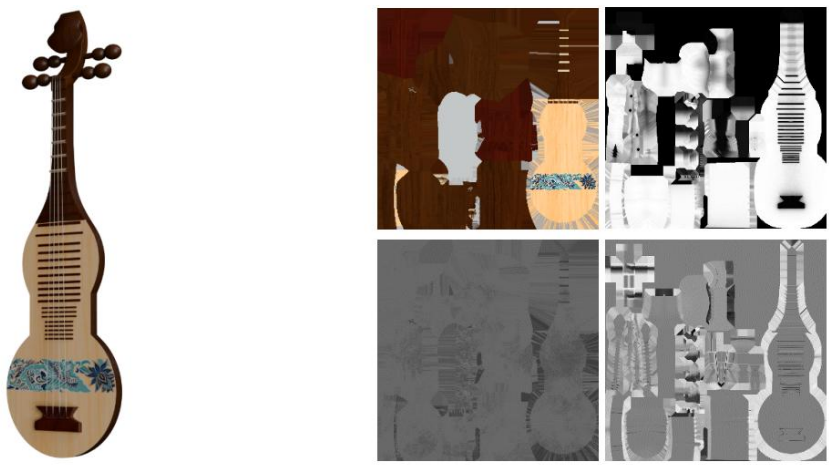

Figure 4 Rendered Images of the Gourd Zither and Four Texture Maps

4.3 Guqin (Mogao Caves 85/172/299)

Long, slender pieces with strong anisotropic reflections suffer local under-sampling and registration drift. We adopted a "segment scan—stitch—global optimize" strategy, using the thirteen hui as a linear baseline to impose sequential constraints. A soft constraint on the top-board camber maintained continuity, and local curvature priors reconstructed under-sampled regions. Measurements reproducibly recovered key indicators (camber profile, yueshan position, tuning-peg to panel proportions), supplying geometric evidence for typological evolution and a baseline for cross-sample comparison (Figure 5).

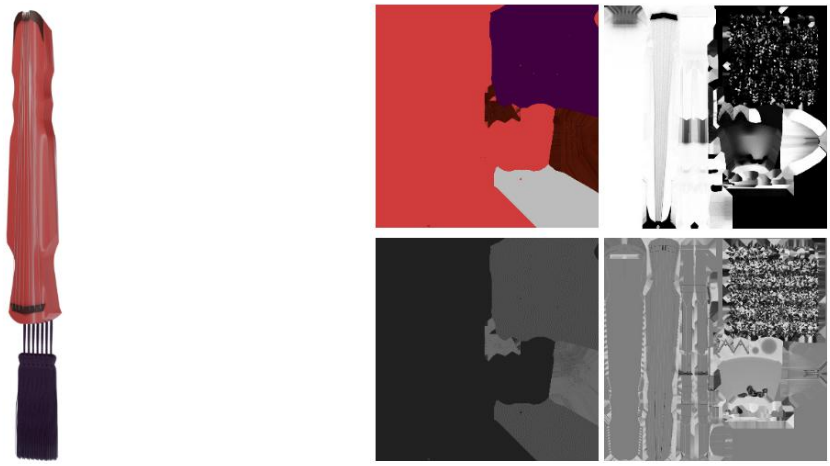

Figure 5 Guqin and Four Rendered Maps

4.4 Bent-Neck and Stick Pipa (Mogao Caves 112/220/313)

Thin curved necks with strong highlights destabilize normals; high curvature gradients at the "straight-neck—pear-body" transition induce artifacts. We employed local polarization imaging and multi-exposure fusion to suppress reflective noise. The bent neck was first fit with segmented NURBS under curvature-continuity constraints, then polygonized; conformal feature lines across the transition stabilized curvature fields. Parameters such as guard (hanbo), "phoenix eye," string spacing, and neck curvature are now comparable across forms. The frontend supports side-by-side comparison with parameter-linked queries, ensuring repeatable observation paths and camera solutions. Across "high-gloss thin walls, weak asymmetry, near-axis symmetry, long slenderness," we achieved (i) geometric consistency and normal stability within preset tolerances; (ii) complete craft-plausible PBR channels; (iii) auditable, replayable FSM-driven interactions; and (iv) traceable asset/process metadata. These validate portability to exhibition-and-teaching scenarios and establish a baseline for quantitative cross-cave, cross-type comparison (Figure 6).

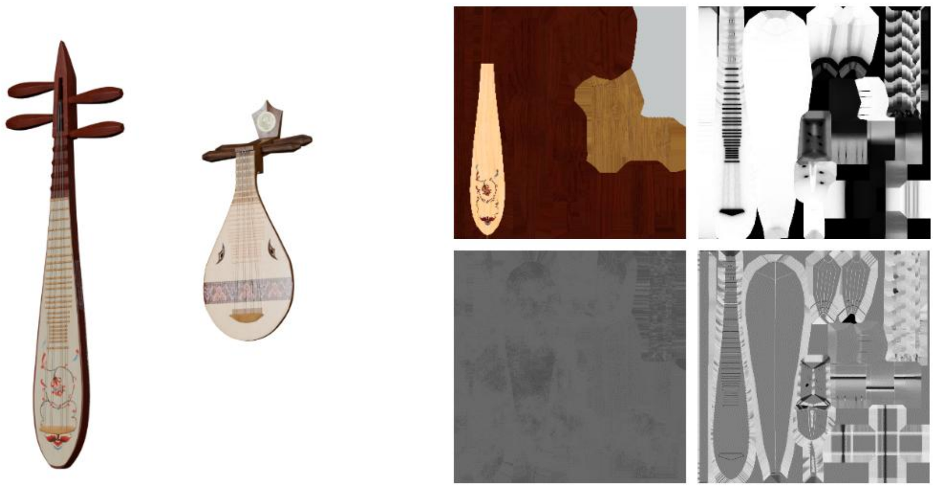

Figure 6 Quxiang Pipa and Rod-shaped Pipa

5 SYSTEM EVALUATION AND DISCUSSION

5.1 Visualization and Interaction Performance

In a 4K viewport, four to six instruments can be displayed side-by-side with stable frame cadence and smooth manipulation. On large touchscreens, gesture rotation/zoom is low-latency and controllable. In VR, in-place scaling and view bookmarks minimize large locomotions and alleviate oculovestibular mismatch. Nanite-based hosting removes manual LOD and baking, lowering asset maintenance and version evolution costs.

5.2 Reconstruction Consistency and Traceability

Geometric consistency. Using mural measurements and physical replica measurements as dual baselines, key parameters stay within engineering tolerances. For thin ornaments and sharp edges, visual and structural legibility is prioritized over acoustics-grade precision.
Pattern consistency. Textures are academically repainted as secondary creations; layer-level metadata and processing lineage enable provenance, correction, and re-editions.
Cognitive readability. Metric overlays, comparison viewports, and exploded diagrams create a closed loop of "observe—analyze—compare—retain," shortening the path from "image → morphology → construction" while improving transparency of the evidence chain.

5.3 Portability and Boundary Conditions

The method is most cost-effective for small- to mid-scale artifacts. Highly specular and ultra-thin parts still require matting plus polarization and close-range rescans to suppress normal instability and under-sampling. Photogrammetry remains sensitive to lighting uniformity; on-site shielding, controlled lighting, and reference calibration are recommended. Texture repainting should follow a "history—craft—style" tri-evidence chain to avoid misreading artistic exaggeration as material property. These boundary conditions should be encoded in acquisition protocols and error budgets when extrapolating to very large or composite artifacts.

6 CONCLUSIONS AND OUTLOOK

We present an integrated geometry–craft–interaction workflow for digital twin reconstruction of Dunhuang mural instruments. Grounded in physical replicas, the pipeline combines multi-sensor 3D capture with semantic retopology and PBR repainting, and delivers measurable, disassemblable, and comparable interactions in UE5 through Nanite/Lumen and data-driven Blueprints. Implementations across multiple plucked and struck instruments demonstrate engineering reusability and dual effectiveness for exhibition and pedagogy.
Future work proceeds along three lines: (1) multi-scale detail fusion using reflectance-transformation imaging and micro-topography to map tool marks and craft textures across scales; (2) spatiotemporal semantic labeling that introduces multidimensional tags ("cave—dynasty—panel—component") and narrative timelines to enhance scholarly retrieval and audience understanding; and (3) open asset management via a versioned digital instrument library and metadata standard to support cross-institution collaboration and long-term preservation.

COMPETING INTERESTS

The authors have no relevant financial or non-financial interests to disclose.

FUNDING

This work was supported by the Lanzhou Science and Technology Plan Project (2023-3-120).

REFERENCES

[1] Zou C, Rhee SY, He L, et al. Sounds of History: A Digital Twin Approach to Musical Heritage Preservation in Virtual Museums. Electronics, 2024, 13(12): 2388. DOI: 10.3390/electronics13122388.
[2] Sun T, Jin T, Huang Y, et al. Restoring Dunhuang Murals: Crafting Cultural Heritage Preservation Knowledge into Immersive Virtual Reality Experience Design. International Journal of Human–Computer Interaction, 2024, 40(8): 2019-2040. DOI: 10.1080/10447318.2023.2232976.
[3] Wheeldon D. 3D Digital Technologies for Studying Historical Musical Instruments. Journal of Musicological Research, 2024, 43(2-3): 197-221. DOI: 10.1080/01411896.2024.2376943.
[4] Liu J. From Traditional Perspectives to Design Innovation: A Study on the Digital Restoration and Redesign of the Konghou in Mural No. 465 of Mogao Grottoes. Highlights in Art and Design, 2024, 7(2): 78-81. DOI: 10.54097/c5f03020.
[5] Chen Y, Peng Y, Tan Y, et al. Achieving Cultural Heritage Sustainability Through Digital Technology: Public Aesthetic Perception of Digital Dunhuang Murals. Sustainability, 2025, 17(17): 7887. DOI: 10.3390/su17177887.
[6] Li Z, Zhang Q, Xu J, et al. Gamification of Virtual Museum Curation: A Case Study of Chinese Bronze Wares. Heritage Science, 2024, 12: 348. DOI: 10.1186/s40494-024-01464-2.
[7] Liu Z, Chen D, Zhang C, et al. Design of a Virtual Reality Serious Game for Experiencing the Colors of Dunhuang Frescoes. Heritage Science, 2024, 12: 370. DOI: 10.1186/s40494-024-01477-x.
[8] Di Iorio F, Fiocco G, Angeloni R, et al. Violins Unveiled: A Photogrammetric Framework Integrating Multiband and Spectroscopic Data for In-Depth Examination of Two Musical Instruments. Sensors, 2025, 25(11): 3278. DOI: 10.3390/s25113278.