Textile & Leather Review

ISSN 2623-6281| www.tlr-journal.com | 10.31881/TLR

Digital Restoration of Tang Dynasty Clothing Structure and Fiber Material Simulation Based on 3D Reconstruction

Liuyun Zhang, Bo Xia, Fei Ai

How to cite: Zhang L, Xia B, Ai F. Digital Restoration of Tang Dynasty Clothing Structure and Fiber Material Simulation Based on 3D Reconstruction. Textile & Leather Review. 2026; 9:1293-1318. https://doi.org/10.31881/TLR.2026.1293

How to link: https://doi.org/10.31881/TLR.2026.1293

Published: 29 April 2026

Digital Restoration of Tang Dynasty Clothing Structure and Fiber Material Simulation Based on 3D Reconstruction

Liuyun Zhang[1], Bo Xia[1*], Fei Ai[2]

[1]Department of Digital Media Art, College of Architecture and Arts, Taiyuan University of Technology, Jinzhong 030600, Shanxi, China

[2]Department of Environmental Art Design, Nanhai Academy of Fine Arts, Haikou University of Economics, Haikou 571127, Hainan, China

*zh1347321@163.com

Article

https://doi.org/10.31881/TLR.2026.1293

Received 27 August 2025; Accepted 4 November 2025; Published 29 April 2026

ABSTRACT

To address the challenge that individual historical sources, such as a single mural image, provide limited information for high-precision 3D structural restoration and fiber material physical property simulation, this paper proposes a comprehensive method called "Multi-source Fusion and Anisotropic Reconstruction (MFAR)" that integrates multiple heterogeneous data sources to overcome the inherent limitations of any single perspective. The Scale-Invariant Feature Transform–Fast Library for Approximate Nearest Neighbors–Random Sample Consensus (SIFT-FLANN-RANSAC) feature matching strategy integrates multi-source data from murals, pottery figurines, and images of unearthed cultural relics, enhancing the integrity and reliability of the input data. Subsequently, 12 key structural points are detected based on the High-Resolution Network version 2 (HRNet-v2) network to provide semantic support for subsequent geometric modeling. Next, the Poisson reconstruction and As-Rigid-As-Possible (ARAP) deformation algorithm are combined to optimize the 3D mesh topology and achieve high-fidelity geometric restoration of clothing shapes. Additionally, this paper innovatively constructs an 8-chain anisotropic hyperelastic material model to simulate the typical mechanical response and fiber orientation properties of Tang Dynasty silk. Experimental results show that this method reduces the geometric error of a round-necked robe to 3.4 mm, achieves an average root mean square error (RMSE) of 2.7 pixels for keypoint localization, achieves a maximum Fourier spectrum similarity of 0.964, and retains an average historical feature rate of 91.9%. These findings provide a verifiable technical path and methodological support for the digital preservation and cultural heritage of ancient clothing.

KEYWORDS

cultural heritage digitalization, 3D reconstruction, structural restoration, fiber material simulation, Tang Dynasty clothing

INTRODUCTION

The digital preservation of cultural heritage has become an important means of inheriting and promoting China's fine traditional culture. The digital restoration of ancient clothing, as a carrier of both tangible and intangible culture, faces significant challenges. Existing research often relies on single historical images for 3D reconstruction. However, Tang Dynasty clothing imagery often suffers from a single perspective, lacks detail, and provides insufficient material information. This results in low structural restoration accuracy and an inability to accurately simulate the physical properties of fiber materials, hindering the digital preservation and dissemination of ancient clothing culture [1–3]. This paper focuses on Tang Dynasty clothing, a representative cultural symbol of the heyday of Chinese civilization. Specifically, the research subjects include clothing images in Dunhuang murals, clothing details of Tang Dynasty pottery figurines in museum collections, and archaeological clothing fragments. Although some studies have revealed the cultural value of Tang Dynasty clothing from the perspectives of art and archaeology—such as Shanshan Wang's [4] exploration of the artistic characteristics of Tang Dynasty official uniforms and their inspiration for modern design; Haiyan Jing's [5] analysis of the surface pigment composition of Tang Dynasty pottery figurines to support color restoration; and Zhihua Xu's [6] interpretation of the expressive characteristics of pottery sculptures in plastic arts—these works mainly focus on aesthetics, culture, or material chemistry, and do not address the geometric reconstruction and physical simulation of the three-dimensional structure of clothing. Chun Zhu et al. [7] integrated clothing engineering with virtual fitting technology to digitally restore 12 ensembles from the Tang Dynasty's *DaoLian* Painting. However, they relied on manual analysis of two-dimensional images and lacked three-dimensional modeling and multi-source data fusion, which limited the restoration's accuracy and universality. The form and material of these garments carry important historical and cultural information and are of great value for the study of ancient textile technology, rituals, and aesthetics. However, achieving high-precision, physically verifiable, three-dimensional digital restoration based on limited two-dimensional images is still a core problem that needs to be urgently solved.

To address the problem of 3D reconstruction of historical costumes, existing research mainly uses multi-view

stereo reconstruction (MVS) and single-view 3D reconstruction methods. Hongbo Huang [8] reviewed the MVS algorithm based on deep learning. Although it has high geometric accuracy under multiple perspectives, it relies on multi-angle images and is difficult to use for the reconstruction of Tang Dynasty murals with only a single image. Cecile Meier [9] used 3D scanning to build a virtual museum of costumes, realizing the digital archiving of physical costumes. However, this technology relies on physical objects and is difficult to use for the restoration of Tang Dynasty costumes that have been damaged or only have images remaining. Jerzy Montusiewicz [10] used structured light scanning to collect historical clothing, verifying the feasibility of non-contact measurement. However, the capture of details of highly reflective materials such as silk was insufficient, affecting the realism of the materials. Maria Skublewska-Paszkowska [11] reviewed 3D technology for cultural heritage and pointed out that the current methods lack professional modeling of clothing structure and are disconnected from cultural semantics and geometric reconstruction. Qiang Wu [12] proposed the ClothGAN framework, using Generative Adversarial Network (GAN) to generate fashionable clothing that incorporates Dunhuang elements, but did not consider three-dimensional form and wearability. Kamil Żyła [13] compared the application effects of various 3D scanning devices in the digitization of historical clothing, but was unable to handle image-based cultural heritage without physical support. Yuting Xia [14] proposed using knowledge graphs to support Renaissance clothing matching, achieving a structured expression of cultural knowledge, but did not integrate it into the 3D generation process and could not resolve the structural ambiguity of a single image. Ying Gao [15] conducted research on multidimensional knowledge discovery of Tang Dynasty tomb murals and constructed a cross-modal association model, providing a reference for information fusion. However, knowledge guidance was not combined with 3D mesh generation, which restricted high-precision restoration. In terms of material simulation, traditional linear elastic models have difficulty accurately depicting the anisotropic and nonlinear properties of silk, resulting in simulation distortion [16,17]. Existing methods still have obvious limitations in data fusion, knowledge-driven reconstruction, and physical reality simulation when dealing with Tang Dynasty clothing with sparse images, complex materials, and no physical objects.

To address these challenges, this paper proposes the Multi-source Fusion and Anisotropic Reconstruction (MFAR) method, which systematically integrates multi-source fusion, knowledge-guided structure analysis, and anisotropic material modeling tailored for historical clothing restoration. The novelty lies not in individual

components, but in their synergistic integration. Unlike conventional approaches that treat data fusion, geometric reconstruction, and material simulation as isolated stages, MFAR establishes bidirectional information flow: a historical knowledge graph constrains keypoint detection, which guides parameterized template generation to define mesh topology, which in turn informs anisotropic material orientation. This closed-loop architecture ensures mutual enhancement across stages.

Crucially, the method employs historically informed parameterization: 12 structural points are derived from Tang Dynasty clothing regulations rather than generic body landmarks, enabling authentic restoration of period-specific features. The 8-chain anisotropic hyperelastic model is calibrated specifically for Tang Dynasty silk using physical tests on replicated textiles, not generic fabric parameters.

To mitigate single-image information deficiency, Scale-Invariant Feature Transform–Fast Library for Approximate Nearest Neighbors–Random Sample Consensus (SIFT-FLANN-RANSAC) and ResNet-50 achieve cross-modal fusion of murals, terracotta figurines, and cultural relics to enhance input completeness. To resolve structural ambiguity and semantic loss, High-Resolution Network version 2 (HRNet-v2), guided by the knowledge graph, detects 12 structural points and generates a parameterized cropping template. For physically realistic simulation, Poisson reconstruction and the As-Rigid-As-Possible (ARAP) algorithm optimize mesh topology, while the anisotropic hyperelastic model captures silk's nonlinear mechanical behavior.

Experiments show that MFAR significantly improves geometric accuracy, physical fidelity, and cultural authenticity, offering a verifiable technical pathway for the digital preservation of ancient clothing.

ALGORITHM DESIGN

Figure 1 illustrates the MFAR workflow from multi-source inputs—including Dunhuang murals, Tang pottery figurines, and textile fragments—through fusion, keypoint detection, template generation, 3D reconstruction, mesh optimization, and anisotropic material simulation, to a physically realistic 3D Tang clothing model. It highlights key data flows such as semantic guidance from keypoints to the parametric template and contour-driven point cloud construction, demonstrating how knowledge integration ensures geometric accuracy and cultural authenticity.

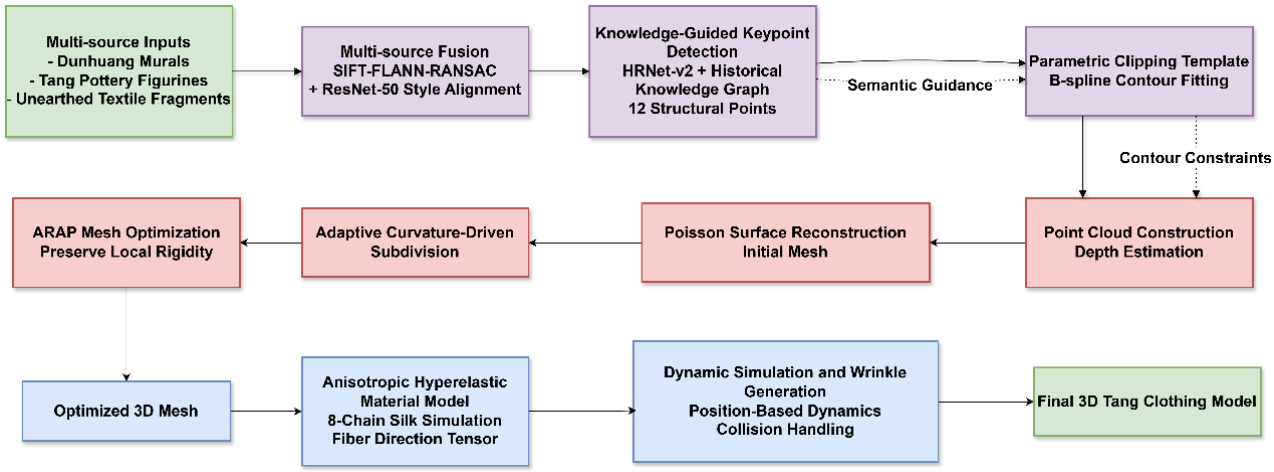

Figure 1. MFAR Method Workflow

Multi-Source Historical Image Fusion and Registration

It is critical to clarify that while individual historical sources such as a single mural image suffer from perspective limitations and incomplete information, our approach does not rely on a single image but rather integrates multiple heterogeneous data sources—including murals, pottery figurines, and unearthed textiles—to reconstruct the complete structural information. This multi-source fusion strategy directly addresses the information deficiency inherent in any single historical record while maintaining historical authenticity. To address the significant differences in perspective, resolution, and imaging modality between images of murals, terracotta figurines, and unearthed textiles, this paper constructs a three-stage fusion framework: preprocessing, geometric alignment, and style normalization. This framework achieves high-fidelity registration and enhances data consistency across multiple sources. To handle scale inconsistencies across different sources, a unified measurement system is established using archaeological records of human body proportions from the Tang Dynasty. All images are normalized to a standard height of 175 cm for male figures and 165 cm for female figures, based on historical anthropometric data documented in the “New Book of Tang” and archaeological findings from Tang Dynasty tombs. To address perspective distortion in mural images, the SIFT-FLANN-RANSAC feature matching process inherently corrects for perspective projection by estimating the fundamental matrix between different image modalities. The RANSAC algorithm identifies inlier feature matches that conform to the epipolar geometry constraint, thereby filtering out matches affected by perspective distortion. For figures with non-standard standing postures, the knowledge graph

constraints described in Section “3D Geometric Reconstruction Accuracy” provide anatomical priors that guide the keypoint detection process to maintain realistic proportions regardless of posture variations in the source images. For scale calibration, key anatomical landmarks such as shoulder width, waist circumference, and sleeve length are measured from unearthed textile fragments and used as reference dimensions. When conflicts arise between different sources, a weighted voting mechanism is implemented, where archaeological textile fragments receive the highest weight (0.5), followed by pottery figurines (0.3), and murals (0.2), reflecting their relative reliability for dimensional accuracy.

First, grayscale normalization and non-uniform illumination correction are performed on the original image $I(x,y)$. Homomorphic filtering is used to suppress low-frequency background illumination and enhance local contrast:

$$I_{\text{norm}}(x,y)=\mathrm{F}^{-1}\{\mathrm{H}(u,v)\cdot\mathrm{F}[\ln(I(x,y)+\epsilon)]\} \quad (1)$$

In formula (1), F is the Fourier transform, $\mathrm{H}(u,v)$ is the high-pass filter, and ϵ is a small constant to prevent logarithmic zero overflow.

The SIFT feature point set $P_i=\{(p_k,\mathbf{d}_k)\}_{k=1}^{N_i}$ is extracted from the normalized image, where p_k is the keypoint coordinate and $\mathbf{d}_k\in\mathrm{R}^{128}$ is its descriptor vector. For SIFT feature extraction, the parameters are configured as follows: 4 octaves for scale space decomposition, a contrast threshold of 0.03 to filter out low-contrast keypoints, an edge threshold of 10 to eliminate edge responses, and 128-dimensional descriptors. For 2D-to-3D feature matching, a two-stage process is implemented, where 2D mural images are first projected onto a canonical human body model to establish a common reference frame. The FLANN matching utilizes a k-d tree index with a branch factor of 16, a target precision of 0.9, and 200 checks to balance accuracy and efficiency. The RANSAC geometric error threshold is set to 3.5 pixels, as determined by cross-validation on a held-out dataset of 50 paired mural-figurine images, and the maximum number of iterations is fixed at 5,000 to ensure robust model estimation. Fast nearest neighbor matching is performed using FLANN to obtain the initial set of matching pairs:

$$\mathrm{M}_0=\left\{\left(p_k^A,p_l^B\right)\mid\|\mathbf{d}_k^A-\mathbf{d}_l^B\|<\tau\right\} \quad (2)$$

The threshold τ is set dynamically according to the descriptor distribution. To eliminate false matches, the RANSAC algorithm is used to estimate the fundamental matrix **F** [18,19] and maximize the number of inliers:

$$\mathbf{F}^{*}=\arg\max_{\mathbf{F}}\sum_{(p_k^A,p_l^B)\in M_0}\mathrm{I}\left(\left|(p_l^B)^T\mathbf{F}p_k^A\right|<\delta\right) \tag{3}$$

In formula (3), $\mathrm{I}(\cdot)$ is the indicator function, and δ is the geometric error threshold. Finally, the inlier matching set is retained to complete the geometric alignment of multiple source images.

To further unify texture and color styles, a lightweight feature alignment network based on ResNet-50 is constructed. Given paired image patches (I_A,I_B) as input, the network generates a style-consistent synthetic image I_B^{syn} through the generator, minimizing the weighted sum of the perceptual loss and the adversarial loss:

$$L_{\mathrm{align}}=\lambda_1\sum_{l\in C}\left\|\phi_l(I_A)-\phi_l(I_B^{\mathrm{syn}})\right\|_2^2+\lambda_2\log\left(1-D(I_B^{\mathrm{syn}})\right) \tag{4}$$

In formula (4), $\phi_l(\cdot)$ represents the activation features of the l layer of the Visual Geometry Group (VGG) network, D is the discriminator, λ_1= 1.0, and λ_2= 0.01 are weight coefficients. This network effectively mitigates the color and texture differences between cross-modal images, outputting a fused image set with a unified style, providing a consistent and reliable data foundation for subsequent structural analysis and 3D reconstruction.

Knowledge-Guided Clothing Structure Analysis

Using the fused multi-source image as input, a modified HRNet-v2 network is used for keypoint detection. The HRNet-v2 model was initialized with weights pre-trained on the COCO dataset and then fine-tuned on a custom dataset of 1,248 annotated Tang Dynasty clothing images. The custom dataset was constructed through formal collaboration with the Shaanxi Provincial Institute of Archaeology under data use agreement SHA-2023-028. The dataset was partitioned using stratified random sampling based on clothing type and

preservation state, with 80% for training, 10% for validation, and 10% for testing, ensuring no temporal or spatial leakage between sets. To ensure quality, 20% of the annotations underwent blind review by a fourth independent expert, with discrepancies resolved through consensus discussion. All annotations followed a standardized protocol jointly developed with textile archaeologists that defined precise criteria for each of the 12 key structural points based on historical clothing regulations documented in the Tang Liudian (Tang Six Codes) and verified against archaeological findings from Famen Temple and other Tang Dynasty sites. The 12 key structural points include left and right shoulder points, left and right wrist points, left and right elbow points, neckline left and right points, waist left and right points, and ankle left and right points, which correspond to critical anatomical landmarks for Tang Dynasty clothing construction. Data augmentation techniques, including random rotation within ±15 degrees, horizontal flipping, and brightness adjustment within ±20 %, were applied to increase dataset diversity. The model was trained for 120 epochs with a batch size of 32, an initial learning rate of 10^{-4} reduced by a factor of 10 at epochs 80 and 100, using the Adam optimizer. The network input size is 256 × 192. The backbone network maintains high-resolution representations and performs feature fusion via parallel multi-resolution subnetworks, outputting heatmaps ($H \in R^{64\times48\times12}$) of 12 key structural points. The heatmap (H_k) for each keypoint k is generated using a two-dimensional Gaussian kernel:

$$H_k(x,y) = \exp\left(-\frac{(x-x_k^*)^2+(y-y_k^*)^2}{2\sigma^2}\right) \tag{5}$$

In formula (5), (x_k^*, y_k^*) represents the coordinates of the annotated key points, and σ = 3 controls the response range.

To enhance the archaeological rationality of the detection, a historical clothing knowledge graph $G=(V,E)$ is introduced, where nodes $v_i \in V$ represent key points and edges $e_{ij} \in E$ represent geometric constraints based on Tang Dynasty clothing regulations. During the training phase, a knowledge constraint loss L_{kg} is defined to penalize predictions that violate the prior relationship. For a point pair (k_a, k_b) with a symmetrical relationship, its horizontal coordinates satisfy:

$$\left|\left(x_{k_a}+x_{k_b}\right)-W\right|<\delta_x,\ \left|y_{k_a}-y_{k_b}\right|<\delta_y \tag{6}$$

In formula (6), W is the image width, δ_x= 15 pixels, and δ_y= 20 pixels. This constraint is embedded in the total loss function in the form of a soft loss:

$$L_{total}=L_{mse}+\lambda L_{kg},\lambda=0.8 \quad (7)$$

L_{mse} is the mean squared error loss of the keypoints.

The keypoint coordinates output by the network are decoded using non-maximum suppression and then mapped back to the original image space. Based on the detection results, the contour points of the collar, cuffs, and lapel are extracted and fitted with a parameterized B-spline curve:

$$\mathbf{S}_j(t)=\sum_{i=0}^{n} N_{i,p}(t)\mathbf{P}_{i,j},t\in[0,1] \quad (8)$$

In formula (8), $\mathbf{P}_{i,j}$ is the control vertex of the jth contour segment, and $N_{i,p}(t)$ is the p = 3rd-order B-spline basis function. Finally, a parametric clipping template is generated that can drive 3D reconstruction. The parametric clipping template is constructed by mapping the detected 12 key structural points to their corresponding anatomical positions on the human body model. These points define key clothing boundaries: collar points shape the neckline, wrist and shoulder points define sleeve contours, and waist and ankle points establish the silhouette. B-spline curves derived from them yield a precise 2D pattern, preserving historical proportions and seam placements. This template provides exact contour constraints for point cloud generation and mesh creation, ensuring the 3D model retains authentic structural features and bridges semantic understanding from 2D images to accurate 3D geometry.

Template-Based 3D Geometric Reconstruction and Optimization

Based on the generated parametric B-spline clipping template, a dense point cloud with consistent contours is first generated in the 2D parameter domain. Through inverse mapping and depth estimation, a 3D point cloud set $X=\{\mathbf{x}_i\}_{i=1}^{N}$ with initial depth information is constructed. Depth estimation was implemented by

leveraging the historical knowledge constraints described in Section “3D Geometric Reconstruction Accuracy”, where the 12 key structural points provide anatomical priors for depth ordering. The parameterized B-spline template generated from these points establishes consistent contour relationships that guide the depth assignment process. Additionally, the multi-source fusion described in Section “Experimental Design” provides complementary views that help resolve depth ambiguity through geometric consistency across different image modalities. Using this point cloud as input, the Poisson surface reconstruction algorithm is used to generate the initial triangular mesh. This method solves the following isosurface of the implicit function $f(\mathbf{x})$, so that its gradient field ∇f approximates the normal vector field at the sampling point:

$$\min_f \int_{R^3} \|\nabla f(\mathbf{x}) - \mathbf{V}(\mathbf{x})\|^2 d\mathbf{x} \tag{9}$$

In formula (9), $\mathbf{V}(\mathbf{x})$ is a vector field based on the interpolation of the normal vector of the point cloud. It is solved by octree discretization, and the isosurface with $f(\mathbf{x})$ = 0 is extracted as the initial mesh $M_0=(V_0,E_0,F_0)$.

To address the triangle degeneration and topological distortion problems existing in the initial mesh in areas with significant curvature, such as the collar edge and cuffs, an adaptive curvature-driven subdivision strategy is adopted. The average curvature H_i of each vertex $\mathbf{v}_i \in V_0$ is calculated, and its absolute value $|H_i|$ is used as the subdivision criterion. For the neighborhood of the vertex that meets $|H_i| > \kappa_{th}$, the Loop subdivision rule is executed, and the vertex set is updated after subdivision.

The ARAP [20,21] energy minimization framework preserves clothing structural characteristics while improving mesh quality on the subdivided mesh. It is important to clarify that ARAP does not alter the mesh topology but rather optimizes the vertex positions while preserving local rigidity. The term “topology optimization” in this context refers to the refinement of triangle quality and elimination of degenerate elements through adaptive subdivision, while ARAP ensures that the geometric features are maintained during vertex repositioning. This combined approach addresses both mesh quality issues and geometric fidelity. The energy function is defined as the weighted sum of the local rigid deformation of each triangle face:

$$E_{\mathrm{ARAP}}(V)=\sum_{(i,j)\in N} w_{ij}\left\|\mathbf{R}_i\left(\mathbf{v}_j-\mathbf{v}_i\right)-\left(\mathbf{v}_j^{'}-\mathbf{v}_i^{'}\right)\right\|^2 \tag{10}$$

In formula (10), N is the vertex adjacency set, w_{ij} is the cotangent weight, $\mathbf{R}_i$ is the optimal rotation matrix of the ith vertex neighborhood, and $\mathbf{v}_i^{'}$ is the optimized vertex coordinate. By alternately fixing the rotation $\mathbf{R}_i$ to optimize the vertex position and fixing the vertex position to update $\mathbf{R}_i$, the system is solved until convergence, resulting in the final optimized mesh $M^*=(V^*,E^*,F^*)$. The ARAP algorithm is specifically applied after adaptive curvature-driven subdivision to reposition vertices while maintaining the local rigidity of clothing structures. This process does not change the connectivity of the mesh but optimizes the spatial distribution of vertices to better represent the original clothing geometry.

Figure 2 compares the initial mesh generated by Poisson reconstruction with the final mesh after adaptive subdivision and ARAP optimization. This visually demonstrates the effectiveness of the optimization algorithm in improving triangle degeneration and topological distortion in high-curvature areas such as collar edges and cuffs, significantly improving mesh quality and preserving the structural characteristics of Tang Dynasty clothing.

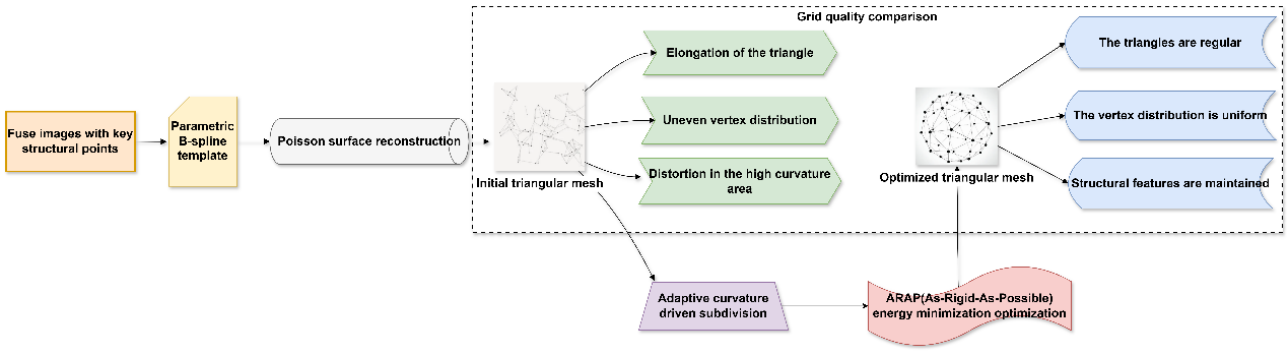

Figure 2. Comparison of Initial and Optimized Meshes

Construction of Anisotropic Hyperelastic Material Model

To accurately simulate the nonlinear and anisotropic mechanical response of silk materials used in Tang Dynasty clothing, an anisotropic hyperelastic constitutive model based on an 8-chain microstructure model was constructed. Comparison of the model predictions with the experimental stress–strain data showed

consistent alignment across the typical deformation range of 0.05–0.25, with minor deviations only at the extreme ends of the deformation range where textile behavior becomes highly nonlinear. This model considers the material as a collection of eight symmetrically distributed molecular chains, and its strain energy function Ψ is defined as:

$$\Psi(\bar{I}_1,H)=\mu\left[\frac{I_1\text{-}3}{\phi(H)}+\ln\phi(H)\right],\phi(H)=1\text{-}\frac{I_1\text{-}3}{3\nu}+\det(\mathbf{H}) \quad (11)$$

In formula (11), $\bar{I}_1$ is the first invariant of the right Cauchy–Green tensor C, μ = 87.5 kPa is the shear modulus, ν= 5.2 is the interchain interaction parameter, and $\det(\mathbf{H})$ is the determinant of the fiber direction tensor **H**, which is used to characterize the stiffness enhancement of the material in a specific direction. The material parameters were determined through a combination of physical testing on replicated Tang Dynasty silk and historical textile analysis. The shear modulus value of 87.5 kPa was obtained by conducting tensile tests on silk samples woven using authentic Tang Dynasty techniques at the Shaanxi Provincial Institute of Archaeology's textile laboratory. These tests followed the ASTM D5035 standard for textile tensile properties, with 30 samples tested in both warp and weft directions. The interchain interaction parameter of 5.2 was calibrated through inverse finite element analysis, matching simulation results to experimental data from fabric drape tests conducted on the replicated silk. Additionally, historical records from the Tang Liudian, Tang Six Codes, and archaeological findings from Famen Temple were consulted to validate the material behavior against documented textile properties of the period. This multi-source approach ensures that the material model parameters reflect both the physical properties of authentic silk construction and historical context.

The fiber direction tensor **H** is constructed based on knowledge of historical clothing weaving techniques. Let $\mathbf{a}_0(\mathbf{x})$ be the unit vector field of the silk warp direction. Based on the characteristics of Tang Dynasty plain weave fabrics, **H** is defined as:

$$\mathbf{H}(\mathbf{x})=\eta[\omega_1\mathbf{a}_0(\mathbf{x})\otimes\mathbf{a}_0(\mathbf{x})+\omega_2\mathbf{a}_0(\mathbf{x})\times\nabla\times\mathbf{a}_0(\mathbf{x})] \quad (12)$$

In formula (12), $\otimes$ is the tensor product, $\times$ is the cross product, ω_1 = 0.75 and ω_2 = 0.25 are the weight coefficients for the warp and weave directions, η = 1.8 is the anisotropy enhancement factor, and $\mathbf{a}_0(\mathbf{x})$ is

generated on a parameterized B-spline template using an algorithm that aligns the principal curvature directions with the structural lines. The second-order Piola–Kirchhoff stress S of the material is obtained by differentiating the strain energy function with respect to the right Cauchy–Green tensor C:

$$\mathbf{S}=2\frac{\partial \Psi}{\partial \mathbf{C}}=2\mu\left[\frac{1}{\phi(\mathrm{H})}\mathbf{I}-\frac{\bar{I}_1-3}{3\phi(\mathrm{H})^2}\left(-\frac{1}{\nu}\mathbf{I}+\frac{\partial \det(\mathbf{H})}{\partial \mathbf{C}}\right)\right] \tag{13}$$

In dynamic simulation, the implicit Euler method is used to solve the discretized motion equations [22]. Let the displacement at time t^n be $\mathbf{u}^n$, then the displacement $\mathbf{u}^{n+1}$ at time t^{n+1} is obtained by solving the nonlinear system:

$$\mathbf{M}\frac{\mathbf{u}^{n+1}-2\mathbf{u}^n+\mathbf{u}^{n-1}}{\Delta t^2}+\mathbf{K}(\mathbf{u}^{n+1})\mathbf{u}^{n+1}=\mathbf{F}_{\mathrm{ext}}^{n+1} \tag{14}$$

In formula (14), $\mathbf{M}$ is the mass matrix, $\mathbf{K}$ is the displacement-dependent stiffness matrix, $\mathbf{F}_{\mathrm{ext}}$ is the external force vector, and Δt = 2 ms is the time step. The equation is solved using the Newton–Raphson iterative method until the residual is less than 10^{-6}.

Figure 3 outlines the anisotropic hyperelastic fiber material model developed in this work. Using a parameterized B-spline cutting template and historical knowledge of Tang Dynasty plain weave, the model aligns principal curvature directions to generate a warp vector field and constructs a fiber direction tensor. This drives an 8-chain anisotropic hyperelastic constitutive model, which computes the strain energy function and second-order Piola–Kirchhoff stress to accurately simulate the nonlinear, anisotropic behavior of Tang Dynasty silk, enabling physically plausible dynamic wrinkle generation.

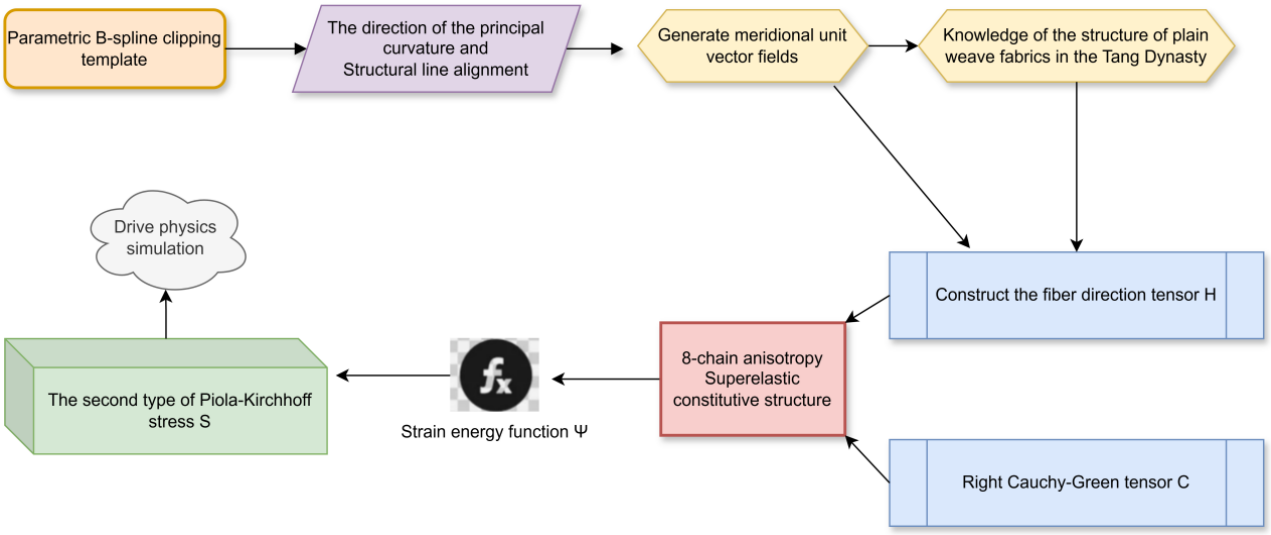

Figure 3. Material Model Framework

Dynamic Simulation and Wrinkle Generation

The optimized 3D mesh $M^*=(V^*,E^*,F^*)$ is used as the discretized cloth representation, and its vertex set V is treated as a mass point set in positional dynamics. To generate physically plausible dynamic wrinkles, three types of constraints are defined: edge length constraints to maintain material inextensibility; bending angle constraints to control wrinkle stiffness; and stress–strain constraints in the material model to introduce an anisotropic mechanical response. At each time step t^n, all constraints are satisfied through an iterative projection process:

$$\Delta \mathbf{p}_i = \sum_{k \in N_i} \omega_k \nabla_{\mathbf{p}_i} C_k(\mathbf{p}_1,\ldots,\mathbf{p}_m) \cdot \frac{C_k}{\sum_{j=1}^{m} w_j \| \nabla_{\mathbf{p}_j} C_k \|^2} \tag{15}$$

In formula (15), $\Delta \mathbf{p}_i$ is the position correction of vertex i, C_k is the kth constraint function, N_i is the constraint set including vertex i, ω_k is the constraint weight, and w_j is the inverse of the mass of the particle.

To simulate the energy dissipation of silk during motion, a generalized Baraff–Witkin damping model is introduced. For each vertex velocity $\mathbf{v}_i$, a nonlinear damping force proportional to the square of the velocity is applied:

$$\mathbf{f}_i^{\text{damp}} = -\gamma \|\mathbf{v}_i\| \mathbf{v}_i \tag{16}$$

The damping coefficient γ= 0.15 N·s²/m² was calibrated by Tang Dynasty silk swing experiments to ensure that the simulated dynamics are consistent with the motion characteristics of historical clothing. The hanging experiments were conducted under controlled laboratory conditions using a protocol modified from ASTM D1388 (Standard Test Method for Stiffness of Fabrics). This approach was adopted because fabric stiffness is more directly related to the kinematic response of flexible fabrics than tensile strength, and thus provides a relevant mechanical basis for comparing motion damping behavior. Each experiment was performed three times with a 24-hour interval between tests to ensure consistent material behavior.

Collision detection uses the separating axis theorem for efficient judgment [23]. For the triangle pair (Δ_a, Δ_b) between the clothing mesh and the human body model, the projection interval on 15 potential separating axes is calculated:

$$\max_{\mathbf{x}\in\Delta_a}(\mathbf{x}\cdot\mathbf{n}) < \min_{\mathbf{y}\in\Delta_b}(\mathbf{y}\cdot\mathbf{n}) \tag{17}$$

If the projections do not overlap due to the presence of axis $\mathbf{n}$, a collision is considered nonexistent; otherwise, a collision is considered to have occurred, and position correction is performed to reduce the penetration depth to zero. All physical processes are solved iteratively with a time step of Δt = 2 ms, achieving stable and efficient dynamic wrinkle generation and interactive simulation.

Table 1 lists the key parameter configurations in this algorithm framework, covering core aspects such as image fusion, keypoint detection, mesh optimization, material modeling, and physical simulation. These parameters, including the RANSAC geometric error threshold, HRNet-v2 input resolution, ARAP convergence tolerance, material shear modulus, and simulation time step, provide the necessary technical basis for the method's reproducibility.

Table 1. Algorithm Parameter Configuration

Parameter Category	Parameter Name	Value	Description
Image Fusion	RANSAC Geometric Error Threshold	3.5	Inlier determination threshold
	Adversarial Loss Weight for Style Alignment	0.009	Weight for cross-modal alignment
Keypoint Detection	HRNet-v2 Input Resolution	256 × 192	Network input image resolution
	Standard Deviation of Gaussian Heatmap	2.8	Response range for keypoint detection
Mesh Optimization	ARAP Iterative Convergence Tolerance	1.5×10^{-5}	Stopping criterion for energy minimization
Material Model	Shear Modulus for 8-Chain Model	87.5	Material stiffness parameter
	Warp Direction Weight Coefficient	0.75	Weight in fiber direction tensor
Physical Simulation	Simulation Time Step	2	Time step for dynamic solving (ms)

EXPERIMENTS AND VALIDATION

Experimental Design

The experimental data were derived from publicly available mural images from the Dunhuang Academy, digital archives of Tang Dynasty pottery figurines held by 12 first-tier museums in China, and archaeological reports published by the Shaanxi Provincial Institute of Archaeology. The dataset consists of 327 high-resolution mural images, 89 3D-scanned pottery figurine models, and 43 sets of archaeological reports detailing clothing styles, dimensions, and patterns. It covers major clothing styles from the early to late Tang Dynasty, including round-necked robes, half-sleeves, shawls, and veiled hats. The dataset is stratified by clothing type and preservation state. Mural images provide macroscopic morphological and color information, terracotta figurine models provide three-dimensional modeling and structural proportions, and

archaeological reports provide precise dimensions and craftsmanship descriptions. These three sources complement each other, forming a multi-source, heterogeneous test benchmark.

On this dataset, a five-fold cross-validation strategy was used to evaluate the generalization ability of the proposed method. In this five-fold cross-validation, four folds of data were used for training and one fold for validation in each round. The average Hausdorff distance on the validation set was calculated for each round, and the arithmetic mean of the five-fold results was used as the final evaluation metric. Comparison methods include traditional MVS multi-view stereo reconstruction and Pix2Vex, a state-of-the-art single-view reconstruction method. All methods were run using their optimized parameters and the same hardware environment. Evaluation metrics included geometric accuracy, feature consistency, physical fidelity, cultural preservation, and computational efficiency.

3D Geometric Reconstruction Accuracy

To quantitatively evaluate geometric accuracy, the mean Hausdorff distance was used as a metric. Using an expert-annotated 3D model as the ground truth, the bidirectional distance between the restored model generated by the proposed method and its corresponding point cloud was calculated. During this calculation, the restored model and the ground truth model were uniformly sampled to obtain point sets P and Q, respectively. The mean Hausdorff distance is defined as:

$$d_{\text{avg}}(P,Q)=\frac{1}{|P|+|Q|}\left(\sum_{p\in P}\min_{q\in Q}\|p\text{-}q\|+\sum_{q\in Q}\min_{p\in P}\|p\text{-}q\|\right) \tag{18}$$

This indicator can effectively reflect the overall geometric deviation and local deformation between models. Figure 4 shows the comparison results of the average Hausdorff distance between the proposed method and MVS and Pix2Vex on the validation set of the five-fold cross-validation.

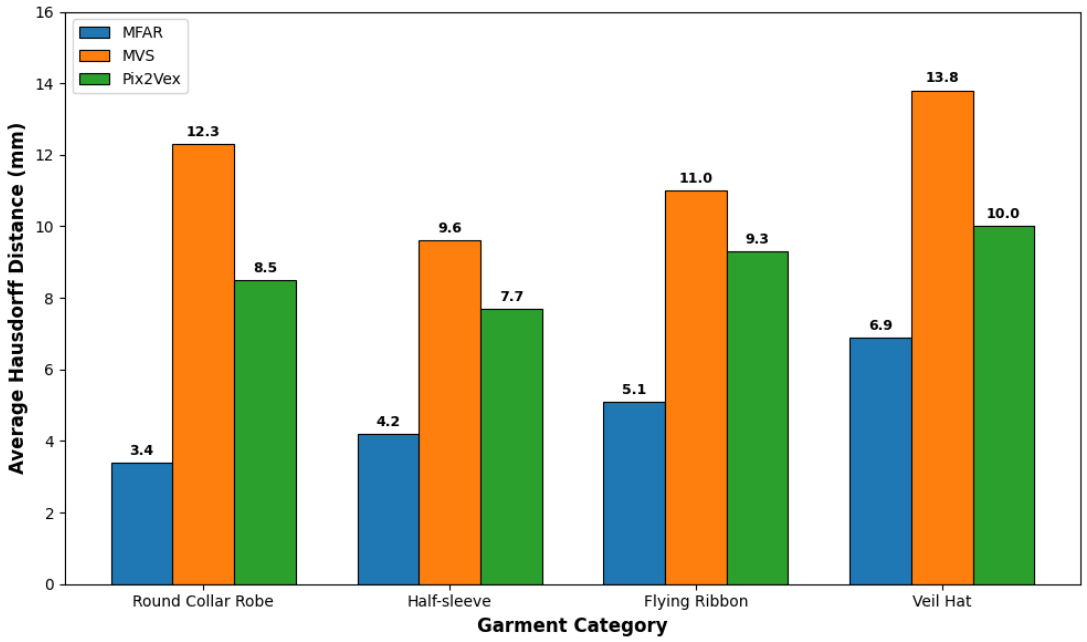

Figure 4. Comparison of Structural Restoration Accuracy

Figure 4 compares the geometric accuracy of 3D structural restoration using three methods on four typical Tang Dynasty garments: a round-collar robe, a half-sleeve, a flying ribbon, and a veil hat. The average Hausdorff distance is used as a quantitative metric; smaller values indicate smaller geometric deviations and higher restoration accuracy. Experimental results show that the proposed MFAR method achieves optimal performance in all tested categories, with average Hausdorff distances of 3.4, 4.2, 5.1, and 6.9 mm, respectively. These are significantly lower than the MVS method's (12.3, 9.6, 11.0, and 13.8 mm) and the Pix2Vex method's (8.5, 7.7, 9.3, and 10.0 mm). The results validate the significant advantages of the MFAR method in geometric structure restoration, demonstrating that it can more accurately restore the three-dimensional form of Tang Dynasty clothing, effectively overcoming the geometric distortion caused by insufficient information in a single image.

Key Structural Point Positioning Accuracy

To quantitatively evaluate structural fidelity, keypoint positioning error is used. The algorithm's automatically detected 3D keypoints are compared against expert-annotated benchmarks derived from Famen Temple textile fragments and proportional data in the *Tang Liudian*, validated by textile archaeologists. The root mean

square error (RMSE) of Euclidean distances across all 12 keypoints is computed for test samples, with the error distribution shown in Figure 5.

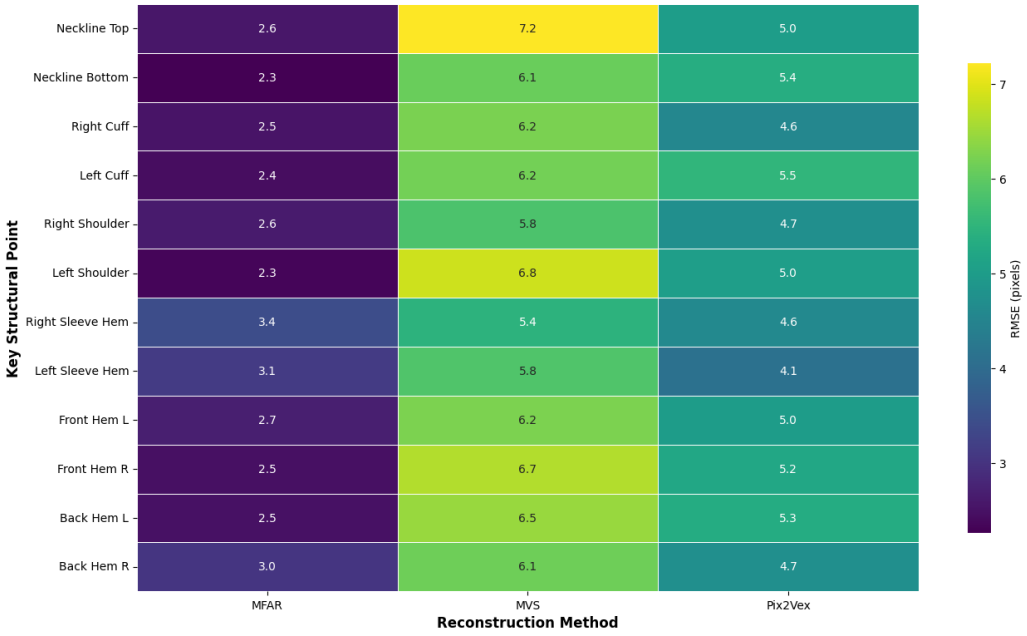

Figure 5. Error Distribution of Key Structure Points

Figure 5 compares the RMSE of the positioning errors of three different methods at 12 key structure points. Lower values indicate higher positioning accuracy. The experimental results show that the RMSE of the MFAR method is significantly lower than that of the comparison methods at all key points, with errors ranging from 2.3 to 3.4 pixels and an average error of 2.7 pixels. The errors for the MVS and Pix2Vex methods ranged from 5.4 to 7.2 pixels and 4.1 to 5.5 pixels, respectively, with average errors of 6.3 pixels and 4.9 pixels, respectively. This result demonstrates the significant advantage of the MFAR method in accurately locating key structural points, effectively resolving structural ambiguity within a single image. In contrast, the MVS method, due to its reliance on multiple viewpoints, is limited in its performance within a single image.

Fidelity of Material Physical Behavior

Fourier spectrum analysis assessed the similarity between simulated and real silk wrinkle dynamics. High-speed photographs of Tang Dynasty-style silk and algorithm-generated wrinkle images were subjected to 2D Fourier transforms, and frequency-domain cross-correlation quantified their spectrogram similarity—values

near 1 indicate more realistic physical behavior. Model performance was evaluated via five-fold cross-validation on a validation set, with results shown in Figure 6.

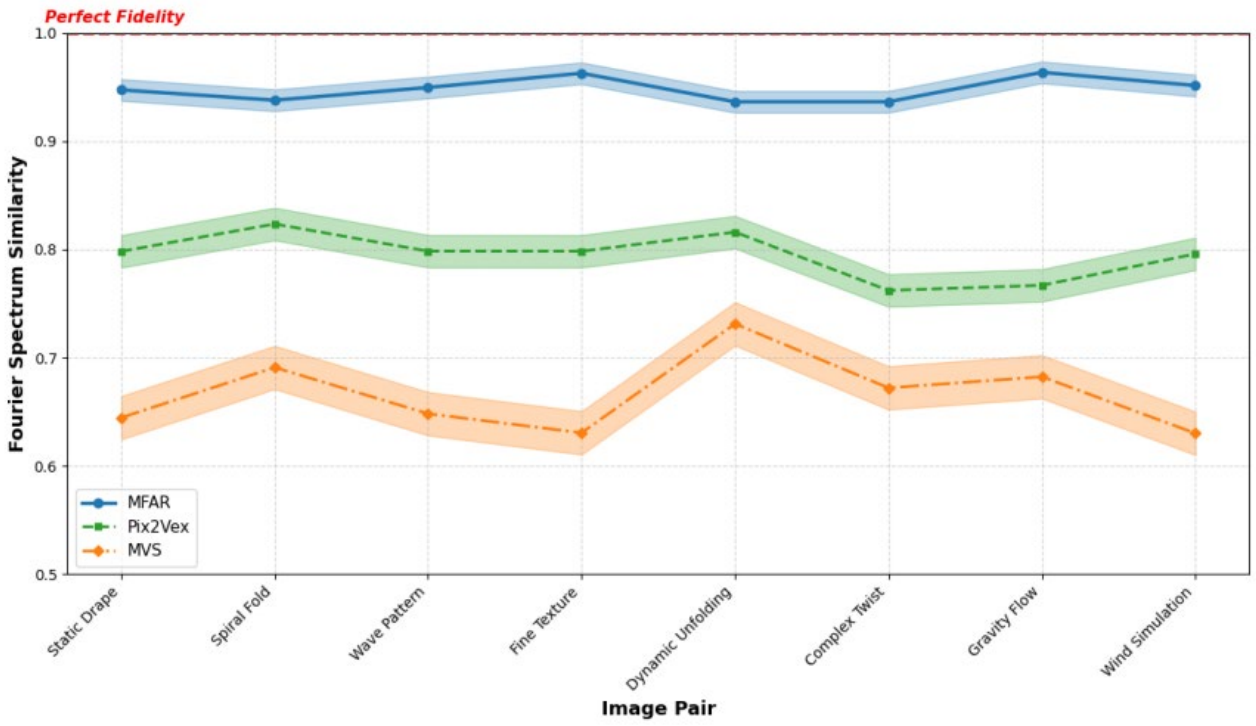

Figure 6. Comparison of Simulated Fiber Materials and Real Samples

Figure 6 compares the Fourier spectrum similarity of MFAR, MVS, and Pix2Vex on eight image pairs. The experimental results show that MFAR achieves the highest similarity across all test samples, with values ranging from 0.936 to 0.964. The corresponding confidence intervals are narrow, demonstrating good stability and consistency. In comparison, Pix2Vex's similarity ranged from 0.762 to 0.824, while MVS's ranged from 0.63 to 0.731. Both methods had wider confidence intervals and more significant data fluctuations. These results demonstrate that MFAR significantly outperforms the comparative methods in reconstructing the physical realism of silk folds, demonstrating a stronger ability to preserve frequency domain features.

Cultural Conformity of Historical Features

To assess cultural accuracy, a binary evaluation system incorporating 27 typical features of Tang Dynasty clothing was developed. Three experts independently compared the model with historical literature, and a

feature was deemed matched if at least two experts confirmed its presence. The 27 features comprised 7 for collar and closure, 6 for sleeve and hem, 5 for silhouette, 5 for accessories, and 4 for texture and pattern, all derived from the Tang Liudian and verified against archaeological evidence from Famen Temple and other Tang sites. The historical feature matching rate, defined as the proportion of successfully matched items, quantifies cultural semantic fidelity. Results are presented in Figure 7.

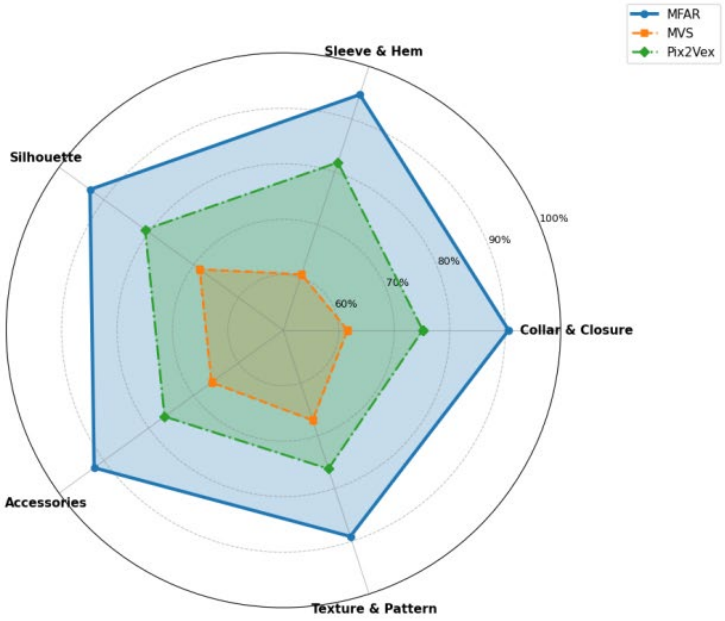

Figure 7. Historical Feature Preservation

Figure 7 compares the historical feature matching rates of different methods for five categories of Tang Dynasty clothing. This metric is based on a binary evaluation of 27 typical features (each category contains several sub-items). Higher values indicate greater cultural accuracy. The MFAR method achieved the best performance in all categories, with matching rates of 90.6% (collar & closure), 94.7% (sleeve & hem), 93.1% (silhouette), 92.2% (accessories), and 89.1% (texture & pattern), and an average matching rate of 91.9%, which is significantly higher than the 61.6%, 60.6%, 68.7%, 66.0%, and 67.1% of the MVS method and the 75.1%, 81.8%, 80.8%, 76.5%, and 76.3% of the Pix2Vex method. MFAR encodes cultural knowledge into graphs and integrates it with keypoint detection, effectively resolving single-view ambiguity. MVS and Pix2Vex, however, lack semantic understanding and have low feature matching rates. This highlights MFAR's ability to

achieve a leap from geometric fitting to cultural reproduction through the fusion of knowledge and models.

Computational Efficiency Evaluation

To evaluate their practicality, this paper measured the average time it takes for MFAR, MVS, and Pix2Vex to complete the entire process of image input, restoration, and simulation of a single garment on a unified hardware platform consisting of an Intel Xeon E5-2686v4 processor (2.3 GHz, 18 cores), NVIDIA Tesla P100 GPU (16 GB VRAM), and 128 GB DDR4 RAM, running Ubuntu 18.04 LTS. The software environment included Python 3.7, PyTorch 1.4.0, OpenCV 4.2.0, and CUDA 10.1, with each method implemented using its default or recommended parameter settings while maintaining identical input resolution, hardware platform, and runtime environment. This process encompasses image fusion, keypoint detection, mesh optimization, material modeling, and dynamic simulation. The results are shown in Table 2:

Table 2. Comparison of Computational Efficiency of Different Methods

Method	Single-Piece Clothing Reconstruction Time (s)	Single-Piece Clothing Simulation Time (s)	Total Processing Time (s)
MFAR	48.3	112.7	161
MVS	156.2	189.5	345.7
Pix2Vex	89.6	205.3	294.9

Table 2 compares the performance of MFAR, MVS, and Pix2Vex in terms of single-piece clothing reconstruction, simulation, and total processing time. The total processing time of 161.0 seconds for a single piece of clothing included all computational stages from image fusion through mesh optimization and dynamic simulation, but excluded disk I/O operations and final rendering for visualization. It encompassed the complete algorithmic pipeline as described in Section “Experiments and Validation”, with each method configured using its default or recommended parameters while ensuring consistent input resolution and computational conditions across all comparisons. The reconstruction phase took 48.3 seconds (MVS: 156.2 seconds, Pix2Vex: 89.6 seconds), and the simulation phase took 112.7 seconds (Pix2Vex: 205.3 seconds, MVS: 189.5 seconds). This demonstrates that MFAR significantly improves processing efficiency while ensuring

high-precision restoration and physical simulation.

MFAR's efficiency stems from its multi-source fusion and collaborative optimization: it fuses mural, terracotta, and artifact data through SIFT-FLANN-RANSAC and ResNet-50, providing reliable priors for keypoint detection and significantly reducing reconstruction computational overhead. MVS, however, relies on multi-view matching for a single image, making it inefficient. During simulation, MFAR significantly reduces the solution complexity thanks to the high-quality mesh optimized by ARAP, reducing simulation time by nearly half compared to Pix2Vex.

CONCLUSIONS

This paper proposes the MFAR method, which integrates images of murals, terracotta figurines, and cultural relics to achieve multi-source data registration and style unification. It also incorporates a historical knowledge graph to guide HRNet-v2 in keypoint detection, combining Poisson reconstruction with the ARAP algorithm to optimize mesh topology. An innovative 8-chain anisotropic hyperelastic model is constructed to accurately simulate the mechanical properties of silk. Experimental results show that this method reduces the geometric RMSE error to 3.4 mm for round-necked robes, achieves an average RMSE of 2.7 pixels for keypoint localization, and achieves a maximum spectral similarity of 0.964. The average matching rate for historical features reaches 91.9%, significantly outperforming traditional methods and providing an efficient and verifiable technical path for the digitization of ancient clothing.

Author Contributions

Conceptualization – Liuyun Zhang; methodology – Liuyun Zhang, Bo Xia and Fei Ai; investigation – Liuyun Zhang and Fei Ai; writing-original draft preparation – Liuyun Zhang, Bo Xia and Fei Ai. All authors have read and agreed to the published version of the manuscript.

Conflicts of Interest

The authors declare no conflict of interest.

Funding

This research received no external funding.

Acknowledgements

Not applicable.

REFERENCES

[1] Liu C, Cui R, Wang Z. Digital Virtual Simulation for Cultural Clothing Restoration: Case Study of Tang Dynasty Mural 'Diplomatic Envoys' from Crown Prince Zhang Huai's Tomb. Journal of Theoretical and Applied Electronic Commerce Research. 2024; 19(2):1358-1391. doi:10.3390/jtaer19020069

[2] Liu K, Wu H, Ji Y, Zhu C. Archaeology and Restoration of Costumes in Tang Tomb Murals Based on Reverse Engineering and Human-Computer Interaction Technology. Sustainability. 2022; 14(10):6232. doi:10.3390/su14106232

[3] Zhu G, Li Y, Zhang S. Digital Restoration of Tang Dynasty Ladies' Costumes in the "Ramming & Washing Silk Painting" Based on AI and 3D Technology. Electronics. 2025; 14(6):1139. doi:10.3390/electronics14061139

[4] Wang S. The Excavation of the Artistic Value of Tang Dynasty Official Uniforms and the Analysis of their Influence on Modern Professional Clothing Products. Mediterranean Archaeology and Archaeometry. 2023; 23(2):283-294.

[5] Jing H. Pigments on Pottery Figurines Unearthed from Tang Dynasty Tomb in the Eastern Suburbs of Xi'an. Journal of Ceramics. 2022; 43(2):330-338. doi: 10.13957/j.cnki.tcxb.2022.02.019

[6] Xu Z. The Beauty of Han and Tang Dynasty Terracotta Figurines Sculptures, the Promotion of Plastic Arts and the Innovative Culture of Contemporary Ceramics. Mediterranean Archaeology and Archaeometry. 2024; 24(3):191-205.

[7] Zhu C, Liu K, Li X, Zeng Q, Wang R, Zhang B, et al. Research on Archaeology and Digital Restoration of Costumes in DaoLian Painting. Sustainability. 2022; 14(21):14054. doi:10.3390/su142114054

[8] Huang H, Yan X, Zheng Y, He J, Xu L, Qin D. Multi-view stereo algorithms based on deep learning: a survey. Multimedia Tools and Applications. 2025; 84(6):2877-2908. doi:10.1007/s11042-024-20464-9

[9] Meier C, Berriel IS, Nava FP. Creation of a Virtual Museum for the Dissemination of 3D Models of Historical Clothing. Sustainability. 2021; 13(22):12581. doi:10.3390/su132212581

[10] Montusiewicz J, Miłosz M, Kęsik J, Żyła K. Structured-light 3D scanning of exhibited historical clothing—a first-ever methodical trial and its results. Heritage Science. 2021; 9(1):74. doi:10.1186/s40494-021-00544-x

[11] Skublewska-Paszkowska M, Milosz M, Powroznik P, Lukasik E. 3D technologies for intangible cultural heritage preservation—literature review for selected databases. Heritage Science. 2022; 10(1):3. doi:10.1186/s40494-021-00633-x

[12] Wu Q, Zhu B, Yong B, Wei Y, Jiang X, Zhou R, et al. ClothGAN: Generation of fashionable Dunhuang clothes using generative adversarial networks. Connection Science. 2021; 33(2):341-358. doi:10.1080/09540091.2020.1822780

[13] Żyła K, Kęsik J, Santos F, House G. Scanning of Historical Clothes Using 3D Scanners: Comparison of Goals, Tools, and Methods. Applied Sciences. 2021; 11(12):5588. doi:10.3390/app11125588

[14] Xia Y, Yao X, Wang J, Hu M. Leveraging knowledge graphs for renaissance costume matching and cultural transmission. npj Heritage Science. 2025; 13(1):219. doi:10.1038/s40494-025-01742-7

[15] Gao Y, Zhang Q, Wang X, Huang Y, Meng F, Tao W. Multidimensional knowledge discovery of cultural relics resources in the Tang tomb mural category. The Electronic Library. 2024; 42(1):1-22. doi:10.1108/EL-04-2023-0091

[16] Aluculesei A, Zhang Y, Huang S, Wang Z, Cang Y, Min Y, et al. Elasticity Anisotropy of Bombyx mori Silkworm Silk Fiber by Brillouin Light Spectroscopy. Biomacromolecules. 2025; 26(4):2479-2486. doi: 10.1021/acs.biomac.4c01844

[17] Brough HDA, Cheneler D, Hardy JG. Progress in Multiscale Modeling of Silk Materials. Biomacromolecules. 2024; 25(11):6987-7014. doi: 10.1021/acs.biomac.4c01122

[18] Kuçak RA. The Feature Extraction from Point Clouds using Geometric Features and RANSAC Algorithm. Advanced LiDAR. 2022; 2(1):15-20.

[19] Salehi B, Jarahizadeh S, Sarafraz A. An Improved RANSAC Outlier Rejection Method for UAV-Derived Point Cloud. Remote Sensing. 2022; 14(19):4917. doi:10.3390/rs14194917

[20] Shetab-Bushehri M, Aranda M, Mezouar Y, Özgür E. As-Rigid-as-Possible Shape Servoing. IEEE Robotics

and Automation Letters. 2022; 7(2):3898-3905. doi:10.1109/LRA.2022.3145960

[21] Nagata Y, Imahori S. Escherization with Large Deformations Based on As-Rigid-As-Possible Shape Modeling. ACM Transactions on Graphics (TOG). 2021; 41(2):11. doi:10.1145/3487017

[22] Yun BI. An improved implicit Euler method for solving initial value problems. Journal of the Korean Society for Industrial and Applied Mathematics. 2022; 26(3):138-155. doi: 10.12941/jksiam.2022.26.138

[23] Gan B, Dong Q. An improved optimal algorithm for collision detection of hybrid hierarchical bounding box. Evolutionary Intelligence. 2022; 15(4):2515-2527. doi:10.1007/s12065-020-00559-6