DIGITAL HERITAGE (2025)
S. Campana, D. Ferdani, H. Graf, G. Guidi, Z. Hegarty, S. Pescarin, and F. Remondino (Editors)

Material Studies For Digital Heritage: Comparative Analysis of Geometric and Photometric 3D Representations

Fabian Töpfer Eliane Christ Zhongyuan Yu Lars Engeln Matthew McGinity

Technische Universität Dresden, Germany

Figure 1: *3D Gaussian Splatting (photometric approach) of Historical Puppets featuring diverse materials and components, such as wood, rods, thin strings, plastics, shiny textiles and coated surfaces.*

Abstract

3D digitisation offers unique opportunities for research, preservation, distribution, presentation, and contextualisation of artefacts in museum and heritage contexts. However, despite advances in the field, no single contemporary capture method or representation can perfectly capture all artefacts for all purposes. For example, heterogenous materials, transparent surfaces, intricate structures, motion, and interactivity all remain significant challenges for contemporary 3D digitisation techniques. From the perspective of digital heritage, we present here an analysis of two prominent techniques: "structure from motion" surface-based "geometric" representations and surface-free "photometric" Gaussian Splat representations. We discuss the advantages and disadvantages of both with respect to digitisation workflows, usability, integrability, and practicality in museum contexts, with a particular focus on performative artefacts, and propose a goal-oriented guide for digital heritage professionals embarking on digitisation projects.

CCS Concepts

• ***Applied computing*** → ***Digital libraries and archives;*** *Arts and humanities;* • ***Computing methodologies*** → *Rendering;* ***Shape modeling;***

1. Introduction

Digital 3D representations play an essential role in the digital preservation of cultural heritage, offering museums new ways to document, research, and present artefacts [GBS14]. Among these, performative artefacts - artefacts whose meaning depends on their motion and contextual use (Fig. 1) - exemplify how digitisation enables advanced and novel forms of contextualisation and interactive presentation [TCY*]. However, the digitisation process poses major challenges, not only due to their possibly fragile condition, complex structures, and diverse materials, but also because of the inherent complexity of the actual digitisation process itself [IM21].

Ongoing research in the field of 3D representations extends the possibilities for digitising artefacts, each involving diverse approaches and yielding distinct results. Conventional digitisation and reconstruction techniques rely on mesh-based methods, as commonly employed in the field [GBS14]. Recent research has increasingly focused on the development and refinement of Neural Radiance Fields (NeRF) [MST*21], and, more recently, Gaussian

DOI: 10.2312/dh.20253050

Splatting (GS) [KKLD23] highlighting their potential and evolving capabilities. These approaches introduce novel characteristics that distinguish them from traditional methods [SWG24].

Building on this foundation, recent studies in digital heritage have begun to explore how these emerging techniques can be applied in practice. Early applications of NeRF have primarily focused on their usability for large-scale architectural sites and monuments [BPCO23, MKR*23, PBP*22]. Similarly, GS was already applied to the reconstruction of architectural structures [WHS*24], emphasising its potential for large-scale, static environments. Broader investigations into the use of NeRF and GS highlight its strengths in capturing objects that are difficult to reconstruct using SfM approaches, particularly due to its ability to represent complex details and realistic material appearances [CCDL*23, LYXC24]. Studies such as [CNAC24, BCG*24] systematically compare NeRF, GS, and Structure-from-Motion (SfM) approaches for architectural heritage, focusing on aspects such as geometric accuracy and visual fidelity. The photorealistic qualities of GS are emphasised [JB25], which determines its immersive potential in representing cultural heritage sites. In addition to spatial considerations, practical aspects of data acquisition and processing have been examined, including the impact of varying the number of source images [BGP*23] and image quality [CBC*24].

Despite the very recent interest in GS, their application within digital heritage remains limited, particularly for smaller, performative artefacts and their use in museums. These objects present unique challenges due to their complex structures and diverse materials, which go beyond those encountered in the digitisation of large-scale architectural subjects. As methods and developments in GS continue to evolve rapidly, this work seeks to explore the following questions:

- What are the specific requirements, workflows, and chances for digital performative artefacts [section 2]?
- What are the basic characteristics of GS representations [section 3] and how do they compare to traditional mesh-based models when applied to performative artefacts [section 4]?
- In what ways can these different 3D representations be both goal-oriented and sustainably embedded within standard museum workflows [section 5]?

To address these questions, this work adopts an object-centered approach, focusing on the performative artefact itself. So here, it starts with examining performative artefacts and their specific requirements. Next, common tasks within museum practice involving digital artefacts are identified and analysed in relation to the corresponding types of 3D data. Subsequently, the fundamental characteristics of geometric and photometric 3D representations are introduced. A comparative analysis is then conducted between SfM-based meshes and GS representations using selected case studies. These are evaluated against key criteria: dataset and processing, geometric accuracy, visual fidelity, and interactivity and dynamics. In the final step, insights from the comparative analysis are synthesised with the identified requirements to map digital representations to relevant museum tasks. The paper concludes by proposing a goal-oriented and object-centered guide tailored to performative artefacts and offering a discussion on the usability and adaptability of these digitisation practices.

2. Digital Performative Artefacts

Performative artefacts are objects whose use involves physical manipulation and motion. Examples include craftsmen's tools, musical instruments, or children's toys. For heritage, such objects pose a challenge, as their true value and purpose are only revealed when they are performed or used. For example, the value or purpose of a scientific instrument or ritual object may only be apparent when it is in motion or used in context. It may be that only when we use something ourselves, with our own hands, do we truly understand it.

In this paper, historical puppets are used as a case study, including for instance marionettes, hand puppets, and rod puppets, see Fig. 1. These objects often consist of multiple components made from various materials, ranging from shiny to matte surfaces and from rigid to soft or deformable parts. These materials may include metal, wood, textiles, thin strings, and various types of plastics. Digitising historical puppets as performative artefacts enables advanced forms of contextualisation, allowing for their use and movement in virtual environments despite their typically fragile physical condition. However, contextualising such artefacts within the scope of a performance introduces additional layers of complexity, involving the performer, their intentions, and the specific settings in which the artefacts were used [TCY*].

For the sake of clarity and focus, the following discussion adopts an object-centered approach, rather than performance-centered, concentrating on the digitisation of the performative artefact itself. Within this scope, various conditions significantly influence the possibilities and design of digitisation workflows. Broadly, performative artefacts can be categorised as follows.

1. Artefacts that can still be moved
2. Artefacts that can no longer be moved due to their fragility
3. Artefacts that have partially or entirely missing parts

A digital replica of a performative artefact enables (1) non-invasive exploration, (2) re-animation, and (3) the reconstruction of missing elements, thereby extending the scope of traditional museum practices. Moreover, a digital workflow facilitates the documentation of holistic performances and opens up new ways for re-contextualisation and artistic reinterpretation of digital historical puppets within contemporary settings—particularly when further involved performative elements are taken into account.

2.1. Museum Workflows for Digitisation

In general, there are two primary approaches to digitisation of artefacts: manual modelling and capture-based methods. Manual modelling offers a high degree of flexibility and can be tailored to specific use cases. However, it is labour-intensive and therefore difficult for large-scale digitisation efforts. Furthermore, it is difficult to verify or guarantee the accuracy of hand-made digital models. In contrast, capture-based approaches, such as SfM-based meshes or GS representations, offer more efficient workflows, enabling up to fully automated 3D digitisation workstations for suitable artefacts [SRD*25]. Within these methods, various techniques (geometric and photometric) exist for capturing and visualising digital artefacts, each suited to different tasks and requirements.

Traditional capture-based pipelines for generating 3D models typically involve capturing raw images, aligning feature points, calculating dense point clouds, integrating mesh geometry, and generating textured models [Bed17]. These methods predominantly result in geometric, mesh-based representations, which are widely used in museum contexts. Museums thus operate on two interconnected yet distinct levels: the physical artefact and its corresponding 3D digital representation. (Fig. 2).

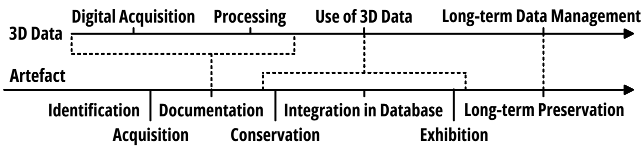


Figure 2: *Artefact and 3D Data Workflows operate on two distinct but interrelated levels.*

At the artefact level, workflows typically involve identifying and selecting objects based on cultural and curatorial value, acquiring them into collections, documenting their provenance and condition (including photography and 3D models), implementing conservation measures, integrating them into digital systems, preparing them for exhibition, often with digital enhancements, and ensuring their long-term preservation [GKK24].

At the data level, workflows adapt to the specifics of digital material and the characteristics of the artefact: selecting objects for 3D digitisation, acquiring data through various capture methods, processing models for specific applications, using them in research, restoration, education, or exhibitions, and ensuring long-term storage and accessibility through standardisation, publication and archiving [GKK24]. The use of the 3D data can vary widely, from analysis and comparison to precise manipulation and immersive contextual experiences [GR24]. Each involves distinct interactions with digital artefacts, ranging from orientational (e.g. targeting, tracking, exploring) to operational (e.g. editing, object manipulation, comparative analysis) [Gro17]. The following chapters explore the suitability of various types of 3D representations for the specific tasks involved in these workflows.

By integrating 3D representations within these workflows, museums can enhance their core tasks: accurate documentation, better preservation strategies, increased access for research and education, and novel forms of visitor engagement such as interactive exhibitions or online platforms [GGH*24]. Despite these opportunities, significant challenges remain in the digitisation process, particularly due to the complex materiality and structure of performative artefacts. Digital skill gaps, questions of authenticity, and the scalability of 3D solutions all affect implementation. Successful applications depend on solutions that are low-tech, easy to use, adaptive, and integrative with existing museum infrastructure and media standards [IM21, GKK24, EUr24].

3. 3D Representations

Digital artefact representations can be categorised as either geometric, focusing on spatial structure, or photometric, capturing visual appearance such as colour and light. These approaches serve different purposes in digital heritage based on their distinct characteristics.

3.1. Geometric 3D Representations

Geometric 3D representations describe an object's geometry using explicit three-dimensional structures. These can be either surface-based such as polygonal meshes or volumetric, like point clouds or voxels (volumetric elements) [FvDHF90]. Point clouds store spatial positions along with colour information for each point, but lack structural connectivity. In contrast, meshes consist of interconnected vertices, edges, and faces that together define a structured surface geometry. This makes them particularly effective for computing precise surface normals. Additionally, meshes support surface texturing, where 2D bitmaps encoding colour, normals, height, specularity, smoothness, occlusion, and metalness are wrapped onto the surface to enhance visual fidelity [GNBL13]. Their structured format also enables efficient geometry simplification, compact data storage, and ease of manipulation, making them especially suitable for animation and simulation tasks [AB17].

One approach for creating geometric representations is manual modelling by a "3D artist", typically yielding polygonal models with materials and textures. Such workflows have a number of advantages. For one, they often employ "physics-based" representations of surface materials, in which materials are modelled in terms of albedo, roughness, metalness and surface-normal deviations, allowing realistic simulation of surface-light interactions in arbitrary contexts. Secondly, the surface topology and segmentation can be tailored specifically for subsequent animation of the artefact, be it through manual key-frame animation or physics-based simulation. However, manual modelling is very time-intensive, demands expert training, and it is difficult to ensure or even verify the accuracy of hand-made models.

Another method is Structure from Motion (SfM). It is a photogrammetric technique that reconstructs three-dimensional models of static objects from overlapping digital images taken from multiple viewpoints. By matching features across these images, SfM estimates both the camera positions and the object's geometry through triangulation and bundle adjustment, initially producing a sparse point cloud which can be densified and converted into a textured mesh [Bed17]. SfM-derived point clouds and meshes have become standard 3D representations for documenting and visualising cultural heritage artefacts [CDR15, CNAC24]. Controlled lighting and systematic image capture further enhance model quality and accuracy [Bed17].

3.2. Photometric 3D Representations

In contrast to geometric approaches such as SfM, where the goal is to extract an explicit description of the geometry of an object, *photometric* approaches seek representations that produce photorealistic *images* of an object, without necessarily generating accurate

Proceedings published by Eurographics - The European Association for Computer Graphics.

representations of the geometry of the object. Examples include photographs, videos, light field rendering, Neural Radiance Fields (NeRFs), and Gaussian Splatting (GS).

A single photo does not represent a 3D structure of an object, rather, it offers a depiction from a specific viewpoint. While a viewer can infer depth, three-dimensional information remains absent. Multiple photos from different angles add view-dependent cues, but projecting them onto 2D surfaces still loses spatial data. Similarly, video recordings introduce temporal dynamics but still lack inherent spatial structure.

Light field rendering adopts a similar principle. The scene is captured through a dense set of images and novel views are synthesised by combining, resampling and interpolating these captured perspectives [Han96]. Rather than reconstructing the scene's structure, it is essentially represented through a 4D function, the light field. In comparison, further advancements in novel view synthesis have introduced neural networks to generate a scene's appearance and support free-viewpoint camera movement. Unlike light field rendering, NeRFs learn the behaviour of light within a scene using a neural network trained on a set of input images [MST*21]. NeRF models scenes as volumetric functions, enabling photorealistic 3D reconstructions and smooth view synthesis, particularly useful for complex materials in digital heritage [CCDL*23]. Despite recent acceleration efforts [MESK22,GKJ*21], it remains more computationally demanding than point-based methods [RWW*25].

Gaussian Splatting (GS) [KKLD23] on the other hand, is a hybrid version. It combines the ideas of view-dependent renderings with a explicit spatial structure. This allows not only for real-time rendering, but also enables realistic and immersive scene depictions. Starting from a set of input images, initial 3D Gaussian ellipsoids are derived from the sparse point cloud generated by SfM, therefore using the same pre-processing steps as for conventional SfM-based meshes [KKLD23]. The second step is the optimisation. Based on differential rendering, the parameters of these Gaussians, including their positions, rotations, scaling vectors, opacity, and spherical harmonics coefficients, are iteratively adjusted until rendered images of the splats appear photometrically similar to the original images. The spherical harmonics coefficients ensure a view-dependent rendering through a rasteriser, which splats the 3D Gaussians onto a 2D image plane, which in comparison to NeRF makes it much more efficient and usable for real-time use. Therefore, in the following we will compare GS with SfM-based meshes. Analogous to extending a photo into a video, GS can be expanded into the temporal domain by incorporating a time-dependent component to the 3D Gaussians [LCLX24]. This enables the reconstruction of immersive dynamic scenes. Here, *Spacetime 3D Gaussians* are interpreted by multi-level perceptron (MLP) rather than spherical harmonics coefficients, making them both time- and view-dependent. Capturing such scenes typically requires complex and carefully synchronised camera setups. However, there are approaches reducing the necessary views, making it more accessible [LLW*25].

From the perspective of digital heritage, the key insight is that photometric approaches like Gaussian Splats can produce highly realistic renderings, without *necessarily* producing accurate geometric structure. They are particularly effective at reproducing material properties in a way that conventional digitisation techniques (e.g. SfM) are unable to achieve [CCDL*23].

4. Comparison of SfM-based Meshes and Gaussian Splatting

The differences between SfM and GS in dataset requirements, processing, geometric accuracy, visual fidelity, and workflow integration are examined using historical puppets with diverse features and materials. Both representations are generated from the same dataset for a direct comparison. Camera registration was performed using Agisoft Metashape Professional Version 2.2.0 build 19890[†]. From there, the masked image sets and corresponding camera poses were exported to ensure consistency when generating the GS representations. The GS representations were created using PostShot[‡] with the MCMC profile and subsequently cleaned using Superspplat[§]. For rendering, Blender[¶] in combination with the Kiri Engine plugin[‖] was used. All models were relit under identical lighting conditions. Additionally, the rendering perspectives were deliberately chosen to be outside the original training views. This is especially relevant for GS, which could not optimise for these unseen angles, offering a fairer basis for comparison.

4.1. Dataset and Processing

The dataset consists of images of a hand puppet with rigid, soft and deformable parts. Images were captured using a Canon EOS 700D with an 18–55 mm lens (focal length: 23 mm, aperture: f/11, ISO 100) and flashlight. The camera is equipped with a CMOS sensor (22.3 × 14.9 mm) at a resolution of 5184 × 3456 pixels. A total of 400 images were captured under constant acquisition parameters, without any changes to the setup or camera settings. To ensure complete coverage, the camera was systematically moved around the stationary object during image acquisition.

In general, the quality of both GS and SfM representations is strongly influenced by the capture process [RWW*25,Bed17]. SfM pipelines usually aim for a high degree of overlapping (60-80%) of high-resolution images to ensure sufficient feature matching and coverage, especially for smaller objects [Bed17]. In GS, the capturing process affect the quality and density of the resulting splats [RWW*25]. Furthermore, GS implementations vary in their image requirements: while some necessitate a similar capture process, recent developments demonstrate promising results even with as few as 2–5 input images [FCW*24]. This raises questions about the performance of both techniques under constrained capture conditions (Fig. 3).

For the high-quality model, $\mathrm{SfM}_{\mathrm{HQ400}}$, 400 images were used as input. The images were aligned with the highest accuracy, followed by gradual selection for tie point filtering and camera optimisation [Agi25,Esc19]. A dense point cloud and model were generated from depth maps (quality: ultra-high, filtering: mild), with

† https://www.agisoft.com/
‡ https://www.jawset.com/
§ https://superspl.at/editor
¶ https://www.blender.org/
‖ https://www.kiriengine.app/blog/announcement/kiri-releases-3dgs-addon-for-blender

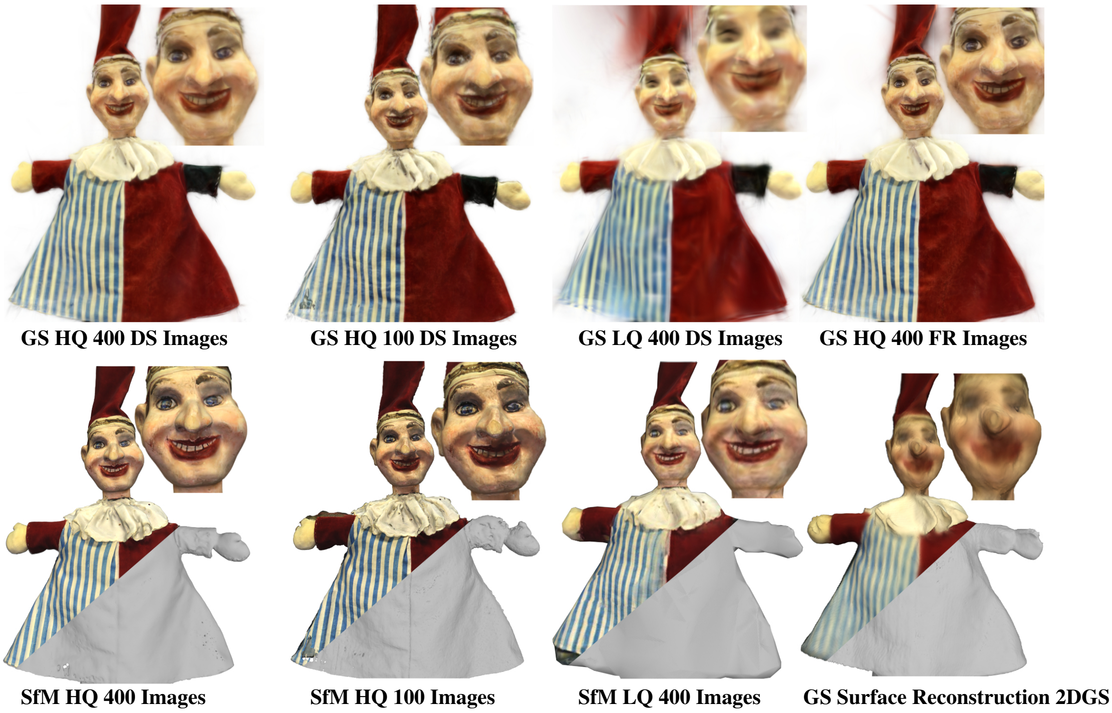


Figure 3: *Comparison of different Capturing and Processing Methods for SfM mesh-based Models and Gaussian Splats (GS). PostShot with 40,000 Steps (HQ) and 10,000 Steps (LQ) as well as Agisoft Metashape was used to create the models. The study also compares GS results using downsampled 1600×1066px (DS) vs. full-resolution 5184×3456px (FR) images, focusing on visual differences and GS surface reconstruction based on 2DGS [HYC*24].*

	SfM			GS		
	HQ_{400}	HQ_{100}	LQ_{400}	HQ_{400}	HQ_{100}	LQ_{400}
Processing Time [h:min:s]	01:08:36	00:22:02	00:06:57	00:41:18	01:11:37	00:08:15
Number Points/Splats	20,337,426	14,185,363	942,711	938,402	2,382,056	61,931
File Size	85.7 MB	45.3 MB	862 KB	232.7 MB	370.2 MB	15.4 MB

Table 1: *CPU: 13th Gen Intel(R) Core(TM) i7-13800H; GPU(s): NVIDIA GeForce RTX 4050 Laptop GPU; HQ/LQ_{400}:398 images, HQ_{100}: 97 images, SfM (texture: 4K, files: *.fbx + diffuseMap.jpg + normalMap.jpg), GS (file: *.ply).*

	$GS_{fullRes}$	$GS_{downsampledRes}$
Processing Time [h:min:s]	00:40:36	00:07:45
Image Resolution [px]	5184 x 3456	1600 x 1066
File Size [MB]	229.8	232.7

Table 2: *CPU: AMD Ryzen 9 7950X 16-Core, GPU: NVIDIA GeForce RTX 4080 Super, Processed using PostShot.*

partial masking applied to exclude unwanted features. A 4K texture was then created. The same workflow was applied to the SfM_{HQ100} model, using a reduced dataset (1/4 of the original). For the low-quality model, SfM_{LQ400}, the full dataset was used, but with adjustments to the alignment settings (low accuracy), followed by gradual selection for tie point filtering and camera optimisation. A dense point cloud and mesh were subsequently created from depth maps (quality: low, filtering: aggressive), and a 4K texture was applied. Masked images of the SfM models were automatically generated in Agisoft Metashape, and the camera poses were exported to be used for generating GS with PostShot. As the masked images were generated from the 3D model, incorrect areas may be carried to appear in the GS, too. Nevertheless, GS can also be generated without the exports from SfM.

For GS, both the output quality and processing time are inherently influenced by the number of training iterations. Using high-resolution images (5184 × 3456 px) significantly increases processing time (by factor 5) compared to downsampled versions (1600 × 1066 px) (Tab. 2), while the visual quality is increased, adding high-resolution details to the face (Fig. 3), for example. Additionally, reducing the number of input images from 400 to 100 images has a lesser impact on the final GS representation than reducing the number of training iterations from 40,000 to 10,000, though it can lead to increased processing time, see Table 1. Fewer images yield fewer matching points and a sparser point cloud, resulting in fewer initialised splats. This reduced coverage limits information for optimisation, often slowing convergence despite the smaller dataset.

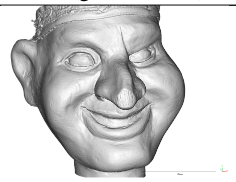
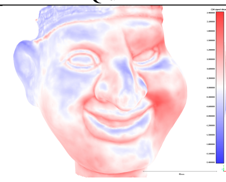
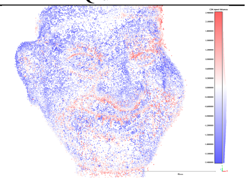
Comparison C2M distance	SLS (ground truth)	$\mathrm{SfM}_{\mathrm{HQfull}}$ to SLS	$\mathrm{GS}_{\mathrm{HQfull}}$ to SLS
			
Mean distance [mm]	—	$\bar{x} = 0.129238$	$\bar{x} = -1.95611$
Standard deviation [mm]	—	$\sigma = \pm 0.579842$	$\sigma = \pm 3.76641$

Table 3: *Comparison cloud-to-mesh (C2M) distance of SfM mesh and GS point cloud to SLS mesh in CloudCompare v2.13.1, displayed distance range ±2,4mm.*

Although still an active area of research, it is possible to derive meshes from GS outputs [SP25, HYC*24], thereby enabling the reconstruction of artefact surfaces that pose significant challenges for conventional SfM approaches, such as highly reflective materials or fine textile structures, see Fig. 3. In comparison, the 2DGS-based mesh [HYC*24] is less detailed but free of holes. However, its point-based shading, unlike texture-based rendering, leads to reduced visual quality.

Different post-processing steps were applied depending on the representation. Despite prior masking of the input images, the GS representations still contained scene splats due to inaccuracies in the masks. Supersplat was used to clean the GS output. In general, post-processing options for GS are limited and mostly restricted to basic operations such as selecting, translating, rotating, and scaling, along with general colour adjustments. For the SfM meshes, Agisoft Metashape was used, which includes several post-processing tools such as mesh cleaning and surface simplifications. Alternatively, Blender can be employed, offering a broader set of post-processing capabilities. These include advanced mesh cleaning, surface simplifications, and the rigging of specific performative artefacts to enable virtual control. In addition, Blender's modelling and animation tools allow for the virtual reconstruction of missing parts based on surface features as well as several animation techniques.

4.2. Geometric Accuracy

The geometric analysis of performative artefacts presents significant challenges. Firstly, ensuring a consistent pose for comparison is essential but difficult to achieve. Secondly, these objects often comprise a diverse range of materials, from rigid to soft components. The presence of soft elements further complicates the comparison process due to their deformability. To address this, we focus exclusively on the rigid parts of an object. This highlights the necessity of standardising the capturing and comparison methodologies for movable and deformable objects. In this approach, only the rigid head of the puppet using SfM and GS is compared. The results are then evaluated against ground truth data obtained via a structured light scanner (Artec Spider), which offers an accuracy of up to 0.1 mm [BGP*23].

To ensure a meaningful and precise comparison, all reconstructed models were first aligned and finely registered to the scaled structured-light-scan (SLS) reference using CloudCompare v2.13.1. Fine registration was performed via an iterative closest point (ICP) algorithm to minimise residual pose differences. For the quantitative evaluation, see Table 3, CloudCompare's "distance mesh to point cloud" tool was applied. In the case of GS, only the positions of the splats were considered, disregarding their radii and anisotropic shape properties, as conventional surface-based measurement approaches may not fully accommodate the splat-based representation.

The comparison between the high-quality SfM-based mesh ($\mathrm{SfM}_{\mathrm{HQ400}}$) and the SLS reference resulted in a mean absolute deviation of $\bar{x} = 0.129238\,\mathrm{mm}$ with a standard deviation of $\sigma = \pm 0.579842\,\mathrm{mm}$. For the GS high-quality reconstruction ($\mathrm{GS}_{\mathrm{HQ400}}$), the comparison yielded a mean deviation of $\bar{x} = -1.95611\,\mathrm{mm}$ with a larger standard deviation of $\sigma = \pm 3.76641\,\mathrm{mm}$. The higher deviation in the $\mathrm{GS}_{\mathrm{HQ400}}$ reconstruction results from splats inside the object's volume. While they increase deviation, they are essential for representing challenging materials with glossy, transparent, or depth-related effects. In contrast, SfM focuses on the surface reconstruction, showing lower deviation but lacking these complex material representations.

4.3. Visual Fidelity

Visual fidelity does not necessarily align with geometric accuracy, rather, it depends on surface textures and material properties such as glossiness, reflectivity, and translucency. Consequently, assessing visual quality extends beyond purely geometric metrics and requires consideration of perceptual factors. Fig. 4 illustrates a range of materials captured and processed as both SfM-based meshes and GS representations.

GS offers highly realistic *images* of a model, particularly excelling in areas where traditional SfM-based meshes often struggle, such as textiles with fine details, reflective materials like metal, monochrome or transparent surfaces, and complex or thin geometries [IM21]. This aligns with the theoretical foundation of geometric and photometric representations (Section 3.2). GS consistently outperforms meshes in visual fidelity, as illustrated in Fig. 4, including specular reflections on glossy surfaces. However, despite its photorealistic rendering, GS is not without limitations: problems like "popping" effects and inconsistent depth perception may occur due to per-camera splat sorting [HYC*24]. Additionally, alpha blending of splats can cause sharp edges to appear blurred at

Figure 4: *Comparison of Ground Truth Photos (GT) with Gaussian Splatting (GS) and SfM mesh-based 3D representations of performative artefacts.*

close zoom levels. That said, recent research has already proposed solutions to many of these rendering issues [HFW*25, HYC*24].

4.4. Interactivity and Dynamics

SfM meshes are more expressive when it comes to the use of performative artefacts and can be seamlessly integrated into established 3D processing pipelines. They support various animation techniques, including physics-based simulations, motion capture, and key framing [Par12]. In this context, SfM meshes provide a flexible foundation for enabling re-contextualisation [TCY*] and control of historical puppets, for instance, in an exhibition setting. While such use sacrifices some degree of historical authenticity, it opens up new ways for artistic expression, especially in cases where the artefact cannot be moved any more or is partially or completely missing (Section 2). Recent advances in dynamic scene reconstruction enable mesh-based approaches for capturing dynamic scenes [SYD*23]. Although this remains an active area of research, these methods exhibit mesh-specific advantages and disadvantages, such as explicit surface structures with less realistic visual appearances.

In contrast, when applied to dynamic scene capture, GS can produce highly realistic results, provided appropriate setups are used, surpassing mesh-based approaches in visual fidelity [LCLX24]. As such, GS is considered a promising solution for the holistic dynamic capturing of moving artefacts. The extension from static to dynamic GS retains many of the same core characteristics. The editing and playback of such dynamic GS introduce new challenges, particularly due to the large volume of data involved. Nevertheless, these methods offer a compelling way to preserve and replay the authentic movement of artefacts, similar to video, especially in cases where the objects can still be physically animated. Despite early approaches [YJM*24, GL24, QKS*24], controlling GS remains an open research topic.

4.5. Evaluation of Mesh-based and GS Approaches

Digitising performative artefacts is inherently individual, as each object features unique materials and structures. Based on a small selection of historical puppets, this comparison highlights general trends despite object-specific variation. Mesh-based 3D representations are well-suited for operational tasks, such as reconstruction, geometric analysis, and animation, due to their structured geometry and compatibility with established workflows. In contrast GS captures visual richness, including material appearance and lighting effects, making it ideal for orientational and immersive interactions. GS also enables realistic dynamic scene capture and, in some cases, outperforms traditional SfM in speed and image efficiency [FCW*24]. Despite their differing paradigms, both methods can be derived from the same dataset, supporting comparative or complementary use. Each offers distinct advantages and can meet different curatorial and institutional needs.

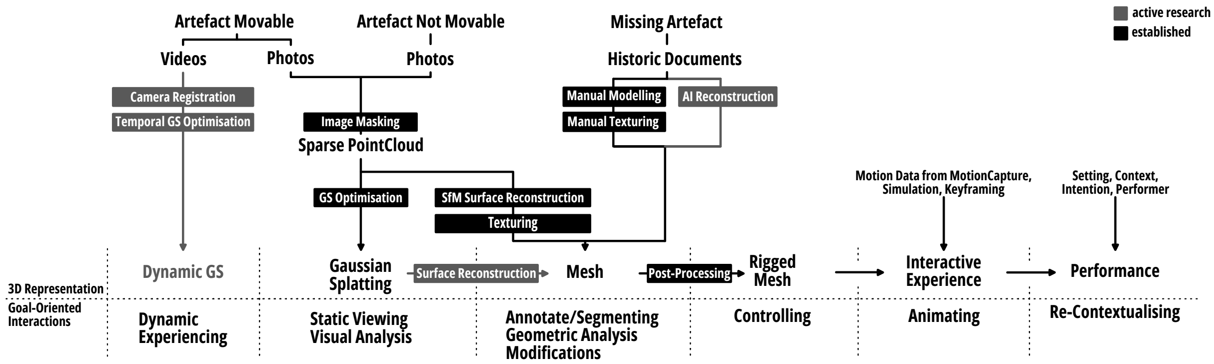


Figure 5: *Goal-oriented and object-centered guide for digitisation of performative artefacts.*

5. Goal-oriented Guide for Digital Performative Artefacts

While digitisation offers transformative potential for museums, its implementation must be intentional, in order to be sustainable for the future [IM21]. Many current digitisation efforts, especially in institutions with limited resources, lack consistent strategies and often overlook 3D-specific considerations [WP22]. Before undertaking 3D digitisation, institutions should clarify the following points of the guidelines [The19]:

- **Goals** – What purpose does the 3D data serve (e.g., documentation, education, research, outreach)?
- **Vision** – How does 3D integrate into broader museum strategies?
- **Resources** – Are trained staff, appropriate equipment, and software available? Will outsourcing reduce internal control?
- **Target Groups** – Who will use or benefit from the 3D data (scholars, educators, general public)?
- **Publication and Access** – What are the access policies? What channels (e.g., online databases, AR/VR applications) will be used?
- **Sustainability** – What systems are in place for the long-term management and updating of 3D data?

The comparative analysis informed a goal-oriented, object-centered guide for digitising performative artefacts. Goals vary, including research and conservation [SBSI24, ABC*18], exhibition [KKB24], education, and participation [PZS*23]. These, in turn, are shaped by specific interactions with the digital material (Section 2). The guide helps heritage professionals choose suitable digitisation techniques, based on the intended interaction with 3D representations and the type and quality of source material available (see Fig. 5). Therefore, performative artefacts can be grouped into three categories by preservation state: (1) movable – suitable for dynamic scene capturing, (2) immovable – requiring static capture, and (3) (partially-) lost – reconstruction from limited or secondary sources.

When the performative **artefact is movable**, dynamic GS is the preferred technique for capturing immersive, motion-rich and authentic representations of its movement and presence in space (see Fig. 5). This approach preserves the authentic performance of artefacts and offers a dynamic, realistic experience. While dynamic GS is still an active research area, future improvements are expected to enhance its quality and usability. Similarly, dynamic mesh reconstruction may become a viable alternative, but currently lacks the quality and accessibility required for this application and is thus excluded.

When a static 3D representation is required, either for **not movable artefacts** or movable ones the primary options are GS representations or SfM-based meshes. Materiality shapes the capture process. Deformable parts, typical of historical puppets, required fixed poses for stable capture, affecting rigging and preventing use of automation like turntables, which introduce motion. Textiles were stuffed to avoid self-occlusion. Rigid artefacts allow therefore more straightforward, automated capture. For immersive viewing and visual analysis, GS is preferable for its realistic, view-dependent rendering. Mesh models remain better suited for structural or interactive tasks requiring precise geometry, such as annotation, segmentation, or modification. In situations where SfM-based reconstruction fails due to challenging artefact properties, GS may be a viable option for achieving results. Based on our experiences, this includes reflective or translucent materials and uniform surfaces (e.g., metal, wood, glass), as well as fine geometries like rods, strings, and textiles. By combining the expressive, editable qualities of mesh-based representations with the photorealistic and real-time rendering strengths of GS, a complementary system can be established. Considering both GS and SfM-based meshes rely on the same source dataset and intermediate processing (camera registration and image masking), as well as the optional surface reconstruction of GS outputs, GS can be seamlessly integrated into existing SfM-based digitisation workflows commonly used in museums. This compatibility positions GS as a powerful extension of current practices, rather than a replacement. If the goal is to create controllable and interactive virtual experiences or even new performances, the use of meshes is appropriate. These reconstructions serve as the base for further development, requiring post-processing steps such as mesh simplifications, rigging and the incorporation of motion data, whether via simulation, motion-capture devices or key framing. In order to create a re-contextualised performance, the additionally involved elements such as narrative contexts, settings, performers, and intentions must be considered [TCY*]. This way new performances can be designed.

In scenarios where the **artefact is partially or entirely lost**, meshes offer a flexible basis for reconstructing and re-imagining missing components. The reconstruction process can be supported by historical documentation, which may inform manual modelling and texturing or the use of AI-driven methods [XCG*24] to synthesise 3D meshes from single or few images [KKS*24]. As with dynamic GS and dynamic mesh-based scene reconstruction, future research is expected to further improve the quality and practicality of these methods.

6. Conclusion

This work explores the specific digitisation requirements of performative artefacts, supported by a comparative analysis of selected historical puppets to understand their material characteristics. Based on these insights, we propose a goal-oriented and object-centered digitisation guide (Fig. 5) tailored to the needs of cultural heritage professionals. For immersive representations, Gaussian Splatting (GS) offers a compelling solution, enabling a photorealistic representation. Furthermore, it allows the digitisation of materials, which are otherwise problematic to digitise with common SfM methods. Additionally, dynamic GS enables holistic and immersive dynamic scene capturing. In contrast, SfM-based meshes are more suitable for analytical tasks and scenarios that require further control over the digital artefact, such as annotating, segmenting, geometric analysis or interactive control. Given that both GS and SfM share similar input requirements and preprocessing steps, a hybrid workflow becomes advantageous. Treating the chosen representation as an interchangeable outcome of a unified dataset encourages a flexible, goal-driven approach to the digitisation of performative artefacts.

Acknowledgments

This project is funded by the EU European Social Fund and Saxon State Parliament (Project No. 100690214). Thanks to Jakob Franke and Staatliche Kunstsammlungen Dresden for providing the artefacts and materials.

References

[AB17] AGOSTO E., BORNAZ L.: 3d models in cultural heritage: Approaches for their creation and use. *International Journal of Computational Methods in Heritage Science (IJCMHS) 1*, 1 (Jan. 2017), 1–9. doi:10.4018/IJCMHS.2017010101. 3

[ABC*18] APOLLONIO F. I., BASILISSI V., CALLIERI M., DELLEPIANE M., GAIANI M., PONCHIO F., RIZZO F., RUBINO A. R., SCOPIGNO R., ET AL.: A 3d-centered information system for the documentation of a complex restoration intervention. *Journal of Cultural Heritage 29* (2018), 89–99. 8

[Agi25] AGISOFT LLC: Agisoft Metashape User Manual - Professional Edition, Version 2.2. URL: https://www.agisoft.com/pdf/metashape-pro_2_2_en.pdf. 4

[BCG*24] BASSO A., CONDORELLI F., GIORDANO A., MORENA S., PERTICARINI M.: Evolution of rendering based on radiance fields: The palermo case study for a comparison between nerf and gaussian splatting. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* (2024), 57–64. 2

[Bed17] BEDFORD J.: *Photogrammetric Applications for Cultural Heritage: Guidance for Good Practice*. Historic England Guidance Series. Historic England, 2017. URL: https://books.google.de/books?id=tdrMtQEACAAJ. 3, 4

[BGP*23] BALLONI E., GORGOGLIONE L., PAOLANTI M., MANCINI A., PIERDICCA R., ET AL.: Few shot photogrametry: A comparison between nerf and mvs-sfm for the documentation of cultural heritage. *International Archives Of The Photogrammetry, Remote Sensing And Spatial Information Sciences 48* (2023), 155–162. 2, 6

[BPCO23] BARAZZETTI L., PREVITALI M., CANTINI L., OTERI A. M.: Digital recording of historical defensive structures in mountainous areas using drones: Considerations and comparisons. *Drones 7*, 8 (2023), 512. 2

[CBC*24] CROCE V., BILLI D., CAROTI G., PIEMONTE A., DE LUCA L., VÉRON P.: Comparative assessment of neural radiance fields and photogrammetry in digital heritage: impact of varying image conditions on 3d reconstruction. *Remote Sensing 16*, 2 (2024), 301. 2

[CCDL*23] CROCE V., CAROTI G., DE LUCA L., PIEMONTE A., VÉRON P.: Neural radiance fields (nerf): Review and potential applications to digital cultural heritage. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences 48* (2023), 453–460. 2, 4

[CDR15] CHIABRANDO F., DONADIO E., RINAUDO F.: Sfm for orthophoto to generation: A winning approach for cultural heritage knowledge. *The international archives of the photogrammetry, remote sensing and spatial information sciences 40* (2015), 91–98. 3

[CNAC24] CLINI P., NESPECA R., ANGELONI R., COPPETTA L.: 3d representation of architectural heritage: a comparative analysis of nerf, gaussian splatting, and sfm-mvs reconstructions using low-cost sensors. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences 48* (2024), 93–99. 2, 3

[Esc19] ESCARZAGA S.: Agisoft Metashape Python Script to Automate Gradual Selection, Camera Optimization and Error Reduction Steps, 2019. Artwork Size: 5985 Bytes Pages: 5985 Bytes. URL: https://figshare.com/articles/Agisoft_Metashape_Python_Script_to_Automate_Gradual_Selection_Camera_Optimization_and_Error_Reduction_Steps/10325909/1, doi:10.6084/M9.FIGSHARE.10325909.V1. 4

[EUr24] EUREKA3D: 3D Digitisation Guidelines, Apr. 2024. URL: https://eureka3d.eu/3d-digitisation-guidelines/. 3

[FCW*24] FAN Z., CONG W., WEN K., WANG K., ZHANG J., DING X., XU D., IVANOVIC B., PAVONE M., PAVLAKOS G., ET AL.: Instantsplat: Unbounded sparse-view pose-free gaussian splatting in 40 seconds. *arXiv preprint arXiv:2403.20309 2*, 3 (2024), 4. 4, 7

[FvDHF90] FOLEY J. D., VAN DAM A., HUGHES J. F., FEINER S. K.: Spatial-partitioning representations; surface detail. *Computer Graphics: Principles and Practice* (1990). 3

[GBS14] GOMES L., BELLON O. R. P., SILVA L.: 3d reconstruction methods for digital preservation of cultural heritage: A survey. *Pattern Recognition Letters 50* (2014), 3–14. 1

[GGH*24] GRÖNE R., GÜNTER B., HEINZELMANN R., KIRCHER-KANNEMANN A., MAHMOUDI Y., RÖMHILD J., SIMON H., STÄRK T., ZEBISCH L., FORUM V. M., BECKER I., BÖHMER O.: 2. STATUS QUO: MUSEUM TASKS AND VIRTUAL MUSEUMS. In *Virtual Museums – A Plea*. Deutscher Kunstverlag (DKV), Oct. 2024, pp. 23–28. Section: Virtual Museums – A Plea. URL: https://www.degruyterbrill.com/document/doi/10.1515/9783422802377-005/html. 3

[GKJ*21] GARBIN S. J., KOWALSKI M., JOHNSON M., SHOTTON J., VALENTIN J.: Fastnerf: High-fidelity neural rendering at 200fps. In *Proceedings of the IEEE/CVF international conference on computer vision* (2021), pp. 14346–14355. 4

[GKK24] GRÖNE R., KIRCHER-KANNEMANN A.: 9. COLLECT. In *Virtual Museums – A Plea*. De Gruyter, Oct. 2024, pp. 55–60. doi:10.1515/9783422802377-012. 3

[GL24] GUÉDON A., LEPETIT V.: Gaussian frosting: Editable complex radiance fields with real-time rendering. In *European Conference on Computer Vision* (2024), Springer, pp. 413–430. 7

[GNBL13] Guidi G., Navarro P. R., Barsanti S. G., Lo L.: Quick textured mesh generation in cultural heritage digitization. 3

[GR24] Günter B., Römhild J.: 3. SUITABILITY CRITERIA OR "WHAT MAKES SENSE AS A VIRTUELL MUSEUM?". In *Virtual Museums – A Plea*. Deutscher Kunstverlag (DKV), Oct. 2024, pp. 29–32. Section: Virtual Museums – A Plea. URL: https://www.degruyterbrill.com/document/doi/10.1515/9783422802377-006/html. 3

[Gro17] Groh R. (Ed.): *An Iconography of Interaction*. 2017. 3

[Han96] Hanrahan M. L. P.: Light field rendering. *SIGGRAPH96, Conputer Graphics Proceeding* (1996). 4

[HFW*25] Hahlbohm F., Friederichs F., Weyrich T., Franke L., Kappel M., Castillo S., Stamminger M., Eisemann M., Magnor M.: Efficient perspective-correct 3d gaussian splatting using hybrid transparency. In *Computer Graphics Forum* (2025), Wiley Online Library, p. e70014. 7

[HYC*24] Huang B., Yu Z., Chen A., Geiger A., Gao S.: 2d gaussian splatting for geometrically accurate radiance fields. In *ACM SIGGRAPH 2024 conference papers* (2024), pp. 1–11. 5, 6, 7

[IM21] Immonen V. A., Malinen I.: 3D Imaging in Museums. In *Museum Studies – Bridging Theory and Practice*, ICOFOM & University of Jyväskylä, Open Science Centre, (Ed.). Jan. 2021, pp. 253–270. URL: https://www.researchgate.net/publication/359115636_3D_Imaging_in_Museums. 1, 6, 8

[JB25] Jamil O., Brennan A.: Immersive heritage through gaussian splatting: A new visual aesthetic for reality capture. *Frontiers in Computer Science 7* (2025), 1515609. 2

[KKB24] Kircher-Kannemann A., Böhmer O.: 11. EXHIBIT AND PRESENT. In *Virtual Museums – A Plea*. De Gruyter, Oct. 2024, pp. 63–66. URL: https://www.degruyter.com/document/doi/10.1515/9783422802377-014/html, doi:10.1515/9783422802377-014. 8

[KKLD23] Kerbl B., Kopanas G., Leimkühler T., Drettakis G.: 3d gaussian splatting for real-time radiance field rendering. *ACM Transactions on Graphics 42*, 4 (July 2023). URL: https://repo-sam.inria.fr/fungraph/3d-gaussian-splatting/. 2, 4

[KKS*24] Kniaz V., Knyaz V., Skrypitsyna T., Moshkantsev P., Bordodymov A.: Deep Learning for Single Photo 3D Reconstruction of Cultural Heritage. *Optical Memory and Neural Networks 33*, 3 (Dec. 2024), S457–S465. doi:10.3103/S1060992X24700723. 9

[LCLX24] Li Z., Chen Z., Li Z., Xu Y.: Spacetime gaussian feature splatting for real-time dynamic view synthesis. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (2024), pp. 8508–8520. 4, 7

[LLW*25] Liu Q., Liu Y., Wang J., Lyu X., Wang P., Wang W., Hou J.: MoDGS: Dynamic gaussian splatting from casually-captured monocular videos with depth priors. In *The Thirteenth International Conference on Learning Representations* (2025). 4

[LYXC24] Li M., Yao S., Xie Z., Chen K.: Gaussianbody: Clothed human reconstruction via 3d gaussian splatting. *arXiv preprint arXiv:2401.09720* (2024). 2

[MESK22] Müller T., Evans A., Schied C., Keller A.: Instant neural graphics primitives with a multiresolution hash encoding. *ACM transactions on graphics (TOG) 41*, 4 (2022), 1–15. 4

[MKR*23] Mazzacca G., Karami A., Rigon S., Farella E., Trybala P., Remondino F., et al.: Nerf for heritage 3d reconstruction. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences 48*, M-2-2023 (2023), 1051–1058. 2

[MST*21] Mildenhall B., Srinivasan P. P., Tancik M., Barron J. T., Ramamoorthi R., Ng R.: Nerf: Representing scenes as neural radiance fields for view synthesis. *Communications of the ACM 65*, 1 (2021), 99–106. 1, 4

[Par12] Parent R.: *Computer animation: algorithms and techniques*. Newnes, 2012. 7

[PBP*22] Palestini C., Basso A., Perticarini M., et al.: Machine learning as an alternative to 3d photomodeling employed in architectural survey and automatic design modelling. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences 48*, 2/W1-2022 (2022), 191–197. 2

[PZS*23] Pattakos A., Zidianakis E., Sifakis M., Roulios M., Partarakis N., Stephanidis C.: Digital Interaction with Physical Museum Artifacts. *Technologies 11*, 3 (Apr. 2023), 65. URL: https://www.mdpi.com/2227-7080/11/3/65, doi:10.3390/technologies11030065. 8

[QKS*24] Qian S., Kirschstein T., Schoneveld L., Davoli D., Giebenhain S., Niessner M.: Gaussianavatars: Photorealistic head avatars with rigged 3d gaussians. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (2024), pp. 20299–20309. 7

[RWW*25] Rangelov D., Waanders S., Waanders K., Keulen M. v., Miltchev R.: Impact of data capture methods on 3d reconstruction with gaussian splatting. *Journal of Imaging 11*, 2 (2025), 65. 4

[SBSI24] Stoean R., Bacanin N., Stoean C., Ionescu L.: Bridging the past and present: AI-driven 3D restoration of degraded artefacts for museum digital display. *Journal of Cultural Heritage 69* (Sept. 2024), 18–26. URL: https://www.sciencedirect.com/science/article/pii/S1296207424001468, doi:10.1016/j.culher.2024.07.008. 8

[SP25] Stuart L. A., Pound M. P.: 3dgs-to-pc: Convert a 3d gaussian splatting scene into a dense point cloud or mesh. *arXiv preprint arXiv:2501.07478* (2025). 6

[SRD*25] Santos P., Ritz M., Domajnko M., Tausch R., Schurig M.: Autonomous 3-d digitization for arbitrary cultural heritage objects. *IEEE Computer Graphics and Applications 45*, 1 (2025), 76–83. 2

[SWG24] Sun J.-M., Wu T., Gao L.: Recent advances in implicit representation-based 3d shape generation. *Visual Intelligence 2*, 1 (2024), 9. 2

[SYD*23] Song C., Yang G., Deng K., Zhu J.-Y., Ramanan D.: Total-recon: Deformable scene reconstruction for embodied view synthesis. In *IEEE International Conference on Computer Vision (ICCV)* (2023). 7

[TCY*] Töpfer F., Christ E., Yu Z., Engeln L., Franke I., McGinity M.: Moving experiences – towards a conceptual framework for performative artefacts in museums. In *KUI Conference Culture and Computer Science | From Humanism to Digital Humanities (KUI 2024), October 03–04, 2024, Florence, Italy*, p. to appear. doi:10.1145/3719236.3719249. 1, 2, 7, 8

[The19] The Swedish National Heritage Board: Guide for publishing 3D models, 2019. URL: https://www.raa.se/in-english/outreach-and-exhibitions/guide-for-publishing-3d-models/. 8

[WHS*24] Wang R., Hua C., Shingys T., Niu M., Yang Q., Gao L., Zheng Y., Yang J., Wang Q.: Enhancement of 3d gaussian splatting using raw mesh for photorealistic recreation of architectures. *arXiv preprint arXiv:2407.15435* (2024). 2

[WP22] Winfried Bergmeyer, Paul Klimpel: Handreichung. Digitale Grunderfassung. 10 Grundsätze, Aug. 2022. URL: https://www.museumsbund.de/wp-content/uploads/2022/12/handreichung-digitale-grunderfassung.pdf. 8

[XCG*24] Xu J., Cheng W., Gao Y., Wang X., Gao S., Shan Y.: Instantmesh: Efficient 3d mesh generation from a single image with sparse-view large reconstruction models. *arXiv preprint arXiv:2404.07191* (2024). 9

[YJM*24] Yu H., Julin J., Milacski Z. Á., Niinuma K., Jeni L. A.: Cogs: Controllable gaussian splatting. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (2024), pp. 21624–21633. 7