Gaussian Splatting for Photorealistic Immersive Virtual Tours

Md Mridul Hossain and SeongKi Kim*

Department of Computer Engineering, Chosun University, Gwangju, Korea
Email: mriduluitsbd@gmail.com (M.M.H.); skkim@chosun.ac.kr (S.K.K.)
*Corresponding author

***Abstract*—To achieve photorealistic Virtual Reality (VR) tours, a balance is required between visual fidelity, responsiveness, and the strict runtime and memory constraints of head-mounted displays. As a more recent explicit view synthesis method, three-Dimensional Gaussian Splatting (3DGS) has been proposed as a real-time alternative to Neural Radiance Fields (NeRF). In this study, we present an overview of 3DGS, demonstrating its application to immersive VR tours, and propose a design philosophy for its implementation on standalone headsets. Specifically, we review state-of-the-art 3DGS variants with respect to photorealism, scalability, efficiency, dynamics, and semantics; analyze VR-specific constraints related to latency, stereo stability, and thermal, memory, and interaction requirements; collate VR style datasets which are suitable for training and benchmarking 3DGS-based pipelines; present case studies in domains such as museums, real estate walkthroughs, and large-scale campus tours; synthesize an evaluation framework that organizes metrics for visual quality, performance, memory usage, and interaction responsiveness. The evaluation framework has four axes: visual fidelity, runtime performance, memory footprint, and interaction responsiveness. Although more recent accelerated NeRF-based systems (e.g., Instant-Neural Graphics Primitives (Instant-NGP) and baked radiance fields) directly reduce the performance difference between 3DGS and explicit methods, 3DGS still provides better trade-offs in interactive tours, such as predictable latency and instant editability. Finally, open challenges in streaming, reflectance, semantics, and authoring ergonomics are outlined, and future directions for 3DGS-based VR tour systems are discussed.**

***Keywords*—gaussian splatting, virtual reality, photorealistic rendering, immersive tours, real-time graphics, evaluative framework**

I. Introduction

The immersion and photorealistic experience of Virtual Reality (VR) is a powerful and increasingly common way for individuals to engage with digital content across many fields [1]. VR applications, such as photorealistic VR tours, require not only high image fidelity but also low latency to provide a realistic look [2]. People have become accustomed to experiencing interactivity in museums, cultural tourism attractions, virtual real estate tours, and college campus tours [3]. The shape of Head-Mounted Displays (HMDs) places system-wide unaffordable latent requirements on motion-to-photon latency (i.e., the time between moving one's head and the device showing the updated image), twin correspondence, consistency across time, and perceptible memory and thermal needs (i.e., the thermal energy produced by the device during operation, which affects comfort, performance, and stability) [4]. A Neural Radiance Field (NeRF) is an efficient implicit model of volumetric light transport in a scene [5]. Early NeRF models required hours of training time and could not be inferred at interactive frame rates, making them infeasible for commercially available, thermally limited VR headsets. Although recent faster NeRF models, such as those using Instant-NGP multiresolution hash encodings and prebaked radiance fields, have significantly reduced the training time, they introduce new challenges in predictably managing latency, memory usage, and editability, especially in standalone devices. In this study, three-Dimensional Gaussian Splatting (3DGS) is presented as an explicit alternative to NeRF models. It features a feed-forward rendering path and offers lifelike control over the number and distribution of primitives.

The novelty of this study lies in its focus on 3DGS for immersive VR tours. In contrast to previous surveys on 3DGS and neural view synthesis, our study focuses on end-to-end pipelines, from capture and optimization to packaging and in-headset rendering. It provides VR-suitable evaluation guidance that correlates 3DGS design choices with headset restrictions, dataset characteristics, and VR-specific tour scenarios (e.g., museums, real estate, and campus).

VR in museums provides global access to heritage objects by reconstructing galleries and expos in detail, enabling visitors to observe these objects remotely [6]. In the real estate business, potential customers can inspect properties and tour an entire building virtually, as if taking a walk around, without ever having to leave the building [7]. VR tours are a powerful tool for preserving cultural heritage, particularlly for at-risk sites. They allow current and future generations to visit these places virtually [6]. Colleges and campuses also use VR to provide prospective students with interactive, guided tours of their campuses, facilities, and student life [8]. At

Manuscript received September 10, 2025; revised November 18, 2025; accepted December 23, 2025; published May 27, 2026.

the center of these VR experiences lies presence: the compelling feeling of being somewhere else [9]. This degree of immersion requires high visual fidelity to display scenes with realistic graphics and low motion-to-photon latency to ensure that head movement is reflected in the display in real time [2]. These requirements are essential to avoid user discomfort, minimize motion sickness, and maintain the illusion of presence [4].

Current solutions to photorealistic VR rendering have significant trade-offs in implementation in real-time, mobile headsets in the wild. Although mesh-based pipelines have the advantage of efficient rasterization, they cannot recreate intricate view-dependent effects without extensive manual modeling or high-quality photogrammetry [10]. Although NeRF techniques enable high-quality volumetric scene reconstruction [5], even their accelerated variants (e.g., Instant-NGP and baked radiance fields) can struggle to run predictably within the strict latency, memory, and thermal constraints of standalone VR hardware. Earlier point-based rendering systems were flexible but lacked stable anisotropic detail representation and exhibited aliasing with temporal instability during fast head-tracked motion [11].

Comparison of NeRF and SL-DGS

Characteristic	NeRF	SL-DGS
Scene Representation	Volumetric grid or implicit MLP	Cloud of ellipsoids (3D Gaussians)
Deformation	Latent Geometry Prior + Deformation Field (MLP)	Latent Geometry Prior + Deformation Field (MLP)
Rendering	Ray marching through volume	Projected ellipses (splats) on image
Speed	Not real-time	Real-time capable
Losses	Photometric, Smoothness, Multi-view Consistency	Photometric, Smoothness, Multi-view Consistency
Input	2D frames from monocular video	2D frames from monocular video
Depth Maps Required	Yes	No
Output	Novel View Images, 4D Scene Playback	Novel View Images, 4D Scene Playback
VR Relevance	Yes	Yes

Fig. 1. Comparison of neural radiance fields and Structured-Light 3D Gaussian Splatting (SL-DGS).

The 3DGS system addresses these constraints by describing a scene as a collection of anisotropic 3D Gaussians. It employs tile-parallel splatting on modern Graphics Processing Units (GPUs) and optimizes the learned parameters for color, opacity, and shape [12, 13]. This explicitly feed-forward representation enables deterministic real-time performance, predictability, and direct scene editing. Therefore, 3DGS is a good candidate for immersive VR tours. However, its suitability in interactive, resource-constrained VR environments is yet to be systematically surveyed. This gap is addressed in this study, which reviews the approaches for understanding 3DGS techniques and splatting workflows through VR tour writing and playback. Fig. 1 shows that 3DGS offers a clear and explicit alternative to implicit radiance field NeRF-like models, inspiring them to use it to render predictable and real-time VR tours.

The incorporation of 3DGS into VR systems presents domain-specific issues. First, latency and stability are critical: motion-to-photon latency should be within 20 ms to avoid user discomfort, and rendering must maintain temporal coherence to prevent shimmering or popping artifacts caused by head-tracked reprojection [2]. Second, stereo consistency is essential: comfortable depth perception requires accurate binocular disparity cues, and even minor geometric inaccuracies on large planar surfaces, such as walls and floors, can induce nausea [4]. Third, standalone VR headsets have strict thermal and memory constraints due to small GPU memory and power dissipation budgets, limiting models and real-time frame rates [14]. Finally, VR experiences require constant interaction: virtual tours typically involve annotations, points of interest, and the possibility to stream new stage pieces dynamically as users move forward [3]. Therefore, a rendering technique should be viable with streaming and incremental updates without obscuring immersion [15].

This review presents the principles of 3DGS and analyzes its viability for immersive, photorealistic VR. Our literature review begins with an overview of Gaussian-based rendering technologies, investigating their speed, quality, and editability. Then, VR-specific limitations, such as latency, stereo consistency, hardware constraints, and interactivity, are analyzed, and the extent to which 3DGS addresses limitations is evaluated. Finally, we outline future directions for integrating 3DGS into VR pipelines. These include hybrid rendering with traditional rasterization, adaptive streaming techniques, and perceptual optimization to improve user presence and comfort.

The primary contributions of this study are as follows:

- To the best of our knowledge, this is the first survey focused on 3DGS for immersive VR tours. It addresses both end-to-end pipelines (from capture and optimization to packaging, streaming, and in-headset rendering) and provides VR-oriented evaluation guidance.
- The development of a taxonomy of 3DGS techniques related to photorealism, scalability, efficiency, dynamics, and semantics.
- A proposal of a structured evaluation framework specific to VR requirements, including fidelity, runtime performance, memory footprint, and responsiveness.
- Illustrations of design patterns with case studies of museums, real estate walkthroughs, and large-scale campus tours.

- Identification of open challenges and future research directions for the deployment of 3DGS in VR-ready systems.

The remainder of this paper is organized as follows: Section II introduces radiance fields and outlines 3DGS rendering and optimization with a particular focus on the VR-specific constraints. Section III analyses the 3DGS method families applicable in VR tours, such as efficiency of data, realism of pictures, scale, dynamics, and meaning. Section IV demonstrates an end-to-end 3DGS process of VR tour capture, optimization, packaging, and in-headset rendering. Section V generally provides a summary of datasets that are typically used in VR-style benchmarking. In Section VI, the application addresses methods of evaluation of visual fidelity, do-not-responsiveness, memory footprint, and in-responsiveness. Section VII gives typical VR tour example studies and patterns of design. Section VIII is about open challenges and research directions. Lastly, Section IX is the final section of the review.

II. Background

A. Radiance Fields

A radiance field describes the path of light passing through each point in a space in any and every direction. It measures the amount of light either emitted or reflected at a point in a certain direction of view [16]. Radiance fields can be represented in two forms: implicit or explicit.

1) Implicit radiance fields

An implicit representation, such as NeRFs, asks a neural network to estimate the color and opacity of a point upon request [17]. To render an image, rays exiting the camera are shot into the scene, and the network is integrated repeatedly along a ray. The outcomes are combined to compute the final pixel color. This process yields high-quality images at a high computational cost, hindering real-time use, except on head-mounted VR devices [18].

$$C(r)=\sum_{i=1}^{N} T_i \propto_i C_i \tag{1}$$

$$\propto_i = 1-e^{-a_i\delta_i} \tag{2}$$

$$T_i=\prod_{j-1}^{i-1}\left(1-a_j\right) \tag{3}$$

2) Explicit radiance fields

In the explicit approach, radiance data are stored in a form that can be rasterized to the image plane or directed into a splat data structure. 3DGS is such a method. It presents a scene using many anisotropic Gaussian kernels [19].

Rendering occurs via projection into an ellipse on the two-Dimensional (2D) picture plane of each 3D Gaussian. A tile is only counted by the splats contained within it (such as a 16×16 pixel block). Blending these splats according to their depths triggers the sorting of the splats, followed by orderly blending via alpha compositing. This tile-parallel architecture implies that every pixel within a tile has the same ordered splat list, thereby enabling efficient GPU memory use. This enables realistic, high-resolution, rendered graphics in real time, even on common consumer-level graphics cards [20].

B. Three-Dimensional Gaussian Splatting (3DGS)

1) Rendering

When rendering a given camera pose, the pipeline simply discards Gaussians of the camera pose that are not in the view frustum. The other Gaussians are imaged on the image plane. The Jacobian of the projection transform transforms the 3D covariance matrix Σ of each Gaussian into a 2D covariance matrix Σ. It determines the form and area of the ellipse that the Gaussian sweeps out in the image space [19].

$$\sum' = J\sum J^T \tag{4}$$

$$G(x')=0\exp\left(-\frac{1}{2}\left(x'-\mu'\right)^T\left(\sum'\right)^{-1}\left(x'-\mu'\right)\right) \tag{5}$$

$$C(p)=\sum_{i\in N} C_i a_i \bar{\prod_{j<1}}\left(1-a_j\right) \tag{6}$$

2) Optimization

The Gaussian parameters are optimized to produce a scene model using calibrated multiview images. The *L1* loss, typically combined with *DSSIM*, is commonly used to optimize pixel precision and perceptual quality. Because the optimal number of Gaussians and positions cannot be known in advance, a density-control loop is performed during training [21].

$$L=\lambda_1\left\|\hat{I}-I\right\|_1+\lambda_2 DSSIM\left(\hat{I},I\right) \tag{7}$$

C. Virtual Reality Tour Requirements

VR headset rendering imposes stringent requirements that do not exist in desktop view synthesis. Fig. 2 provides an overview of the major VR headset limitations: latency/stability, binocular consistency, memory/thermal limits, and interactivity, which have to be taken into account when implementing 3DGS in immersive tours.

Latency and stability: Headsets require a motion-to-photon latency of less than approximately 20 ms. Any delay, shimmer, popping, or distortion becomes obvious with constant head movement [22].

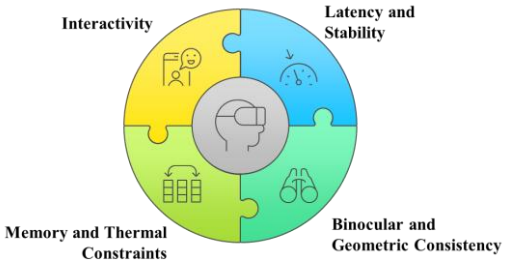

Fig. 2. Components of VR headset performance.

Binocular and geometric consistency: The images produced by the two eyes should have the right disparities so that the sense of depth is natural. Even small geometric errors are evident on large surface areas [23].

Memory and thermal constraints: As standalone systems, VR devices are GPU memory and power-constrained. Therefore, techniques such as culling and Level-of-Detail (LoD) rendering are essential to remain within these limits [24].

Interactivity: VR tours commonly involve 3D definitions, teleport stations, audio tours, and occasionally live editing. The manifestation of these interactions can be accommodated because 3DGS is explicitly represented, and the model does not need to be retrained [25].

In summary, VR tours require high performance and high-quality rendering at low latency in a stable, interactive environment with narrow memory requirements. The explicit, parallelizable nature of 3DGS is suited to this need. However, it requires careful consideration of optimization, geometry accuracy, and resource cost to create a comfortable and compelling experience [20].

D. Literature Survey

During our literature review, we also conducted a citation audit to select 3DGS methods and VR/graphics sources that are relevant to immersive tours. This process ensured that our reference specifically addresses headset constraints, presence, and VR assessment, superseding tangential or generic work.

TABLE I. TAXONOMY OF 3DGS METHODS FOR VR TOURS

Work (Year)	Category	Core idea	Relevance to VR tours	Key limitation for tours	Ref.
3DGS (2023)	Baseline representation	Millions of learnable anisotropic 3D Gaussians + differentiable splatting; tile-based parallel rasterization.	Real-time rendering enables comfortable navigation; explicit representation eases editing.	Memory footprint increases with scene size; aliasing and reflections require further research.	[19]
Scaffold-GS (2023)	Structureaware/ Memory efficiency	Sparse grid of anchor points; local Gaussians tied to anchors; pruning by gradient policies.	Better scalability for room-scale and multiroom scenes.	Extra complexity; dependence on SfM quality.	[25]
CompGS/EAGLES (2024)	Compression/ Vector quantization	Quantize Gaussian attributes into codebooks to shrink storage with minimal loss.	Reduces download time and device memory for standalone VR headsets.	Quantization may introduce small artifacts; training time overhead.	[26]
LightGaussian/ Octree-GS (2024)	Pruning and LoD	Prune insignificant Gaussians; octree-based LoD to feed only necessary splats to rasterizer.	View-dependent LoD keeps FPS high as the user moves.	LoD transitions can cause popping if not smoothed.	[27]
Analytic/Multscale splatting (2023, 2024)	Photorealism/ Antialiasing	Approximate pixel integral analytically; represent the scene at multiple Gaussian scales.	Sharper details without shimmering across headset resolutions.	More complex kernels; potential runtime overhead.	[28]
GaussianShader /GS-IR (2024)	Photorealism/ Reflectance and Inverse rendering	Shading models and normal regularization; decompose lighting to handle specular materials.	Improves realism for glossy floors, glass, and metals commonly found in interiors.	Extra supervision/assumptions; visibility/occlusion handling remains difficult.	[29]
FreGS/GeoGaussian (2024)	Optimization and Geometry	Frequency-based densification and geometry-aware regularization to preserve structure.	Crisp walls/furniture; fewer floaters in walkable areas.	Heavier regularization tuning; good initialization is still required.	[30]
VR SPlat (2025)	Data efficiency	VR-tuned 3DGS pipeline.	Practical HMD guidance; stability.	Engine-specific tweaks; limited public VR.	[31]
GS SLAM/Photo SLAM (2023, 2024)	Tracking and mapping	Use Gaussians for dense mapping and hybrid features for localization.	On-device mapping, drift reduction across a long tour.	Dynamic objects and non-Lambertian surfaces remain challenging.	[32]
VastGaussian/LoD hierarchies (2024)	Large-scale scenes	Divide and conquer training and hierarchical rendering of city/estate scale spaces.	Scales to campuses, museums, and outdoor parks.	GPU memory peaks during training; consistency across cells.	[33]
DrivingGaussian/4DGS/Deformable GS (2024)	Dynamics/4D	Model motion with deformable or 4D Gaussians and learned deformation fields.	Animated elements (fountains, crowds) without re-meshing.	Temporal artifacts and runtime costs; motion blur.	[34]
Feature "3DGS/Language'augmented GS	Semantics and Interaction	Distill 2D foundation features or compress language embeddings into Gaussians.	Clickable points of interest; multimodal search within a tour.	Feature drift across views; storage overhead for embeddings.	[35]

3DGS techniques for immersive tours can be categorized into six groups: baseline models, efficiency-driven methods, compression methods, photorealism-driven methods, geometry refinements, and semantic or dynamic extensions. This taxonomy highlights the trade-offs between scalability, quality, and interactivity.

As shown in Table I, 3DGS techniques for VR tours have distinct categories, including baseline, efficiency, compression, photorealism, geometry, dynamics, and semantics. Baseline models are used to produce high memory and real-time rendering. Scalability is enhanced by efficiency and compression techniques, whereas quality is improved by photorealism and geometry techniques. Interactivity is provided by dynamic and semantic extensions at the expense of cost. Therefore, hybrid pipelines are essential because no single category is sufficient alone.

E. Mapping Three-Dimensional Gaussian Splatting (3DGS) Families to VR Constraints and Datasets

To make the taxonomy feasible, the 3DGS families (Section II. D) are related to VR constraints (Section II. C) and dataset types (Section V). Variants based on 3DGS and photorealism are largely focused on visual fidelity and temporal stability, typically evaluated on challenging room/building and orbit datasets (Section V. A and B). Compression-based and efficiency-driven approaches focus on memory, storage, and streaming constraints on individual headsets and should be assessed on large multiroom, forward-facing, and outdoor datasets (Section V. A, C and D). Geometry-based techniques are well-suited for real estate and campus touring because they excel with planar geometry and stable parallax. Furthermore, the organized, architectural nature of these datasets provides adequate benchmarks (Section V. A and D). Dynamic/4D GS variants are content motion-focused and can be most effectively evaluated using dynamic and driving-style data (Section V. B and D). Semantic and interaction-oriented 3DGS extensions assist in search and annotations and narrative experiences and provide the advantages of semantics-rich datasets with either labels or text (Section V. A–C). This mapping between 3DGS families, predominant VR constraints, and indicative datasets is summarized in Table II.

TABLE II. MAPPING 3DGS FAMILIES TO VR CONSTRAINTS

GS Family	Strengths	VR Constraints Addressed	Weaknesses/Caveats
Baseline 3DGS	High FPS, explicit, editable	Low latency, stable rendering	High memory footprint for large tours.
Efficiency/Pruning/Scaffold-GS	Fewer splats, scalable	Memory limits on standalone HMDs	Quality drops in reflective or thin geometry.
Compression (CompGS/EAGLES)	Smaller assets, faster loading	Storage + streaming constraints	Quantization artifacts may appear in VR stereo.
Photorealism-focused GS (Mip, Multiscale, GS-IR)	Better edges, fewer shimmer	Stereo comfort, temporal stability	Slightly higher compute per frame.
Geometry-consistent GS (FreGS/GeoGS)	Clean walls, stable parallax	Avoids nausea from geometric inconsistencies	More regularization tuning.
4D/Dynamic GS	Handles motion	Crowds, fountains, trees	Difficult to maintain temporal stability.
Semantic/Language GS (Feature3DGS)	Interactive annotations, search	Enhances VR POIs, narratives	Feature drift; increased memory.

III. METHODS RELEVANT TO VR TOURS

A. Data Efficiency

Comprehensive imagery capture from all possible angles of a view may not be possible in some VR tour projects due to temporal, financial, and spatial constraints. Field acquisition normally requires some photos of each room or scene, posing a basic problem in generating accurate geometry and realistic textures [36]. This shortcoming can be addressed by two counterbalancing measures.

The first strategy is depth regularized optimization, which introduces geometry priors during training by exploiting ground-truth depth maps or depths calculated using monocular or stereo images [37]. This procedure shrinks the set of solutions that result in floating artifacts and enforces structure compatibility even when using the optimization in underconstrained regions by constraining the optimization based on depth information. The second strategy uses pretrained prior-based initialization, where pretrained training on massive data learns a mapping to an initial Gaussian set based on minimal input imagery [38]. This initialization provides a good opportunity for further optimization and reduces the required number of iterations and computation.

Although both approaches provide obvious benefits in minimizing the on-site capture time and subsequent processing time, they are sensitive to scenes with high frequencies of specular surfaces (e.g., mirrors and glass) or repeat patterns. In such cases, the learned priors can be incompatible with scene-specific data, resulting in reconstruction errors. These problems can be avoided using postprocessing regularization, which imposes planarity or smooth surface constraints on the surface [39].

Table III is a survey of specific 3DGS techniques of particular importance in VR tours, including baseline splatting, anti-aliasing using less memory, scaffolding/pruning, dynamic (4D) extensions of splatting, interactive editing, and Simultaneous Localization and Mapping (SLAM)-style online mapping. These approaches are taken together to put emphasis on the practical VR trade-offs that are visual stability and runtime/memory cost versus the capability to support interaction and dynamic content.

TABLE III. 3DGS METHODS RELEVANT TO VR TOURS

Method (Year)	Key Idea	Speed/Memory	Photorealism	VR Relevance	Ref.
3DGS (2023)	Explicit 3D Gaussians + rasterization	Real-time (> 100 FPS on scenes)	High with SH and densification	Baseline for photorealistic tours.	[19]
Mip Splatting (2023–2024)	Mipmaps/anti-aliasing for 3DGS	Efficient multiscale splats	Sharper far field detail	Reduces shimmer in walkthroughs.	[26]
Scaffold-GS (2023)	SfM guided scaffold pruning	Lower memory, fast training	Maintains quality with fewer splats	Lighter assets for standalone VR.	[29]
4DGS (2024)	Extends 3DGS to time with 4D covariance	Interactive 4D rendering	Good dynamic fidelity	Supports dynamic VR tours.	[33]
DynMF (2024)	Basis motion fields for 3DGS	Compact motion representation	Physically plausible trajectories	Stable moving elements in tours.	[40]
GaussianEditor/Point'n Move (2024)	Interactive selection + inpainting	Real-time editing	View consistent edits	Content cleanup for tours.	[41]
GS SLAM/SplaTAM (2023–2024)	3DGS mapping for tracking	Online mapping at VR rates	Consistent scene updates	Authoring and live capture.	[42]

B. Enhancing Photorealism in Virtual Reality Environments

Photorealism is essential for user immersion, especially in VR, where stereo displays exaggerate visual imperfections compared with conventional flat-screen monitors [43]. Even minor artifacts, e.g., temporal shimmer and geometric distortion, can severely affect the perceived quality of the experience. Three closely related improvement areas are essential to enhance realism.

First, anti-aliasing and resolution control should be employed to avoid undersampling artifacts appearing as flickers on head motion. Techniques where splatting is performed over a pixel footprint, or those using a multiscale Gaussian representation, have been demonstrated to de-alias fine structure and reduce temporal aliasing [44]. Second, shading-aware models are necessary to accurately render presented reflections and non-Lambertian surfaces such as polished floors, metallic railings, and glass. Rougher bidirectional reflectance distribution function models, combined with constraints on smooth normal fields, can enhance the realism of highlights only seen in the screen space [45]. Finally, geometry consistency on large texture-less surfaces (e.g., ceilings and walls) should be maintained, because inaccuracy results in deformation artifacts. Frequency-progressive densification plans and plane-aware constraints ensure structurally stable outputs and clean parallax effects, which are critical for comfortable VR viewing [46].

C. Memory Optimization and Scalability

Millions or tens of millions of Gaussian primitives are typically used in large-scale VR tours, resulting in optimization-inhibitive load times and memory consumption. Three major approaches address this scalability challenge.

First, computations of Gaussian contributions via reconstruction error measurement prune and remove Gaussians with little influence on the rendered image [47]. Second, compression saves disk storage and runtime memory requirements with a reduction in the data size by quantizing Gaussian parameters, encoding them in compact codebooks, and maintaining perceptual quality [48]. Finally, LoD rendering with streaming dynamically adds or removes scene detail based on users' current viewpoint, optimizing the resolution accordingly. Divisions of scenes into streamable cells and introductions of smooth LoD transitions adequately address artifact stalling (popping), enabling real-time navigation in large spaces [49].

D. Dynamics and 4D Gaussian Splatting

Most VR tours are static, which limits their realism and engagement. This can be significantly improved by introducing dynamic environmental elements, such as wind-swayed vegetation, flowing water, and moving groups of people. This improvement is achieved in dynamic four-Dimensional (4D) Gaussian splatting, which leverages a thin representation of the deformation field to unify motion and space [50]. This approach must overcome technical challenges such as achieving temporal smoothness (to eliminate jitter), incorporating motion blur (to prevent sharpness loss), and handling the memory demands of long-term or repeating animation. Data compression is an important consideration in improving real-time performance without impairing scene quality [51].

E. Semantics and Interaction

To enhance visual verisimilitude, VR tours are increasing their semantic content through richer narrative and interactive experiences. Features extracted by 2D foundation models can be used directly for general-vocabulary search (e.g., "find the Monet painting") or repurposed for scene editing, such as recoloring or occluding objects. This enables progressively detailed interactions, accelerating the execution of natural language requests such as "recolor that car" or "hide that cup" [52].

Although the usability of these features has significantly improved, this advancement introduces new challenges: The storage demands of semantic metadata and the need to control feature drift, which occurs when semantic consistency diminishes across different views. This drift can be addressed via multiview semantic alignment and confidence-weighted feature fusion [53].

IV. 3DGS Workflow for VR Tours

A. Scene Capture

When creating a 3DGS-based VR tour, the capture process must be carefully planned to cover as many spatial areas as possible. In room-scale spaces, the recommended method is to sweep the camera over walls and floors with various heights, achieving coverage of both vertical and horizontal surfaces efficiently [54]. In orbit-type scenes, e.g., 360° capture around a statuary or any object of focus, there should be no change in the camera-to-topic distance, whereas the difference between elevations should be minimal to achieve the resultant balanced angular coverage [55]. Fig. 3 shows a general layout of the end-to-end VR tour-development process with the use of 3DGS, including capture and calibration, all the way to packaging and in-headset rendering.

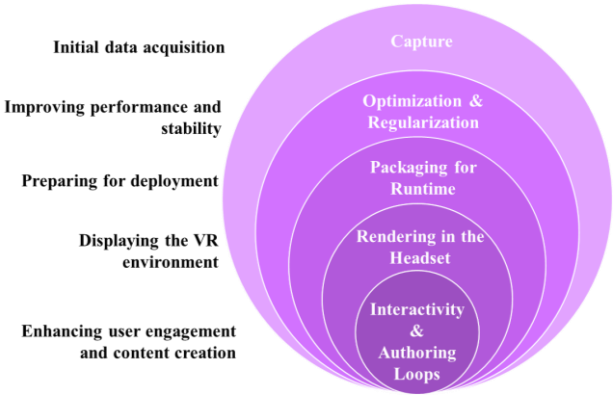

Fig. 3. VR development process.

Doorway and portal shots enable space interconnection because the shots provide frames of space upon scene stitching. Exposure and white balance must be fixed across the dataset to avoid illumination and color disparities. To retain the microstructural fidelity of fine surface details, such as carpets, paintings, and textures, they should be captured using close-range non-flash imagery. Where available, LiDAR or time-of-flight sensors should be used during reconstruction to capture sparse depth maps, which provide precise geometric constraints [56].

B. Calibration and Initialization

Proper camera pose estimation is essential for high-quality scene reconstruction. This is generally achieved using Structure-from-Motion (SfM) algorithms, e.g., those provided in COLMAP [57]. When the captured data is continuous video, structure-aware capture techniques use temporal continuity that may be helpful [58].

After pose estimation, the first Gaussian can be calculated using point clouds in SfM or a learned prediction model. When captures lie in the low-view-count regime, incorporating depth- or normal-based regularization early in the optimization process enhances optimization stability. Content can have less impact on VR realism than sparse coverage [59].

C. Optimization and Regularization

The optimization stage further adjusts the spatial placement, scale, opacity, and color of each Gaussian to reduce the difference between the rendered and captured images. The most commonly used loss function combination balances the L1 reconstruction error and the Structural Similarity Index Measure (SSIM) [60].

A densification algorithm multiplies or splits Gaussians into areas of low details, whereas pruning removes those that have insignificant influence on reconstruction quality. Geometry-specific simplifications, e.g., the enforcement of planarity on walls and floors, normal regularization on smooth surfaces, and frequency-progressive densification in gradual detail addition, are implemented to verify structural reality [61]. Regarding reflective or refractive surfaces, view dependency can be rendered less expensively in terms of cost using shading-consistent Gaussian models or inverse rendering methods.

D. Packaging for Runtime

The process following optimization is dataset preparation to ensure efficient loading and playback during VR runtime. This process entails an end convolution surgery of eliminating superfluous Gaussians and attribution slims to save memory. A LoD hierarchy is built such that geometry that is far away is described using fewer and simpler Gaussians, resulting in more efficient rendering [62].

At the large scale, on-demand streaming is supported by partitioning the space into cells, which are aligned (usefully) with architectural partitions such as rooms or visibility clusters. Interactive metadata, such as points of interest, captions, and safety boundaries, can be embedded in the structure. Furthermore, versioning such metadata enables updates to a scene package without completely regenerating it [63].

E. Rendering in Headsets

Runtime stability and speed are paramount. This is achieved using techniques such as frustum culling, which culls out objects beyond the user frustum, and portal culling, which culls occluded parts in the scene, e.g., behind closed doors. The LoD is dynamically adapted based on the projected Gaussian dimensions in the picture plane [64].

Appendage-free kernels and depth-sorted splatting, which are resistant to flicker and popping effects, are used to reduce those effects. Foveated rendering saves computation power by rendering only a central area of the vision at full pixels, and asynchronous reprojection prevents a smooth motion caused by varying frame rates. Predictive cell streaming streams the next portions of a scene in advance, and cache prewarming during teleportation prevents the loading process from stalling mid-experience. Using stereo disparity checks on large planar sections is useful to avoid discomfort caused by binocular inconsistencies [65].

F. *Authoring and Interactivity: Feedback Loops*

Beyond simple static playback, VR tour systems are increasingly accommodating interactive and diagnostic tools that can be used directly within headsets. Park and Kim [66] may alter the Gaussians or regions of the scene, allowing real-time concealment, colorization, and semantic validation through the headset interface.

Automated diagnostics detect potential quality problems, e.g., oversampled splats, low-sampled areas, or noisy surface normals, and make specific capture or constraint advisories. Such iterative loops during authoring reduce full pipeline reprocessing needs, thereby accelerating the refinement process and supporting additional nimble production cycles [67].

V. Datasets for Virtual Reality Style Evaluation

The development of successful VR tours requires data, as shown in Table IV, that capture the conditions users experience when immersed in the environment. Consequently, the dataset design should account for headset navigation limitations and incorporate plausible scene coverage, varied navigation styles, and various spatial setup patterns. Common VR viewing tasks involve indoor room-scale exploration, forward-looking paths, full 360° orbit capture, and outdoor traversal sequences [68]. Fig. 4 visualises the vast scale difference in the dataset (log scale), and it highlights the importance of scene count and capture diversity on benchmarking conclusions.

TABLE IV. Commonly Used Datasets for VR Style

Dataset	Samples_or_Scenes	Views	Type	VR Relevance/Notes	Ref.
Mip-NeRF 360	9	360° viewpoints	Videos	Unbounded indoor/outdoor scenes; standard for NVS in VR-like flythroughs.	[69]
Plenoptic Video	6	21 views	Videos	Dynamic real-world multiview sequences; used widely for 4D NVS.	[70]
DyNeRF dataset	6	20 views	Videos	Time-synchronized & calibrated multiview dynamic scenes.	[71]
NVOS	7	20–62 images	Images	Novel view object segmentation; front-facing real scenes with masks.	[72]
SPIn-NeRF	10	100 images	Images	Multiview segmentation and 3D inpainting dataset (with/without object).	[73]
LERF	13	—	Videos	Language-embedded radiance fields; long-tail household scenes.	[74]
ActorsHQ	16	160 views	Videos	12-MP multiview human sequences; useful for VR character inserts.	[75]
Tensor4D	4	1/4/12 views	Videos	Efficient dynamic 4D decomposition; small set of synchronized captures.	[76]
DL3DV-10K	10,510	variable	Videos	Large-scale real-world scenes at 4K; strong benchmark for NVS/VR tours.	[77]
SceneSplat-7K	7,916	—	3DGS scenes	Indoor 3DGS scenes aggregated across sources.	[78]

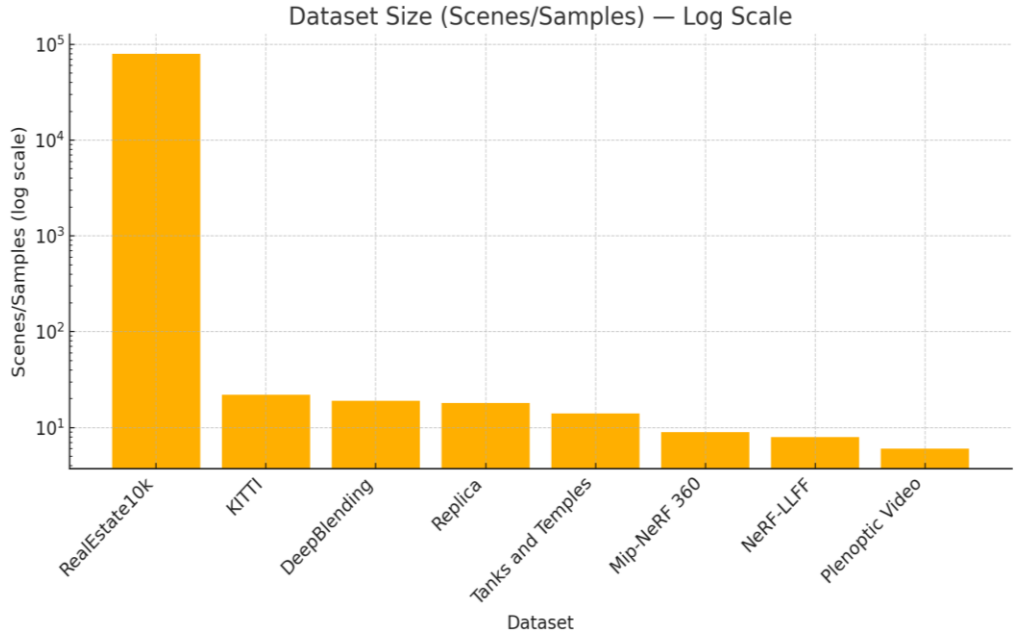

Fig. 4. Dataset size (scenes/samples)–log scale.

The choice of the dataset directly affects the evaluation outcomes because the geometry of the capture, or spatial resolution, and scene variability influence constancy, reconstruction complexity, and apparent realism [79]. To enable valid comparison across studies, it is essential to use carefully selected, replicable datasets as Table V.

TABLE V. CURATED DATASETS FOR VR TOURS

Dataset	Description and VR Relevance	Sources
Replica	18 photorealistic indoor scenes with high geometry density, HDR textures, semantic labels, and glass/mirror reflectors, making them suitable for room/building scale VR tours.	[80]
Mip-NeRF 360	Orbit, no horizon display, open scenery, designed for immersive, 360° navigation around central objects.	[81]

A. Room/Building Datasets

The Indoors dataset is essential for evaluating room-scale VR tour systems. This dataset generally exhibits good geometrical structures, sharp textures, and an interrelation of a few rooms.

The Replica dataset offers photorealism, fine meshes, and accurate geometry, along with abundant in-depth texture maps for various indoor environments [82]. The lifelike lighting and multiple layouts of this dataset are suitable for large-scale building tours. Similarly, Matterport3D has several training characteristics with many proliferating interior scenes, both commercial and residential [83]. The diversity of architectural styles and materials necessitates generalization testing across various dataset layouts. The Indoors and Replica datasets are useful for approaches that require exact structural geometry and high-textured surfaces and for evaluating multiroom navigation performance.

B. Orbit and 360° Datasets

Some VR tours use an orbit form of viewing in which a user moves around a central object or place while varying the user's elevation and line of sight. Such a presentation is widely used in museum displays and cultural heritage exhibitions.

The Mip-NeRF 360 dataset offers a complete 360° view of objects and scenes, enabling the assessment of the stability of rendering and its consistency in the context of progressive shifts in viewpoints [84]. Heritage360 records high-resolution 360° scans of culturally relevant artifacts and sites, enabling the creation of high-fidelity (object-based) VR experiences [85]. These datasets can also be used to validate methods that capture all-angle, accurate view-dependent reflectance while preserving fine geometric and material details.

C. Forward-Facing Datasets

In most VR applications, such as real estate walkthrough or street walking, the user tends to move mostly in a forward direction without a complete rotational orientation. This type of navigation expects datasets with forward-facing capturing geometry.

Structured forward-facing captures are available in the Local Light Field Fusion (LLFF) dataset. Thus, this dataset can be used to measure "3-DoF+" experiences, where lateral motion and depth are available but restricted [86]. The RealEstate10K dataset consists of tens of thousands of video frames obtained based on real estate announcements. This dataset provides various architectural and lighting styles and interior designs [87]. The LLFF dataset is particularly useful for pretraining and evaluating the robustness and performance of constrained-path navigation systems.

D. Driving and Outdoor Datasets

In the case of VR tours with large outdoor scenes, including campus and city tours, the dataset needs to support extended range sizes, changing lighting, and different geometry complexity levels.

The Tanks and Temples dataset comprises very detailed and rich models (indoors and outdoors) of complicated buildings and structures suitable for evaluating geometric precision [88]. The KITTI dataset was initially designed for autonomous driving studies. It consists of stereo imagery and depth data in the form of long driving routes [89]. The extensive motion sequences of the KITTI dataset are particularly interesting for investigating streaming capabilities, LoD control, and streaming stability during constant movement.

E. Benchmarking Recommendation

VR tour evaluation studies are recommended to explicitly define the dataset(s) used, capture the geometry, and align the data with the intended VR experience [90]. The integration of data that consider indoor, orbit style, guard-facing, and outdoor modalities enables a holistic evaluation of 3DGS-based VR methods, including their advantages and limitations across various functionality areas [91].

VI. EVALUATION METHODOLOGY

This survey evaluates 3DGS-based VR tour systems across four axes: visual fidelity, runtime performance, memory footprint, and interaction responsiveness. This section explains in detail the metrics and protocols that can be used to determine each axis during headset-based immersion [92]. Fig. 5 shows the steps of evaluation of 3DGS-based VR tour solutions, starting with rendering and along representative motion paths and moving on to profiling the runtime, memory, and the latency of interaction.

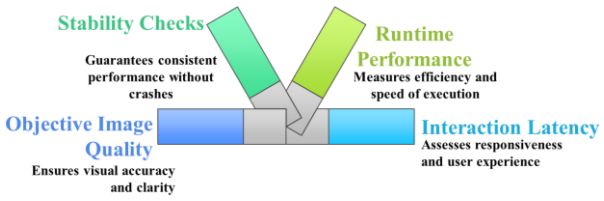

Fig. 5. Evaluation stages for 3DGS-based VR tour solutions.

A. Visual Fidelity

Visual fidelity evaluation uses objective image quality metrics combined with VR-related stability simulations.

It uses three traditional quantitative measures: Peak Signal-to-Noise (PSNR) for pixel-scale reconstruction accuracy; SSIM for luma and structure preservation; Learned Perceptual Image Patch Similarity (LPIPS) for perceptual similarity based on deep feature embeddings [93–95].

To assess temporal stability, which is vital to the headset sound, the video is always rendered along camera paths simulating head motion, including jittered, sinusoidal, and other resonant tracks. Artifacts such as shimmer, popping, and rolling-shutter distortions are quantified [96]. Reflective and refractive surfaces are examined in terms of view consistency, evaluating the correlation of the positions of the specular lobes with small shifting parallaxes [97, 98].

B. Runtime Performance and Memory Footprint

The key runtime performance metrics are median FPS and 1%-low FPS, measured on representative head motion sequences. These capture the overall and worst-case frame rates of the system [99]. The memory usage of the GPU in a steady state is measured to ensure compliance with standalone VR headset limitations.

Load time analysis measures the latency between scene selection and the first rendered frame, including the time for asset retrieval and shader compilation. The latency of streaming across the cells in space is logged to measure the availability of the asset on constant navigation [100]. To prevent performance overestimation, testing is conducted not only on mainstream desktop GPUs but also on headset-level hardware, which is subject to mobile memory bandwidth constraints and thermal throttling [101].

C. Interaction Responsiveness

Response time is a critical component for assessing systems that use semantic querying or in-headset authoring. A labeled collection of interest and freeform text retrieval tasks (e.g., “find the Monet painting”) can be used to test the accuracy of open vocabulary retrieval in terms of both precision and recall [102]. Edit latency measurements are used in everyday authoring tasks, including region visibility toggle, color changes, and textually driven recoloring.

These responsiveness measures are essential for evaluating the feasibility of 3DGS-based VR platforms in creative and production environments. Table VI quantifies how latency impacts narrative tasks, impedes real-time interactivity, and degrades the overall user experience [103].

TABLE VI. TYPE SIZES FOR FINAL PAPERS

Dimension	Metrics	Method	VR Relevance
Visual Fidelity	PSNP, SSIM, LPIPS; stability checks	Head motion paths; shimmer/popping detection.	Maintains immersion, prevents sickness.
Runtime and Memory	FPS (median/1%-low); GPU memory; load/stream times	Frame-rate logs; GPU usage; load/stream tests.	Smooth navigation on limited hardware.
Interaction Responsiveness	Query accuracy; edit latency	Retrieval and editing tasks in headset.	Supports real-time inaction and authoring.

Fig. 6 illustrates that VR-tour assessment is commonly conducted through a tradeoff between a triangle of antithetical objectives: visual fidelity, runtime/memory performance, and interaction responsiveness, which are limited to headset capacities.

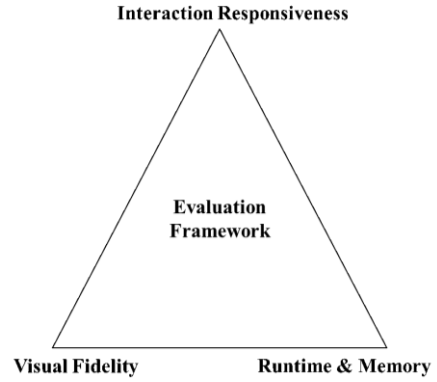

Fig. 6. Triangle of evaluation for GS-based VR tours.

D. Summary of Evaluation Framework

Presence and headset comfort cannot be measured using objective measures alone. Therefore, we augment PSNR/SSIM/LPIPS with subjective measures. These include presence ratings, simulator sickness questionnaire scores, and motion-to-photon latency targets: approximately 15 ms for PC VR, 20–22 ms for standalone headsets, and 25–28 ms for mobile VR. The proposed methodology integrates existing objective visual quality measures, time evolution integrity tests, runtime execution profiling, and interaction lag metrics to provide a comprehensive assessment of 3DGS-based VS tour systems. This multimodal solution strategy ensures both high reconstruction accuracy and smooth, comfortable, and responsive user experiences that are within the operation constraints of target VR platforms [104].

VII. CASE STUDIES AND DESIGN PATTERNS

A. Museum Gallery (Multiroom Indoor)

This case study aims to create a walkable, six Degrees-of-Freedom (6-DoF) VR tour of a museum (Figs. 7 and 8), focusing on challenging surfaces such as glass showcases and shiny floors, which are difficult to render due to their reflective nature [105].

Table VII provides a summary of the advised capture strategy, 3DGS approach selections, anticipated issues, and workaround solutions to museum gallery VR tours.

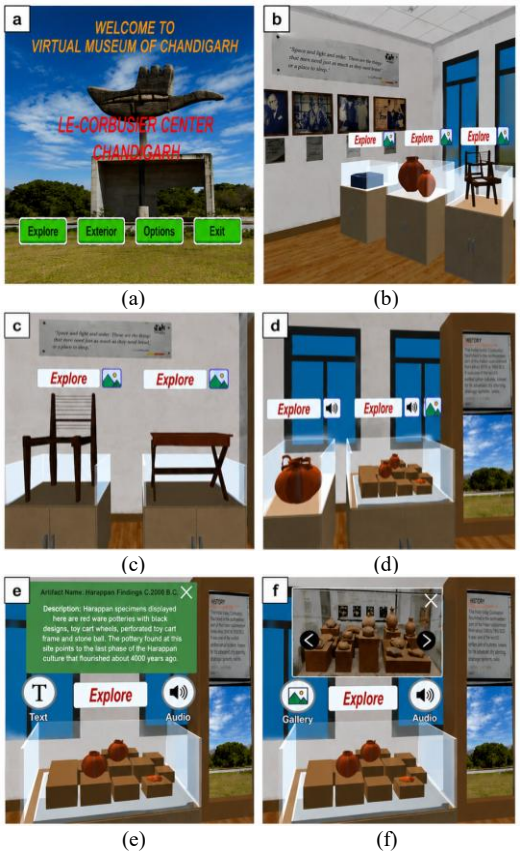

Fig. 7. Impact of VR on museum experiences: Visitor immersion and experience consequences. (a). VR tour main page; (b). Virtual museum view; (c). Artifact UI interaction; (d). Dual-channel interaction; (e). Artifact text & audio details; (f). Artifact gallery interaction [105].

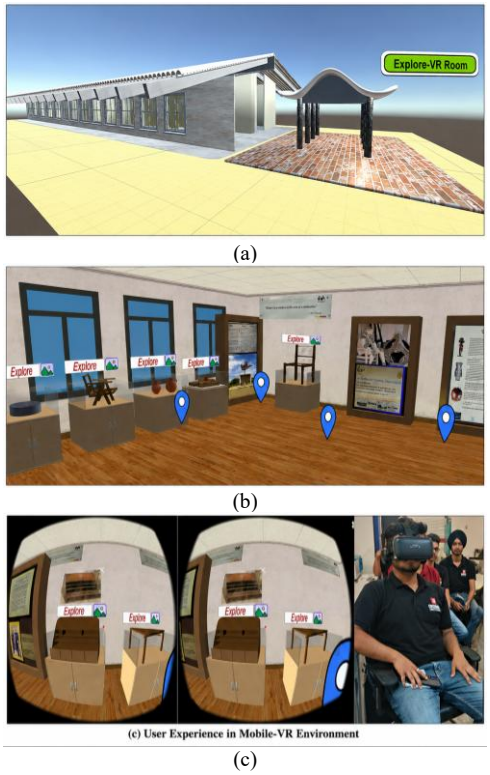

Fig. 8. Impact of VR on museum experiences: Visitor immersion and experience consequences. (a) Exterior view of the virtual museum; (b) Interior virtual room with interactive artifacts; (c) User experience in mobile-VR environment [105].

TABLE VII. DESIGN PATTERNS FOR MUSEUM GALLERY TOURS

Aspect	Details
Capture Strategy	Tight image sweeps (wall-to-wall, floor-to-ceiling); additional passes in glass cases and on metallic surfaces; detailed image sweep of corridors and doorways [106].
Methods	Geometry-sensitive regularization; shading-aware 3DGS; portal culling; LoD; semantic labels on Artworks for guided tours and annotations [107–110].
Challenges	Mirrored rooms with duplicate geometry; repeat textures; occluded corners lose detail.
Mitigation	Use laser-based depth scanning; increase viewpoints on reflective surfaces [111].

B. Real Estate Apartment (Medium Scale, Fast Turnaround)

The case study implies a rapid creation of a VR tour of a multi-II room apartment (Fig. 9) with adequate lighting reproduction [112].

A summary of capturing assumptions, preferred 3DGS methodologies, frequent failure modes (e.g., texture-less walls and mixed lighting), and mitigation strategies of fast-turnaround real-estate VR tours is summarized in Table VIII.

TABLE VIII. DESIGN PATTERNS FOR VR REAL ESTATE TOURS

Aspect	Details
Capture Strategy	1–2 images per room, with smooth doorway transitions and consistent illumination [113].
Methods	Depth regularized optimization; frequency-progressive densification; aggressive pruning and compression for efficient storage and fast loading [114–116].
Challenges	Texture-less walls distort geometry, mixed lighting results in uneven colors, and large windows create overexposed areas.
Mitigation	Apply planarity-constraint, white balance calibration, and sky probes for lighting correction [117].

Fig. 9. Online VR apartment tour demo in the browser using 360° images (on smartphone) [118].

C. Outdoor Campus Quad (Large Scale, Walkthrough)

This state of affairs presupposes a wide-ranging open-air VR tour (Fig. 10), with free movement within open areas with facilities, trees, and people (Fig. 11) [119].

Fig. 10. The scene manager UI [120].

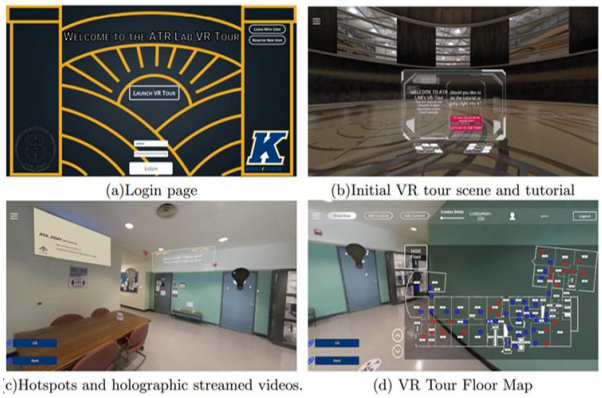

Fig. 11. Walkthrough demonstration of the KSU VR tour system [120].

Table IX establishes an overview of the end-to-end design patterns of VR tours of campuses on a large scale, such as capture strategy, scalable training/LoD render processes, common artefacts, and solutions.

TABLE IX. DESIGN PATTERNS FOR VR CAMPUS TOURS

Aspect	Details
Capture Strategy	Path-based video capture using LiDAR scanning for accurate geometry; primary walkways, quads, and landmarks covered [121].
Methods	Divide and conquer training by scene subdivision; octree-based LoD rendering; deformable Gaussians for dynamic objects (trees, flags); navigation graph for preloading assets [122–125].
Challenges	Flickering/ghosting from dynamic crowds; visible LoD transitions in long sightlines.
Mitigation	Hierarchical LoD control; smooth switching policies; masking and time-based blending of moving objects [126].

VIII. CHALLENGES

A. Streaming and Multi-User Scale

Large-scale VR tours require streaming architectures for prefetching and prioritizing assets ahead of unpredictable user movements. Motion prediction is critical to minimize stalling [127]. Multi-user editing in the same space remains an unresolved challenge, limited by latency, bandwidth constraints, and the difficulty of synchronizing annotations and interactions [126].

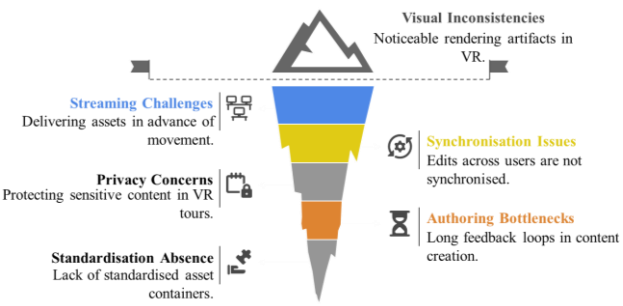

Fig. 12. Challenges in photorealistic VR tours.

In Fig. 12, a brief summary of the key open challenges to photorealistic VR tours streaming, reflectance, privacy, ergonomics, and standardisation discussed in this section is shown.

B. Lighting and Reflectance

3DGS pipelines are sensitive to mixed illumination, reflective surfaces, and transparency. Visual artifacts from reflections and refractions are magnified in VR, which breaks immersion. Future work should investigate inverse rendering methods for physically accurate material and visibility reasoning under complex lighting conditions [126].

C. Safety and Privacy

VR tours may include sensitive content, such as recognizable people, valuable artwork, and private places. Protecting user and content privacy requires automatic masking, blurring, or access control. Semantic segmentation in 3DGS pipelines can support redaction and visibility based on content type [127].

D. Ergonomics

Current authoring pipelines include lengthy capture, processing, and review cycles, making the iterative processes slow. In-headset diagnostics for coverage gaps, alignment errors, and reflective surface artifacts may reduce feedback loops. Improved ergonomics will facilitate faster and more reliable content creation [128].

E. Standardization and Benchmarks

There is no cohesive asset container for 3DGS that integrates geometry, appearance, semantics, and LoD compression. The lack of standardized formats makes it challenging to create interoperability between existing tools. Portable benchmarks and standard datasets are essential for comparing performance and facilitating content exchange [129].

IX. CONCLUSION

3DGS is redefining the potential for photorealistic VR tours by combining explicit scene representation, real-time rendering, and direct editability. Most neural view synthesis models are computationally expensive in both training and inference, particularly when deployed on mobile VR platforms. 3DGS is a feed-forward implementation, capable of supporting high frame rates, and its latency can be controlled, which is particularly advantageous in HMDs.

This survey demonstrates the unique advantages of 3DGS, such as shading-aware splats, anti-aliasing techniques, hierarchical LoD engines, pruning strategies, and learned priors, which reduce capture requirements. Collectively, these advances facilitate the creation of scalable and immersive workflows for museums, real estate, and campus environments.

Key challenges remain in streaming, reflective surfaces, ergonomics, privacy, and multi-user scalability. Addressing these issues via inverse rendering, semantic integration, compression, and standardization will improve performance and interoperability. With the rise of shared datasets and authoring toolchains, 3DGS is poised to be the building block for scalable, efficient, and high-quality VR tour systems.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

AUTHOR CONTRIBUTIONS

Md Mridul Hossain and SeongKi contributed to the study conception and design. Md Mridul Hossain conducted the research and wrote the paper; SeongKi Kim revised this paper, managed this project, and funded this work; all authors had approved the final version.

FUNDING

This research was supported by the Regional Innovation System & Education (RISE) program through the (Gwangju RISE Center), the Ministry of Education (MOE) and the (Gwangju Metropolitan City), Republic of Korea (2025-RISE-05-013). This research work was also supported by a National Research Foundation of Korea (NRF) grant, funded by the Korean government (MSIT) (NRF-2023R1A2C1005950).

REFERENCES

[1] A. M. Al-Ansi, M. Jaboob, A. Garad, and A. Al-Ansi, "Analyzing Augmented Reality (AR) and Virtual Reality (VR) recent development in education," *Social Sciences & Humanities Open*, vol. 8, no. 1, 100532, 2023. https://doi.org/10.1016/j.ssaho.2023.100532

[2] J. J. LaViola, "A discussion of cybersickness in virtual environments," *ACM SIGCHI Bulletin*, vol. 32, no. 1, pp. 47–56, 2000. https://doi.org/10.1145/333329.333344

[3] G. Premsankar, M. D. Francesco, and T. Taleb, "Edge computing for the internet of things: A case study," *IEEE Internet of Things Journal*, vol. 5, no. 2, pp. 1275–1284, 2018. https://doi.org/10.1109/jiot.2018.2805263

[4] A. S. Fernandes and S. K. Feiner, "Combating VR sickness through subtle dynamic field-of-view modification," in *Proc. 2016 IEEE Symposium on 3D User Interfaces (3DUI)*, 2016, pp. 201–210. https://doi.org/10.1109/3dui.2016.7460053

[5] B. Mildenhall, P. P. Srinivasan, M. Tancik, J. T. Barron, R. Ramamoorthi, and R. Ng, "NeRF: Representing scenes as neural radiance fields for view synthesis," *Communications of the ACM*, vol. 65, no. 1, pp. 99–106, 2021. https://doi.org/10.1145/3503250

[6] X. Chen, "Virtual Reality (VR) for cultural heritage preservation in China," *Asian Journal of Computing and Engineering Technology*, vol. 5, no. 1, 2024. https://doi.org/10.47604/ajcet.2807

[7] S. Kim and I. Nasim, "Human electromagnetic field exposure in 5G at 28 GHz," *IEEE Consumer Electronics Magazine*, vol. 9, no. 6, pp. 41–48, 2020. https://doi.org/10.1109/mce.2019.2956223

[8] E. Wang, Q. Dong, Y. Li, and Y. Zhang, "Estimation of node cache occupancy in satellite storage network," *IEEE Access*, vol. 9, pp. 122039–122050, 2021. https://doi.org/10.1109/access.2021.3109876

[9] . Slater and S. Wilbur, "A Framework for Immersive Virtual Environments (FIVE): Speculations on the role of presence in virtual environments," *Presence: Teleoperators and virtual environments*, vol. 6, no. 6, pp. 603–616, 1997. https://doi.org/10.1162/pres.1997.6.6.603

[10] M. Levoy, "Efficient ray tracing of volume data," *ACM Transactions on Graphics*, vol. 9, no. 3, pp. 245–261, 1990. https://doi.org/10.1145/78964.78965

[11] B. Kerbl, G. Kopanas, T. Leimkuehler, and G. Drettakis, "3D Gaussian splatting for real-time radiance field rendering," *ACM Transactions on Graphics*, vol. 42, no. 4, pp. 1–14, 2023. https://doi.org/10.1145/3592433

[12] Q. Hou and F. Liu, "Auxiliary features-guided super resolution for monte carlo rendering," *Computer Graphics Forum*, vol. 43, no. 1, e14987, 2023. https://doi.org/10.1111/cgf.14987

[13] T. Wu, Y. J. Yuan, L. X. Zhang *et al*., "Recent advances in 3D Gaussian splatting," *Computational Visual Media*, vol. 10, no. 4, pp. 613–642, 2024. https://doi.org/10.1007/s41095-024-0436-y

[14] A. Bunge, P. Herholz, O. Sorkine-Hornung, M. Botsch, and M. Kazhdan, "Variational quadratic shape functions for polygons and polyhedra," *ACM Transactions on Graphics*, vol. 41, no. 4, pp. 1–14, 2022. https://doi.org/10.1145/3528223.3530137

[15] L. Lin, X. Liao, G. Tan *et al*., "LiveRender: A cloud gaming system based on compressed graphics streaming," in *Proc. 22nd ACM International Conference on Multimedia (MM 2014)*, 2014, pp. 347–356.

[16] R. L. Cook, T. Porter, and L. Carpenter, "Distributed ray tracing," *ACM SIGGRAPH Computer Graphics*, vol. 18, no. 3, pp. 137–145, 1984. https://doi.org/10.1145/964965.808590

[17] K. Zhang, G. Riegler, N. Snavely, and V. Koltun, "NERF++: Analyzing and improving neural radiance fields," arXiv Preprint, arXiv:2010.07492, 2020.

[18] F. D. Goes, A. Butts, and M. Desbrun, "Discrete differential operators on polygonal meshes," *ACM Transactions on Graphics*, vol. 39, no. 4, 110, 2020. https://doi.org/10.1145/3386569.3392389

[19] B. Lee, H. Lee, X. Sun, U. Ali, and E. Park, "Deblurring 3D Gaussian splatting," in *Proc. Lecture Notes in Computer Science*, 2024, pp. 127–143. https://doi.org/10.1007/978-3-031-73636-0_8

[20] G. Wu, T. Yi, J. Fang *et al*., "4D Gaussian splatting for real-time dynamic scene rendering," in *Proc. IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2023, pp. 20310–20320. https://doi.org/10.48550/arxiv.2310.08528

[21] H. Zhang, J. Guo, J. Zhang *et al*., "Deep fourier-based arbitrary-scale super-resolution for real-time rendering," in *Proc. ACM SIGGRAPH 2024 Conference Papers*, 2024, pp. 1–11. https://doi.org/10.1145/3641519.3657439

[22] K. M. Stanney and R. S. Kennedy, "The psychometrics of cybersickness," *Communications of the ACM*, vol. 40, no. 8, pp. 66–68, 1997. https://doi.org/10.1145/257874.257889

[23] H. Buhler, S. Misztal, and J. Schild, "Reducing VR sickness through peripheral visual effects," in *Proc. 2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, 2018, pp. 517–519. https://doi.org/10.1109/vr.2018.8446346

[24] D. Zhu, R. Xu, X. Di *et al*., "Cache strategy for information center satellite networks based on node importance," in *Proc. International Conference on Emerging Networking Architecture and Technologies*, 2023, pp. 585–597. https://doi.org/10.1007/978-981-19-9697-9_47

[25] E. Ververas, R. A. Potamias, J. Song, J. Deng, and S. Zafeiriou, "SAGS: Structure-aware 3D Gaussian splatting," in *Proc. European Conference on Computer Vision*, 2024, pp. 221–238. https://doi.org/10.1007/978-3-031-72655-2_13

[26] R. Chandra and J. Simmons, "Bayesian neural networks via MCMC: A python-based tutorial," *IEEE Access*, vol. 12, pp. 70519–70549, 2024. https://doi.org/10.1109/access.2024.3401234

[27] L. Lyu, X. Ren, W. Cao, J. Zhu, E. Wu, and Z. Yang, "Wavelet potentials: An efficient potential recovery technique for pointwise incompressible fluids," *Computer Graphics Forum*, vol. 43, no. 2, 2024. https://doi.org/10.1111/cgf.15023

[28] F. Zhong, H. Zhao, H. Li *et al*., "VASCO: Volume and surface co-decomposition for hybrid manufacturing," *ACM Transactions on Graphics*, vol. 42, no. 6, pp. 1–17, 2023. https://doi.org/10.1145/3618324

[29] Z. Liang, Q. Zhang, Y. Feng, Y. Shan, and K. Jia, "GS-IR: 3D Gaussian splatting for inverse rendering," in *Proc. 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2024, pp. 21644–21653. https://doi.org/10.1109/cvpr52733.2024.02045

[30] P. Rottmann, P. Rodgers, X. Yan, D. Archambault, B. Wang, and J. Haunert, "Generating euler diagrams through combinatorial optimization," *Computer Graphics Forum*, vol. 43, no. 3, e15089, 2024. https://doi.org/10.1111/cgf.15089

[31] X. Tu, L. Radl, M. Steiner, M. Steinberger, B. Kerbl, and F. D. L. Torre, "VRSPlat: Fast and robust Gaussian splatting for virtual reality," *ACM on Computer Graphics and Interactive Techniques*, vol. 8, no. 1, pp. 1–22, 2025. https://doi.org/10.1145/3728311

[32] C. Yan, D. Qu, D. Xu *et al*., "GS-SLAM: Dense visual SLAM with 3D Gaussian splatting," in *Proc. 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2024, pp. 19595–19604. https://doi.org/10.1109/cvpr52733.2024.01853

[33] J. Lin, Z. Li, X. Tang *et al*., "Vastgaussian: Vast 3D Gaussians for large scene reconstruction," in *Proc. IEEE/CVF Conf. on Computer Vision and Pattern Recognition*, 2024, pp. 5166–5175.

[34] J. Zhao, J. Xu, P. Di *et al*., "Modeling the interplay between loop tiling and fusion in optimizing compilers using affine relations," *ACM Transactions on Computer Systems*, vol. 41, no. 1–4, pp. 1–45, 2023. https://doi.org/10.1145/3635305

[35] S. Zhou, H. Chang, S. Jiang *et al*., "Feature 3DGS: Supercharging 3D gaussian splatting to enable distilled feature fields," in *Proc. 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2024, pp. 21676–21685. https://doi.org/10.1109/cvpr52733.2024.02048

[36] Y. Zhang, S. Song, E. Yumer *et al*., "Physically-based rendering for indoor scene understanding using convolutional neural networks," in *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, 2016, pp. 5287–5295. https://doi.org/10.48550/arxiv.1612.07429

[37] J. Chung, J. Oh, and K. M. Lee, "Depth-regularized optimization for 3D gaussian splatting in few-shot images," in *Proc. 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, 2024, pp. 811–820. https://doi.org/10.1109/cvprw63382.2024.00086

[38] S. Ji, S. Xu, Q. Cheng, N. Xiao, C. Zhou, and M. Xiong, "A masked-pre-training-based fast deep image prior denoising model," *Applied Sciences*, vol. 14, no. 12, 5125, 2024. https://doi.org/10.3390/app14125125

[39] A. Vlachokostas and N. Madamopoulos, "Daylight and thermal harvesting performance evaluation of a liquid filled prismatic façade using the radiance five-phase method and energyplus," *Building and Environment*, vol. 126, pp. 396–409, 2017. https://doi.org/10.1016/j.buildenv.2017.10.017

[40] A. Kratimenos, J. Lei, and K. Daniilidis, "DynMF: Neural motion factorization for real-time dynamic view synthesis with 3D Gaussian splatting," in *Proc. European Conference on Computer Vision*, 2023, pp. 252–269.

[41] J. Huang, H. Yu, J. Zhang *et al*., "Point'n move: Interactive scene object manipulation on gaussian splatting radiance fields," *IET Image Processing*, vol. 18, no. 12, pp. 3507–3517, 2023.

[42] K. Zibrek, S. Martin, and R. McDonnell, "Is photorealism important for perception of expressive virtual humans in virtual reality?" *ACM Transactions on Applied Perception*, vol. 16, no. 3, pp. 1–19, 2019. https://doi.org/10.1145/3349609

[43] C. Kleinbeck, H. Schieber, K. Engel, R. Gutjahr, and D. Roth, "Multi-layer Gaussian splatting for immersive anatomy visualization," *IEEE Transactions on Visualization and Computer Graphics*, vol. 31, no. 5, pp. 2353–2363, 2025. https://doi.org/10.1109/tvcg.2025.3549882

[44] M. Colbert, S. Pattanaik, and J. Krivanek, "BRDF-Shop: Creating physically correct bidirectional reflectance distribution functions," *IEEE Computer Graphics and Applications*, vol. 26, no. 1, pp. 30–36, 2006. https://doi.org/10.1109/mcg.2006.13

[45] L. Xu, V. Agrawal, W. Laney *et al*., "VR-NeRF: High-fidelity virtualized walkable spaces," in *Proc. SIGGRAPH Asia 2023 Conference Papers*, 2023, pp. 1–12. https://doi.org/10.1145/3610548.3618139

[46] A. Hanson, A. Tu, V. Singla, M. Jayawardhana, M. Zwicker, and T. Goldstein, "PUP 3D-GS: Principled uncertainty pruning for 3D Gaussian splatting," in *Proc. Computer Vision and Pattern Recognition Conference*, 2024, pp. 5949–5958.

[47] S. Qiu, C. Wu, Z. Wan, and S. Tong, "High-fold 3D Gaussian splatting model pruning method assisted by opacity," *Applied Sciences*, vol. 15, no. 3, 1535, 2025. https://doi.org/10.3390/app15031535.

[48] X. Liu, X. Wu, P. Zhang, S. Wang, Z. Li, and S. Kwong, "CompGS: Efficient 3D scene representation via compressed gaussian splatting," in *Proc. 32nd ACM International Conference on Multimedia*, 2024, pp. 2936–2944. https://doi.org/10.1145/3664647.3681468

[49] P. V. Sander and J. L. Mitchell, "Progressive buffers: View-dependent geometry and texture lod rendering" *Third Eurographics Symposium on Geometry Processing*, 2006, pp. 1–18. https://doi.org/10.1145/1185657.1185826

[50] D. Li, S. S. Huang, Z. Lu, X. Duan, and H. Huang, "ST-4DGS: Spatial-temporally consistent 4D Gaussian splatting for efficient dynamic scene rendering," in *Proc. ACM SIGGRAPH 2024 Conference Papers*, 2024, pp. 1–11. https://doi.org/10.1145/3641519.3657520

[51] O. Gertsiy, "Research on graphic data formats for compact representation and comparison of images," *Transport Systems and Technologies*, no. 43, pp. 173–187, 2024. https://doi.org/10.32703/2617-9059-2024-43-14

[52] C. Ouerghemmi, M. Ertz, N. Bouslama, and U. Tandon, "The impact of Virtual Reality (VR) tour experience on tourists' intention to visit," *Information*, vol. 14, no. 10, 546, 2023. https://doi.org/10.3390/info14100546

[53] C. Y. Lin, T. L. Tseng, and T. H. Tsai, "A multi-machine and Multi-Modal Drift Detection (M2D2) framework for semiconductor manufacturing," *Applied Sciences*, vol. 15, no. 12, 6500, 2025. https://doi.org/10.3390/app15126500

[54] S. Pokuciński and D. Mrozek, "Methods and applications of space understanding in indoor environment—A decade survey," *Applied Sciences*, vol. 14, no. 10, 3974, 2024. https://doi.org/10.3390/app14103974

[55] T. Ely, S. Bhaskaran, N. Bradley, T. J. W. Lazio, and T. Martin-Mur, "Comparison of deep space navigation using optical imaging, pulsar Time-of-Arrival tracking, and/or radiometric tracking," *The Journal of the Astronautical Sciences*, vol. 69, no. 2, pp. 385–472, 2022. https://doi.org/10.1007/s40295-021-00290-z

[56] S. K. Choi, R. A. Ramirez, and T. H. Kwon, "Acquisition of high-resolution topographic information in forest environments using integrated UAV-LiDAR system: System development and field demonstration," *Heliyon*, vol. 9, no. 9, e20225, 2023. https://doi.org/10.1016/j.heliyon.2023.e20225

[57] S. Jiang, C. Jiang, and W. Jiang, "Efficient structure from motion for large-scale UAV images: A review and a comparison of SfM tools," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 167, pp. 230–251, 2020. https://doi.org/10.1016/j.isprsjprs.2020.04.016

[58] X. Chen, A. Wu, and Y. Han, "Capturing the spatio-temporal continuity for video semantic segmentation," *IET Image Processing*, vol. 13, no. 14, pp. 2813–2820, 2019. https://doi.org/10.1049/iet-ipr.2018.6479

[59] H. Luo, J. Zhang, X. Liu, L. Zhang, and J. Liu, "Large-scale 3D reconstruction from multi-view imagery: A comprehensive review," *Remote Sensing*, vol. 16, no. 5, 773, 2024. https://doi.org/10.3390/rs16050773.

[60] J. Terven, D. M. Cordova-Esparza, J. A. Romero-González, A. Ramírez-Pedraza, and E. A. Chávez-Urbiola, "A comprehensive survey of loss functions and metrics in deep learning," *Artificial Intelligence Review*, vol. 58, no. 7, 2025. https://doi.org/10.1007/s10462-025-11198-7

[61] A. Gruen, O. Kuebler, and P. Agouris, *Automatic Extraction of Man-Made Objects from Aerial and Space Images (II)*, Springer Science & Business Media, 1995.

[62] B. Fei, J. Xu, R. Zhang, Q. Zhou, W. Yang, and Y. He, "3D Gaussian splatting as new era: A survey," *IEEE Transactions on Visualization and Computer Graphics*, vol. 31, no. 8, pp. 4429–4449, 2024. https://doi.org/10.1109/tvcg.2024.3397828

[63] A. Polyviou and I. O. Pappas, "Chasing metaverses: Reflecting on existing literature to understand the business value of metaverses," *Information Systems Frontiers*, vol. 25, no. 6, pp. 2417–2438, 2023. https://doi.org/10.1007/s10796-022-10364-4

[64] K. Ren, L. Jiang, T. Lu *et al*., "Octree-GS: Towards consistent real-time rendering with LOD-structured 3D Gaussians," arXiv Preprint, arXiv:2403.17898, 2025. doi.org/10.1109/tpami.2025.3568201

[65] A. M. Aizenman, G. A. Koulieris, A. Gibaldi, V. Sehgal, D. M. Levi, and M. S. Banks, "The statistics of eye movements and binocular disparities during VR gaming: Implications for headset design," *ACM Transactions on Graphics*, vol. 42, no. 1, pp. 1–15, 2022. https://doi.org/10.1145/3549529

[66] S. M. Park and Y. G. Kim, "A metaverse: Taxonomy, components, applications, and open challenges," *IEEE Access*, vol. 10, pp. 4209–4251, 2022. https://doi.org/10.1109/access.2021.3140175

[67] A. F. Nogueira and M. Zenha-Rela, "Process mining software engineering practices: A case study for deployment pipelines," *Information and Software Technology*, vol. 168, 107392, 2023. https://doi.org/10.1016/j.infsof.2023.107392

[68] M. Rostami, P. Pradhan, N. Karki *et al*., "A comprehensive review of extended reality and its application in aerospace engineering,"

Progress in Aerospace Sciences, vol. 157, 101118, 2025. https://doi.org/10.1016/j.paerosci.2025.101118

[69] J. T. Barron, B. Mildenhall, D. Verbin, P. P. Srinivasan, and P. Hedman, "MIP-NERF 360: Unbounded anti-aliased neural radiance fields," in *Proc. IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2021, pp. 5470–5479.

[70] S. C. Chan, K. T. Ng, Z. F. Gan *et al*., "The plenoptic video," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 15, no. 12, pp. 1650–1659, 2005. https://doi.org/10.1109/tcsvt.2005.858616

[71] T. Li, M. Slavcheva, M. Zollhoefer *et al*., "Neural 3D video synthesis from multi-view video," in *Proc. IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2021, pp. 5521–5531.

[72] Z. Ren, A. Agarwala, B. Russell *et al*., "Neural volumetric object selection," in *Proc. IEEE/CVF Conf. on Computer Vision and Pattern Recognition*, 2022, pp. 6133–6142.

[73] A. Mirzaei, T. Aumentado-Armstrong, K. G. Derpanis *et al*., "SPIn-NeRF: Multiview segmentation and perceptual inpainting with neural radiance fields," in *Proc. IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2022, pp. 20669–20679.

[74] J. Kerr, C. M. Kim, K. Goldberg, A. Kanazawa, and M. Tancik, "LERF: Language embedded radiance fields," in *Proc. IEEE/CVF International Conference on Computer Vision*, 2023, pp. 19729–19739.

[75] Synthesiaresearch. GitHub—synthesiaresearch/humanrf: Official code for 'HumanRF: High-Fidelity Neural Radiance Fields for Humans in Motion'. *GitHub*. [Online]. Available: https://github.com/synthesiaresearch/humanrf

[76] R. Shao, Z. Zheng, H. Tu, B. Liu, H. Zhang, and Y. Liu, "Tensor4D: Efficient neural 4D decomposition for high-fidelity dynamic reconstruction and rendering," in *Proc. IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2022, pp. 16632–16642.

[77] L. Ling, Y. Sheng, Z. Tu *et al*., "DL3DV-10K: A Large-Scale scene dataset for deep learning-based 3D vision," in *Proc. IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2023, pp. 22160–22169.

[78] Y. Li, Q. Ma, R. Yang *et al*., "SceneSplat: Gaussian splatting-based scene understanding with vision-language pretraining," arXiv Preprint, arXiv:2503.18052, 2025. https://arxiv.org/abs/2503.18052v2

[79] P. Martin, A. Rodrigues, J. Ascenso, and M. P. Queluz, "NERF view synthesis: Subjective quality assessment and objective metrics evaluation," *IEEE Access*, vol. 13, pp. 26–41, 2024. https://doi.org/10.1109/access.2024.3522768

[80] J. Straub, T. Whelan, L. Ma *et al*., "The replica dataset: A digital replica of indoor spaces," arXiv Preprint, arXiv:1906.05797, 2019.

[81] D. Verbin, P. Hedman, B. Mildenhall *et al*., "Ref-NERF: Structured view-dependent appearance for neural radiance fields," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 47, no. 11, pp. 9426–9437, 2025. https://doi.org/10.1109/cvpr52688.2022.00541

[82] S. K. Ramakrishnan, A. Gokaslan, E. Wijmans *et al*., "Habitat-Matterport 3D Dataset (HM3D): 1000 large-scale 3D environments for embodied AI," arXiv Print, arXiv:2109.08238, 2021. https://doi.org/10.48550/arxiv.2109.08238

[83] A. Chang, A. Dai, T. Funkhouser *et al*., "Matterport3D: Learning from RGB-D data in indoor environments," in *Proc. 2021 International Conference on 3D Vision (3DV)*, 2017, pp. 667–676. https://doi.org/10.1109/3dv.2017.00081

[84] Z. X. Yin, J. Qiu, M. M. Cheng, and B. Ren, "Multi-space neural radiance fields," in *Proc. 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2023, pp. 12407–12416. https://doi.org/10.1109/cvpr52729.2023.01194

[85] T. Xue, A. E. Ali, T. Zhang, G. Ding, and P. Cesar, "CEAP-360VR: A continuous physiological and behavioral emotion annotation dataset for 360° VR videos," *IEEE Transactions on Multimedia*, vol. 25, pp. 243–255, 2021. https://doi.org/10.1109/tmm.2021.3124080

[86] B. Mildenhall, P. P. Srinivasan, R. Ortiz-Cayon *et al*., "Local light field fusion: Practical view synthesis with prescriptive sampling guidelines," *ACM Transactions on Graphics*, vol. 38, no. 4, pp. 1–14, 2019. https://doi.org/10.1145/3306346.3322980

[87] RealEstate10K: A large dataset of camera trajectories from video clips. [Online]. Available: https://google.github.io/realestate10k/

[88] Y. He, T. Foley, T. Hofstee, H. Long, and K. Fatahalian, "Shader components: Modular and high performance shader development," *ACM Transactions on Graphics*, vol. 36, no. 4, pp. 1–11, 2017. https://doi.org/10.1145/3072959.3073648

[89] A. Geiger, P. Lenz, C. Stiller, and R. Urtasun, "Vision meets robotics: The KITTI dataset," *The International Journal of Robotics Research*, vol. 32, no. 11, pp. 1231–1237, 2013. https://doi.org/10.1177/0278364913491297

[90] P. Strojny and N. Dużmańska-Misiarczyk, "Measuring the effectiveness of virtual training: A systematic review," *Computers & Education X Reality*, vol. 2, 100006, 2023. https://doi.org/10.1016/j.cexr.2022.100006

[91] S. F. Alfalah, W. Chan, S. Khan *et al*., "Gait Analysis Data Visualisation in Virtual Environment (GADV/VE)," in *Proc. 2014 Science and Information Conference*, 2014, pp. 742–751. https://doi.org/10.1109/sai.2014.6918270

[92] M. T. Bagdasarian, P. Knoll, Y. Li *et al*., "3DGS.zip: A survey on 3D Gaussian splatting compression methods," *Computer Graphics Forum*, vol. 44, no. 2, e70078, 2025. https://doi.org/10.1111/cgf.70078

[93] R. Zhang, P. Isola, A. A. Efros, E. Shechtman, and O. Wang, "The unreasonable effectiveness of deep features as a perceptual metric," in *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2018)*, 2018, pp. 586–595. https://doi.org/10.1109/cvpr.2018.00068

[94] H. R. Sheikh and A. C. Bovik, "Image information and visual quality," *IEEE Transactions on Image Processing*, vol. 15, no. 2, pp. 430–444, 2006. https://doi.org/10.1109/tip.2005.859378

[95] Z. Li, S. Lu, Z. Yuan, B. Hou, and J. Bian, "Interactive instance search: User-centered enhanced image retrieval with learned perceptual image patch similarity," *Electronics*, vol. 14, no. 9, 1766, 2025. https://doi.org/10.3390/electronics14091766

[96] M. D. B. Machuca and J. Fotouhi, "IEEE VR 2022," *IEEE Transactions on Visualization and Computer Graphics*, vol. 28, no. 5, p. xiii, 2022. https://doi.org/10.1109/tvcg.2022.3156769

[97] A. Cannavò, F. G. Pratticò, A. Bruno, and F. Lamberti, "AR-MoCap: Using augmented reality to support motion capture acting," in *Proc. 2023 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, 2023, pp. 318–327. https://doi.org/10.1109/vr55154.2023.00047

[98] I. Santesteban, E. Garces, M. A. Otaduy, and D. Casas, "SOFTSMPL: Data-driven modeling of nonlinear soft-tissue dynamics for parametric humans," *Computer Graphics Forum*, vol. 39, no. 2, pp. 65–75, 2020. https://doi.org/10.1111/cgf.13912

[99] R. Sasaki, M. Yasumi, and T. Matsuda, "Performance evaluation of the low-cost virtual reality system for real-time visualization of underwater environments," in *Proc. 2025 IEEE Underwater Technology*, 2025, pp. 1–4. https://doi.org/10.1109/ut61067.2025.10947405

[100] H. Gao, L. Hasenbein, E. Bozkir, R. Göllner, and E. Kasneci, "Evaluating the effects of virtual human animation on students in an immersive VR classroom using eye movements," in *Proc. 28th ACM Symposium on Virtual Reality Software and Technology*, 2022, pp. 1–11. https://doi.org/10.1145/3562939.3565623

[101] Tom, GitHub—authorTom/notes-on-VR-performance: How to optimise VR performance. *GitHub*. [Online]. Available: https://github.com/authorTom/notes-on-VR-performance

[102] J. Jang, W. Frier, and J. Park, "Multimodal volume data exploration through mid-air haptics," in *Proc. 2022 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, 2022, pp. 243–251. https://doi.org/10.1109/ismar55827.2022.00039

[103] S. V. Damme, J. Sameri, S. Schwarzmann *et al*., "Impact of latency on QoE, performance, and collaboration in interactive multi-user virtual reality," *Applied Sciences*, vol. 14, no. 6, 2290, 2024. https://doi.org/10.3390/app14062290

[104] M. A. AlGerafi, Y. Zhou, M. Oubibi, and T. T. Wijaya, "Unlocking the potential: A comprehensive evaluation of augmented reality and virtual reality in education," *Electronics*, vol. 12, no. 18, 3953, 2023. https://doi.org/10.3390/electronics12183953

[105] S. Jangra, G. Singh, A. Mantri, Z. Ahmed, T. W. Liew, and F. Ahmad, "Exploring the impact of virtual reality on museum experiences: Visitor immersion and experience consequences,"

Virtual Reality, vol. 29, no. 2, 84, 2025. https://doi.org/10.1007/s10055-025-01140-1

[106] J. M. Morcillo, F. Schaaf, R. H. Schneider, and C. Y. R. V. Trotha, "Authenticity through VR-based documentation of cultural heritage. A theoretical approach based on conservation and documentation practices," *Virtual Archaeology Review*, vol. 8, no. 16, pp. 35–43, 2017. https://doi.org/10.4995/var.2017.5932

[107] X. Zhao, "Playfulness, realism and authenticity in cultural presence: A case study of virtual heritage players," *Body Space & Technology*, vol. 20, no. 1, pp. 106–115, 2021. https://doi.org/10.16995/bst.352

[108] M. Slater and M. V. Sanchez-Vives, "Enhancing our lives with immersive virtual reality," *Frontiers in Robotics and AI*, vol. 3, 74, 2016. https://doi.org/10.3389/frobt.2016.00074

[109] A. Wang, M. Thompson, C. Uz-Bilgin, and E. Klopfer, "Authenticity, interactivity, and collaboration in virtual reality games: Best practices and lessons learned," *Frontiers in Virtual Reality*, vol. 2, 734083, 2021. https://doi.org/10.3389/frvir.2021.734083

[110] Z. Lu, Y. Wang, W. Yu, and Q. An, "Research on the exhibition design process of cultural elements in virtual reality interaction: Taking the grand canal culture as an example," in *Proc. Lecture Notes in Computer Science*, 2025, pp. 91–97. https://doi.org/10.1007/978-3-031-93227-4_7

[111] H. Mara, S. Krömker, S. Jakob, and B. Breuckmann, "GigaMesh and gilgamesh 3D multiscale integral invariant cuneiform character extraction," in *Proc. 11th International Conference on Virtual Reality, Archaeology and Cultural Heritage (VAST 2010)*, ACM, 2010, pp. 131–138. https://doi.org/10.2312/vast/vast10/131-138

[112] M. Carrozzino and M. Bergamasco, "Beyond virtual museums: Experiencing immersive virtual reality in real museums," *Journal of Cultural Heritage*, vol. 11, no. 4, pp. 452–458, 2010. https://doi.org/10.1016/j.culher.2010.04.001

[113] L. P. Berg and J. M. Vance, "Industry use of virtual reality in product design and manufacturing: A survey," *Virtual Reality*, vol. 21, no. 1, pp. 1–17, 2016. https://doi.org/10.1007/s10055-016-0293-9

[114] M. Bolognesi, A. Furini, V. Russo, A. Pellegrinelli, and P. Russo, "Accuracy of cultural heritage 3D models by RPAS and terrestrial photogrammetry," *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 40, pp. 113–119, 2014. https://doi.org/10.5194/isprsarchives-xl-5-113-2014

[115] P. Fuchs, G. Moreau, and P. Guitton, *Virtual Reality: Concepts and Technologies*, CRC Press. 2011. https://doi.org/10.1201/b11612

[116] J. M. P. V. Waveren, "The asynchronous time warp for virtual reality on consumer hardware," in *Proc. 22nd ACM Conference on Virtual Reality Software and Technology*, 2016, pp. 37–46. https://doi.org/10.1145/2993369.2993375

[117] B. Rieck, H. Mara, and S. Krömker, "Unwrapping highly-detailed 3D meshes of rotationally symmetric man-made objects," in *Proc. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 2013, pp. 259–264. https://doi.org/10.5194/isprsannals-ii-5-w1-259-2013

[118] B. A. Deaky and A. L. Parv, "Virtual reality for real estate—A case study," *IOP Conference Series Materials Science and Engineering*, vol. 399, no. 1, 012013, 2018. https://doi.org/10.1088/1757-899x/399/1/012013

[119] Z. Allam, A. Sharifi, S. E. Bibri, D. S. Jones, and J. Krogstie, "The metaverse as a virtual form of smart cities: Opportunities and challenges for environmental, economic, and social sustainability in urban futures," *Smart Cities*, vol. 5, no. 3, pp. 771–801, 2022. https://doi.org/10.3390/smartcities5030040

[120] S. Hendricks, A. Shaker, and J. H. Kim, "Design of a VR-based campus tour platform with a user-friendly scene asset management system," in *Lecture Notes in Computer Science*, 2022, pp. 337–348. https://doi.org/10.1007/978-3-030-98404-5_32

[121] A. Campiani, S. McAvoy, N. Lercari *et al*., "Developing an interoperable cloud-based isualization workflow for 3D archaeological heritage data: The Palenque 3D Archaeological Atlas," *Digital Applications in Archaeology and Cultural Heritage*, vol. 31, e00293, 2023. https://doi.org/10.1016/j.daach.2023.e00293

[122] J. Hong and C. Kim, "Animating smoke with dynamic balance," *Computer Animation and Virtual Worlds*, vol. 16, no. 3–4, pp. 405–414, 2005. https://doi.org/10.1002/cav.87

[123] J. Shade, D. Lischinski, D. H. Salesin, T. DeRose, and J. Snyder, "Hierarchical image caching for accelerated walkthroughs of complex environments," in *Proc. 23rd Annual Conference on Computer Graphics and Interactive Techniques*, 1996, pp. 75–82. https://doi.org/10.1145/237170.237209

[124] E. Kalogerakis, S. Chaudhuri, D. Koller, and V. Koltun, "A probabilistic model for component-based shape synthesis," *ACM Transactions on Graphics*, vol. 31, no. 4, pp. 1–11, 2012. https://doi.org/10.1145/2185520.2185551

[125] D. Luebke, M. Reddy, J. D. Cohen, A. Varshney, B. Watson, and R. Huebner, *Level of Detail for 3D Graphics*, Elsevier, 2012.

[126] Y. K. Dwivedi, L. Hughes, A. M. Baabdullah *et al*., "Metaverse beyond the hype: Multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice and policy," *International Journal of Information Management*, vol. 66, 102542, 2022. https://doi.org/10.1016/j.ijinfomgt.2022.102542

[127] H. Kiya, T. Nagamori, S. Imaizumi, and S. Shiota, "Privacy-preserving semantic segmentation using vision transformer," *Journal of Imaging*, vol. 8, no. 9, 233, 2022. https://doi.org/10.3390/jimaging8090233

[128] Y. Shen, B. Li, J. Huang *et al*., "GaussianShopVR: Facilitating immersive 3D authoring using Gaussian splatting in VR," in *Proc. 2022 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*, 2025, pp. 1292–1293. https://doi.org/10.1109/vrw66409.2025.00292

[129] L. Franke, L. Fink, and M. Stamminger, "VR-Splatting: Foveated radiance field rendering via 3D gaussian splatting and neural points," in *Proc. ACM on Computer Graphics and Interactive Techniques*, vol. 8, no. 1, pp. 1–21, 2025. https://doi.org/10.1145/3728302