PanoGS: Foundation-Depth-Guided High-Resolution Gaussian Splatting from Sparse Panoramic Images for Interactive Web3D Exploration

Uzair Shah
College of Science & Engineering,
Hamad Bin Khalifa University
Qatar
ushah@hbku.edu.qa

Giovanni Pintore
Visual and Data-Intensive Computing,
CRS4
Italy
giovanni.pintore@crs4.it

Muhammad Tukur
College of Science & Engineering,
Hamad Bin Khalifa University
Qatar
mtukur@hbku.edu.qa

Wala El-Sharif
College of Science & Engineering,
Hamad Bin Khalifa University
Qatar
welsharif@hbku.edu.qa

Jens Schneider
College of Science & Engineering,
Hamad Bin Khalifa University
Qatar
jeschneider@hbku.edu.qa

Alberto Jaspe-Villanueva
CEMSE, KAUST
Saudi Arabia
alberto.jaspe@kaust.edu.sa

Enrico Gobbetti
Visual and Data-Intensive Computing,
CRS4
Italy
enrico.gobbetti@crs4.it

Marco Agus*
College of Science & Engineering,
Hamad Bin Khalifa University
Qatar
magus@hbku.edu.qa

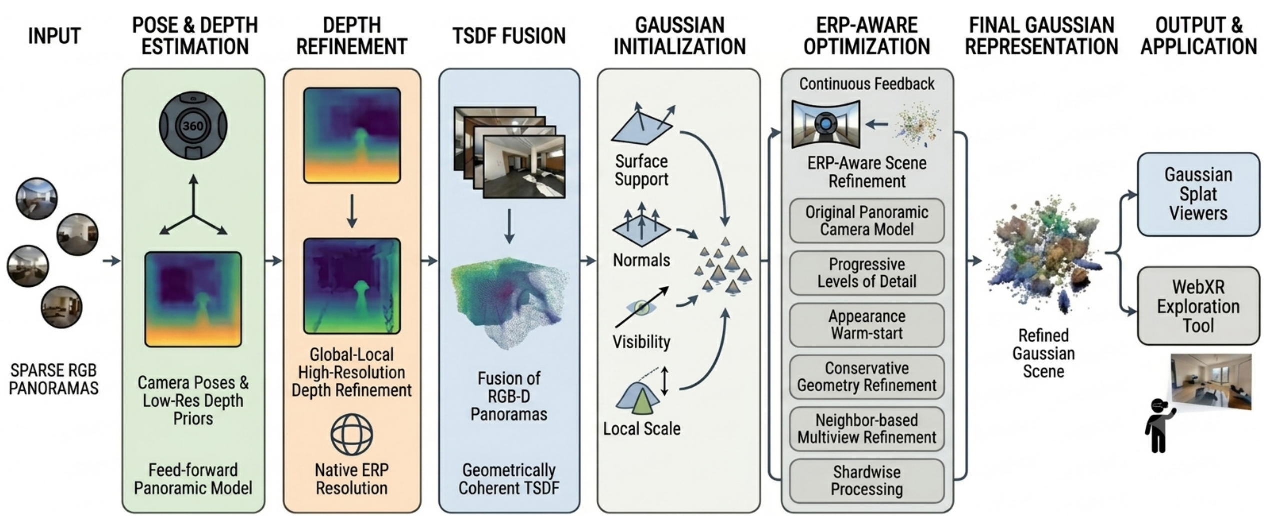

Figure 1: PanoGS takes as input a sparse set of indoor equirectangular images and produces a web-based high-resolution Gaussian Splatting scene for interactive inspection.

*Corresponding author.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

ACM Siggraph Web3D '26, Doha, Qatar

© 2026 Copyright held by the owner/author(s). Publication rights licensed to ACM.

ACM ISBN /26/01
https://doi.org/10.1145/3563045.3563055

Abstract

Sparse panoramic capture enables efficient indoor digitization, but transforming a few 360° images into web-based, high-resolution 3D representations remains difficult. Limited viewpoint overlap, equirectangular (ERP) distortions, uncertain depth scale, and the perspective-camera assumptions of most image-to-3D pipelines make direct reconstruction unreliable. This paper introduces PanoGS, a pipeline that uses inferred depth from vision foundation models to guide the generation of high-resolution 3D Gaussian Splatting scenes from sparse panoramic images. Given a small set of

RGB panoramas, we first estimate camera poses and low-resolution depth priors using a feed-forward panoramic reconstruction model. We then refine these priors at native ERP resolution through global-local depth refinement and fuse the resulting RGB-D panoramas into a Truncated Signed Distance Function (TSDF) geometric substrate, which provides stable surface support, normals, visibility information, and local scale estimates for initializing anisotropic Gaussian primitives. The initialized scene is further refined using an ERP-aware Gaussian optimization strategy with progressive levels of detail, appearance warm-start, conservative geometry updates, neighbor-based multiview refinement, and shardwise processing for large scenes. The resulting assets can be exported to Gaussian splat viewers and browser-based inspection tools. Results on indoor panoramic scenes demonstrate that the proposed guided formulation produces more stable and complete Gaussian scenes than direct depth-based seeding. At the same time, high-resolution depth refinement improves fine geometric and visual detail. PanoGS shows how sparse 360° captures can be transformed into practical web-based Gaussian representations for Web3D applications. [1]

CCS Concepts

• **Computing methodologies** → **Computer graphics**; **Computer vision**; **Machine learning**; **Reconstruction**.

Keywords

360, panoramic images, surround-view images, omnidirectional images, exploration, Web3D, Gaussian Splatting, browser-based visualization

ACM Reference Format:

Uzair Shah, Giovanni Pintore, Muhammad Tukur, Wala El-Sharif, Jens Schneider, Alberto Jaspe-Villanueva, Enrico Gobbetti, and Marco Agus. 2026. PanoGS: Foundation-Depth-Guided High-Resolution Gaussian Splatting from Sparse Panoramic Images for Interactive Web3D Exploration. In *The 31st International Conference on 3D Web Technology (Web3D '26), October 13–15, 2026, Doha, Qatar*. ACM, New York, NY, USA, 9 pages. https://doi.org/10.1145/3563045.3563055

1 Introduction

Panoramic imagery has become a practical capture modality for indoor digitization, virtual tours, interactive visualization, and web-based 3D applications [Pintore et al. 2026]. A single equirectangular (ERP) panorama captures the full surrounding environment from one viewpoint, making it possible to cover large indoor spaces with far fewer images than conventional perspective capture. This acquisition model is particularly attractive in architectural documentation, cultural heritage, real-estate visualization, industrial inspection, and digital twin creation, where fast capture and interactive exploration are often more important than dense image acquisition [Tukur et al. 2025]. However, transforming a sparse set of 360° images into a high-quality navigable 3D scene remains difficult. Sparse panoramic captures typically exhibit wide baselines, limited overlap, large textureless regions, occlusions, and strong ERP distortions, all of which complicate camera estimation, depth recovery, and view-consistent rendering [Guo et al. 2026].

Recent neural scene representations have significantly improved novel-view synthesis and real-time rendering. In particular, 3D Gaussian Splatting (3DGS) represents scenes with explicit anisotropic Gaussian primitives and enables high-quality rasterization-based rendering at interactive rates [Bao et al. 2025a]. This makes Gaussian Splatting especially appealing for interactive web-based applications, where responsiveness and visual fidelity are both critical. Nevertheless, standard 3DGS pipelines are usually designed around dense perspective image collections and reliable Structure-from-Motion initialization [Jiang et al. 2025]. When applied to sparse panoramic inputs, they face two major issues: first, accurate geometry is hard to recover from few wide-baseline observations; second, panoramic images are not naturally handled by perspective-camera assumptions and require careful treatment of spherical projection, distortion, and visibility [Pintore et al. 2026].

Several recent works have started to address these limitations by combining Gaussian Splatting with feed-forward reconstruction [Jiang et al. 2025]. Generalizable 3DGS methods predict Gaussian primitives directly from sparse images, reducing or eliminating per-scene optimization. Panoramic Gaussian Splatting has also received increasing attention [Chen et al. 2025], and they have started to obtain high-quality results. However, such approaches are primarily learned, feed-forward, novel-view synthesis systems. They depend on training data, operate within the resolution and memory constraints of their networks, and output Gaussian primitives through learned prediction. In contrast, many Web3D and digital-twin scenarios involve scene-specific reconstruction from high-resolution panoramas, where one may prefer a modular pipeline that explicitly leverages pose/depth foundation models, native-resolution depth refinement, geometric fusion, and controllable Gaussian optimization.

Another important observation is that high-resolution depth matters. Existing panoramic depth estimators often operate at reduced resolutions because holistic 360° processing is computationally demanding. This creates a gap between the resolution of modern panoramic cameras and the geometry available to downstream reconstruction systems. Recent methods addressed this gap by combining coarse global ERP depth priors with high-resolution local perspective refinement on gnomonic projections [Shah et al. 2026]. For Gaussian scene construction, this is particularly relevant: the density, position, anisotropic scale, and visibility of Gaussian primitives depend directly on the quality and resolution of the underlying geometric support. Low-resolution or naively upsampled depth tends to blur depth discontinuities, remove thin structures, and produce unstable Gaussian initialization.

In this paper, we present *PanoGS*, a pipeline for reconstructing high-resolution 3D Gaussian Splatting scenes from sparse panoramic images, using guidance from inferred depth. Rather than treating the feed-forward depth or Gaussian prediction as the final representation, we use panoramic vision-foundation-model outputs as initialization for an explicit reconstruction pipeline. Given a sparse set of RGB panoramas, PanoGS first estimates camera poses and low-resolution depth priors using a feed-forward panoramic reconstruction model. These depth priors are then refined at native ERP resolution using a global-local high-resolution depth refinement stage. The resulting RGB-D panoramas are fused

[1]A demonstration video accompanying this paper is available at https://drive.google.com/file/d/1oj243hKgK-gF6oB2wRSqtByCG0BrmIkD/view.

into a Truncated Signed Distance Function (TSDF) geometric substrate, which supplies coherent surface support, normals, visibility cues, and local scale estimates for initializing anisotropic Gaussian primitives. The initialized Gaussian scene is then refined using an ERP-aware optimization strategy. This design keeps the original panoramic camera model during rendering and optimization, avoiding unnecessary loss of information due to early conversion into multiple perspective views. To support large scenes and high native resolutions, the optimization is organized through progressive levels of detail, appearance warm-start, conservative geometry refinement, neighbor-based multiview refinement, and shardwise processing. The final scene can be exported to Gaussian splat viewers and browser-based inspection tools, enabling interactive interactive exploration of reconstructed indoor environments. In summary, we introduce:

(1) A sparse panoramic reconstruction pipeline guided by inferred depths from vision foundation models that converts a small set of 360° RGB images into high-resolution Gaussian Splatting scenes suitable for browser-based interactive exploration.
(2) A native-resolution depth-to-Gaussian initialization strategy that uses high-resolution ERP RGB-D panoramas and TSDF fusion to derive stable surface support, normals, visibility cues, and anisotropic Gaussian scales.
(3) An ERP-aware Gaussian optimization workflow combining progressive LODs, appearance warm-start, conservative geometry updates, neighbor-based multiview refinement, and shardwise processing for large panoramic scenes.

By combining panoramic foundation models, native-resolution depth refinement, TSDF-based geometric substrate generation, and explicit Gaussian optimization, PanoGS aims to bridge the gap between efficient 360° capture and practical web-based 3D scene exploration, especially in Web3D settings.

2 Related Work

We focus here on the work most closely related to ours and refer readers to comprehensive surveys for broader context on sparse panoramic Web3D scene creation [Ai et al. 2025; Bao et al. 2025a; Chen and Wang 2026; da Silveira et al. 2022; Fei et al. 2025; Pintore et al. 2026].

Interactive Exploration from Sparse Panoramic Input. Panoramic 3D representations offer efficient full-scene capture for applications like digital twins and virtual tours [Tukur et al. 2025]. To support local viewer motion, classic pipelines pair panoramas with inferred depth maps to synthesize disoccluded regions [Huang et al. 2017; Tukur et al. 2023; Wei et al. 2018; Zhao et al. 2013], frequently using end-to-end networks [Pintore et al. 2023; Xu et al. 2021] or precomputed structures like MSI [Attal et al. 2020], MCI [Waidhofer et al. 2022], and stereo-optimized views [Pintore et al. 2024]. Because these representations restrict novel views to small local volumes, larger translations rely on path interpolation between capture points [Chen et al. 2023; Li et al. 2022b; Pintore et al. 2025; Song and Zhang 2025; Zhao et al. 2013]. Generating unified 3D scenes directly from sparse captures bypasses these navigation constraints.

3D Gaussian Splatting for Interactive Exploration and Panoramic Rendering. 3D Gaussian Splatting (3DGS) has emerged as a dominant representation for real-time neural rendering due to its explicit anisotropic primitives and fast rasterization [Bao et al. 2025a; Chen and Wang 2026; Fei et al. 2025]. To bypass dense capture and scene-specific optimization, recent sparse feed-forward methods leverage geometric priors: DepthSplat [Xu et al. 2025] links depth prediction and Gaussian splatting via mutual supervision, while AnySplat [Jiang et al. 2025] jointly predicts poses and primitives using VGGT geometry priors [Wang et al. 2025]. These approaches demonstrate the importance of geometry priors for Gaussian scene generation, but they are mainly designed for perspective or general unstructured image collections rather than native high-resolution panoramic captures. Panoramic GS has recently attracted attention because 360° sparse capture introduces wide-baseline geometry, projection distortions, and large image resolutions that are not naturally handled by perspective 3DGS pipelines [Gu et al. 2026]. For instance, Splatter-360 [Chen et al. 2025] builds a spherical cost volume with a combined ERP-cubemap encoder; CylinderSplat [Wang et al. 2026] uses a cylindrical Triplane representation to capture the unique geometric properties of panoramic images and adhere to the Manhattan-world assumption; and UniTriSplat[Zhu et al. 2026] reformulates Gaussian splatting on the unit sphere using HEALPix discretizationx [Gorski et al. 2005] to create a unified framework for universal cameras. Other recent works address layout guidance, spherical representations, unbounded view synthesis, seamless omnidirectional splatting, and high-resolution panorama synthesis [Bai et al. 2025; Bao et al. 2025b; Li et al. 2025; Ren et al. 2025; Shin et al. 2025; Zhang et al. 2025; Zhuang et al. 2026]. Unlike these generalizable or feed-forward view synthesis methods, our modular, scene-specific Web3D pipeline transforms sparse panoramas into Gaussian scenes through explicit pose/depth initialization, native-resolution refinement, TSDF geometric substrate generation, and controllable ERP-aware optimization. TSDF combined with 3DGS has previously been employed for surface reconstruction [Lyu et al. 2024; Noda et al. 2026].

3D Reconstruction from Sparse Panoramic Images. Surveys on omnidirectional vision emphasize the benefits of 360° capture, but also identify persistent challenges like ERP distortions, sparse viewpoints, limited overlap, and scale ambiguity [Ai et al. 2025; da Silveira et al. 2022], For indoor modeling, layout, gravity alignment, and high-resolution geometry have also key roles [Pintore et al. 2026]. To recover geometry from sparse captures, classic/hybrid pipelines (e.g., IM360 [Jung et al. 2025a], EDM [Jung et al. 2025b]) and recent feed-forward models (such as VGGT [Wang et al. 2025], DUSt3R [Wang et al. 2024], PanoVGGT [Guo et al. 2026], Wid3R [Jung et al. 2026], and MapAnything [Keetha et al. 2026]) leverage learned priors for initial pose, depth, and structural estimation. While essential for bootstrapping our pipeline, these methods do not produce high-resolution, web-ready Gaussian assets. To balance global consistency and local detail in high-resolution panoramic depth, existing approaches employ projection fusion (OmniFusion [Li et al. 2022a], 360MonoDepth [Rey-Area et al. 2022]), regional fusion (HRDFuse [Ai et al. 2023]), registration [Peng and Zhang 2023], or foundation models [Cao et al. 2025; Lin et al. 2025]. FRED [Shah et al. 2026] specifically refines low-resolution priors into native ERP depth via gnomonic local refinement and seam-robust fusion. Unlike FRED and related depth-only methods, we treat refined depth as an intermediate step, fusing it into a TSDF representation [Werner et al. 2014] as a geometric substrate for Gaussian scene construction. Our pipeline, therefore, extends beyond depth maps, point clouds, or neural view synthesis. We use sparse panoramic geometry to provide surface support, normals, visibility, and local scale estimates for anisotropic Gaussian initialization—reducing floaters, duplicate structures, and inconsistent primitive scales during Web3D asset creation.

3 Methods

PanoGS converts a sparse set of high-resolution ERP panoramas into a renderable 3D Gaussian Splatting scene for web-based interactive exploration. In the following, we first provide a general overview of the pipeline (Sec. 3.1), and then illustrate the detail of each of the pipeline steps (Sec. 3.2-Sec. 3.6).

3.1 Pipeline Overview

The input is a set of RGB panoramas,

$$\mathcal{I} = \{I_k\}_{k=1}^{N}, \tag{1}$$

together with camera poses and depth maps produced by an upstream panoramic reconstruction stage. In our implementation, PanoVGGT [Guo et al. 2026] provides the initial camera poses and low-resolution depth priors, while FRED [Shah et al. 2026] refines these priors at native ERP resolution. PanoGS then converts the resulting high-resolution RGB-D panoramas into a Gaussian scene

$$\mathcal{G} = \{G_i\}_{i=1}^{M}, \qquad G_i = (\boldsymbol{\mu}_i, \mathbf{q}_i, \mathbf{s}_i, \alpha_i, \mathbf{c}_i), \tag{2}$$

where $\boldsymbol{\mu}_i$ is the Gaussian center, $\mathbf{q}_i$ its orientation, $\mathbf{s}_i$ its anisotropic scale, α_i its opacity, and $\mathbf{c}_i$ its color.

The pipeline follows five stages: panoramic geometry preparation (Sec. 3.2), native seed generation (Sec. 3.3), guided Gaussian initialization (Sec. 3.4), progressive ERP-aware optimization (Sec. 3.5), and export to splat-viewer formats (Sec. 3.6). A central design choice is to preserve the ERP camera model during reconstruction and optimization, rather than immediately decomposing each panorama into perspective subviews. This allows PanoGS to exploit the complete spherical observation while still producing standard Gaussian assets for interactive inspection.

3.2 Panoramic Geometry Preparation

The upstream reconstruction results are first converted into a canonical scene representation. Each view is stored as an RGB panorama, a metric depth panorama, and a camera-to-world pose $(R_k, \mathbf{t}_k)$. For each view k, a generic point $\mathbf{x}_k^w$ in camera coordinates is mapped to the scene frame in world coordinates as

$$\mathbf{x}_k^w = R_k \mathbf{x}_k^c + \mathbf{t}_k. \tag{3}$$

The scene manifest records, for each view, the RGB path, depth path, image resolution, pose, depth scale, and camera model. The default camera model is equirectangular, although the same representation can also store pinhole intrinsics for perspective-view baselines.

3.3 Native Seed Generation

For an ERP image of width W and height H, pixel coordinates (u, v) are mapped to longitude and latitude as

$$\lambda = \frac{u}{W-1} 2\pi - \pi, \qquad \phi = \frac{v}{H-1}\pi - \frac{\pi}{2}. \tag{4}$$

The corresponding unit ray is

$$\mathbf{r}(u, v) = \begin{bmatrix} \cos\phi\cos\lambda \\ \sin\phi \\ \cos\phi\sin\lambda \end{bmatrix}. \tag{5}$$

For each valid depth sample $D_k(u, v)$, PanoGS lifts the pixel into camera space and then transforms it into world space:

$$\mathbf{x}_k^w(u, v) = R_k \left(D_k(u, v)\mathbf{r}(u, v)\right) + \mathbf{t}_k. \tag{6}$$

Only finite depths inside a valid metric range are retained. Each accepted sample stores its world position, RGB color, source-view index, estimated normal, local support radius, and initial opacity. Normals are estimated from local 3D neighborhoods, while support radii are derived from nearest-neighbor distances. Since native-resolution panoramas can produce tens of millions of samples, seeds are written into compressed shards, and a seed manifest stores the global bounding box, total seed count, shard paths, and depth/normal-estimation parameters.

3.4 Gaussian Initialization Guided by a Geometric Substrate

The native RGB-D seeds are converted into optimization-ready Gaussian primitives. The center and color are initialized directly from the lifted seed:

$$\boldsymbol{\mu}_i = \mathbf{x}_i, \qquad \mathbf{c}_i = \mathbf{c}_i^{RGB}. \tag{7}$$

The estimated normal $\mathbf{n}_i$ defines the local Gaussian frame: PanoGS computes a quaternion $\mathbf{q}_i$ that rotates the canonical z-axis onto $\mathbf{n}_i$. The Gaussian is then initialized as a flattened, surface-aligned primitive rather than an isotropic blob. Given the local support radius r_i, the scale is set to

$$\mathbf{s}_i = \left[\eta_t r_i, \eta_t r_i, \eta_n r_i\right], \qquad \eta_n < \eta_t, \tag{8}$$

where η_t and η_n control tangential and normal support. In our default configuration, the normal scale is smaller than the tangential scale, encouraging primitives to behave as local surface elements. Opacity is stored both directly and in logit form to support later optimization.

PanoGS also builds a geometric substrate from the registered RGB-D panoramas using TSDF-style fusion. The substrate is not the final output representation; instead, it provides a consolidated geometric support for diagnosing and regularizing the Gaussian scene. This is important because sparse panoramic RGB-D lifting may create duplicate or inconsistent supports near depth discontinuities, small foreground objects, and weakly observed regions. In our current implementation, the Gaussian initialization preserves native per-view samples, while the TSDF substrate is used to expose misregistration, ghosting, and duplicate structures. Inspired by recent hybrid implicit-explicit approaches [Hu and Han 2026; Ye et al. 2025; Zhang et al. 2026], the geometric substrate can also be used to weight opacity or penalize off-surface primitives through a surface-consistency term:

$$\mathcal{L}_{\text{scaf}} = \lambda_\Phi \sum_i \rho\left(|\Phi(\boldsymbol{\mu}_i)|\right), \tag{9}$$

where Φ is the TSDF, ρ is a robust penalty, and λ_Φ is a tunable weight. This term encourages Gaussians to remain close to the fused surface support and is especially useful for reducing ghosting around small objects and contact regions.

3.5 Progressive Optimization

The initialized Gaussian shards are organized into a progressive workspace with three levels of detail: a preview LOD with up to one million Gaussians, a practical warm-start LOD with up to five million Gaussians, and a full-native LOD containing all generated primitives. Sampling is performed proportionally across shards so that each LOD preserves the distribution of the native seed set. This design supports both fast preview and high-resolution export, while avoiding the need to load all native Gaussians during every optimization stage.

Optimization uses an ERP-aware differentiable renderer. Given a Gaussian center in world coordinates, PanoGS first transforms it into the camera frame:

$$\mathbf{x}_i^c = R_k^T(\boldsymbol{\mu}_i - \mathbf{t}_k). \tag{10}$$

The normalized direction is then mapped to ERP coordinates through the inverse spherical projection:

$$d_i = \|\mathbf{x}_i^c\|, \tag{11}$$

$$\mathbf{r}_i = \frac{\mathbf{x}_i^c}{d_i}, \tag{12}$$

$$\phi_i = \arcsin(r_{i,y}), \tag{13}$$

$$\lambda_i = \operatorname{atan2}(r_{i,z}, r_{i,x}), \tag{14}$$

$$u_i = (\lambda_i + \pi)\frac{W-1}{2\pi}, \tag{15}$$

$$v_i = (\phi_i + \pi/2)\frac{H-1}{\pi}. \tag{16}$$

Horizontal coordinates wrap around the panorama seam, while vertical coordinates are clamped to the valid latitude range. Instead of assigning each Gaussian to a single pixel, the renderer splats it over a local soft footprint. The footprint size is proportional to tangential scale divided by depth, providing differentiable image-space support and more stable gradients than hard rasterization.

The first optimization stage is an appearance warm-start. Since the lifted RGB-D geometry is already metrically meaningful, PanoGS initially keeps geometry fixed and updates appearance-related parameters, mainly color and opacity. Optional conservative geometry refinement can then be enabled by allowing Gaussian centers to move with a reduced learning rate, delayed activation, and bounded displacement. This prevents photometric gradients from pulling primitives away from the high-resolution geometric initialization.

For multiview refinement, PanoGS restricts supervision to nearby panoramas rather than optimizing every view against every other view. This neighbor-based strategy is important for sparse wide-baseline captures, where distant panoramas may share little valid support and can introduce unstable visibility constraints. For large scenes, shardwise optimization processes subsets of the Gaussian scene independently, allowing native-resolution assets to be refined under GPU memory constraints. Optimized checkpoints are appended to the workspace as additional LODs, preserving the raw initialization, warm-start results, and refined variants for comparison and export.

3.6 Export for Web-Based Exploration

The final stage converts any selected LOD into a packed `.splat` asset. Each primitive is stored using a compact layout containing position, scale, RGBA color, and packed quaternion orientation. This separates the reconstruction backend from the viewer frontend: the same workspace can produce raw full-native splats, optimized practical splats, and additional LODs for interactive inspection. The exported assets can be loaded in standard Gaussian splat viewers or integrated into browser-based exploration tools, enabling the sparse panoramic capture to be inspected as an interactive 3D scene rather than as a point cloud, mesh, or collection of disconnected panoramas.

3.7 Web-Based Splat Viewer

For interactive inspection, exported assets can be loaded directly in lightweight WebGL-based Gaussian splat viewers, avoiding any dependency on heavier neural rendering frameworks. In our implementation, we rely on the open-source viewer Antimatter15 [Kwok 2023], which renders 3D Gaussian Splatting scenes in real time using only standard WebGL primitives, without external dependencies. Since Gaussian primitives can be treated as an extension of point-based rendering, this allows native-resolution PanoGS scenes to be displayed efficiently on commodity graphics hardware, in contrast to implicit neural representations such as NeRFs, which typically require substantially more compute for comparable interactive navigation.

Each primitive is projected to screen space in a vertex shader: the Gaussian center is transformed by the camera projection matrix, while its scale and orientation are used to derive projected eigenvectors that define a screen-aligned bounding quadrilateral. A fragment shader then evaluates the Gaussian falloff within this quadrilateral, converting distance from the splat center into an opacity contribution. Since compositing translucent primitives is order-dependent, correct blending requires depth sorting before rasterization; the viewer performs this sort asynchronously on the CPU in a web worker, which keeps the sort decoupled from the main render loop and supports progressive loading, so that scenes become visible and navigable before all primitives have finished loading. This design favors broad compatibility and dependency-free deployment over the GPU-based sorting strategies used in some reference implementations, which is well-suited to the lightweight, browser-based inspection scenario targeted by PanoGS.

4 Results

We evaluate PanoGS as a web-based Gaussian asset generation pipeline rather than as a benchmark-only novel-view synthesis model. Recent feed-forward panoramic Gaussian methods obtain very strong scores on Replica and HM3D under their own evaluation protocols [Chen et al. 2025; Ren et al. 2025; Zhang et al. 2025]; our

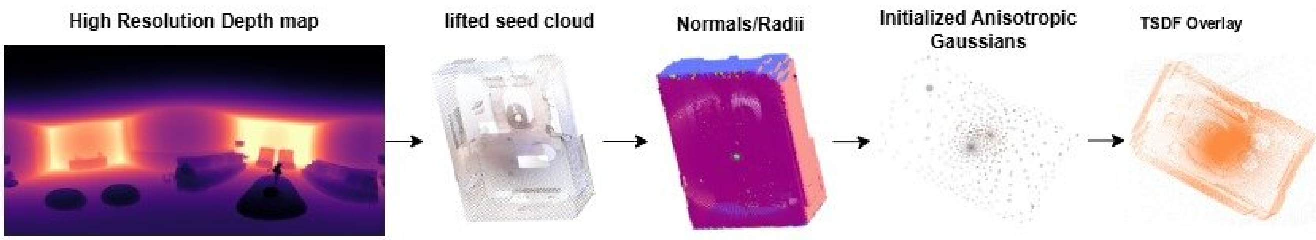

Figure 2: Geometry-to-Gaussian initialization. Starting from a native-resolution ERP depth map, PanoGS lifts valid RGB-D samples into a world-space seed cloud, estimates local normals and support radii, initializes flattened anisotropic Gaussian primitives, and uses the fused TSDF geometric substrate as a geometric support for diagnosing and regularizing the resulting scene.

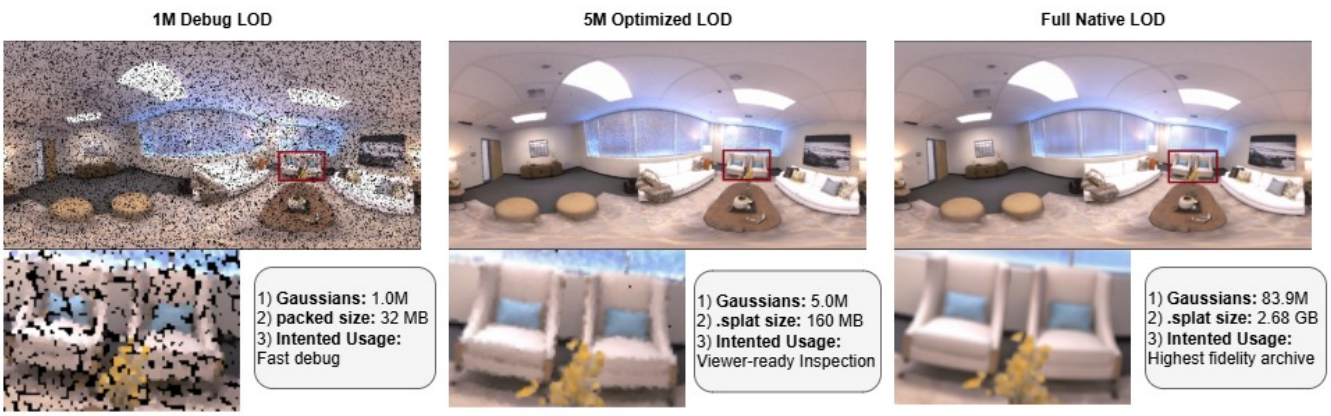

Figure 3: Progressive Gaussian asset construction. The same scene is shown at the 1M preview LOD, the 5M appearance-optimized LOD, and the full-native LOD. The lower-density levels support rapid inspection and viewer-ready export, while the full-native level preserves the highest amount of geometry produced by the native RGB-D lifting stage.

current goal is complementary: to assess whether sparse panoramic pose/depth estimates can be converted into inspectable, multi-resolution Gaussian assets that run in standard browser-based splat viewers. We therefore emphasize qualitative exploration frames, LOD behavior, export statistics, and visible artifact analysis. Quantitative image-quality comparisons with specialized panoramic NVS systems are left for a later version once the constrained refinement is fully integrated in all stages. A short video showcasing an example of interactive exploration of one reconstructed scene can be found at https://drive.google.com/file/d/1oj243hKgK-gF6oB2wRSqtByCG0BrmIkD/view.

Datasets and setup. We use two indoor panoramic datasets, Replica360 and HM3D. Replica360 provides controlled synthetic rooms with dense indoor structure and high-resolution panoramas, while HM3D contains more diverse and cluttered indoor layouts with variable overlap and scene complexity. Tab. 1 summarizes the processed data branch. For all scenes, the pipeline produces a native seed set, a 1M preview LOD, a 5M practical LOD when enough seeds are available, and a full-native packed `.splat` export. The four scenes used for the qualitative exploration figure are the same representative scenes reported in Tab. 2.

Qualitative web-based exploration. Fig. 4 shows frames extracted from browser-based exploration videos for four representative scenes: two from Replica360 and two from HM3D. The figure intentionally includes both favorable views and difficult regions. Large surfaces such as walls, ceilings, floors, sofas, beds, and tables are generally reconstructed as coherent Gaussian support, and the 5M LOD is usually sufficient for routine web-based inspection. At the same time, the frames show the main remaining limitations of the current pipeline: ghosting around small furniture and decorative objects, duplicate structures near depth discontinuities, and unstable thin geometry for plants, vases, chair legs, and cluttered table regions. These artifacts are consistent with the sparse panoramic setting: depth estimates around object boundaries are less stable, multiview support is limited, and native per-view seed lifting can preserve conflicting observations.

Asset statistics and LOD behavior. Tab. 2 and Tab. 3 report representative and average asset sizes. The results show that native panoramic lifting produces several million Gaussians per scene even for sparse inputs. The full-native exports preserve the densest representation, but they can be heavy for frequent web-based inspection. The 5M LOD provides a more practical compromise: it is substantially easier to view and share while preserving most

Table 1: Dataset and configuration summary used in the PanoVGGT-to-Gaussian processing branch. Replica360 is reported for the processed subset currently available in the workspace; HM3D is reported for the full test directory inventory.

Dataset	Scenes	Input panoramas	Resolution	Output LODs	Notes
Replica360	8	10 per scene	4096×2048	1M, 5M, full native	PanoVGGT depth/pose priors; indoor panoramas with dense coverage
HM3D test	423	2–10 per scene	1024×512	1M, 5M, full native	variable overlap; 333 scene instances contain 10 panoramas

Figure 4: Frames extracted from browser-based exploration of Replica360 example scene. The frames are selected to show both successful viewpoints and remaining artifacts, including local ghosting, duplicated small objects, and unstable thin structures around clutter, plants, chairs, and table regions.

dominant surfaces and furniture-scale structures. The 1M LOD is mainly useful for previewing geometry, checking scene orientation, and rapidly identifying severe pose or depth failures.

Observed artifacts and diagnostic variants. Tab. 4 summarizes the behavior of the main asset variants. Raw native seeding preserves detail but also preserves conflicting per-view observations, so ghosting remains high. Appearance optimization improves visual coherence and viewer usability, especially at 5M Gaussians, but it does not fully solve duplicate structures around small objects. Full-native exports improve completeness but are less practical for routine browser-based use because of asset size and sorting cost. Local overlap-first subscenes are currently the most stable configuration for difficult objects because they restrict the reconstruction to strongly supported regions, at the cost of less global coverage. These observations motivate the constrained pruning and opacity weighting discussed in Sec. 3.4.

5 Conclusions

We presented PanoGS, a pipeline guided by an inferred geometric substrate for converting sparse panoramic indoor captures into high-resolution Gaussian Splatting assets for web-based 3D exploration. The method addresses challenges caused by limited viewpoint overlap, ERP distortions, and depth-scale uncertainty by combining feed-forward panoramic pose/depth initialization, native-resolution depth refinement, TSDF-based geometric constraints, surface-aware anisotropic Gaussian initialization, progressive LOD construction, and ERP-aware optimization. The resulting assets can be exported to standard browser-compatible `.splat` viewers, supporting practical inspection in a lightweight web-based runtime.

Our current results on Replica360 and HM3D show that PanoGS can produce usable browser-viewable Gaussian scenes from sparse panoramic inputs, especially for dominant indoor structures and furniture-scale geometry. The qualitative evaluation also exposes remaining limitations: ghosting and duplicate support persist around small objects, thin structures, and depth discontinuities. These artifacts motivate the next development step, namely stronger geometrically-constrained opacity weighting, pruning, and geometry regularization. Overall, PanoGS provides a modular bridge between sparse panoramic reconstruction and deployable Web3D Gaussian assets, with explicit control over resolution, LODs, export size, and viewer behavior.

Acknowledgments

Authors from Hamad Bin Khalifa University (US, MT, JS, WE, and MA) acknowledge the support of the Office of Vice Presidence of Research of HBKU under the grant HBKU-OVPR-TG-03-8 *3G Gaze, Guide, and Grasp: LLM-Augmented Immersive Experiences for Qatar's Heritage Sites*. Authors from CRS4 (GP and EG) acknowledge the contribution of Sardinian Regional Authorities under grants XDATA and HUBIAS. AJ acknowledges the contribution of KAUST under the TRG2025 grant 3D-TURATH.

References

Hao Ai, Zidong Cao, Yan-Pei Cao, Ying Shan, and Lin Wang. 2023. HRDFuse: Monocular 360° Depth Estimation by Collaboratively Learning Holistic-With-Regional Depth Distributions. In *Proc. CVPR*. 13273–13282. doi:10.1109/CVPR52729.2023.01275

Hao Ai, Zidong Cao, and Lin Wang. 2025. A Survey of Representation Learning, Optimization Strategies, and Applications for Omnidirectional Vision. *Int. J. Comput. Vision* 133, 8 (2025), 4973–5012. doi:10.1007/s11263-025-02391-w

Benjamin Attal, Selena Ling, Aaron Gokaslan, Christian Richardt, and James Tompkin. 2020. MatryODShka: Real-time 6DoF video view synthesis using multi-sphere images. In *Proc. ECCV*. 441–459. doi:10.1007/978-3-030-58452-8_26

Jiayang Bai, Letian Huang, Jie Guo, Wen Gong, Yuanqi Li, and Yanwen Guo. 2025. 360-GS: Layout-Guided Panoramic Gaussian Splatting for Indoor Roaming. In *2025 International Conference on 3D Vision (3DV)*. 1042–1053. doi:10.1109/3DV66043.2025.00101

Table 2: Measured asset statistics for representative Replica360 and HM3D scenes. These are the four scenes used in Fig. 4. The `.splat` size and export time columns correspond to the final full-native packed export produced by the PanoGS export stage.

Scene	Seeds	1M LOD	5M LOD	Full native	.splat size	Export time
Replica apartment_0_rand	5,343,342	1,000,000	5,000,000	5,343,342	170.99 MB	1.048 s
Replica apartment_1_rand	5,366,446	1,000,000	5,000,000	5,366,446	171.73 MB	0.974 s
HM3D 00800-TEEsavR23oF_0000	2,647,055	1,000,000	2,647,055	2,647,055	84.71 MB	0.562 s
HM3D 00801-HaxA7YrQdEC_0000	5,351,088	1,000,000	5,000,000	5,351,088	171.23 MB	1.016 s

Table 3: Average asset statistics over the completed processed scenes in each dataset branch. Replica360 averages are computed over 8 completed scenes; HM3D averages are computed over 8 completed scene instances currently processed in the workspace. Export time is the measured average runtime of the full-native `.splat` export.

Dataset	Scenes	Avg. seeds	Avg. 1M LOD	Avg. 5M LOD	Avg. full native	Avg. .splat size	Avg. export time
Replica360	8	5,363,195	1,000,000	5,000,000	5,363,195	171.62 MB	0.979 s
HM3D test	8	4,481,177	1,000,000	4,305,084	4,481,177	143.40 MB	0.859 s

Table 4: Qualitative system behavior across the main Gaussian asset variants. Ratings use a fixed rubric: visual completeness measures missing regions and holes; ghosting measures duplicated or shifted structures; small-object stability measures consistency of thin objects and clutter; viewer usability measures practicality for routine browser-based inspection.

Variant	Visual completeness	Ghosting	Small-object stability	Viewer usability
Raw native seeding	Medium	High	Low	Medium
5M appearance-optimized LOD	Medium–High	Medium	Medium	High
Full-native LOD export	High	Medium	Medium	Low
Overlap-first local subscenes	High	Low	High	Medium

Rubric. Visual completeness: Low = large missing regions, Medium = most major surfaces present with visible gaps, High = scene largely complete. Ghosting: Low = little duplicated structure, Medium = local duplicates around hard objects, High = repeated walls or furniture. Small-object stability: Low = thin or cluttered objects frequently fragment, Medium = large furniture stable but fine structure inconsistent, High = thin and cluttered objects remain coherent. Viewer usability: Low = impractical for routine browser-based viewing, Medium = usable with selective inspection, High = practical for frequent inspection and sharing.

Chong Bao, Xiyu Zhang, Zehao Yu, Jiale Shi, Guofeng Zhang, Songyou Peng, and Zhaopeng Cui. 2025b. Free360: Layered Gaussian Splatting for Unbounded 360-Degree View Synthesis from Extremely Sparse and Unposed Views. In *2025 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. 16377–16387. doi:10.1109/CVPR52734.2025.01527

Yanqi Bao, Tianyu Ding, Jing Huo, Yaoli Liu, Yuxin Li, Wenbin Li, Yang Gao, and Jiebo Luo. 2025a. 3D Gaussian Splatting: Survey, Technologies, Challenges, and Opportunities. *IEEE Transactions on Circuits and Systems for Video Technology* 35, 7 (2025), 6832–6852. doi:10.1109/TCSVT.2025.3538684

Zidong Cao, Jinjing Zhu, Hao Ai, Lutao Jiang, Yuanhuiyi Lyu, and Hui Xiong. 2025. ST^{2}360D: Spatial-to-Temporal Consistency for Training-free 360 Monocular Depth Estimation. In *Proc. NeurIPS*. https://neurips.cc/virtual/2025/loc/san-diego/poster/117039

Guikun Chen and Wenguan Wang. 2026. A Survey on 3D Gaussian Splatting. *ACM Comput. Surv.* 58, 12, Article 313 (May 2026), 39 pages. doi:10.1145/3807511

Zheng Chen, Yan-Pei Cao, Yuan-Chen Guo, Chen Wang, Ying Shan, and Song-Hai Zhang. 2023. PanoGRF: generalizable spherical radiance fields for wide-baseline panoramas. In *Proc. NeurIPS*. doi:10.48550/arXiv.2306.01531

Zheng Chen, Chenming Wu, Zhelun Shen, Chen Zhao, Weicai Ye, Haocheng Feng, Errui Ding, and Song-Hai Zhang. 2025. Splatter-360: Generalizable 360° Gaussian Splatting for Wide-baseline Panoramic Images. In *2025 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. 21590–21599. doi:10.1109/CVPR52734.2025.02011

Thiago LT da Silveira, Paulo GL Pinto, Jeffri Murrugarra-Llerena, and Cláudio R Jung. 2022. 3D scene geometry estimation from 360° imagery: A survey. *Comput. Surveys* 55, 4 (2022), 1–39.

Ben Fei, Jingyi Xu, Rui Zhang, Qingyuan Zhou, Weidong Yang, and Ying He. 2025. 3D Gaussian Splatting as a New Era: A Survey. *IEEE Transactions on Visualization and Computer Graphics* 31, 8 (2025), 4429–4449. doi:10.1109/TVCG.2024.3397828

Krzysztof M Gorski, Eric Hivon, Anthony J Banday, Benjamin D Wandelt, Frode K Hansen, Mstvos Reinecke, and Matthia Bartelmann. 2005. HEALPix: A framework for high-resolution discretization and fast analysis of data distributed on the sphere. *The Astrophysical Journal* 622, 2 (2005), 759–771.

Meiying Gu, Jiawei Zhang, Jiahe Li, Xiaohan Yu, Haonan Luo, Jin Zheng, and Xiao Bai. 2026. SparseSurf: Sparse-view 3D gaussian splatting for surface reconstruction. In *Proc. AAAI*, Vol. 40. 4311–4319.

Yijing Guo, Mengjun Chao, Luo Wang, Tianyang Zhao, Haizhao Dai, Yingliang Zhang, Jingyi Yu, and Yujiao Shi. 2026. PanoVGGT: Feed-Forward 3D Reconstruction from Panoramic Imagery. In *Proc. CVPR to appear.*

Pengchong Hu and Zhizhong Han. 2026. SGAD-SLAM: Splatting Gaussians at Adjusted Depth for Better Radiance Fields in RGBD SLAM. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 18934–18945.

Jingwei Huang, Zhili Chen, Duygu Ceylan, and Hailin Jin. 2017. 6-DOF VR videos with a single 360-camera. In *Proc. IEEE VR*. 37–44. doi:10.1109/VR.2017.7892229

Lihan Jiang, Yucheng Mao, Linning Xu, Tao Lu, Kerui Ren, Yichen Jin, Xudong Xu, Mulin Yu, Jiangmiao Pang, Feng Zhao, Dahua Lin, and Bo Dai. 2025. AnySplat: Feed-forward 3D Gaussian Splatting from Unconstrained Views. *ACM Trans. Graph.* 44, 6, Article 257 (Dec. 2025), 16 pages. doi:10.1145/3763326

Dongki Jung, Jaehoon Choi, Yonghan Lee, Somi Jeong, Taejae Lee, Dinesh Manocha, and Suyong Yeon. 2025b. EDM: Equirectangular Projection-Oriented Dense Kernelized Feature Matching. In *Proceedings of the Computer Vision and Pattern Recognition Conference*. 6337–6347.

Dongki Jung, Jaehoon Choi, Yonghan Lee, and Dinesh Manocha. 2025a. IM360: Large-scale Indoor Mapping with 360 Cameras. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*. 29040–29050.

Dongki Jung, Jaehoon Choi, Adil Qureshi, Somi Jeong, Dinesh Manocha, and Suyong Yeon. 2026. Wid3R: Wide Field-of-View 3D Reconstruction via Camera Model Conditioning. *arXiv:2602.05321* (2026).

Nikhil Keetha, Norman Müller, Johannes Schönberger, Lorenzo Porzi, Yuchen Zhang, Tobias Fischer, Arno Knapitsch, Duncan Zauss, Ethan Weber, Nelson Antunes, Jonathon Luiten, Manuel Lopez-Antequera, Samuel Rota Bulò, Christian Richardt, Deva Ramanan, Sebastian Scherer, and Peter Kontschieder. 2026. MapAnything: Universal Feed-Forward Metric 3D Reconstruction. In *International Conference on 3D Vision (3DV)*. IEEE.

Kevin Kwok. 2023. splat: WebGL 3D Gaussian Splat Viewer. https://github.com/antimatter15/splat. Accessed 2026.

David Li, Yinda Zhang, Christian Häne, Danhang Tang, Amitabh Varshney, and Ruofei Du. 2022b. OmniSyn: Synthesizing 360 videos with wide-baseline panoramas. In *Proc. VRW*. 670–671. doi:10.48550/arXiv.2202.08752

Junbo Li, Florian Hahlbohm, Timon Scholz, Martin Eisemann, Jan-Philipp Tauscher, and Marcus Magnor. 2025. SPaGS: Fast and Accurate 3D Gaussian Splatting for Spherical Panoramas. *Computer Graphics Forum* 44, 4 (Jun 2025). doi:10.1111/cgf.70171

Yuyan Li, Yuliang Guo, Zhixin Yan, Xinyu Huang, Ye Duan, and Liu Ren. 2022a. Omnifusion: 360 monocular depth estimation via geometry-aware fusion. In *Proc. CVPR*. 2801–2810. doi:10.1109/CVPR52688.2022.00282

Haotong Lin, Sida Peng, Jingxiao Chen, Songyou Peng, Jiaming Sun, Minghuan Liu, Hujun Bao, Jiashi Feng, Xiaowei Zhou, and Bingyi Kang. 2025. Prompting Depth Anything for 4K Resolution Accurate Metric Depth Estimation. In *Proc. CVPR*. 17070–17080. doi:10.1109/CVPR52734.2025.01591

Xiaoyang Lyu, Yang-Tian Sun, Yi-Hua Huang, Xiuzhe Wu, Ziyi Yang, Yilun Chen, Jiangmiao Pang, and Xiaojuan Qi. 2024. 3DGSR: Implicit surface reconstruction with 3d Gaussian splatting. *ACM TOG* 43, 6 (2024), 1–12.

Takeshi Noda, Yu-Shen Liu, and Zhizhong Han. 2026. 3D Gaussian Splatting with Self-Constrained Priors for High Fidelity Surface Reconstruction. In *Proc. CVPR*.

Chi-Han Peng and Jiayao Zhang. 2023. High-Resolution Depth Estimation for 360° Panoramas through Perspective and Panoramic Depth Images Registration. In *Proc. WACV*. 3115–3124. doi:10.1109/WACV56688.2023.00313

Giovanni Pintore, Marco Agus, Jens Schneider, and Enrico Gobbetti. 2026. State-of-the-art in deep learning approaches for automatic single-panorama indoor modeling and exploration. *Computer Graphics Forum* 45, 2 (2026). doi:10.1111/cgf.70396

Giovanni Pintore, Fabio Bettio, Marco Agus, and Enrico Gobbetti. 2023. Deep scene synthesis of Atlanta-world interiors from a single omnidirectional image. *IEEE TVCG* 29, 11 (2023), 4708–4718. doi:10.1109/TVCG.2023.3320219

Giovanni Pintore, Sara Jashari, Marco Agus, and Enrico Gobbetti. 2025. PanoFloor: Reconstruction and Immersive Exploration of Large Multi-Room Scenes from a Minimal Set of Registered Panoramic Images Using Denoised Density Maps. In *2025 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. 414–424. doi:10.1109/ISMAR67309.2025.00052

Giovanni Pintore, Alberto Jaspe-Villanueva, Markus Hadwiger, Jens Schneider, Marco Agus, Fabio Marton, Fabio Bettio, and Enrico Gobbetti. 2024. Deep synthesis and exploration of omnidirectional stereoscopic environments from a single surround-view panoramic image. *Computers & Graphics* 119 (2024), 103907. doi:10.1016/j.cag.2024.103907

Jiahui Ren, Mochu Xiang, Jiajun Zhu, and Yuchao Dai. 2025. PanoSplatt3R: Leveraging Perspective Pretraining for Generalized Unposed Wide-Baseline Panorama Reconstruction. In *2025 IEEE/CVF International Conference on Computer Vision (ICCV)*. 28959–28969. doi:10.1109/ICCV51701.2025.02689

Manuel Rey-Area, Mingze Yuan, and Christian Richardt. 2022. 360MonoDepth: High-Resolution 360° Monocular Depth Estimation. In *Proc. CVPR*. 3762–3772. doi:10.1109/CVPR52688.2022.00374

Uzair Shah, Giovanni Pintore, Muhammad Tukur, Zahoor Ahmad, Jens Schneider, Matteo Sgrenzaroli, Giorgio Vassena, Pere-Pau Vázquez, Enrico Gobbetti, and Marco Agus. 2026. FRED: Full-Resolution Equirectangular Depth Estimation. In *Proceedings of the Spanish Computer Graphics Conference (CEIG)*.

Changha Shin, Woong Oh Cho, and Seon Joo Kim. 2025. Seam360GS: Seamless 360deg Gaussian Splatting from Real-World Omnidirectional Images. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*. 28970–28979.

Kai Song and Lei Zhang. 2025. Novel view synthesis with wide-baseline stereo pairs based on local–global information. *Computers & Graphics* 126 (2025), 104139. doi:10.1016/j.cag.2024.104139

Muhammad Tukur, Sara Jashari, Mahmood Alzubaidi, Babatunde Abiodun Salami, Yehia Boraey, Sindy Yong, Dina Saleh, Giovanni Pintore, Enrico Gobbetti, Jens Schneider, Noora Fetais, and Marco Agus. 2025. Panoramic Imaging in Immersive Extended Reality: A Scoping Review of Technologies, Applications, Perceptual Studies, and User Experience Challenges. *Frontiers in Virtual Reality* 6 (2025). doi:10.3389/frvir.2025.1622605

Muhammad Tukur, Giovanni Pintore, Enrico Gobbetti, Jens Schneider, and Marco Agus. 2023. SPIDER: A framework for processing, editing and presenting immersive high-resolution spherical indoor scenes. *Graphical Models* 128 (2023), 101182:1–101182:11. doi:10.1016/j.gmod.2023.101182

John Waidhofer, Richa Gadgil, Anthony Dickson, Stefanie Zollmann, and Jonathan Ventura. 2022. PanoSynthVR: Toward Light-weight 360-Degree View Synthesis from a Single Panoramic Input. In *Proc. ISMAR*. 584–592. doi:10.1109/ISMAR55827.2022.00075

Jianyuan Wang, Minghao Chen, Nikita Karaev, Andrea Vedaldi, Christian Rupprecht, and David Novotny. 2025. VGGT: Visual Geometry Grounded Transformer. In *Proc. CVPR*.

Qiwei Wang, Xianghui Ze, Jingyi Yu, and Yujiao Shi. 2026. CylinderSplat: 3D Gaussian Splatting with Cylindrical Triplanes for Panoramic Novel View Synthesis. In *Proc. ICLR*.

Shuzhe Wang, Vincent Leroy, Yohann Cabon, Boris Chidlovskii, and Jerome Revaud. 2024. DUSt3R: Geometric 3D Vision Made Easy. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. 20697–20709.

Zijun Wei, Jianming Zhang, Xiaohui Shen, Zhe Lin, Radomir Mech, Minh Hoai, and Dimitris Samaras. 2018. Good view hunting: Learning photo composition from dense view pairs. In *Proc. CVPR*. 5437–5446. doi:10.1109/CVPR.2018.00570

Diana Werner, Ayoub Al-Hamadi, and Philipp Werner. 2014. Truncated Signed Distance Function: Experiments on Voxel Size. In *Image Analysis and Recognition*, Aurélio Campilho and Mohamed Kamel (Eds.). Springer International Publishing, Cham, 357–364.

Haofei Xu, Songyou Peng, Fangjinhua Wang, Hermann Blum, Daniel Barath, Andreas Geiger, and Marc Pollefeys. 2025. DepthSplat: Connecting Gaussian Splatting and Depth. In *2025 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. 16453–16463. doi:10.1109/CVPR52734.2025.01534

Jiale Xu, Jia Zheng, Yanyu Xu, Rui Tang, and Shenghua Gao. 2021. Layout-guided novel view synthesis from a single indoor panorama. In *Proc. CVPR*. 16438–16447. doi:10.1109/CVPR46437.2021.01617

Keyang Ye, Tianjia Shao, and Kun Zhou. 2025. When Gaussian Meets Surfel: Ultra-fast High-fidelity Radiance Field Rendering. *ACM Trans. Graph.* 44, 4, Article 113 (July 2025), 15 pages. doi:10.1145/3730925

Cheng Zhang, Haofei Xu, Qianyi Wu, Camilo Cruz Gambardella, Dinh Phung, and Jianfei Cai. 2025. PanSplat: 4K Panorama Synthesis with Feed-Forward Gaussian Splatting. In *2025 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. 11437–11447. doi:10.1109/CVPR52734.2025.01068

Xiaoxue Zhang, Xiaoxu Zheng, Yixuan Yin, Tiao Zhao, Kaihua Tang, Michael Bi Mi, Zhan Xu, and Dave Zhenyu Chen. 2026. AnchorSplat: Feed-Forward 3D Gaussian Splatting with 3D Geometric Priors. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 18924–18933.

Qiang Zhao, Liang Wan, Wei Feng, Jiawan Zhang, and Tien-Tsin Wong. 2013. Cube2Video: Navigate between cubic panoramas in real-time. *IEEE Trans. Multimedia* 15, 8 (2013), 1745–1754. doi:10.1109/TMM.2013.2280249

Yipeng Zhu, Huajian Huang, Tristan Braud, and Sai-Kit Yeung. 2026. UniTriSplat: A Unified 3D Gaussian Splatting Framework with Uniform Spherical Rasterization for Universal Cameras. In *Proc. ECCV*.

Chuanqing Zhuang, Xin Lu, Zehui Deng, Zhengda Lu, Yiqun Wang, Junqi Diao, and Jun Xiao. 2026. Pose-Free Omnidirectional Gaussian Splatting for 360-Degree Videos with Consistent Depth Priors. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. 4881–4890.