

Reconstructing 3D Scenes in Native High Dynamic Range

Kaixuan Zhang¹ Minxian Li^{1*} Mingwu Ren¹
 Jiankang Deng² Xiatian Zhu³

¹Nanjing University of Science and Technology ²Imperial College London
³University of Surrey

Abstract

High Dynamic Range (HDR) imaging is essential for professional digital media creation, e.g., filmmaking, virtual production, and photorealistic rendering. However, 3D scene reconstruction has primarily focused on Low Dynamic Range (LDR) data, limiting its applicability to professional workflows. Existing approaches that reconstruct HDR scenes from LDR observations rely on multi-exposure fusion or inverse tone-mapping, which increase capture complexity and depend on synthetic supervision. With the recent emergence of cameras that directly capture native HDR data in a single exposure, we present the first method for 3D scene reconstruction that directly models native HDR observations. We propose **Native High dynamic range 3D Gaussian Splatting (NH-3DGS)**, which preserves the full dynamic range throughout the reconstruction pipeline. Our key technical contribution is a novel luminance-chromaticity decomposition of the color representation that enables direct optimization from native HDR camera data. We demonstrate on both synthetic and real multi-view HDR datasets that NH-3DGS significantly outperforms existing methods in reconstruction quality and dynamic range preservation, enabling professional-grade 3D reconstruction directly from native HDR captures. Code and datasets will be made available.

1. Introduction

Novel View Synthesis (NVS) has achieved remarkable progress in reconstructing 3D scenes from multi-view images [3, 26], enabling photorealistic rendering for gaming [25], AR/VR [28], and autonomous driving [19]. However, existing methods [7, 23] remain confined to Low Dynamic Range (LDR) imagery, whose limited radiometric precision fails to capture the full spectrum of scene radiance. Under

*Minxian Li (minxianli@njust.edu.cn) is the corresponding author with School of Computer Science and Engineering, Nanjing University of Science and Technology.

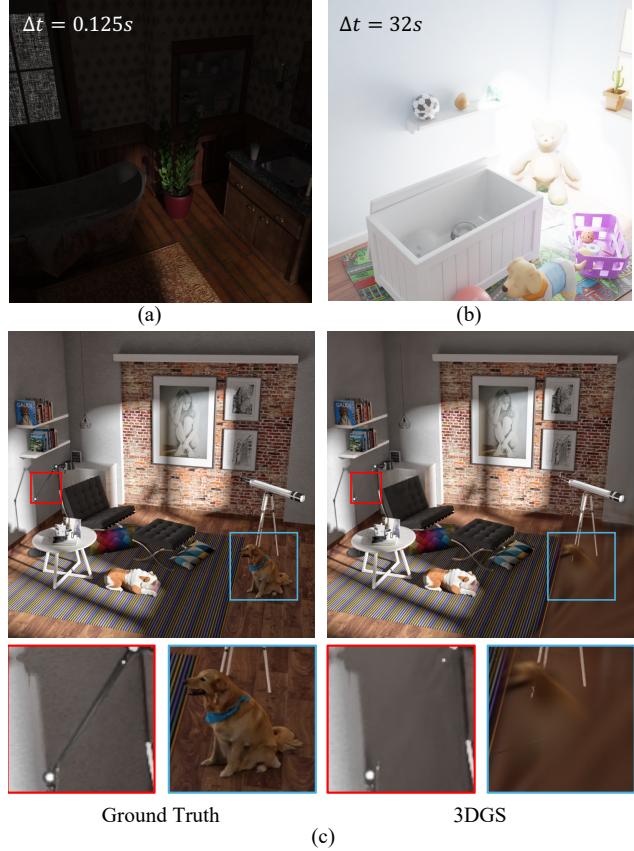


Figure 1. Examples of (a) underexposure and (b) overexposure. (c) 3DGS [15] exhibits blurring artifacts in dark regions when applying HDR supervision directly. Δt : Exposure time.

high-contrast illumination (e.g., direct sunlight or nighttime scenes), these models suffer from severe overexposure and underexposure artifacts (see Fig. 1).

High Dynamic Range (HDR) imaging captures substantially broader luminance ranges, preserving details across highlights and shadows. Yet current HDR reconstruction methods rely on multi-exposure LDR fusion [4, 13, 18], requiring carefully captured exposure brackets that introduce capture complexity and storage overhead. Other approaches

leverage RAW sensor data [14, 17, 21, 27] but target only low-light scenarios, overlooking RAW’s potential for high-dynamic-range reconstruction.

The emergence of cameras that directly capture native HDR data in a single exposure opens new possibilities [29]. Unlike multi-exposure methods, these sensors provide genuine HDR observations without fusion artifacts. This motivates our central question: *Can we reconstruct 3D radiance fields directly from native HDR data (including RAW sensor data) without LDR supervision or tone-mapping?*

Our analysis reveals that applying 3D Gaussian Splatting (3DGS) [15], a state-of-the-art framework for NVS, to HDR inputs leads to catastrophic failure, *yielding radiance fields heavily biased toward bright regions and failing to capture details in darker areas* (Fig. 1 and Fig. 1 in the Supplementary Material). Standard 3DGS pipelines are designed for LDR rendering, with each Gaussian’s appearance being typically modeled using low-order spherical harmonics (SHs) (*e.g.*, degree 3, yielding 16 coefficients). This efficiently encodes *modest view-dependent color variations* under the assumption of a bounded and approximately linear mapping between radiance and RGB intensity. Crucially, we note this formulation *entangles luminance and chromaticity* within the SH coefficients, which works well for LDR data. But for HDR observations, radiance can *vary dramatically across viewing directions*, spanning orders of magnitude even within the same surface point. Low-order SHs lack the representational capacity to disentangle such extreme view-dependent luminance changes from true chromatic variations, causing optimization biased towards high-radiance views and poor reconstruction fidelity in darker/low-exposure regions. Increasing SH order is computationally prohibitive and prone to overfitting (Fig. 2 and Tab. 1). NeRF [20] presents similar limitation due to entanglement of geometric density and radiance within a single neural representation (see Tab. 1).

To overcome these fundamental limitations, we propose **NH-3DGS** (Native High dynamic range 3D Gaussian Splatting), a novel variant that fundamentally rethinks the color representation for HDR scenes. At its core, NH-3DGS introduces a physically and perceptually grounded *luminance–chromaticity decomposition* that disentangles intensity from color information. Specifically, we replace the conventional entangled SH representation with two complementary components: (1) an explicit, per-Gaussian luminance co-efficiency to model the radiance magnitude conditioned on the viewpoint, and (2) view-dependent chromatic coefficients represented by low-order SHs that exclusively encode color ratios without intensity scaling. This design eases reconstruction of drastic HDR observations, from dimly lit interiors to direct sunlight exposures, while maintaining the real-time rendering efficiency.

Our **contributions** are as follows: **(I)** We present the

first systematic study of 3D scene reconstruction from native HDR imagery. This work embraces the emergence of high-end imaging devices and bridges the gap with the requirements of professional digital content creation. **(II)** We provide a comprehensive analysis of why existing NVS methods fail on HDR data, revealing fundamental limitations under extreme luminance variations. **(III)** We propose a novel 3DGS variant for modeling native HDR data via luminance-chromaticity decomposition. **(IV)** Experiments show that NH-3DGS outperforms all alternatives on both synthetic and real-world benchmarks.

2. Related Work

Novel view synthesis (NVS) is a fundamental task in 3D vision, with applications in AR/VR [28], gaming [25], and autonomous driving [19]. Classical methods like Structure-from-Motion (SfM) [24] and Multi-View Stereo (MVS) [12] rely on geometric cues from multi-view images but struggle with occlusions, textureless regions, and high computational costs [8]. Recent learning-based approaches model scenes as continuous differentiable representations. NeRF [20] pioneered this paradigm by using neural networks to map 3D coordinates and viewing directions to volume density and color via differentiable volume rendering. Subsequent works [9, 10, 22, 32, 34, 37] aim to improve efficiency and quality. Alternatively, 3DGS [15] and its variations [6, 16, 33, 35] represent scenes as learnable 3D Gaussians, bypassing volumetric rendering and enabling real-time performance. Despite progress, such existing state-of-the-art NVS methods are predominantly focused on LDR sRGB inputs, limiting their ability to reconstruct or render true HDR radiance from real-world scenes.

HDR novel view synthesis (HDR NVS) aims to reconstruct scenes with high dynamic range from multi-view observations. Huang *et al.* [13] introduced HDR-NeRF, the first HDR-NVS framework, which extends standard NeRF [20] to learn mappings from physical radiance to HDR color using multi-exposure LDR inputs. However, its reliance on the NeRF architecture results in prohibitively slow inference. HDR-GS [4] addressed this limitation using 3DGS [4], achieving significantly faster rendering and improved visual quality. GaussHDR [18] further introduced a unified tone-mapping strategy that enables HDR novel view synthesis without requiring HDR supervision. Despite these advances, all aforementioned methods fundamentally rely on multi-exposure LDR image stacks, which entail complex capture protocols and substantial storage overhead. Moreover, several studies [14, 17, 21, 27] have explored leveraging RAW sensor data for HDR reconstruction. Owing to their higher bit depth and linear radiometric response, RAW images can capture scene radiance more faithfully than conventional LDR formats. However, existing approaches predominantly target low-light denoising and refocusing sce-

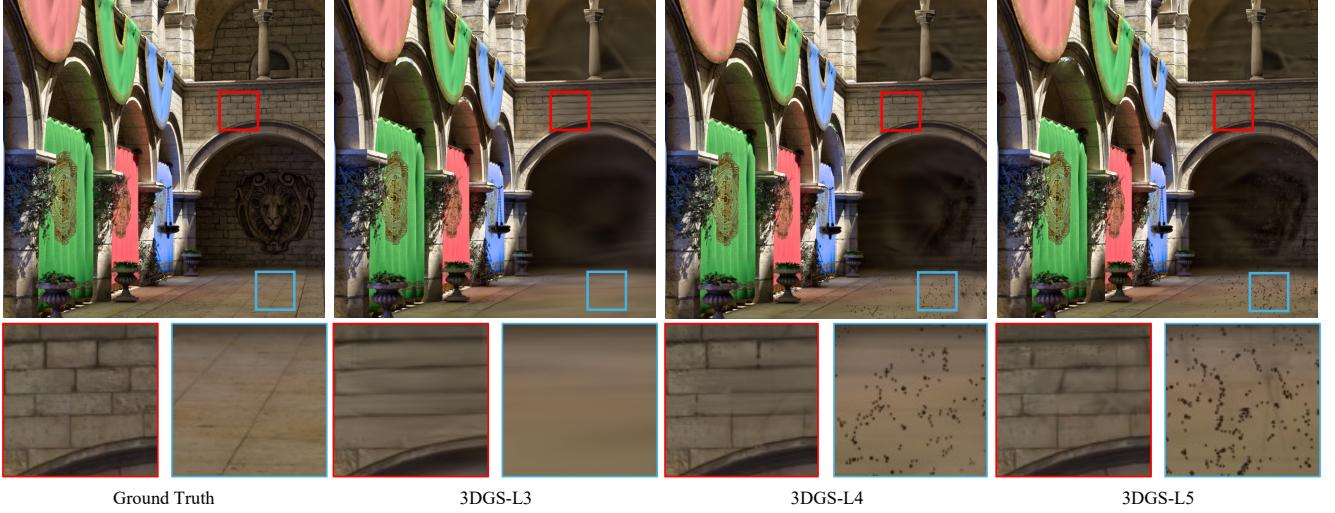


Figure 2. Elevating the SH order L (3 by default) mitigates blurring artifact while simultaneously inducing additional artifacts.

narios, overlooking the potential of RAW data for general high-dynamic-range scene reconstruction.

The recent availability of native HDR cameras that capture high dynamic range in a single exposure [29] motivates our approach. We present the first study that directly operates on native HDR imagery (including RAW sensor data) without multi-exposure fusion or LDR-based supervision. Our NH-3DGS employs luminance-chromaticity decomposition to enable superior HDR-grade scene reconstruction and high-fidelity novel-view synthesis.

3. Method

Problem. We aim to learn an HDR 3D model \mathcal{F} for a target scene from posed HDR images, $\mathcal{F} : (V) \rightarrow \mathbf{I}_V$, that can render an HDR image \mathbf{I}_V for any viewpoint V . To that end, we capture a set of HDR training images $\mathbf{I} = \{\mathbf{I}_1, \dots, \mathbf{I}_n, \dots, \mathbf{I}_N\}$, with \mathbf{I}_n the n -th view V_n .

3.1. Preliminary

In 3DGS [15], each Gaussian point stores color parameters represented by a set of low-order SHs with a set of coefficients $\mathbf{k} = \{k_l^m | 0 \leq l \leq L, -l \leq m \leq l\} \in \mathbb{R}^{(L+1)^2 \times 3}$ that model view-dependent appearance, and each $\mathbf{k}_l^m \in \mathbb{R}^3$ is a set of three coefficients corresponding to the RGB components where the index m denotes the azimuthal order of the spherical harmonic basis function Y_l^m , controlling its angular frequency and phase around the polar axis for a given degree l . L is the degree of SH. Formally, the color of the i -th Gaussian under viewing direction $\mathbf{d} = (\theta, \phi)$ is expressed as

$$\mathbf{c}_i(\mathbf{d}, \mathbf{k}) = \sum_{l=0}^L \sum_{m=-l}^l \mathbf{k}_l^m Y_l^m(\theta, \phi), \quad (1)$$

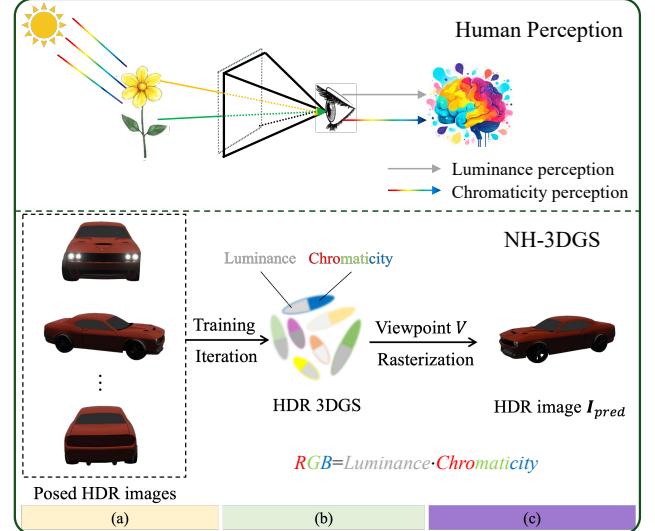


Figure 3. The pipeline of NH-3DGS draws inspiration from human visual perception. (a) With a set of HDR training images with corresponding camera poses, NH-3DGS learns a native HDR 3DGS representation. To that end, (b) we reformulate the conventional SH color representation through *luminance-chromaticity decomposition*. (c) The final HDR color is reconstructed through multiplicative composition of luminance and chromaticity, enabling physically consistent novel view synthesis across the full dynamic range.

where $Y_l^m : \mathbb{S}^2 \rightarrow \mathbb{R}$ is the SH function that maps 3D points on the sphere to real numbers [4, 15].

Remarks. The formulation of 3DGS above implicitly assumes that input images are normalized LDR RGB values within $[0, 1]$, where *radiance variations across views are moderate* [30]. Under such cases, the SH expansion provides a compact low-frequency approximation of the bidirectional color variation. However, when applied to HDR

imagery, where radiance can vary by *orders of magnitude* across different viewing directions, the SH representation becomes limited and less accurate, even after normalization, because the coefficients \mathbf{k}_l^m must jointly encode both overall brightness (luminance) and color (chromaticity) simultaneously.

More specifically, this entanglement leads to two critical issues. *First*, during optimization, gradient updates $\nabla \mathbf{k}_l^m$ tend to be dominated by high-radiance observations, causing dark or low-exposure regions to receive negligible learning signal and resulting in inferior reconstruction. *Second*, as \mathbf{k}_l^m conflates luminance and chromaticity, the relative scaling among RGB channels becomes inconsistent under extreme view-dependent radiance variations, manifesting as perceptible hue shifts (*e.g.*, neutral highlights rendered with greenish or magenta tints), as shown in Fig. 4.

3.2. Elevating SH Order in 3DGS

As discussed before, vanilla 3DGS pipelines are designed for LDR rendering, modeling each Gaussian’s appearance with low-order SHs (*e.g.*, degree 3, yielding 16 coefficients). This efficiently encodes modest view-dependent color variations under the assumption of a bounded, approximately linear radiance-to-RGB mapping. However, HDR observations exhibit extreme radiance variations, spanning orders of magnitude across viewing directions at identical surface points. Low-order SHs lack the capacity to disentangle such intense luminance shifts from true chromatic variations, biasing optimization toward high-radiance views and degrading reconstruction fidelity in darker regions. This raises a natural question: *Can increasing SH order alleviate these limitations?*

Under this consideration, an intuitive approach is to elevate the SH degree in 3DGS [15], *e.g.*, increasing from the default degree of 3 (16 coefficients) to 5 (36 coefficients). We find while higher-order SHs marginally improve metrics (see rows 2-4 of Tab. 1), they fail to enhance perceptual rendering quality (Fig. 2), while slowing down. This indicates fundamental limitations of 3DGS’s color design under HDR optimization, going beyond SH order. The inherent formulation of SHs, rather than the degree, ultimately impedes the performance under HDR supervision.

3.3. NH-3DGS

To address the above limitation, inspired by the observations as discussed earlier, we propose NH-3DGS, as shown in Fig. 3, where a luminance–chromaticity decomposition strategy of the color representation is newly introduced. We explicitly factorize each Gaussian’s color into a scalar luminance term $Lum \in \mathbb{R}^+$ and a chromaticity function $f_{SH}(\theta, \phi) \in [0, 1]$ represented by SHs:

$$\begin{aligned}\mathbf{c}(\mathbf{d}, \mathbf{b}) &= Lum \cdot f_{SH}(\theta, \phi), \\ f_{SH}(\theta, \phi) &= \sum_{l=0}^L \sum_{m=-l}^l \mathbf{b}_l^m Y_l^m(\theta, \phi),\end{aligned}\quad (2)$$

where $\mathbf{b}_l^m \in \mathbb{R}^3$ are chromatic coefficients. Here, Lum controls the overall radiance magnitude, while $f_{SH}(\theta, \phi)$ captures only view-dependent color variations.

This decomposition can be seamlessly integrated into the vanilla 3DGS pipeline without architectural changes. We simply replace Eq. (1) with Eq. (2) during training and rendering, maintaining full compatibility with existing 3DGS implementations.

Discussion. This proposed decomposition yields critical advantages: *First*, by isolating luminance into a dedicated scalar parameter, optimization gradients are no longer dominated by high-radiance regions, and dark areas receive proportionate learning signals, enabling faithful reconstruction of shadow details. *Second*, the chromatic SH coefficients operate on limited color vectors (*e.g.*, $f_{SH}(\theta, \phi) \in [0, 1]$), constraining their domain to a compact manifold where low-order harmonics suffice to model view-dependent hue variations, even under extreme radiance gradients. *Third*, this representation aligns with human visual perception, where luminance and chromaticity are processed separately in the visual cortex [31], yielding more stable training dynamics and eliminating hue shifts in high-intensity regions.

3.4. Model Optimization

Considering the high dynamic range of HDR images, we adopt a μ -law function [4, 13] to the predicted HDR images and the ground truth images:

$$\hat{\mathbf{I}} = \frac{\log(1 + \mu * \mathbf{I})}{\log(1 + \mu)} \quad (3)$$

where μ is a compression factor and \mathbf{I} denotes rendered images or HDR ground truth. This design allows preserving the absolute luminance scale and maintaining radiometric consistency across scenes without any view-dependent normalization. Moreover, the logarithmic form of the μ -law compresses high-intensity regions while expanding the relative contrast of dark regions, increasing the perceptual and gradient-level weight of low-luminance pixels. This allows the model to better attend to dark areas that are easily neglected during optimization. Here, the μ -law serves as a differentiable radiometric compression operator for training, rather than a perceptual tone mapper for display.

The objective of training NH-3DGS is formed as:

$$\mathcal{L} = \lambda \cdot \mathcal{L}_1(\hat{\mathbf{I}}_{\text{pred}}, \hat{\mathbf{I}}_{\text{gt}}) + (1 - \lambda) \cdot \mathcal{L}_{SSIM}(\hat{\mathbf{I}}_{\text{pred}}, \hat{\mathbf{I}}_{\text{gt}}) \quad (4)$$

where $\hat{\mathbf{I}}_{\text{pred}}$ and $\hat{\mathbf{I}}_{\text{gt}} \in \mathbb{R}^{H \times W \times C}$ represent rendered and HDR ground truth, respectively, both transformed by the μ -

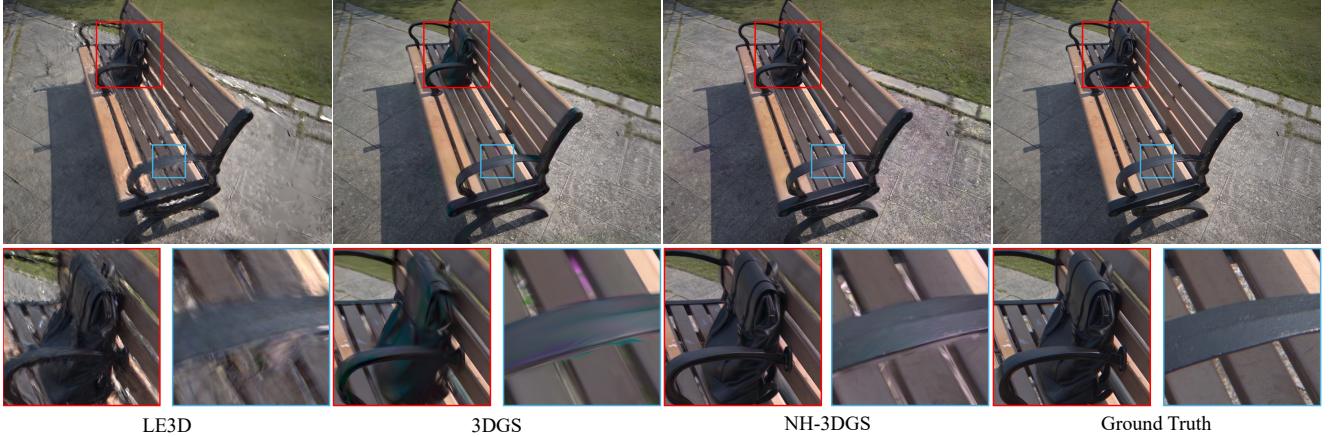


Figure 4. HDR rendering on our collected RAW-4S dataset. 3DGS [15] fails to learn a well-calibrated HDR representation, exhibiting spatial blurring in low-illumination regions and chromatic aberrations in neutral highlights (manifesting as greenish/magenta tints).

law compression, and λ controls the trade-off between the two loss terms.

Training on RAW data. For native Bayer-pattern RAW images (*e.g.*, RAW-4S), we operate directly in the sensor-native Bayer domain rather than demosaicing to RGB space as in existing approaches [14, 17, 27]. This eliminates reconstruction artifacts and chromatic noise introduced by conventional demosaicing algorithms.

During optimization, we apply the physical Bayer sampling pattern to rendered outputs and compute photometric loss in this subsampled color space. Specifically, given a rendered HDR RGB image $\mathbf{I}_{\text{pred}} = [\mathbf{I}_R, \mathbf{I}_G, \mathbf{I}_B]$, we simulate the native sensor response by masking each channel according to the Bayer filter layout $\mathcal{M} = \{\mathbf{M}_R, \mathbf{M}_G, \mathbf{M}_B\}$, yielding a single-channel Bayer image:

$$\mathbf{I}_{\text{pred}} \leftarrow \mathbf{M}_R \odot \mathbf{I}_R + \mathbf{M}_G \odot \mathbf{I}_G + \mathbf{M}_B \odot \mathbf{I}_B, \quad (5)$$

where \odot denotes element-wise multiplication, and each mask $\mathbf{M}_c \in \{0, 1\}^{H \times W}$ ($c \in \{R, G, B\}$) encodes the spatial locations of the corresponding color filter in the Bayer pattern (*e.g.*, BGGR). The \mathcal{L}_1 loss is computed between μ -law compressed predictions and ground truth.

For SSIM, direct computation on Bayer images is ill-defined due to spatially varying spectral sensitivity. We therefore define a Bayer-pattern-consistent SSIM loss as the mean across four constituent color channels:

$$\mathcal{L}_{\text{SSIM}}(\hat{\mathbf{I}}_{\text{pred}}, \hat{\mathbf{I}}_{\text{gt}}) = 1 - \frac{1}{4} \sum_{i=1}^4 \text{SSIM}(\hat{\mathbf{I}}_{\text{pred}}^i, \hat{\mathbf{I}}_{\text{gt}}^i) \quad (6)$$

where $\hat{\mathbf{I}}^i \in \mathbb{R}^{\frac{H}{2} \times \frac{W}{2}}$, $i \in \{1, 2, 3, 4\}$ denotes the monochromatic sub-image for each color channel, extracted via pixel subsampling according to the Bayer pattern. This formulation preserves native sensor characteristics while avoiding premature demosaicing artifacts.

4. Experiments

Datasets. To validate the performance of our NH-3DGS, we conduct extensive experiments on both synthetic and real-world data. We adopt the *synthetic* HDR dataset from HDR-NeRF [13], comprising 8 scenes rendered at 800×800 resolution in Blender [2], named as *Syn-8S*. Each scene contains 35 multi-view HDR images, with every HDR image explicitly paired with 5 corresponding LDR captures taken at distinct shutter speeds. Critically, this multi-exposure LDR acquisition protocol which demands precise exposure bracketing, vibration-free camera rigs, and per-scene recalibration represents a costly and tedious process that severely limits real-world deployment scalability.

For real-world evaluation, we introduce **RAW-4S**, a new multi-view RAW dataset comprising four indoor and outdoor scenes with geometrically and photometrically complex characteristics. Capturing this dataset required careful calibration and synchronization of multiple RAW-capable cameras under challenging illumination conditions, including strong direct sunlight (Bag scene) and severe back-lighting with extreme dynamic range silhouetting (Chair scene). Unlike existing datasets that rely on synthetic HDR data or simple exposure bracketing, RAW-4S provides native single-exposure RAW captures with precise camera poses, making it a valuable resource for HDR-aware 3D reconstruction research. The acquisition process involved significant effort in multi-camera calibration, pose estimation under varying lighting, and quality control to ensure radiometric consistency across views. Please refer to Sec. 2 in the Supplementary Material for detailed capture protocols and dataset statistics.

Implementation details. NH-3DGS is trained with Adam optimizer. The learning rate of luminance attribute is set to 0.05, and luminance is learned in the log space for synthetic dataset. While for RAW-4S dataset, luminance is



Figure 5. Visual comparisons on the Syn-8S dataset.

directly trained in the linear space. All the experiments are conducted with a single NVIDIA RTX 4090 GPU. λ and μ are set to 0.2 and 5000 during all the experiments. When adapting NeRF [20] for HDR image training, we replace its output activation function from sigmoid to soft-plus, enabling the model to predict unbounded radiance values while preserving the original architecture and training pipeline. For fair comparison, some compared methods [4, 36] are granted access to LDR inputs on the Syn-8S dataset per their original implementations.

Evaluation metrics. We adopt PSNR and SSIM as primary metrics, supplemented by LPIPS as a perceptual similarity measure. Inference speed (fps) is also reported for efficiency analysis. Following established practice in prior work [4, 13, 18], we use Photomatix Pro [11] to apply tone mapping to our HDR renderings, converting them into display-referred LDR images for qualitative visualization and fair comparison with existing methods.

For the RAW-4S dataset, we further report PSNR com-

puted directly on rendered novel-view RAW images, without any tone mapping or color space conversion, thereby preserving their native radiometric fidelity. Given our focus on high-fidelity HDR reconstruction, all metric evaluations, including PSNR, SSIM, and LPIPS, are performed between the rendered HDR outputs and the corresponding ground-truth novel-view RAW images after bilinear demosaicing. *All reported results are averaged across all scenes.* Importantly, unless otherwise stated, evaluations are conducted at the **full resolution** of the original training images, and **no** quantization to the 8-bit range is applied, ensuring an unbiased and physically meaningful assessment of HDR reconstruction performance. Please refer to Sec. 2.2 in the Supplementary Material for more details.

4.1. Quantitative Evaluation

Competitors. On the synthetic Syn-8S dataset, we evaluate against canonical LDR-focused NVS approaches, namely NeRF [20], and 3DGS [15] (with different orders of SH),

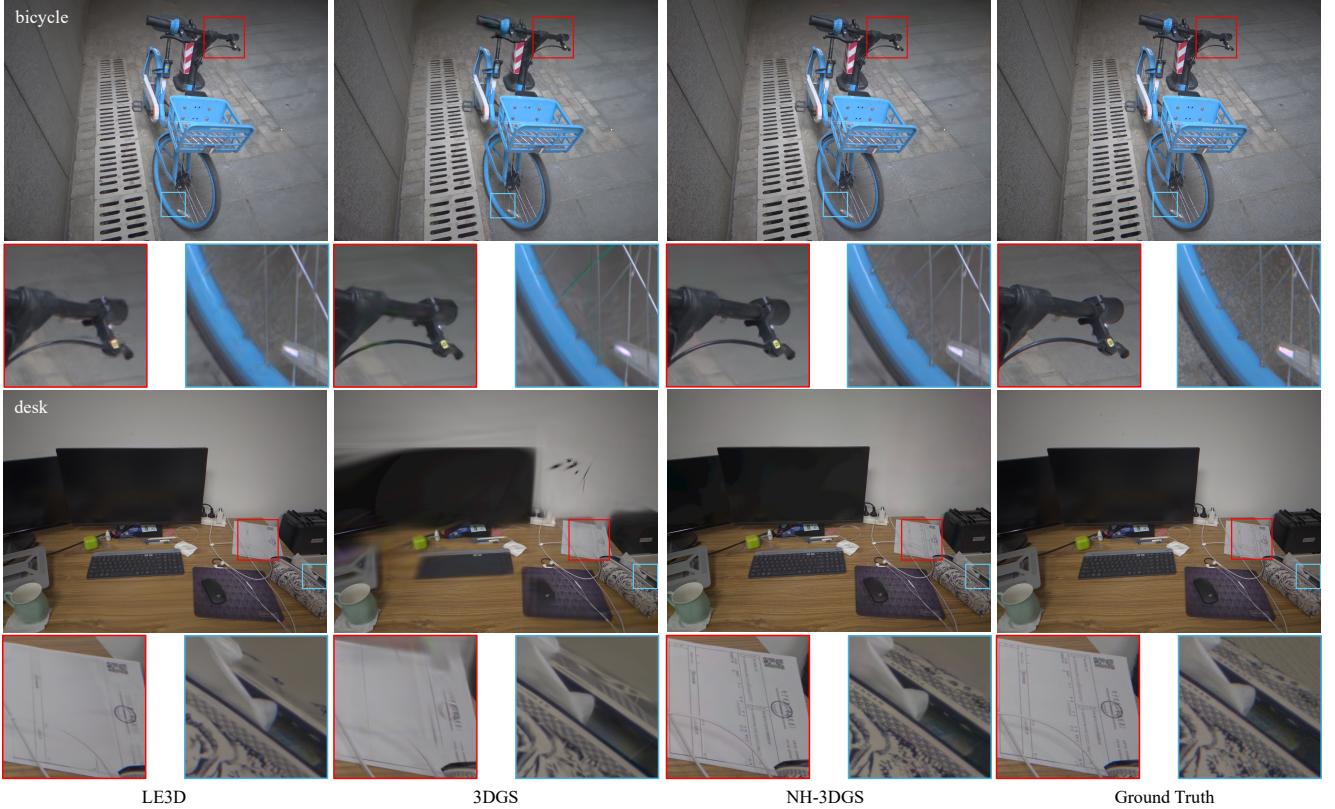


Figure 6. Visual comparisons on the RAW-4S dataset.

which we adapt to accept HDR inputs directly, thereby serving as strong competitors for evaluating the necessity of HDR-specific modeling. Additionally, we compare NH-3DGS against four state-of-the-art HDR NVS methods: (1) HDR-NeRF [13], the pioneering approach that leverages the implicit NeRF framework to synthesize HDR novel views from multi-exposure LDR inputs; (2) HDR-GS [4], which adapts the efficient explicit representation of 3DGS to model HDR radiance, also trained with multi-exposure LDR images and HDR ground-truth supervision; (3)–(4) Mono-HDR-GS and Mono-HDR-NeRF [36] that aim to reconstruct HDR scenes from single-exposure LDR images and similarly utilize HDR supervision during training on synthetic scenes.

On the real-world RAW-4S dataset, we compare our method against three state-of-the-art approaches: (1) RAW-NeRF [21], the first method to reconstruct HDR scenes from multi-view low-light RAW images by extending Mip-NeRF [1]; (2) LE3D [14], a 3DGS-based framework that employs an MLP to model HDR appearance and is trained on bilinearly demosaiced RAW images. (3) We adapt 3DGS [15] to operate directly on HDR inputs.

Tab. 1 presents the results on the Syn-8S dataset. We highlight the following key points:

(I) Superior HDR reconstruction quality. Our model

Table 1. Results on the Syn-8S dataset. 3DGS-L4/5 denotes 3DGS with orders of SH 4/5.

Method	Supervision	Inference		
		PSNR↑	SSIM↑	LPIPS↓
NeRF [20]	HDR	15.20	0.388	0.753
3DGS [15]	HDR	33.19	0.914	0.095
3DGS-L4	HDR	34.93	0.839	0.075
3DGS-L5	HDR	35.37	0.948	0.067
HDR-NeRF [13]	LDR	36.40	0.936	0.018
HDR-GS [4]	HDR+LDR	38.29	0.968	0.014
Mono-HDR-GS [36]	HDR+LDR	38.57	0.970	0.013
Mono-HDR-NeRF [36]	HDR+LDR	32.86	0.940	0.068
NH-3DGS (Ours)	HDR	39.77	0.972	0.011

significantly outperforms all LDR-focused competitors [15, 20]. It is observed that directly applying NeRF [20] to HDR imagery leads to poor HDR reconstruction performance since its radiance field formulation implicitly assumes bounded color intensities, making it incapable of representing the wide, nonlinear luminance variations inherent in HDR data. Relative to 3DGS, NH-3DGS achieves a substantial +6.58 dB PSNR gain, attributable to our luminance-chromaticity decomposition that resolves the fundamental limitation of entangled SH representations in HDR regimes. While increasing SH orders (*e.g.*, to fifth-order) marginally improves 3DGS’s HDR performance, such approaches in-

Table 2. Results on the RAW-4S dataset. GM: GPU Memory

Method	RAW		HDR RGB		Inference (fps)↑	GM (GB)↓
	PSNR ↑	PSNR ↑	SSIM ↑	LPIPS ↓		
3DGS [15]	33.17	30.05	0.838	0.405	91	5
LE3D [14]	33.78	29.12	0.844	0.283	0.033	11
RAW-NeRF [21]	18.94	18.46	0.750	0.331	0.005	19
NH-3DGS	34.98	31.30	0.864	0.275	28	7

introduce significant computational overhead (41% slower inference) and may exhibit pronounced overfitting artifacts, as shown in Fig. 2 and discussed in Sec. 3.2. In contrast, our approach preserves the explicit, optimization-friendly structure of 3DGS while decoupling intensity from chromaticity, enabling seamless integration into any SH-based 3DGS framework without architectural overhaul.

(II) Advantages of native HDR representation. NH-3DGS surpasses HDR NVS methods that reconstruct HDR scenes with additional HDR supervision [4, 36] as well as HDR-NeRF [13] across all metrics, which confirms that native HDR observations preserve radiometric fidelity irrecoverable from multi-exposure LDR sequences. Crucially, our framework eliminates the need for complex LDR-to-HDR domain translation modules (*e.g.*, tone-mapping networks [4, 13]), operating directly in the physical radiance domain, avoiding error propagation and architectural complexity inherent in cross-domain approaches. While our model achieves marginally lower LPIPS scores (*e.g.*, 0.011 *vs.* 0.014 for HDR-GS [4]), this metrics cannot fully reflect critical perceptual improvements in HDR space [5], as evidenced in Fig. 5. These experiments demonstrate that multi-exposure LDR may not be necessary when HDR images are available, as directly learning from HDR data can yield more accurate radiance estimation and more stable color reconstruction without relying on exposure calibration or tone-mapping consistency.

(III) Computational efficiency. NH-3DGS achieves inference speeds comparable to vanilla 3DGS, demonstrating that luminance-chromaticity decomposition introduces negligible computational overhead. Compared to prior HDR methods, NH-3DGS is 2,000× faster than HDR-NeRF and 1.85× faster than HDR-GS [4]. This efficiency stems from our explicit parametric representation: all HDR attributes are optimized as direct Gaussian parameters rather than latent network outputs, eliminating the need for auxiliary neural networks while preserving 3DGS’s real-time rendering capability and extending its dynamic range coverage.

Tab. 2 reports HDR NVS results on our captured real-world RAW-4S dataset. We make several observations.

(I) Superior RAW radiance fidelity. NH-3DGS achieves state-of-the-art RAW-PSNR at native sensor resolution, outperforming all the compared methods. This gain stems from our *Bayer-native optimization*: by applying loss directly in the sensor subsampled domain (without demosaicing) and decoupling luminance from chromaticity,

we preserve the physical linearity of photon counts while avoiding reconstruction artifacts inherent to demosaicing algorithms. Crucially, our explicit luminance parameter *Lum* models absolute radiance magnitude independent of color sampling patterns, whereas LE3D’s MLP renderer fails to disentangle sensor noise from true radiance, resulting in blurred textures and inaccurate highlight reproduction in RAW space.

(II) High-fidelity HDR RGB reconstruction from non-demosaiced inputs. Remarkably, NH-3DGS achieves the highest PSNR/SSIM/LPIPS in *RGB space* despite training exclusively on non-demosaiced RAW images. This demonstrates our decomposition’s ability to implicitly learn color interpolation physics through geometric constraints: the chromaticity SH coefficients reconstruct full spectral response by leveraging multi-view consistency across Bayer patterns, while luminance *Lum* provides absolute scale calibration. Consequently, we outperform methods trained on demosaiced RGB data (*e.g.*, LE3D [14] *by* +2.18 dB PSNR), proving that native RAW optimization when combined with explicit radiance modeling can recover more accurate colorimetric and photometric properties.

(III) Computational efficiency and memory footprint. NH-3DGS demonstrates exceptional inference efficiency, achieving 5,600× and 840× speedups over RAW-NeRF [21] and LE3D [14], respectively. While NH-3DGS is marginally slower than vanilla 3DGS, this stems from the need for additional Gaussians to faithfully reconstruct details in dark regions—areas where 3DGS produces blurred renderings due to its entangled luminance-chromaticity representation. In terms of memory, NH-3DGS requires only 7 GB of VRAM for training, substantially less than LE3D (11 GB) and RAW-NeRF (19 GB), with only a modest increase over vanilla 3DGS (5 GB).

4.2. Qualitative Results

Numerical metrics such as PSNR, SSIM, and LPIPS may not fully reflect the perceived quality of images. Therefore, a qualitative evaluation through visual comparison is essential. For results on the Syn-8S dataset, as shown in Fig. 5, 3DGS tends to produce blurry and visually unappealing results in dark areas, as discussed before, and even HDR-GS tends to incur blurring in some typically dark regions. In contrast, our NH-3DGS can successfully recover smoother color details and present the brightness properly. For results on the captured RAW-4S dataset, as shown in Fig. 6 (also see Fig. 3 in the Supplementary Material), LE3D and 3DGS tend to render blurry content in dark areas (*e.g.*, the characters on the paper are missing). In contrast, our NH-3DGS achieves superior color consistency and detail preservation.

5. Conclusion

We introduce NH-3DGS, the first framework for high-fidelity 3D scene reconstruction directly from native HDR observations including single-exposure RAW captures without multi-exposure sequences or domain translation. By decoupling luminance and chromaticity in 3DGS, our approach eliminates three fundamental limitations of conventional HDR novel view synthesis: multi-shot motion artifacts, computational overhead, and exposure-stack calibration dependency. Comprehensive evaluations demonstrate state-of-the-art radiometric accuracy and visual fidelity across synthetic and real-world scenes captured by single-exposure RAW images.

References

- [1] Jonathan T. Barron, Ben Mildenhall, Matthew Tancik, Peter Hedman, Ricardo Martin-Brualla, and Pratul P. Srinivasan. Mip-nerf: A multiscale representation for anti-aliasing neural radiance fields, 2021. 7
- [2] Blender Foundation. Blender, 2025. 5
- [3] Jintong Cai and Huimin Lu. Nerf-based multi-view synthesis techniques: A survey. In *2024 International Wireless Communications and Mobile Computing (IWCMC)*, pages 208–213. IEEE, 2024. 1
- [4] Yuanhao Cai, Zihao Xiao, Yixun Liang, Minghan Qin, Yulin Zhang, Xiaokang Yang, Yaoyao Liu, and Alan L Yuille. Hdr-gs: Efficient high dynamic range novel view synthesis at 1000x speed via gaussian splatting. *Advances in Neural Information Processing Systems*, 37:68453–68471, 2024. 1, 2, 3, 4, 6, 7, 8
- [5] Peibei Cao, Rafal K Mantiuk, and Kede Ma. Perceptual assessment and optimization of hdr image rendering. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 22433–22443, 2024. 8
- [6] Kai Cheng, Xiaoxiao Long, Kaizhi Yang, Yao Yao, Wei Yin, Yuexin Ma, Wenping Wang, and Xuejin Chen. Gaussianpro: 3d gaussian splatting with progressive propagation. In *Forty-first International Conference on Machine Learning*, 2024. 2
- [7] Anurag Dalal, Daniel Hagen, Kjell G Robbersmyr, and Kristian Muri Knausgård. Gaussian splatting: 3d reconstruction and novel view synthesis: A review. *IEEE Access*, 12: 96797–96820, 2024. 1
- [8] Shengjie Feng, Xiaoqun Wu, and Jian Cao. A survey of multi-view stereo 3d reconstruction algorithms based on deep learning. *Digital Signal Processing*, page 105291, 2025. 2
- [9] Yang Fu, Sifei Liu, Amey Kulkarni, Jan Kautz, Alexei A. Efros, and Xiaolong Wang. Colmap-free 3d gaussian splatting. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 20796–20805, 2024. 2
- [10] Xuan Gao, Wei Li, and Baojie Fan. Mip-nerf+: Multi-scale 3d scene synthesis. In *International Conference on Intelligent Robotics and Applications*, pages 174–188. Springer, 2024. 2
- [11] HDRsoft Team. Photomatix pro, 2025. 6
- [12] Hongbo Huang, Xiaoxu Yan, Yaolin Zheng, Jiayu He, Longfei Xu, and Dechun Qin. Multi-view stereo algorithms based on deep learning: a survey. *Multimedia Tools and Applications*, 84(6):2877–2908, 2025. 2
- [13] Xin Huang, Qi Zhang, Ying Feng, Hongdong Li, Xuan Wang, and Qing Wang. Hdr-nerf: High dynamic range neural radiance fields. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 18398–18408, 2022. 1, 2, 4, 5, 6, 7, 8
- [14] Xin Jin, Pengyi Jiao, Zheng-Peng Duan, Xingchao Yang, Chong-Yi Li, Chun-Le Guo, and Bo Ren. Lighting every darkness with 3dgs: Fast training and real-time rendering for hdr view synthesis. In *NIPS*, 2024. 2, 5, 7, 8
- [15] Bernhard Kerbl, Georgios Kopanas, Thomas Leimkühler, and George Drettakis. 3d gaussian splatting for real-time radiance field rendering. *ACM Transactions on Graphics*, 42(4), 2023. 1, 2, 3, 4, 5, 6, 7, 8
- [16] Jonas Kulhanek, Songyou Peng, Zuzana Kukelova, Marc Pollefeys, and Torsten Sattler. Wildgaussians: 3d gaussian splatting in the wild. *arXiv preprint arXiv:2407.08447*, 2024. 2
- [17] Zhihao Li, Yufei Wang, Alex Kot, and Bihan Wen. From chaos to clarity: 3dgs in the dark. *Advances in Neural Information Processing Systems*, 37:94971–94992, 2024. 2, 5
- [18] Jinfeng Liu, Lingtong Kong, Bo Li, and Dan Xu. Gaussldr: High dynamic range gaussian splatting via learning unified 3d and 2d local tone mapping. In *Proceedings of the Computer Vision and Pattern Recognition Conference*, pages 5991–6000, 2025. 1, 2, 6
- [19] Xin Ma, Jiguang Zhang, Peng Lu, Shibiao Xu, and Chengwei Pan. Novel view synthesis under large-deviation viewpoint for autonomous driving. In *Proceedings of the AAAI Conference on Artificial Intelligence*, pages 6000–6008, 2025. 1, 2
- [20] Ben Mildenhall, Pratul P. Srinivasan, Matthew Tancik, Jonathan T. Barron, Ravi Ramamoorthi, and Ren Ng. Nerf: Representing scenes as neural radiance fields for view synthesis. In *ECCV*, 2020. 2, 6, 7
- [21] Ben Mildenhall, Peter Hedman, Ricardo Martin-Brualla, Pratul P Srinivasan, and Jonathan T Barron. Nerf in the dark: High dynamic range view synthesis from noisy raw images. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 16190–16199, 2022. 2, 7, 8
- [22] Thomas Müller, Alex Evans, Christoph Schied, and Alexander Keller. Instant neural graphics primitives with a multiresolution hash encoding. *ACM Trans. Graph.*, 41(4):102:1–102:15, 2022. 2
- [23] Yuandong Niu, Limin Liu, Fuyu Huang, Siyuan Huang, and Shuangyou Chen. Overview of image-based 3d reconstruction technology. *Journal of the European Optical Society-Rapid Publications*, 20(1):18, 2024. 1
- [24] Linfei Pan, Dániel Baráth, Marc Pollefeys, and Johannes L Schönberger. Global structure-from-motion revisited. In *European Conference on Computer Vision*, pages 58–77. Springer, 2024. 2

- [25] Xiaonan Pan, Qilei Sun, Jia Wang, and Eng Gee Lim. Game engine based multi-view video dataset synthesis for pedestrian detection and tracking. In *2024 IEEE International Conference on Metaverse Computing, Networking, and Applications (MetaCom)*, pages 259–264. IEEE, 2024. [1](#) [2](#)
- [26] Santosh Reddy, H Abhiram, and KS Archish. A survey of 3d gaussian splatting: Optimization techniques, applications, and ai-driven advancements. In *2025 International Conference on Intelligent and Innovative Technologies in Computing, Electrical and Electronics (IITCEE)*, pages 1–6. IEEE, 2025. [1](#)
- [27] Shreyas Singh, Aryan Garg, and Kaushik Mitra. Hdrsplat: Gaussian splatting for high dynamic range 3d scene reconstruction from raw images. *BMVC*, 2024. [2](#) [5](#)
- [28] Laurie Van Bogaert, Daniele Bonatto, Sarah Fernandes Pinto Fachada, and Gauthier Lafruit. Novel view synthesis in embedded virtual reality devices. *Electronic Imaging*, 34:1–6, 2022. [1](#) [2](#)
- [29] Jian-Gang Wang, Lubing Zhou, Zhiwei Song, and Miaolong Yuan. Real-time vehicle signal lights recognition with hdr camera. In *2016 IEEE International Conference on Internet of Things (iThings) and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData)*, pages 355–358. IEEE, 2016. [2](#) [3](#)
- [30] Rui Wang, John Tran, and David Luebke. All-frequency relighting of glossy objects. *ACM Transactions on Graphics (TOG)*, 25(2):293–318, 2006. [3](#)
- [31] Dajun Xing, Ahmed Ouni, Stephanie Chen, Hinde Sahmoud, James Gordon, and Robert Shapley. Brightness–color interactions in human early visual cortex. *Journal of Neuroscience*, 35(5):2226–2232, 2015. [4](#)
- [32] Xingting Yao, Qinghao Hu, Fei Zhou, Tielong Liu, Zitao Mo, Zeyu Zhu, Zhengyang Zhuge, and Jian Cheng. Spin-erf: Direct-trained spiking neural networks for efficient neural radiance field rendering. *Frontiers in Neuroscience*, 19:1593580, 2025. [2](#)
- [33] Zongxin Ye, Wenyu Li, Sidun Liu, Peng Qiao, and Yong Dou. Absgs: Recovering fine details in 3d gaussian splatting. In *Proceedings of the 32nd ACM International Conference on Multimedia*, pages 1053–1061, 2024. [2](#)
- [34] Alex Yu, Ruilong Li, Matthew Tancik, Hao Li, Ren Ng, and Angjoo Kanazawa. PlenOctrees for real-time rendering of neural radiance fields. In *ICCV*, 2021. [2](#)
- [35] Zehao Yu, Anpei Chen, Binbin Huang, Torsten Sattler, and Andreas Geiger. Mip-splatting: Alias-free 3d gaussian splatting. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 19447–19456, 2024. [2](#)
- [36] Kaixuan Zhang, Hu Wang, Minxian Li, Mingwu Ren, Mao Ye, and Xiatian Zhu. High dynamic range novel view synthesis with single exposure. *arXiv preprint arXiv:2505.01212*, 2025. [6](#) [7](#) [8](#)
- [37] Lin Zhu, Kangmin Jia, Yifan Zhao, Yunshan Qi, Lizhi Wang, and Hua Huang. Spikenerf: Learning neural radiance fields from continuous spike stream. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 6285–6295, 2024. [2](#)