2026 IEEE International Conference on Robotics and Automation (ICRA 2026)
June 1-5, 2026. Vienna, Austria

TI-3DGS: 3D Thermal Reconstruction via Thermal Imaging-guided 3D Gaussian Splatting

Yong Tang[1*], Yunhao Li[2*], Xiaodong Wang[2], Qiang Song[3], Bing Qin[1†], Xiaocheng Feng[1†] and Xin Yuan[2†]

***Abstract*— Thermal imaging, with its all-weather capabilities and strong penetration, enables 3D reconstruction in low-light and adverse conditions. In this paper, we investigate RGB-independent pure 3D thermal reconstruction, aiming to overcome the challenges of 3D reconstruction in extreme environments where RGB images are unavailable. However, directly applying visible-light 3D reconstruction methods to thermal images often leads to severe artifacts due to two key challenges: (i) thermal images lack rich textures, hindering detail reconstruction, and (ii) heat conduction causes intensity diffusion, resulting in blurred edges. To address these issues, we propose TI-3DGS, a novel 3D Gaussian Splatting framework guided by thermal imaging. We introduce a Thermal Imaging Field (TIF) to model radiance in thermal domains and a Thermal Attenuation-aware Density Control (TADC) strategy to densify sparse point clouds from low-texture thermal inputs. Additionally, we incorporate an edge-enhancement constraint to mitigate blur from heat diffusion. Extensive experiments on the TI-NSD dataset, covering indoor and outdoor scenarios, show that our TI-3DGS achieves state-of-the-art performance, effectively overcoming texture sparsity and edge degradation in thermal reconstruction.**

I. INTRODUCTION

Thermal imaging technology is a non-contact, passive temperature measurement method that captures thermal radiation from objects [1]. Due to its high sensitivity to thermal energy and robust performance in low-light or obscured environments, this modality has become indispensable for critical applications such as surveillance, industrial equipment diagnostics, and medical thermography. Building upon these capabilities, 3D thermal reconstruction—the process of reconstructing multi-view thermal images into thermally accurate 3D scenes or objects—has emerged as a pivotal task for thermal analysis.

This work was supported by the National Natural Science Foundation of China (NSFC) (grant 6252200908, 62276078, U22B2059, U2541205, 62271414), the Fundamental Research Funds for the Central Universities (XNJKKGYDJ2024013), the State Key Laboratory of Micro-Spacecraft Rapid Design and Intelligent Cluster (MS01240122), National Key R&D Program of China (2024YFF0505603), "Pioneer" and "Leading Goose" R&D Program of Zhejiang (grant 2024SDXHDX0006, 2024C03182), the Key Project of Westlake Institute for Optoelectronics (grant number 2023GD007) and the 2023 International Sci-tech Cooperation Projects under the purview of the "Innovation Yongjiang 2035" Key R&D Program (grant number 2024Z126).

*Equal contribution

†Corresponding authors

[1]Yong Tang, Bing Qin and Xiaocheng Feng are with Harbin Institute of Technology, Harbin, Heilongjiang, China. {yongt, qinb, xcfeng}@ir.hit.edu.cn

[2]Yunhao Li, Xiaodong Wang and Xin Yuan are with Westlake University, Hangzhou, Zhejiang, China. {liyunhao, wangxiaodong, xyuan}@westlake.edu.cn

[3]Qiang Song is with Hunan University, Changsha, Hunan, China. songqiangshanghai@foxmail.com

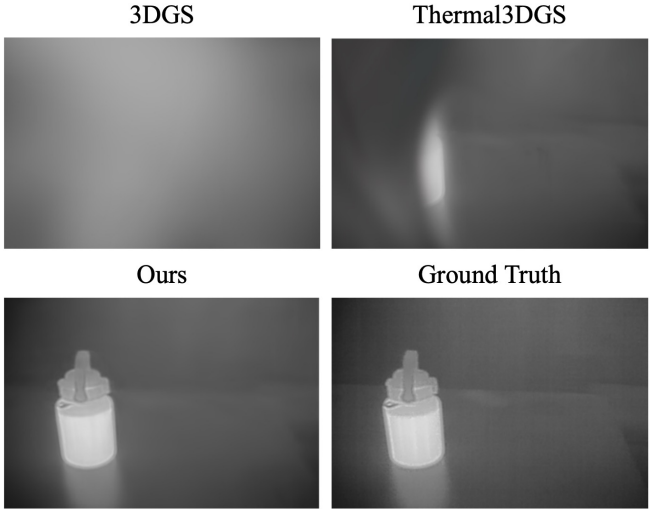

Fig. 1: 3D-GS and Thermal3DGS find it challenging to accurately model 3D thermal scenes from low texture multi-view thermal images, which leads to severe artifacts and even failures. The proposed TI-3DGS, taking into account the low texture thermal images caused by thermal imaging and heat conduction, achieves higher reconstruction quality.

In this paper, we study **RGB-independent pure 3D thermal reconstruction**, focusing on scenarios where RGB imagery is unavailable, such as darkness, fog, or smoke. Compared with conventional RGB-based reconstruction, thermal-only 3D reconstruction introduces two fundamental challenges. First, thermal cameras capture low-texture radiation patterns rather than rich visual textures, significantly limiting the recovery of fine geometric details. Second, heat conduction causes spatially varying radiation intensity around heat sources, leading to blurred boundaries and degraded structural fidelity in the reconstructed results.

To address these challenges, early methods [2], [3] adopted a two-stage strategy: reconstructing geometry from multi-view RGB images and subsequently mapping thermal data onto the recovered model. However, such pipelines inherit the limitations of visible-light reconstruction, resulting in suboptimal thermal modeling. More recently, adaptations of Neural Radiance Fields (NeRF) [4], [5] and 3D Gaussian Splatting (3D-GS) [6] for thermal scenes [7]–[9] have been proposed. Nevertheless, most approaches still depend on RGB cues to facilitate reconstruction and exhibit three key limitations. First, they often neglect thermal radiation attenuation, producing inaccurate temperature distributions. Second, due to the inherently low-texture nature of thermal imagery, common structure-from-motion (SfM) tools such

as COLMAP [10] yield sparse and incomplete point clouds, limiting geometric fidelity. Third, excessive smoothing in texture-deficient regions introduces noticeable artifacts, blurring thermal edges and obscuring structural boundaries.

To address the above challenges, we propose a novel thermal imaging-guided 3D Gaussian splatting framework, termed **TI-3DGS**, which tackles the three issues as follows. (1) For inaccurate 3D thermal reconstruction, we introduce a **Thermal Imaging Field (TIF)** that employs a multilayer perceptron (MLP) to model the thermal radiation capture process, predict imaging parameters, and compensate for attenuation, thereby improving thermal fidelity. (2) To enhance thermal scene representation, we design a **Thermal Attenuation-aware Density Control (TADC)** strategy that combines mesh subdivision [11] with radiation attenuation modeling [1]. A directed acyclic graph is constructed to connect each thermal Gaussian with its K nearest neighbors, enabling effective densification and initialization. (3) To mitigate visible artifacts, we propose an **Edge Enhancement Loss (EE-Loss)** that penalizes discontinuous thermal gradients, encouraging boundary consistency and improving edge quality in reconstructed thermal scenes. TI-3DGS is applicable to scenarios such as nighttime surveillance, autonomous driving, and fire rescue.

The main contributions of this paper are summarized as follows:

- We propose TI-3DGS, a novel thermal imaging-guided 3D thermal reconstruction framework that simulates a full-chain thermal imaging process by jointly learning emissivity, atmospheric attenuation coefficients, sensor sensitivity, and thermal radiation propagation distance.
- TI-3DGS integrates a Thermal Attenuation-aware Density Control (TADC) strategy to adaptively adjust the density of thermal Gaussians, mitigating the limitations of the sparse point cloud generated by the traditional SfM algorithm from low texture thermal images.
- TI-3DGS incorporates an edge enhancement loss to improve the reconstruction accuracy of edges in thermal images, due to the blurred edges caused by heat conduction.
- Extensive experiments conducted on TI-NSD, the largest real 3D thermal reconstruction dataset, demonstrate that the proposed method outperforms baseline approaches by achieving state-of-the-art (SOTA) performance in indoor scenes and comparable performance in outdoor scenes.

II. RELATED WORK

A. 3D Gaussian Splatting

3D Gaussian Splatting (3D-GS) [6] is an emerging explicit 3D representation that enables real-time novel view synthesis from multi-view images by modeling scenes with Gaussian primitives, each parameterized by position, scale, orientation, opacity, and spherical harmonics (SH) coefficients. Rendering is achieved through EWA splatting [12], projecting 3D Gaussians onto 2D image planes.

Due to its high rendering speed and photorealistic quality, 3D-GS has gained widespread attention across domains such as AIGC, autonomous driving, VR/AR, and robotics [13]–[15]. Various works focus on compressing 3D-GS via vector quantization [16]–[18], improving rendering fidelity [19]–[21], or incorporating additional cues like depth [22], [23] and diffusion priors [24]. Strategies such as progressive propagation [19] and geometry-aware initialization [25], [26] have further enhanced reconstruction robustness under sparse views. Surface-based methods [27]–[29] extract geometry via signed distance functions (SDF), though challenges remain in unbounded scenes and high computational cost. For semantic tasks, recent work integrates 2D semantic priors into the 3D-GS framework for detection, segmentation, and editing [30]–[33].

While most research focuses on RGB, geometry, and semantics, 3D-GS has the potential to encode broader physical modalities such as heat, pressure, and material. We build upon this by integrating the thermal modality into the 3D-GS pipeline and designing task-specific strategies based on thermal imaging principles.

B. 3D Thermal Reconstruction

3D thermal reconstruction aims to recover both geometry and thermal properties from thermal image sequences. Early works combined thermal and depth sensors to reconstruct thermal point clouds or meshes [2], [3], [34], [35], relying on multi-sensor calibration, stereo matching, or thermal-invariant features. However, these traditional geometry-based methods are limited in novel view synthesis quality compared to neural rendering techniques.

Neural Radiance Fields (NeRF) [36] brought significant advances in high-fidelity view synthesis and were quickly adopted for thermal scene reconstruction [4], [5], [37]–[39]. Approaches like ThermoNeRF [39], ThermalNeRF [4], and Thermal-NeRF [5] integrate thermal imagery into the NeRF framework. Others explore joint reconstruction of RGB and thermal modalities [38], [40]. However, NeRF-based methods remain limited by slow rendering and implicit representations, hindering practical deployment.

Recently, 3D-GS has emerged as an efficient alternative for thermal reconstruction due to its explicit scene modeling and real-time rendering. Several studies apply 3D-GS to thermal data [7]–[9]. ThermalGaussian [8] achieves photorealistic rendering of both RGB and thermal views but depends on accurate RGB-thermal camera calibration. NTR-Gaussian [9] incorporates temporal cues to address environmental effects on thermal readings. Thermal3DGS [7] simulates physical principles such as atmospheric transmission and heat conduction to guide thermal Gaussian parameter learning.

Despite these efforts, two critical limitations remain unaddressed: (1) the attenuation of thermal radiation introduced by hardware imaging systems, and (2) the low-texture nature of thermal images, which degrades reconstruction quality.

To overcome these challenges, we propose **TI-3DGS**, a thermal imaging-guided 3D Gaussian splatting framework

that introduces a *Thermal Imaging Field* for radiation modeling, a *Thermal Attenuation-aware Density Control* strategy for enhanced Gaussian initialization, and an *Edge Enhancement Loss* to recover fine thermal boundaries.

III. METHOD

A. Preliminaries on 3D Gaussian Splatting

3D Gaussian Splatting [6] represents scenes via anisotropic 3D Gaussians, enabling high-fidelity novel view synthesis (NVS) with ultra-low rendering latency. A 3D Gaussian includes a center position $\mu \in \mathbb{R}^3$ and a scale factor $s \in \mathbb{R}^3$, and a rotation quaternion $q \in \mathbb{R}^4$, such that the Gaussian distribution is

$$G(x) = e^{-\frac{1}{2}(x-\mu)^T \Sigma^{-1}(x-\mu)}, \tag{1}$$

where Σ is defined as

$$\Sigma = \mathbf{RSS}^T\mathbf{R}^T, \tag{2}$$

where S is the scaling matrix determined by s and R is the rotation matrix determined by q. To additionally model the view-dependent appearance, and an $\alpha \in \mathbb{R}$ for opacity. The color and opacity are also calculated by the Gaussian distribution illustrated in Eq.(1).

For rendering, all 3D Gaussians are projected onto the 2D camera planes via a differentiable Gaussian splatting pipeline [12]. Given the viewing transform matrix W and Jacobian matrix J of the affine approximation of the projective transformation, the covariance matrix Σ in camera coordinates is calculated as

$$\Sigma' = JW\Sigma W^T J^T, \tag{3}$$

The differentiable splatting method efficiently projects the 3D Gaussians to 2D Gaussian distributions, ensuring fast α-blending for rendering and color supervision. Pixel color C is rendered by depth-sorted Gaussians overlapping the pixel:

$$C = \sum_{i \in N} \mathcal{C}_i \alpha_i \prod_{j=1}^{i-1}(1-\alpha_i). \tag{4}$$

B. Thermal Imaging Model

Here, we briefly introduce the theory of thermal radiation, how thermal cameras capture thermal radiation, and the thermal imaging model [1]. Planck's law defines blackbody spectral emittance M_e at temperature T:

$$M_e(\lambda; T) = \frac{2\pi hc^2}{\lambda^5}(e^{hc/\lambda kT} - 1)^{-1} \tag{5}$$

where λ denotes wavelength, c light speed, h Planck's constant, and k Boltzmann's constant. For non-blackbody objects:

$$E(\lambda; \epsilon, T) = \epsilon M_e(\lambda; T) \tag{6}$$

with ϵ as object-dependent emissivity.

A thermographic camera is designed to measure the intensity of thermal radiation. A typical sensor measures the intensity of Long-Wave Infrared radiation(8–14 μm) as

$$E(\epsilon, T) = \int_{\lambda_1}^{\lambda_2} E(\lambda; \epsilon, T) d\lambda, \tag{7}$$

where E is the intensity of the radiated Long-Wave Infrared radiation light.

The air is an attenuating medium and some Long-Wave Infrared light is absorbed through the air as it travels from the objects to the thermal camera [1]. Using the Lambert-Beer law, the transmission of the medium is expressed as

$$I_{\text{out}}(\lambda) = I_{\text{in}}(\lambda) e^{-(\mu_{\text{a}}(\lambda)+\mu_{\text{s}}(\lambda))d}, \tag{8}$$

where I_{in} and I_{out} are the input and transmitted intensity of light, respectively. $\mu_{\text{a}}(\lambda)$ and $\mu_{\text{s}}(\lambda)$ represents the medium absorption attenuation coefficient and scattering attenuation coefficient, respectively, and d is the thickness of the medium.

Based on Eqs. (5), (6), and (8), the observed intensity I_{cam} by thermal camera is expressed as (i.e., Thermal Imaging Model)

$$I_{cam} = R_v e^{-(\mu_{\text{a}}+\mu_{\text{s}})d} E, \tag{9}$$

where R_v is the sensitivity that associates the energy density with the observed intensity [41]. The sensitivity R_v depends on the imaging system (e.g., image sensor, lens, and spectral filter). E is the thermal radiation of objects.

Object thermal radiation is attenuated by atmosphere and imaging systems before capture. This process (Eq. (9)) reveals two issues: 1) Existing 3D thermal reconstruction methods inadequately model the full imaging process, causing inaccuracies; 2) Thermal imaging captures only radiation intensity, lacking RGB's color and texture richness. Our method specifically addresses these limitations.

C. Proposed Method

Based on the thermal imaging process described in Section III-B, we present the TI-3DGS framework where each Gaussian (termed thermal Gaussian) encodes scene thermal radiation intensity via spherical harmonics. The thermal radiation intensity attenuates during atmospheric transmission and subsequently captured by the thermal camera, as mathematically modeled in Eq. (9). To simulate this propagation process, predict imaging parameters, and optimize thermal Gaussians, we developed a Thermal Imaging Field. Subsequently, to enhance scene representation through Gaussian densification, we designed the Thermal Attenuation-induced Density Control strategy. Additionally, to learn edge features while alleviating blur and artifacts, we introduce an Edge Enhancement Loss.

1) Thermal Imaging Field: In Sec.III-B, we systematically formalize the thermal imaging of the thermal cameras, and demonstrate that Thermal3DGS [7] solely accounts for atmospheric scattering and absorption attenuation, while neglecting the sensor-induced radiation attenuation inherent to thermal imaging system. This limitation introduces representational bias in thermal radiation field modeling. Our framework holistically incorporates four key attenuation mechanisms governing thermal intensity during sensor acquisition: atmospheric absorption attenuation coefficient μ_a, atmospheric scattering attenuation coefficient μ_s, geometric distance d, and thermal camera sensitivity R_v.

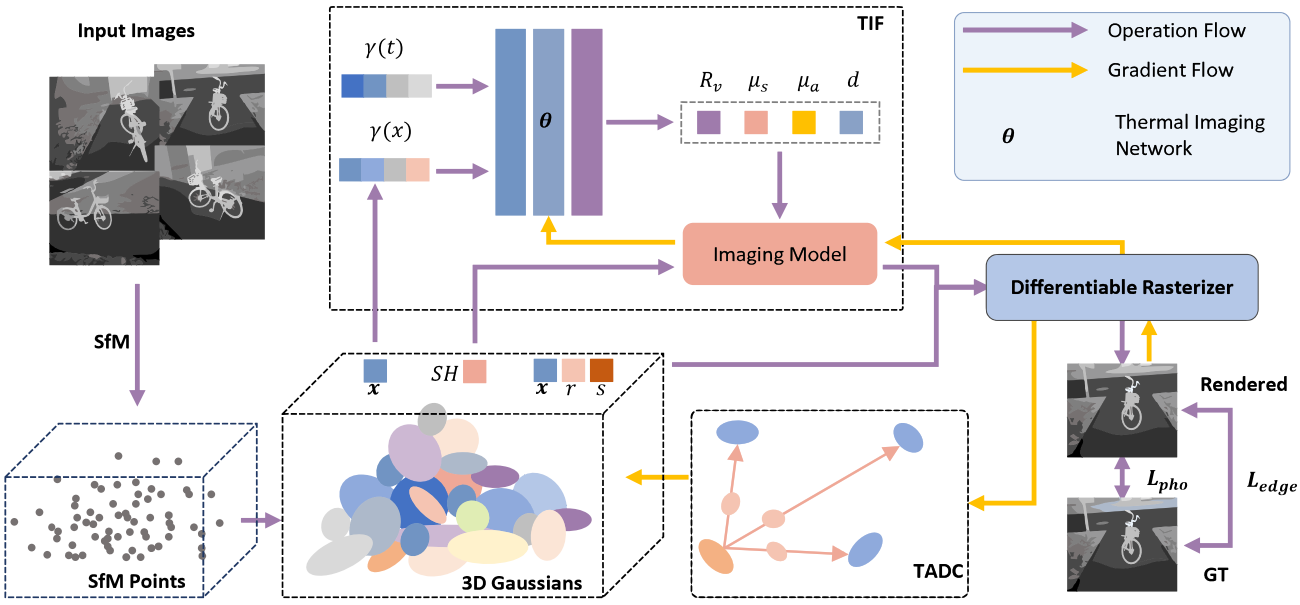

Fig. 2: An overview of our TI-3DGS: We model the thermal radiation field of the scene using a 3D thermal Gaussian model initialized from point clouds obtained via SfM algorithms such as COLMAP [10]. (1) We designed a thermal imaging field to simulate the thermal camera's radiation capture process, facilitating accurate thermal scene modeling. (2) For sparsely placed Gaussians, we introduced the Thermal Attenuation-induced Density Control strategy to effectively increase thermal Gaussian density while ensuring newly generated Gaussians inherit attributes from existing ones, thereby enhancing scene coverage. (3) By incorporating an Edge Enhancement Loss, we guide the Gaussian model to prioritize scene edge features, enabling salient feature learning in thermal radiation fields and mitigating artifacts.

We construct an MLP network to perform decoupling modeling on the coupling effects of the above four thermal radiation attenuation mechanisms, achieving independent learning of these coefficients (as shown in Fig. 2 TIF). This method effectively mitigates the coupling interference of atmospheric transmission, distance, and thermal camera on the thermal radiation, significantly improving the representation accuracy of the thermal radiation field. Specifically, the TIF network takes the positional encoded 3D Gaussian coordinate $x \in \mathbb{R}^3$ and the shooting time $t \in \mathbb{R}$ as inputs, dynamically predicting the attenuation coefficients of 3D Gaussians through the following formula

$$(\mu_a, \mu_s, d, R_v) = \mathcal{F}_\theta(\gamma(sg(\mathbf{x})), \gamma(t)), \tag{10}$$

Here, $\mathcal{F}_\theta$ represents our MLP network, $sg(\cdot)$ denotes the gradient truncation operation, and $\gamma(\cdot)$ represents a positional encoding function with frequency parameter $L = 10$.

$$\gamma(p) = \left(\sin(2^k \pi p), \cos(2^k \pi p)\right)_{k=0}^{L-1}. \tag{11}$$

Here, p serves as the input for the function $\gamma(\cdot)$.

Then, these four coefficients are applied to the normalized thermal radiation intensity SH of each 3D thermal Gaussian, computing the attenuated true thermal radiation intensity through the Thermal Imaging Model(i.e., Eq.(9)):

$$SH_t = R_v e^{-(\mu_a+\mu_s)d} SH, \tag{12}$$

where SH_t corresponds to the thermal radiation intensity captured by the thermal camera, and SH represents the scene's thermal radiation intensity. The true spherical harmonic coefficients SH_t with other Gaussian attributes are fed into a differentiable rasterizer to render a high-fidelity thermal image.

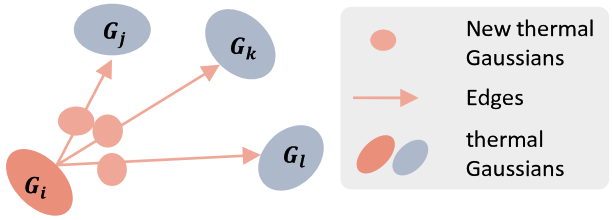

Fig. 3: Thermal Attenuation-induced Gaussian Density Control Illustration. G_i is the source node of the directed graph, referred to as the source thermal Gaussian, while G_j, G_k, G_l are the nearest neighbor nodes of G_i, termed destination thermal Gaussians. During the thermal Gaussian densification process, we generate new thermal Gaussians $G_{i,j}, G_{i,k}, G_{i,l}$ at a distance of Δd from the source thermal Gaussian along the edge connecting the source thermal Gaussian to its nearest neighboring thermal Gaussian.

2) Thermal Attenuation-induced Density Control: Existing 3D-GS-based methods suffer from limited reconstruction quality in thermal scenes due to sparse and inaccurate initialization from SfM (e.g., COLMAP [10]), especially under weak-texture conditions. Although Gaussian densifi-

cation improves coverage to some extent, it is insufficient in thermally homogeneous regions. Since the granularity of scene modeling directly depends on the spatial distribution of Gaussians, an effective density control mechanism is essential for robust thermal reconstruction under sparse-view settings.

To this end, we propose a **Thermal Attenuation-induced Density Control (TADC)** strategy, which consists of two components: (1) construction of a directed acyclic graph (*Proximity Graph*) to capture local spatial relations among thermal Gaussians, and (2) a thermal-aware densification mechanism that generates new Gaussians guided by thermal attenuation principles.

Proximity Graph Construction. We represent the thermal scene using a set of thermal Gaussians, each approximated as a heat-emitting source. During optimization, we build a Proximity Graph (Fig. 3) for each source thermal Gaussian, where vertices correspond to Gaussians and directed edges point to their K nearest neighbors. Each edge is weighted by the Euclidean distance between centers. The *Proximity Score* P_i of a Gaussian G_i is computed as:

$$P_i = \frac{1}{K} \sum_{k=1}^{K} d_{i,k}, \tag{13}$$

where $d_{i,k}$ denotes the distance between G_i and its k-th nearest neighbor. The Proximity Graph is dynamically updated during optimization to reflect changes from densification and pruning.

Thermal-aware Gaussian Densification. Inspired by mesh subdivision strategies [11], [42] and thermal radiation decay, we design a thermal-guided Gaussian densification approach. If the proximity score P_i exceeds a threshold t_{prox}, we insert new Gaussians (termed *slave Gaussians*) along each edge toward neighboring nodes, at a distance Δd from the source.

The attributes (scale, rotation, opacity, SH coefficients) of new Gaussians are derived from their source counterparts. In particular, spherical harmonics SH_n are computed using thermal attenuation:

$$SH_n = SH_s \cdot e^{-\mu \Delta d}, \tag{14}$$

where SH_s and SH_n denote the SH coefficients of the source and new Gaussians respectively; μ is the attenuation coefficient; Δd is defined as the distance between the new Thermal Gaussian and the source Thermal Gaussian. This strategy encourages new Gaussians to distribute adaptively in low-density regions, improving spatial coverage and thermal fidelity during optimization.

3) Edge Enhancement Loss: Thermal images typically exhibit smooth intensity distributions with few high-frequency features, resulting in weak texture and ambiguous boundaries. This leads to blurred reconstructions due to model over-smoothing in homogeneous regions. To preserve sharp thermal edges that often correspond to material boundaries or heat transfer discontinuities, we introduce a boundary-aware constraint in the form of an **Edge Enhancement Loss (EE-Loss)**.

Thermal edges are crucial for scene understanding, as they encode physical and geometric transitions. To capture these, we employ the Sobel operator to extract gradient magnitudes from both ground-truth and rendered thermal images. The edge loss $\mathcal{L}_{\text{edge}}$ is defined as:

$$\mathcal{L}_{\text{edge}} = \frac{1}{N} \sum_{i=1}^{N} \left(|\nabla E_{\text{source}}(i)| - |\nabla E_{\text{recon}}(i)| \right)^2, \tag{15}$$

where $\nabla E_{\text{source}}(i)$ and $\nabla E_{\text{recon}}(i)$ are the Sobel gradients of the original and reconstructed thermal images at pixel i.

We combine $\mathcal{L}_{\text{edge}}$ with standard photometric losses for end-to-end training:

$$\mathcal{L}_{\text{total}} = \lambda_{\text{edge}} \mathcal{L}_{\text{edge}} + \lambda \mathcal{L}_{\text{D-SSIM}} + (1 - \lambda_{\text{edge}} - \lambda) \mathcal{L}_1, \tag{16}$$

where $\mathcal{L}_1$ denotes the L1 photometric loss and $\mathcal{L}_{\text{D-SSIM}}$ represents a structural similarity loss. The combination ensures both global appearance accuracy and fine-grained edge preservation.

IV. EXPERIMENTS

A. Dataset

We compare our TI-3DGS against prior SOTA methods on TI-NSD [7] dataset, a large-scale real thermal infrared novel view synthesis dataset. Each scenario was captured in video format for approximately 90 seconds, yielding an excess of 50,000 frames of raw data. The scenes within the dataset are categorized into indoor and outdoorscenarios, consisting of 7 indoor scenes and 7 outdoor scenes. Specifically, the names of these scenes are Apples, Basketball-Court, Bicycle, Bridge, Car, Chair, Corridor, Heated, Merge, Sitting, Soccer-Goal, Standing and Tall-Building. Additionally, we captured two additional real datasets which contains two dynamic scenes Milk-Tea and Pacing using a Guide Infrared PLUG1212 thermal camera. We primarily assess our experimental results using various metrics, encompassing peak-signal-to-noise ratio (PSNR), learned perceptual image patch similarity (LPIPS) [43], and structural similarity index (SSIM).

B. Implementation Details

All of our code is implemented using PyTorch [44], and all experiments are conducted on an NVIDIA RTX 4090 GPU with 24GB of memory. We retain the differential Gaussian rasterization proposed by 3D-GS [6]. Training is performed for 7K and 30K iterations. For optimization, a single Adam optimizer [45] is used, but each component has a different learning rate: the 3D Gaussian maintains the same learning rate as the official implementation, while the learning rate of the TIN network in TIF undergoes exponential decay, ranging from (8×10^{-4}) to (1.6×10^{-6}). The beta values of Adam are set to $(0.9, 0.999)$.

C. Results

We use NeRF-based methods including InstantNGP [46], Plenoxels [47], and 3DGS-based methods including vanilla 3D-GS [6], and Thermal3DGS [7] as baselines. To ensure

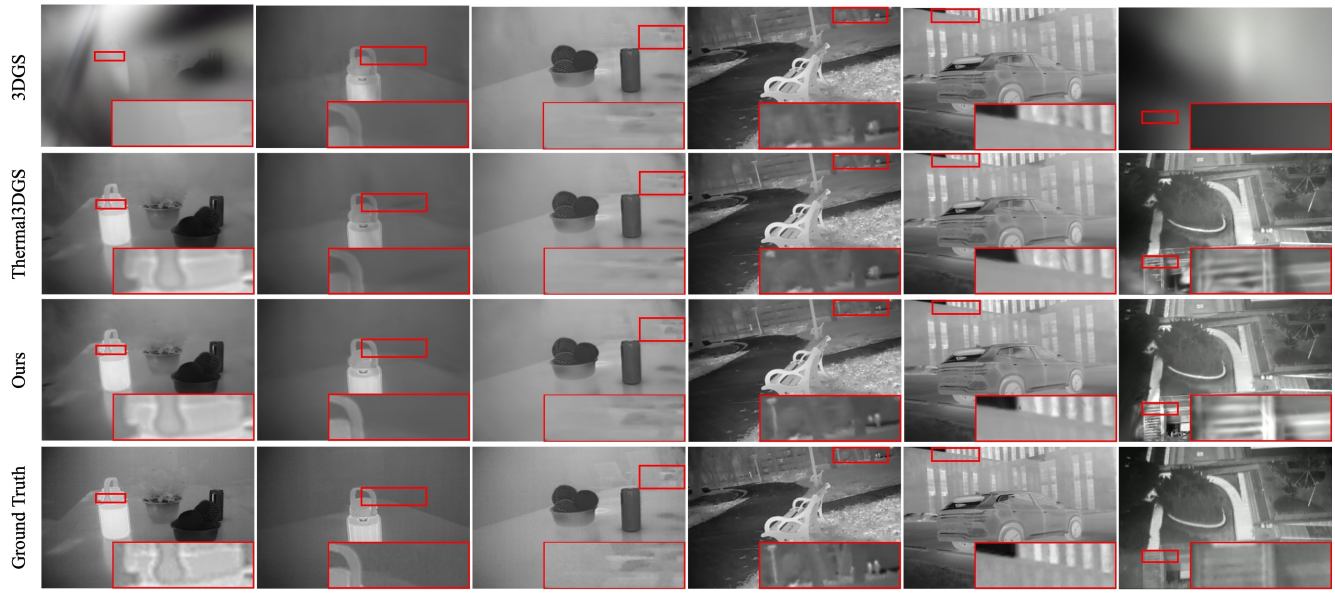

Fig. 4: We show comparisons of ours to previous methods and the corresponding ground truth images from test views. Left to right shows different scenes, with discrepancy regions highlighted by red bounding boxes.

TABLE I: Quantitative 3D thermal reconstruction comparisons on indoor scenes including Apples, Heated, Standing, Sitting, Corridor, Wall, and Merge. The best results are shown in bold and the second-best results are underlined. The results show that our method exhibits superior performance compared to state-of-the-art methods in indoor 3D thermal scene reconstruction.

Method	Apples			Heated			Standing			Sitting			Corridor			Wallk			Merge		
	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓
Plenoxels	20.11	0.890	0.373	24.73	0.908	0.335	19.82	0.825	0.441	18.23	0.797	0.442	23.42	0.898	0.381	28.99	0.909	0.342	19.62	0.843	0.378
Instant-NGP	30.52	0.934	0.278	30.81	0.938	0.283	30.55	0.917	0.267	31.94	0.937	0.243	20.38	0.897	0.358	25.52	0.933	0.265	22.52	0.868	0.331
3D-GS-7k	30.67	0.949	0.292	30.05	0.950	0.282	34.71	0.963	0.235	33.59	0.962	0.229	27.53	0.940	0.308	34.12	0.953	0.264	28.28	0.936	0.282
3D-GS-30k	32.05	0.952	0.284	30.58	0.951	0.279	36.14	0.967	0.219	36.51	0.968	0.211	30.63	0.942	0.294	35.28	0.952	0.256	29.70	0.936	0.274
Thermal3DGS-7k	32.32	0.957	0.282	33.17	0.959	0.274	35.06	0.965	0.236	34.92	0.966	0.228	28.33	0.940	0.305	34.66	0.960	0.259	30.40	0.950	0.270
Thermal3DGS-30k	34.99	0.961	0.272	36.62	0.963	0.265	37.50	0.968	0.222	37.94	0.970	0.212	32.10	0.949	0.290	**39.63**	**0.968**	0.244	33.27	**0.955**	0.257
Ours-7k	32.56	0.955	0.063	35.23	0.960	0.074	35.54	0.964	0.074	35.00	0.966	0.066	29.25	0.947	0.092	36.66	0.959	0.097	32.28	0.950	0.086
Ours-30k	**35.70**	**0.961**	**0.053**	**37.92**	**0.964**	**0.064**	**37.56**	**0.968**	**0.068**	**38.48**	**0.971**	**0.061**	**32.94**	**0.950**	**0.080**	39.40	0.967	**0.083**	**34.28**	0.952	**0.074**

TABLE II: Quantitative 3D thermal reconstruction comparisons on outdoor scenes. The best results are shown in bold and the second-best results are underlined. The results demonstrate that our method can render thermal images with high quality from represented thermal 3D scenes.

Method	Chair			Bicycle			Car			Soccer-Goal			Basketball-Court			Tall-Building			Bridge		
	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓
Plenoxels	22.48	0.732	0.447	20.99	0.758	0.423	22.98	0.794	0.420	22.92	0.803	0.418	19.53	0.803	0.405	23.28	0.741	0.462	22.84	0.743	0.455
Instant-NGP	26.79	0.818	0.301	26.33	0.856	0.260	27.63	0.852	0.319	29.19	0.888	0.265	24.16	0.877	0.308	23.98	0.751	0.290	27.05	0.834	0.304
3D-GS-7k	30.12	0.919	0.202	30.13	0.941	0.195	27.54	0.910	0.249	30.91	0.937	0.222	17.73	0.809	0.346	21.54	0.838	0.278	26.71	0.892	0.257
3D-GS-30k	32.32	0.934	0.170	31.86	0.948	0.177	31.74	0.936	0.201	33.11	0.948	0.194	19.29	0.810	0.346	24.96	0.847	0.273	28.91	0.906	0.225
Thermal3DGS-7k	29.46	0.921	0.203	30.70	0.944	0.194	28.03	0.917	0.242	30.96	0.940	0.219	24.21	0.909	0.260	24.19	0.873	0.235	27.66	0.900	0.251
Thermal3DGS-30k	**33.03**	**0.937**	0.172	33.15	**0.954**	0.174	32.08	**0.942**	0.199	**34.61**	**0.952**	0.191	30.98	0.945	0.212	33.85	**0.944**	0.151	30.52	**0.918**	0.212
Ours-7k	30.60	0.919	0.104	31.35	0.944	0.067	28.19	0.915	0.115	31.49	0.938	0.081	26.55	0.918	0.133	28.84	0.920	0.096	28.21	0.901	0.128
Ours-30k	32.73	0.932	**0.076**	**33.40**	0.952	**0.053**	**32.34**	0.941	**0.077**	34.52	0.951	**0.060**	**31.54**	**0.947**	**0.080**	**33.87**	0.943	**0.065**	**31.09**	0.917	**0.095**

fair comparisons, we adhere to the experimental settings of Thermal3DGS, employing a train/test split strategy where every 8th photo is selected for the test set. This approach guarantees consistent and meaningful comparisons when generating error metrics. We compare our method and Thermal3DGS under two configurations—7K and 30K iterations—across 7 indoor scenes (Apples, Heated, Standing, Sitting, Corridor, Wall, Merge) and 7 outdoor scenes (Chair, Bicycle, Car, Soccer-Goal, Basketball-Court, Tall-Building, Bridge). The results among indoor and ourdoor scenes from different methods are presented in Fig. 4, Tab. I and Tab. II. The experimental results demonstrate that our TI-3DGS achieves a 0.6dB higher PSNR than the SOTA method Thermal3DGS in indoor scenes and 0.18dB superiority in outdoor scenes. Specifically, our TI-3DGS achieves PSNR improvements of 1.3 dB, 1.01 dB, and 0.84 dB over the state-of-the-art (SOTA) method in the Heated, Merge, and Corridor scenes, respectively, demonstrating its effectiveness in indoor 3D thermal reconstruction. We also noticed that our method performs significantly better than SOTA methods on 7K-iteration scenarios, with $>$**1dB** in terms of PSNR in indoor and outdoor scenes. These results demonstrate that our method can achieve high performance in a very short time after training begins, which is highly advantageous for time-sensitive applications and situations with limited computing resources, such as autonomous driving, surveillance robots and drones. Fig. 5 qualitatively compares TI-3DGS with baseline methods on our dynamic thermal dataset,

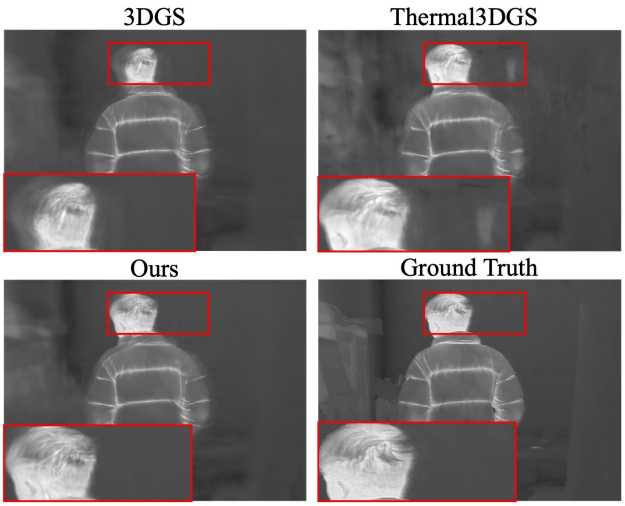

Fig. 5: Qualitative comparisons of our method against prior SOTA 3D thermal scene reconstruction methods on our captured dynamic scenes. Discrepancy regions highlighted by red bounding boxes.

while Tab. III reports quantitative results against state-of-the-art approaches. Together, these comparisons show that TI-3DGS better preserves scene details and suppresses artifacts in dynamic thermal 3D reconstruction, validating its effectiveness for time-varying thermal 3D scene modeling and cross-dataset generalization. These findings comprehensively validate the overall effectiveness of our method.

Fig.1 and Fig.4 illustrate the differences in visual quality. Our analysis reveals that conventional 3D-GS exhibits limited representational capacity in low-texture scenes, frequently manifesting severe artifacts and even complete reconstruction failures. Although Thermal3DGS demonstrates improved performance over basic 3D-GS, its reconstructions remain constrained by noticeable blurring effects. In contrast, our proposed methodology achieves superior 3D thermal scene reconstruction fidelity, effectively addressing both blurred edges and reconstruction collapse observed in previous approaches. We also recorded the training time and rendering/inference speed of methods. For 30K-iteration training, our TI-3DGS achieves an average training time of 36.12 minutes and a rendering speed of 82.65 FPS, enabling real-time performance. Compared to Thermal3D-GS, which requires 38.27 minutes for training and rendered at 85.54 FPS, TI-3DGS is faster in training while maintaining comparable rendering efficiency.

D. Ablation Study

To prove the effectiveness of the components of our TI-3DGS, including Thermal Imaging Field (TIF), Thermal Attenuation-aware Density Control strategy (TADC), and Edge Enhancement Loss $\mathcal{L}_{edge}$, we conducted an ablation study of our model on the indoor and outdoor datasets. Tab. 3 shows the quantitative results, in which are the average image quality metrics on indoor and outdoor datasets. As is shown in Tab. IV, it has been verified that TADC, TIF and

TABLE III: Quantitative 3D thermal reconstruction comparisons on dynamic thermal scenes in our captured datasets.

Method	Milk-Tea			Pacing		
	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓
3DGS	15.93	0.792	0.386	23.18	0.895	0.296
Thermal3DGS	22.11	0.880	**0.301**	24.31	0.889	0.318
Ours	**22.23**	**0.883**	0.304	**25.64**	**0.905**	**0.287**

TABLE IV: Quantitative results of ablation study on TI-NSD dataset.

	Indoor		Outdoor		Average	
	PSNR↑	SSIM↑	PSNR↑	SSIM↑	PSNR↑	SSIM↑
Ours w/o TIF	32.12	0.951	28.09	0.897	31.63	0.933
Ours w/o TADC	35.14	0.957	32.69	**0.945**	34.79	0.954
Ours w/o $\mathcal{L}_{edge}$	34.88	0.957	31.00	0.934	34.12	0.951
Ours (full)	**36.61**	**0.962**	**32.78**	0.942	**35.36**	**0.955**

$\mathcal{L}_{edge}$ are effective and all of them contribute to improving the quality of reconstructed 3D thermal scenes. Specifically, TIF significantly reduces ambiguous regions, while TADC mitigates insufficient reconstruction, demonstrating that this strategy effectively enhances the model's representational capability for 3D thermal scenes.

V. CONCLUSIONS

In this paper, we present TI-3DGS, a novel approach for 3D thermal scene representation learning from static multi-view thermal images. TI-3DGS simulates a full-chain thermal imaging process by jointly learning emissivity, atmospheric attenuation coefficients, sensor sensitivity, and thermal radiation propagation distance, thereby effectively increasing the detailed representation of the thermal scenes. TI-3DGS integrates a Thermal-aware Density Control (TADC) strategy to adaptively adjust the density of thermal Gaussians, thus mitigating the limitations of the sparse point cloud generated by the traditional Structure-from-Movement algorithm from low-texture thermal images. Due to the blurred edges caused by heat conduction, TI-3DGS also incorporates an Edge Enhancement Loss into the optimization objective to improve the reconstruction accuracy of edges in thermal images. Extensive experimental results demonstrate both the superior performance of our method compared with existing methods and the necessity of considering the full-chain thermal imaging process, low-texture regions within thermal images, as well as heat conduction for 3D thermal reconstruction.

REFERENCES

[1] M. Vollmer and K.-P. Möllmann, *Infrared thermal imaging: fundamentals, research and applications*. John Wiley & Sons, 2018.

[2] A. R. O. Müller, *Close range 3D thermography: real-time reconstruction of high fidelity 3D thermograms*. kassel university press GmbH, 2019, vol. 7.

[3] H. Li, S. Wang, Z. Bai, H. Wang, S. Li, and S. Wen, "Research on 3d reconstruction of binocular vision based on thermal infrared," *Sensors*, vol. 23, no. 17, p. 7372, 2023.

[4] Y. Y. Lin, X.-Y. Pan, S. Fridovich-Keil, and G. Wetzstein, "Thermal-nerf: Thermal radiance fields," in *2024 IEEE International Conference on Computational Photography (ICCP)*. IEEE, 2024, pp. 1–12.

[5] T. Ye, Q. Wu, J. Deng, G. Liu, L. Liu, S. Xia, L. Pang, W. Yu, and L. Pei, "Thermal-nerf: Neural radiance fields from an infrared camera," in *2024 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2024, pp. 1046–1053.

[6] B. Kerbl, G. Kopanas, T. Leimkühler, and G. Drettakis, "3d gaussian splatting for real-time radiance field rendering," *ACM Transactions on Graphics*, vol. 42, no. 4, July 2023. [Online]. Available: https://repo-sam.inria.fr/fungraph/3d-gaussian-splatting/

[7] Q. Chen, S. Shu, and X. Bai, "Thermal3d-gs: Physics-induced 3d gaussians for thermal infrared novel-view synthesis," in *European Conference on Computer Vision*. Springer, 2024, pp. 253–269.

[8] R. Lu, H. Chen, Z. Zhu, Y. Qin, M. Lu, L. Zhang, C. Yan, and A. Xue, "Thermalgaussian: Thermal 3d gaussian splatting," *arXiv preprint arXiv:2409.07200*, 2024.

[9] K. Yang, Y. Liu, Z. Cui, Y. Liu, M. Zhang, S. Yan, and Q. Wang, "Ntr-gaussian: Nighttime dynamic thermal reconstruction with 4d gaussian splatting based on thermodynamics," *arXiv preprint arXiv:2503.03115*, 2025.

[10] J. L. Schönberger and J.-M. Frahm, "Structure-from-Motion Revisited," in *Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016.

[11] D. Zorin, P. Schröder, and W. Sweldens, "Interpolating subdivision for meshes with arbitrary topology," in *Proceedings of the 23rd annual conference on Computer graphics and interactive techniques*, 1996, pp. 189–192.

[12] M. Zwicker, H. Pfister, J. Van Baar, and M. Gross, "Ewa volume splatting," in *Proceedings Visualization, 2001. VIS'01.* IEEE, 2001, pp. 29–538.

[13] S. Zhu, G. Wang, X. Kong, D. Kong, and H. Wang, "3d gaussian splatting in robotics: A survey," *arXiv preprint arXiv:2410.12262*, 2024.

[14] B. Fei, J. Xu, R. Zhang, Q. Zhou, W. Yang, and Y. He, "3d gaussian as a new vision era: A survey," *arXiv e-prints*, pp. arXiv–2402, 2024.

[15] Y. Bao, T. Ding, J. Huo, Y. Liu, Y. Li, W. Li, Y. Gao, and J. Luo, "3d gaussian splatting: Survey, technologies, challenges, and opportunities," *IEEE Transactions on Circuits and Systems for Video Technology*, 2025.

[16] S. Girish, K. Gupta, and A. Shrivastava, "Eagles: Efficient accelerated 3d gaussians with lightweight encodings," in *European Conference on Computer Vision*. Springer, 2024, pp. 54–71.

[17] K. Navaneet, K. P. Meibodi, S. A. Koohpayegani, and H. Pirsiavash, "Compact3d: Compressing gaussian splat radiance field models with vector quantization," *arXiv preprint arXiv:2311.18159*, vol. 4, 2023.

[18] S. Niedermayr, J. Stumpfegger, and R. Westermann, "Compressed 3d gaussian splatting for accelerated novel view synthesis," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2024, pp. 10 349–10 358.

[19] K. Cheng, X. Long, K. Yang, Y. Yao, W. Yin, Y. Ma, W. Wang, and X. Chen, "Gaussianpro: 3d gaussian splatting with progressive propagation," in *Forty-first International Conference on Machine Learning*, 2024.

[20] J. Zhang, F. Zhan, M. Xu, S. Lu, and E. Xing, "Fregs: 3d gaussian splatting with progressive frequency regularization," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2024, pp. 21 424–21 433.

[21] L. Radl, M. Steiner, M. Parger, A. Weinrauch, B. Kerbl, and M. Steinberger, "Stopthepop: Sorted gaussian splatting for view-consistent real-time rendering," *ACM Transactions on Graphics (TOG)*, vol. 43, no. 4, pp. 1–17, 2024.

[22] J. Chung, J. Oh, and K. M. Lee, "Depth-regularized optimization for 3d gaussian splatting in few-shot images," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2024, pp. 811–820.

[23] J. Li, J. Zhang, X. Bai, J. Zheng, X. Ning, J. Zhou, and L. Gu, "Dngaussian: Optimizing sparse-view 3d gaussian radiance fields with global-local depth normalization," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2024, pp. 20 775–20 785.

[24] X. Liu, J. Chen, S.-H. Kao, Y.-W. Tai, and C.-K. Tang, "Deceptive-nerf/3dgs: Diffusion-generated pseudo-observations for high-quality sparse-view reconstruction," in *European Conference on Computer Vision*. Springer, 2024, pp. 337–355.

[25] A. Swann, M. Strong, W. K. Do, G. S. Camps, M. Schwager, and M. Kennedy, "Touch-gs: Visual-tactile supervised 3d gaussian splatting," in *2024 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2024, pp. 10 511–10 518.

[26] Z. Liu, J. Su, G. Cai, Y. Chen, B. Zeng, and Z. Wang, "Georgs: Geometric regularization for real-time novel view synthesis from sparse inputs," *IEEE Transactions on Circuits and Systems for Video Technology*, 2024.

[27] A. Guédon and V. Lepetit, "Sugar: Surface-aligned gaussian splatting for efficient 3d mesh reconstruction and high-quality mesh rendering," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2024, pp. 5354–5363.

[28] H. Chen, C. Li, and G. H. Lee, "Neusg: Neural implicit surface reconstruction with 3d gaussian splatting guidance," *arXiv preprint arXiv:2312.00846*, 2023.

[29] M. Yu, T. Lu, L. Xu, L. Jiang, Y. Xiangli, and B. Dai, "Gsdf: 3dgs meets sdf for improved rendering and reconstruction," *arXiv preprint arXiv:2403.16964*, 2024.

[30] M. Ye, M. Danelljan, F. Yu, and L. Ke, "Gaussian grouping: Segment and edit anything in 3d scenes," in *European Conference on Computer Vision*. Springer, 2024, pp. 162–179.

[31] S. Zhou, H. Chang, S. Jiang, Z. Fan, Z. Zhu, D. Xu, P. Chari, S. You, Z. Wang, and A. Kadambi, "Feature 3dgs: Supercharging 3d gaussian splatting to enable distilled feature fields," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2024, pp. 21 676–21 685.

[32] K. Lan, H. Li, H. Shi, W. Wu, L. Wang, and Y. Liao, "2d-guided 3d gaussian segmentation," in *2024 Asian Conference on Communication and Networks (ASIANComNet)*. IEEE, 2024, pp. 1–5.

[33] B. Dou, T. Zhang, Y. Ma, Z. Wang, and Z. Yuan, "Cosseggaussians: Compact and swift scene segmenting 3d gaussians," *arXiv e-prints*, pp. arXiv–2401, 2024.

[34] J. Rangel, S. Soldan, and A. Kroll, "3d thermal imaging: Fusion of thermography and depth cameras," in *International Conference on Quantitative InfraRed Thermography*, vol. 3, 2014.

[35] S. Zhao, Z. Fang, and S. Wen, "A real-time handheld 3d temperature field reconstruction system," in *2017 IEEE 7th Annual International Conference on CYBER Technology in Automation, Control, and Intelligent Systems (CYBER)*. IEEE, 2017, pp. 289–294.

[36] B. Mildenhall, P. P. Srinivasan, M. Tancik, J. T. Barron, R. Ramamoorthi, and R. Ng, "Nerf: Representing scenes as neural radiance fields for view synthesis," in *European Conference on Computer Vision*. Springer, 2020, pp. 405–421.

[37] M. Poggi, P. Z. Ramirez, F. Tosi, S. Salti, S. Mattoccia, and L. Di Stefano, "Cross-spectral neural radiance fields," in *2022 International Conference on 3D Vision (3DV)*. IEEE, 2022, pp. 606–616.

[38] H. Zhu, Y. Sun, C. Liu, L. Xia, J. Luo, N. Qiao, R. Nevatia, and C.-H. Kuo, "Multimodal neural radiance field," in *2023 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2023, pp. 9393–9399.

[39] M. Hassan, F. Forest, O. Fink, and M. Mielle, "Thermonerf: Joint rgb and thermal novel view synthesis for building facades using multimodal neural radiance fields," *arXiv preprint arXiv:2403.12154*, 2024.

[40] J. Xu, M. Liao, R. P. Kathirvel, and V. M. Patel, "Leveraging thermal modality to enhance reconstruction in low-light conditions," in *European Conference on Computer Vision*. Springer, 2024, pp. 321–339.

[41] Y. Nagase, T. Kushida, K. Tanaka, T. Funatomi, and Y. Mukaigawa, "Shape from thermal radiation: Passive ranging using multi-spectral lwir measurements," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2022, pp. 12 661–12 671.

[42] Z. Zhu, Z. Fan, Y. Jiang, and Z. Wang, "Fsgs: Real-time few-shot view synthesis using gaussian splatting," in *European conference on computer vision*. Springer, 2024, pp. 145–163.

[43] R. Zhang, P. Isola, A. A. Efros, E. Shechtman, and O. Wang, "The unreasonable effectiveness of deep features as a perceptual metric," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018, pp. 586–595.

[44] A. Paszke, "Pytorch: An imperative style, high-performance deep learning library," *arXiv preprint arXiv:1912.01703*, 2019.

[45] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980*, 2014.

[46] T. Müller, A. Evans, C. Schied, and A. Keller, "Instant neural graphics primitives with a multiresolution hash encoding," *ACM transactions on graphics (TOG)*, vol. 41, no. 4, pp. 1–15, 2022.

[47] S. Fridovich-Keil, A. Yu, M. Tancik, Q. Chen, B. Recht, and A. Kanazawa, "Plenoxels: Radiance fields without neural networks," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2022, pp. 5501–5510.