

MSGS: Multispectral 3D Gaussian Splatting

Iris Zheng* Guojun Tang† Alexander Doronin‡ Paul Teal§ Fang-Lue Zhang¶
Victoria University of Wellington

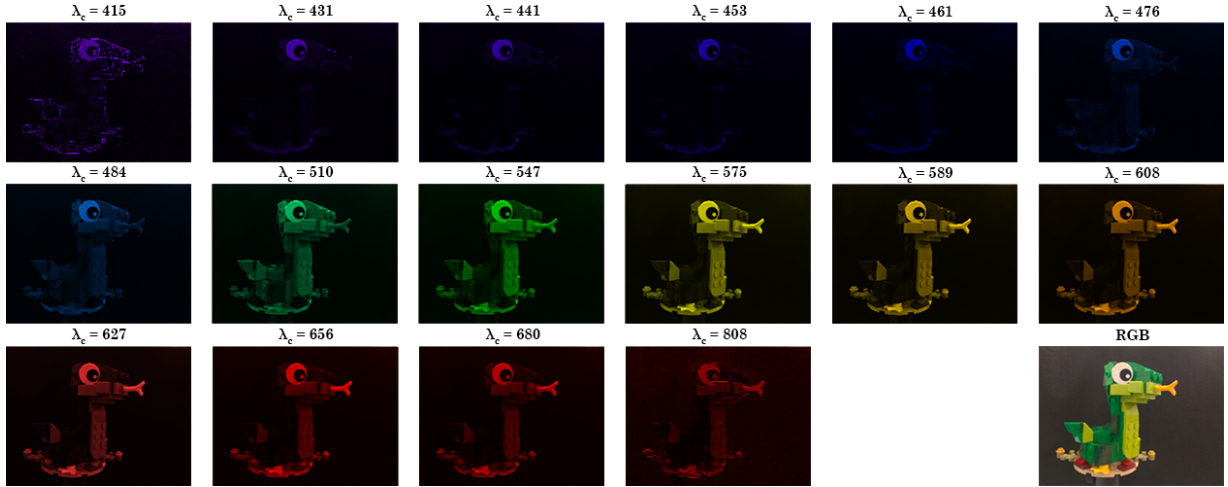


Figure 1: Real-world multispectral (MS) dataset *Snake* captured by our multispectral camera system. λ_c denotes the corresponding wavelength. The images are shown for visualization purposes: the original reflectance values (ranging from 0 to 1) in each MS grayscale image are mapped to RGB space for display. The final RGB image is obtained by converting multiple MS grayscale images into RGB using Equation 2.

ABSTRACT

We present a multispectral extension to 3D Gaussian Splatting (3DGS) for wavelength-aware view synthesis. Each Gaussian is augmented with spectral radiance, represented via per-band spherical harmonics, and optimized under a dual-loss supervision scheme combining RGB and multispectral signals. To improve rendering fidelity, we perform spectral-to-RGB conversion at the pixel level, allowing richer spectral cues to be retained during optimization. Our method is evaluated on both public and self-captured real-world datasets, demonstrating consistent improvements over the RGB-only 3DGS baseline in terms of image quality and spectral consistency. Notably, it excels in challenging scenes involving translucent materials and anisotropic reflections. The proposed approach maintains the compactness and real-time efficiency of 3DGS while laying the foundation for future integration with physically based shading models.

Index Terms: Multispectral Imaging, Gaussian Splatting, Spectral Rendering, Novel View Synthesis.

1 INTRODUCTION

Multispectral information plays a crucial role in a wide range of real-world graphics and vision applications where material properties, lighting conditions, and appearance must be accurately captured and reproduced. Fields such as cultural heritage preservation,

biomedical imaging, and realistic AR/VR rendering rely on spectral fidelity beyond the visible RGB channels to perform wavelength-aware analysis and physically plausible rendering. For example, subtle differences in surface reflectance can reveal material composition, enable non-invasive diagnostics, or support realistic rendering under varied illuminants—all of which demand accurate multispectral representations.

Despite recent advances in neural and point-based rendering, most existing learning-based 3D reconstruction methods—including the widely adopted 3D Gaussian Splatting (3DGS) [6]—are fundamentally limited to RGB inputs. These models are typically optimized for photorealistic novel view synthesis in the visible spectrum but are incapable of modeling wavelength-dependent radiance variations essential for spectral analysis and physically grounded rendering. This restricts their applicability in scenarios where color alone is insufficient to distinguish between visually similar but spectrally distinct materials.

In the original 3DGS pipeline, each Gaussian encodes view-dependent color via low-order spherical harmonics (SH) coefficients across the RGB channels, with color parameters often initialized from structure-from-motion (SfM) pipelines. While this formulation is effective for RGB reconstruction, it lacks the flexibility to represent high-dimensional wavelength-dependent radiance distributions, which are essential for multispectral rendering.

To address this limitation, we propose a wavelength-aware extension of the 3DGS framework that enables rendering from a broad range of spectral dimensions while preserving its compactness and efficiency. Specifically, we introduce a spectral SH-based radiance representation that generalizes the original RGB encoding, allowing each Gaussian to store and render rich spectral information. In addition, we propose a dual-loss supervision strategy that jointly optimizes multispectral and RGB reconstruction, facilitating robust training. We validate our approach on both public and self-captured MS datasets, demonstrating significant improvements in spectral fi-

*iris.zheng@vuw.ac.nz

†guojun.tang@vuw.ac.nz

‡alex.doronin@vuw.ac.nz

§paul.teal@vuw.ac.nz

¶fanglue.zhang@vuw.ac.nz, the corresponding author

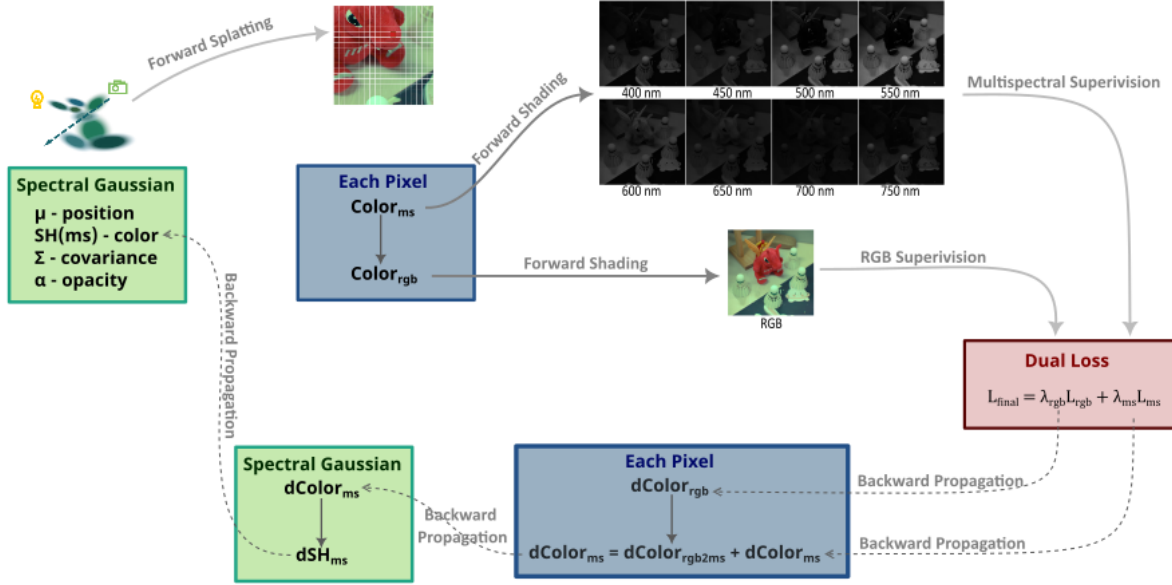


Figure 2: Overview of the proposed Multispectral Gaussian Splatting pipeline. Dual supervision from RGB and multispectral images guides SH parameter updates through both spectral and RGB color spaces.

delity and rendering quality over the RGB-only baselines.

2 RELATED WORK

Multispectral imaging (MSI) has been widely applied in remote sensing, medical diagnostics, and material analysis, enabling radiance capture across wavelengths beyond the visible RGB range. Traditional MSI systems typically rely on CIE-defined spectral-to-RGB pipelines, involving color matching functions and XYZ-to-sRGB transformations [14]. Recent learning-based approaches have explored data-driven solutions for multispectral reconstruction and view synthesis [9, 7, 12, 4].

Neural Radiance Fields (NeRF) [8] have become a dominant framework for novel view synthesis by learning continuous volumetric scene properties via differentiable rendering. Several NeRF variants target the multispectral domain. X-NeRF [9] addresses cross-spectral synthesis from heterogeneous sensors (e.g., RGB, MS, IR) using joint optimization and a Normalized Cross-Device Coordinate (NXDC) system, rather than improving rendering quality itself. In contrast, SpectralNeRF [7] targets physically based spectral rendering by predicting per-wavelength radiance and employing a Spectrum Attention UNet (SAUNet) to fuse outputs into RGB. While flexible, these methods suffer from high computational cost and slow inference due to their reliance on implicit fields and MLP-based rendering.

3D Gaussian Splatting (3DGS) [6] offers a compelling alternative to NeRF, replacing implicit volumetric fields with explicit 3D Gaussians for real-time rendering and improved geometric clarity. Recent efforts have sought to adapt 3DGS to the multispectral setting. SpectralGaussians [12] enhances the original 3DGS by integrating two key components: the Gaussian Shader [5], which augments each Gaussian with additional attributes such as surface normals and physically based reflectance (PBR) parameters; and Gaussian Grouping [17], which clusters Gaussians based on semantic or material similarity to promote structural coherence and reuse. These enhancements enable more expressive and physically plausible appearance modeling. However, the method creates an independent set of Gaussians for each spectral channel, which breaks geometric consistency across wavelengths and significantly increases memory consumption. Meanwhile, SOC-GS [4], addresses cross-

spectral pose misalignment by enforcing spatial occupancy consistency across spectra and adopting a two-stage optimization strategy for camera poses and scene representation. Similar in motivation to X-NeRF, SOC-GS focuses primarily on improving alignment rather than enhancing rendering fidelity. It lacks explicit modeling of wavelength-dependent reflectance or support for physically-based spectral shading, limiting its effectiveness in scenes with complex material properties.

3 METHODOLOGY

We propose a spectral 3D Gaussian Splatting representation where each Gaussian encoded multispectral radiance. We also propose a dual-loss supervision strategy that jointly leverages RGB and multispectral data for high-fidelity scene reconstruction. This design preserves the efficiency and geometric compactness of 3DGS while enabling wavelength-aware rendering. An overview of the framework is shown in Figure 2.

3.1 Spectral-Aware Gaussian

We generalize the radiance representation of each Gaussian from a 3-channel RGB spherical harmonics (SH) [10, 13] coefficients to an N -channel multispectral Spherical-Harmonics (MS-SH) coefficients, where N denotes the number of wavelength bands captured by the imaging system. Specifically, the zeroth-order MS-SH coefficients are initialized from the base colors extracted via COLMAP [11], while higher-order terms are zero-initialized and optimized during training to capture angular variations. Meanwhile, each Gaussian retains a single set of geometric attributes—position, scale, and rotation—shared across all spectral channels. This ensures spectral coherence while avoiding redundant memory usage.

Before the forward rendering stage, we first convert the MS-SH coefficients associated with each Gaussian into its corresponding spectral values, following the approach used in vanilla 3DGS [6]. Based on these spectral values, we perform alpha blending along each camera ray using the standard Gaussian Splatting compositing model, formulated as:

$$C(p) = \sum_{i=1}^N c_i \cdot \alpha_i(p) \cdot \prod_{j=1}^{i-1} (1 - \alpha_j(p)), \quad (1)$$

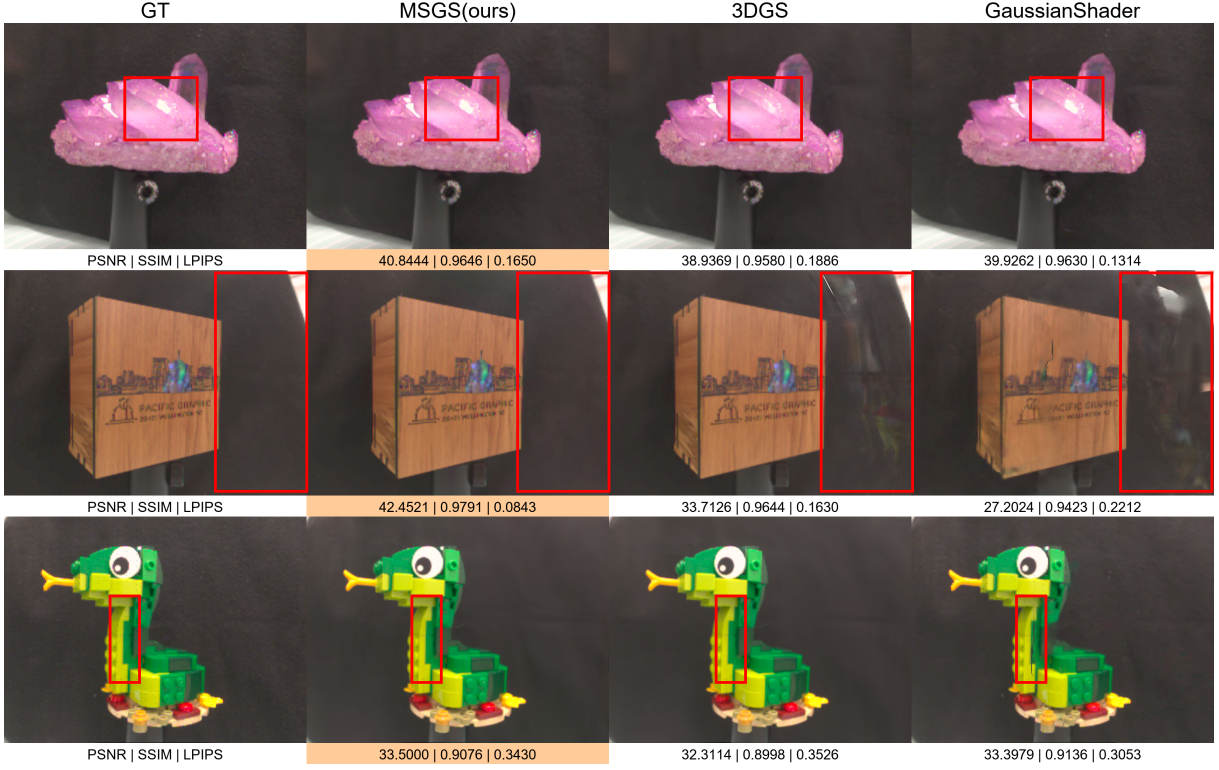


Figure 3: Qualitative Comparison. Qualitative comparison. In the first row, our method better preserves high-frequency reflections compared to 3DGS. In the second row, both 3DGS and Gaussian Shader fails to reconstruct details in dark background regions, while our method maintains structural consistency. In the third row, despite GaussianShader achieving the highest scores on the snake dataset, noticeable artifacts often occur — for example, a long visible crack appears along the snake’s neck.

where overlapping Gaussians contribute to the final pixel value according to their learned opacity and 2D projected footprint. This process produces a per-pixel spectral value map, which serves as the input for the subsequent spectral-to-RGB color conversion stage.

Following the deferred shading concept [2] and its recent adaptation to 3DGS [16], we convert the per-pixel multispectral color into RGB color using CIE 1931 [14] color matching functions [3]. Specifically, the spectral radiance $L(\lambda)$ is converted to tristimulus values (X, Y, Z) as:

$$X = \sum_{\lambda} L(\lambda) \bar{x}(\lambda), \quad Y = \sum_{\lambda} L(\lambda) \bar{y}(\lambda), \quad Z = \sum_{\lambda} L(\lambda) \bar{z}(\lambda) \quad (2)$$

where $\bar{x}(\lambda)$, $\bar{y}(\lambda)$, $\bar{z}(\lambda)$ represent the CIE 1931 standard observer functions sampled at the input wavelengths.

To further ensure perceptual fidelity, we apply D65 white-balance correction, aligning the rendering with daylight illumination. Gamma correction is then applied to map the linear XYZ values to display-friendly RGB values.

3.2 Dual-loss Supervision

To effectively train the proposed multispectral 3D Gaussian Splatting model, we introduce a dual-loss strategy that jointly optimizes both multispectral accuracy and RGB rendering consistency. This is achieved by combining two complementary loss terms: a multispectral reconstruction loss $\mathcal{L}_{\text{MS}}$, computed directly between rendered and ground-truth spectral data, and an RGB reconstruction loss $\mathcal{L}_{\text{RGB}}$, computed after converting the spectral outputs to RGB.

A critical design decision lies in choosing the appropriate stage for spectral-to-RGB conversion, as it directly affects gradient propagation and the fidelity of reconstruction. We explore two strategies: the Gaussian-level conversion, which transforms each Gaussian’s spectral radiance to RGB before rasterization, and the pixel-

level conversion, which defers this conversion until after pixel accumulation. While the former simplifies RGB supervision, it discards spectral details early in the pipeline. In contrast, the latter retains full spectral resolution throughout rendering and supervision, leading to better reconstruction quality.

The final objective combines both loss terms using weighted coefficients:

$$\mathcal{L}_{\text{total}} = \lambda_{\text{MS}} \mathcal{L}_{\text{MS}} + \lambda_{\text{RGB}} \mathcal{L}_{\text{RGB}} \quad (3)$$

where λ_{MS} and λ_{RGB} balance the relative contributions of each loss component.

For a more detailed evaluation of the dual-loss design and conversion strategy, please refer to subsection 4.3.

Table 1: Quantitative comparison with the original 3DGS [6] and PBR-based Gaussian Shader [5]. All three methods are trained for 30K iterations under identical resolution and supervision settings. Our method achieves better scores on most datasets. The best results are colored in red.

Method \ Dataset	Projector	Dragon	Onion	Mushroom	Snake	Crystal	Box
PSNR ↑							
3DGS	36.7580	38.1470	37.8868	34.5161	33.0125	37.4461	37.9612
GaussianShader	30.0921	34.4725	35.8232	32.8344	34.3955	37.6846	34.1192
Ours	36.8304	38.5589	38.3313	33.9437	33.8852	37.6949	39.9916
SSIM ↑							
3DGS	0.9607	0.9715	0.9254	0.9139	0.9042	0.9571	0.9615
GaussianShader	0.9210	0.9550	0.9308	0.9098	0.9221	0.9576	0.9508
Ours	0.9619	0.9717	0.9293	0.9233	0.9121	0.9634	0.9686
LPIPS ↓							
3DGS	0.0967	0.0745	0.1634	0.2214	0.3478	0.1852	0.1538
GaussianShader	0.1911	0.1149	0.1515	0.2582	0.2818	0.1642	0.1911
Ours	0.0915	0.0747	0.1491	0.2093	0.3392	0.1628	0.1302

4 RESULTS

We evaluate our method on both public and self-captured multi-spectral data to validate its effectiveness in spectral-aware NVS and RGB rendering quality. Results are reported using both quantitative

metrics and qualitative comparisons. Additionally, we conduct ablation studies to analyze the impact of our dual-loss supervision strategy and pixel-level spectral-to-RGB conversion design.

4.1 Datasets

Our experiments are conducted on a combination of public real-world multispectral datasets and self-captured data. Specifically, we use two real-world scenes from the SpectralNeRF [7] dataset—*Dragon* and *Project*—captured with RGB cameras and rotating filter wheels. Each scene contains images sequentially captured across eight narrow spectral bands spanning 400–750 nm.

To further evaluate generalizability across diverse material properties and geometries, we capture five additional scenes—*Onion*, *Mushroom*, *Snake*, *Crystal*, and *Box*—using a multispectral camera system (S1-EVK2b [1]) that directly acquires spectral images across 16 bands in the 415–808 nm range, as illustrated in Figure 1. These scenes were carefully selected to cover a wide range of material and appearance characteristics. *Onion* and *Mushroom* represent smooth, homogeneous diffuse surfaces with low-frequency texture, ideal for evaluating color consistency and shading. *Snake* is a colorful LEGO object composed of high-frequency patterns and synthetic materials. *Crystal* features strong translucency and partial transparency, challenging both geometry and appearance reconstruction. *Box* is a wooden container made from abalone shell, with complex, view-dependent reflectance due to its iridescent surface. All datasets provide per-pixel spectral measurements along with corresponding RGB images, enabling both spectral and RGB supervision for joint evaluation.

4.2 Quantitative and Qualitative Analysis

Although recent works like X-NeRF [9], SpectralNeRF [7], SpectralGaussians [12], and SOC-GS [4] have explored multispectral view synthesis, none provide official implementations for direct comparison. We therefore compare against the original 3DGS baseline [6] and PBR-based Gaussian Shader [5], focusing on RGB reconstruction accuracy. Our method consistently outperforms 3DGS in most scenes (see Table 1).

Qualitative comparisons are provided for both RGB and multispectral outputs: for RGB, we visually compare our results with those from 3DGS and Gaussian Shader, as presented in Figure 3; for multispectral outputs, we showcase rendered spectral images produced by our model, as 3DGS cannot represent or synthesize spectral data. Notably, *Crystal* presents a challenging case due to its translucent material and complex optical effects, while *Box* contains an anisotropic reflective icon that highlights view-dependent appearance variations. Our method demonstrates improved fidelity and robustness in both cases.

Table 2: Comparison between conversion strategies at Gaussian and Pixel Level across 7K and 30K iterations. The results are evaluated on our self-captured *crystal* dataset. The best/second-best results are colored in red / orange, respectively.

Method	Iteration	7000				30000			
		Time ↓	PSNR ↑	SSIM ↑	LPIPS ↓	Time ↓	PSNR ↑	SSIM ↑	LPIPS ↓
Conversion at Gaussian Level	RGB	04:40	34.0308	0.9505	0.2074	22:08	37.2594	0.9572	0.1866
	MS	06:39	29.1552	0.9352	0.2295	29:04	29.3512	0.9387	0.2200
	MS & RGB	09:47	29.0900	0.9374	0.2276	54:01	26.9797	0.9382	0.2315
Conversion at Pixel Level	RGB	04:23	33.9366	0.9505	0.2072	18:54	36.9047	0.9569	0.1877
	MS	07:05	34.0480	0.9472	0.2145	30:40	36.6371	0.9528	0.1984
	MS & RGB	09:00	34.4375	0.9543	0.1936	41:56	38.6076	0.9624	0.1662

4.3 Ablation Study

We perform ablation studies on three key components: (1) the dual-loss supervision strategy, (2) the level of spectral-to-RGB conversion, and (3) the choice of spectral band range. See Table 2 and Table 3 for detailed results.

1. **Loss Supervision Strategy.** We evaluate three loss settings: RGB-only loss, MS-only loss, and full dual-loss. The dual-loss strategy consistently achieves the best performance across different scenes and iteration checkpoints (7K and 30K), providing

a balanced trade-off between perceptual quality and spectral accuracy. RGB-only supervision lacks spectral fidelity, while MS-only training, though spectrally accurate, can result in perceptual color distortions due to the absence of visual priors.

2. **Spectral-to-RGB Conversion Strategy.** We further compare two different strategies for spectral-to-RGB conversion: one at the Gaussian level (before rasterization), and the other at the pixel level (after rasterization). Results show that pixel-level conversion consistently yields better performance. Interestingly, when applying dual-loss supervision with Gaussian-level conversion, the performance at 30K iterations is even worse than at 7K, and lower than using either loss alone. This suggests that converting spectral data to RGB before rasterization—an inherently nonlinear projection—may introduce conflicting gradient directions during back-propagation, especially after blending in the rasterization step. Such conflicts could hinder optimization and reduce convergence stability.

3. **Spectral Band Selection.** To investigate the difference between visible and near-infrared performance, we conduct experiments on four band configurations using our self-captured *crystal* dataset. The 808 nm band, which lies in the near-infrared region beyond the visible spectrum, provides valuable structural information and contributes positively to reconstruction quality. Interestingly, however, the full spectral range from 415 to 808 nm does not achieve the best overall performance. In fact, the 415–680 nm configuration yields the lowest scores across all metrics, suggesting that the 415 nm band—despite being within the visible range—may introduce excessive noise due to sensor limitations, thus degrading the results in both cases where it is included. Among the remaining settings, the 431–808 nm range achieves the highest PSNR, while the 431–680 nm range performs better in terms of SSIM and LPIPS. This indicates a trade-off: near-infrared bands improve structural accuracy, whereas visible-only bands offer better perceptual quality, depending on the rendering objective.

Table 3: Comparison of different spectral band selections using our self-captured *crystal* dataset across 7K and 30K iterations. The best/second-best results are colored in red / orange, respectively.

Band Range	Iteration	7K Iterations				30K Iterations			
		Time ↓	PSNR ↑	SSIM ↑	LPIPS ↓	Time ↓	PSNR ↑	SSIM ↑	LPIPS ↓
431–680 nm		13:45	34.8798	0.9542	0.1949	1:07:07	38.1892	0.9610	0.1739
415–680 nm		14:51	34.5260	0.9522	0.1986	1:12:28	37.5136	0.9588	0.1778
431–808 nm		14:58	35.2441	0.9528	0.1993	1:12:22	38.6689	0.9598	0.1765
415–808 nm		15:01	34.8456	0.9507	0.2037	1:10:41	38.5316	0.9586	0.1797

5 CONCLUSIONS AND FUTURE WORK

We proposed a multispectral extension to 3D Gaussian Splatting (3DGS) that enables wavelength-aware view synthesis while maintaining the efficiency and geometric compactness of the original framework. By augmenting the color representation with spectral SH-based radiance and introducing a dual-loss supervision strategy, our method incorporates multispectral information without duplicating geometry and achieves superior rendering performance across both RGB and spectral domains. Experimental results highlight the benefits of leveraging richer spectral data, especially in scenes with complex materials or lighting conditions.

For future work, we aim to improve the physical accuracy of spectral radiance representation to achieve higher visual fidelity in NVS. While the current SH-based model [10, 13] is efficient, it remains limited in capturing anisotropic reflectance and fine angular details. To address this, future extensions may adopt more expressive basis functions such as ASG [15], or incorporate physically grounded models like BRDFs and BSRDFs for better handling of complex light transport. Additionally, learning adaptive weights for multispectral channels and losses could further improve optimization and generalization.

ACKNOWLEDGMENTS

This work was supported by Marsden Fund Council managed by the Royal Society of New Zealand under Grant MFP-20-VUW-180. The authors would like to thank Xuan Zhu and Simin Kou for their helpful discussions and valuable suggestions during the development of this work.

REFERENCES

- [1] S1-EVK2b Evaluation Kit. 4
- [2] M. Deering, S. Winner, B. Schemiwy, C. Duffy, and N. Hunt. The triangle processor and normal vector shader: a VLSI system for high performance graphics. *ACM SIGGRAPH Computer Graphics*, 22(4):21–30, Aug. 1988. Publisher: Association for Computing Machinery (ACM). doi: 10.1145/378456.378468 3
- [3] K. Devlin, A. Chalmers, A. Wilkie, and W. Purgathofer. Tone Reproduction and Physically Based Spectral Rendering. 2002. Publisher: Eurographics Association. 3
- [4] H. Guo, H. Liu, J. Wen, and J. Li. Cross-Spectral Gaussian Splatting with Spatial Occupancy Consistency. 2, 4
- [5] Y. Jiang, J. Tu, Y. Liu, X. Gao, X. Long, W. Wang, and Y. Ma. GaussianShader: 3D Gaussian Splatting with Shading Functions for Reflective Surfaces. In *2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 5322–5332, Nov. 2023. ISSN: 2575-7075. doi: 10.1109/CVPR52733.2024.00509 2, 3, 4
- [6] B. Kerbl, G. Kopanas, T. Leimkuehler, and G. Drettakis. 3D Gaussian Splatting for Real-Time Radiance Field Rendering. *ACM transactions on graphics*, 42(4):1–14, 2023. Place: New York, NY, USA Publisher: ACM. doi: 10.1145/3592433 1, 2, 3, 4
- [7] R. Li, J. Liu, G. Liu, S. Zhang, B. Zeng, and S. Liu. Spectral-NeRF: Physically Based Spectral Rendering with Neural Radiance Field. *Proceedings of the AAAI Conference on Artificial Intelligence*, 38(4):3154–3162, Mar. 2024. Number: 4. doi: 10.1609/aaai.v38i4.28099 2, 4
- [8] 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 A. Vedaldi, H. Bischof, T. Brox, and J.-M. Frahm, eds., *Computer Vision – ECCV 2020*, pp. 405–421. Springer International Publishing, Cham, 2020. doi: 10.1007/978-3-030-58452-8_24 2
- [9] M. Poggi, P. Z. Ramirez, F. Tosi, S. Salti, S. Mattoccia, and L. D. Stefano. Cross-Spectral Neural Radiance Fields. pp. 606–616. IEEE Computer Society, Sept. 2022. doi: 10.1109/3DV57658.2022.00071 2, 4
- [10] R. Ramamoorthi and P. Hanrahan. An Efficient Representation for Irradiance Environment Maps. In *Proceedings of the 28th annual conference on Computer graphics and interactive techniques*, pp. 497–500. ACM, Aug. 2001. doi: 10.1145/383259.383317 2, 4
- [11] J. L. Schönberger and J.-M. Frahm. Structure-from-Motion Revisited. pp. 4104–4113. IEEE, 2016. Book Title: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) ISSN: 1063-6919. doi: 10.1109/CVPR.2016.445 2
- [12] S. N. Sinha, H. Graf, and M. Weinmann. SpectralGaussians: Semantic, spectral 3D Gaussian splatting for multi-spectral scene representation, visualization and analysis, Aug. 2024. arXiv:2408.06975. doi: 10.48550/arXiv.2408.06975 2, 4
- [13] P.-P. Sloan, J. Kautz, and J. Snyder. Precomputed Radiance Transfer for Real-Time Rendering in Dynamic, Low-Frequency Lighting Environments. *ACM Trans. Graph.*, 21(3):527–536, July 2002. doi: 10.1145/566654.566612 2, 4
- [14] G. Wyszecki and W. S. Stiles. *Color Science: Concepts and Methods, Quantitative Data and Formulae*. John Wiley & Sons, Aug. 2000. Google-Books-ID: dOKdEAAAQBAJ. 2, 3
- [15] K. Xu, W.-L. Sun, Z. Dong, D.-Y. Zhao, R.-D. Wu, and S.-M. Hu. Anisotropic Spherical Gaussians. *ACM Trans. Graph.*, 32(6):209:1–209:11, Nov. 2013. doi: 10.1145/2508363.2508386 4
- [16] K. Ye, Q. Hou, and K. Zhou. 3D Gaussian Splatting with Deferred Reflection. In *Special Interest Group on Computer Graphics and Interactive Techniques Conference Conference Papers '24*, pp. 1–10. ACM, Denver CO USA, July 2024. doi: 10.1145/3641519.3657456 3
- [17] M. Ye, M. Danelljan, F. Yu, and L. Ke. Gaussian Grouping: Segment and Edit Anything in 3D Scenes. In *Computer Vision – ECCV 2024: 18th European Conference, Milan, Italy, September 29–October 4, 2024, Proceedings, Part XXIX*, pp. 162–179. Springer-Verlag, Berlin, Heidelberg, Dec. 2023. doi: 10.1007/978-3-031-73397-0_10 2