LumiState: Retinex-inspired Illumination Decoupling in 4D Gaussian Splatting for Endoscopic Reconstruction

Dai Sun[1⋆], Huhao Guan[1⋆], Kun Zhang[1], Yanzhe Liu[2], Xike Xie[1✉], and S. Kevin Zhou[1✉]

[1] School of Biomedical Engineering, Division of Life Sciences and Medicine, University of Science and Technology of China (USTC), Hefei Anhui, 230026, China
[2] Medical School of Chinese People's Liberation Army (PLA), Faculty of Hepato Biliary-Pancreatic Surgery, Chinese PLA General Hospital, Beijing
sundai@ustc.edu {guanhuhao, kkzhang}@mail.ustc.edu.cn liuyanzhe31@163.com
xkxie@ustc.edu.cn s.kevin.zhou@gmail.com

Abstract. Reconstructing deformable 3D anatomy from endoscopic video is essential for surgical navigation and computer-assisted diagnosis, yet remains challenging under the extreme illumination variations inherent to endoscopic imaging. Co-axial near-field illumination, specular reflections on wet tissue, and automatic exposure changes violate the photometric consistency assumption underlying 3D/4D Gaussian Splatting (3DGS/4DGS), causing per-frame illumination variations to be entangled with Gaussian attributes, which degrades both geometry and rendered appearance. Existing approaches either assume distant illumination incompatible with the endoscopic near-field setting, or apply learned appearance corrections without physically grounded decomposition. We propose **LumiState**, which integrates Retinex-inspired reflectance illumination decomposition into 4D Gaussian Splatting for dynamic endoscopic reconstruction. LumiState decomposes each frame into a view-independent tissue reflectance image and a per-frame illumination map at the per-Gaussian level, ensuring that base Gaussians encode reflectance alone. A two-stage training strategy and bright-channel prior further regularize the inherently ill-posed decomposition. Experiments on the RLLS and EndoNeRF-EC benchmarks demonstrate state-of-the-art novel-view synthesis quality, with the largest gains under severe illumination variation, while enabling controllable illumination editing and temporally consistent reflectance rendering. Our code will be available here[3].

Keywords: Endoscopic Reconstruction · 4D Gaussian Splatting · Illumination Decomposition · Retinex Theory.

1 Introduction

Reconstructing deformable 3D anatomy from endoscopic video is a longstanding research direction with direct clinical relevance to surgical navigation, augmented

[3] https://github.com/oDaiSuno/LumiState

Table 1: **Comparison of illumination-aware reconstruction paradigms.**

Method	**Explicit decomposition**	**Dynamic (4D)**	**Near-field illumination**
PBR-based [10,5,6]	✓	×	×[†]
Learned correction [4,8]	×	✓[‡]	×
LumiState (Ours)	✓	✓	✓

[†] PR-ENDO [6] models co-located illumination but does not support dynamic deformation. [‡] Endo-4DGX [4] supports 4D; WildGaussians [8] is static only.

reality overlay, and robot-assisted intervention [12,14,16]. Neural Radiance Field (NeRF) [13] and its temporal extensions have demonstrated the feasibility of modeling dynamic surgical scenes from monocular sequences [17,21,24]. More recently, 3D Gaussian Splatting (3DGS) [7] and its dynamic variant 4DGS [19] have been adapted to endoscopic reconstruction, achieving notable quality under relatively stable illumination [22,26,2,11,20].

However, these methods fundamentally rely on *photometric consistency* [15,3], the assumption that the same surface point exhibits a stable appearance across frames. Endoscopic imaging intrinsically violates this assumption due to co-axial near-field illumination, specular reflections on wet tissue, automatic exposure variations, and instrument-induced shadowing. Under such inconsistent supervision, 4DGS absorbs illumination variations into Gaussian attributes, producing floating artifacts around highlights and structural degradation in shadows.

Recent works have explored two paradigms for illumination-aware reconstruction (Tab. 1). *Physically-based rendering (PBR)* methods [10,5,6] decompose appearance via explicit bidirectional reflectance distribution function (BRDF) estimation, achieving material-illumination separation; however, they either assume distant environment-map illumination incompatible with endoscopic near-field illumination [10,5], or operate on static scenes without temporal deformation support [6]. *Learned appearance correction* methods [4,8,23] adjust Gaussian colors through per-frame affine or nonlinear tone mapping within dynamic frameworks, yet these learned parameters lack physical correspondence to reflectance or illumination, leaving illumination entangled in base Gaussian attributes.

We propose **LumiState**, which addresses these limitations by integrating a Retinex-inspired reflectance-illumination decomposition into 4D Gaussian Splatting, modeling each frame as $\mathbf{I}_t = \mathbf{R}_t \odot \mathbf{L}_t$ where $\mathbf{R}_t$ encodes a view-independent tissue reflectance image and $\mathbf{L}_t$ captures a per-frame near-field illumination map. Rather than correcting appearance in a physically ungrounded manner, LumiState modulates per-Gaussian SH coefficients conditioned on a compact illumination embedding, structurally constraining base Gaussian attributes to encode reflectance alone. A two-stage training strategy and bright-channel prior further stabilize the inherently ill-posed decomposition. Our main contributions are:

- We introduce Retinex-inspired reflectance-illumination decomposition into 4D Gaussian Splatting, addressing photometric inconsistency under extreme endoscopic near-field illumination.
- We realize this decomposition via spatially adaptive SH coefficient modulation driven by a compact illumination embedding, combined with two-stage training and bright-channel prior to prevent degenerate solutions.
- Experiments on two endoscopic benchmarks demonstrate state-of-the-art reconstruction quality with the largest gains under severe illumination variation, while enabling controllable illumination editing and temporally consistent reflectance rendering.

2 Methodology

LumiState decomposes the observed endoscopic image into view-independent tissue reflectance image and per-frame illumination map. As illustrated in Fig. 1, it consists of two components: an *Illumination Estimation Network* that recovers the per-frame illumination map from the observed image, and an *Illumination-Disentangled Reflectance Rendering* module that propagates the estimated illumination to each Gaussian via spatially adaptive SH coefficient modulation.

2.1 Retinex-inspired Reflectance-Illumination Decomposition

We extend Retinex theory [9] to sequential reconstruction: for a tissue surface observed across an endoscopic video, the reflectance image R_t is an intrinsic material property which is view-independent and consistent across frames, whereas the illumination map L_t varies with the co-axial light source position as the endoscope moves. The observed image at timestep t thus decomposes as:

$$\mathbf{I}_t = \mathbf{R}_t \odot \mathbf{L}_t, \tag{1}$$

where $\mathbf{R}_t$ is the reflectance image rendered from deformed Gaussians and $\mathbf{L}_t \in [0,1]^{H\times W}$ is the per-frame illumination map.

This formulation explicitly decouples intrinsic reflectance from frame-varying illumination. In contrast, affine appearance correction [8,4] introduces learnable scaling and shifting parameters without scene-property correspondence, merely compensating appearance variations in a latent manner.

Illumination Estimation Network. To recover spatially smooth near-field illumination image $\mathbf{L}_t$ from observed image while avoiding interference from reflectance-dependent color variations, we design a lightweight Illumination Estimation Network operating on a single-channel brightness projection. Specifically, we compute

$$\mathbf{S}_t = \max(R_t^{\text{ch}}, G_t^{\text{ch}}, B_t^{\text{ch}}), \tag{2}$$

which preserves the illumination spatial structure while suppressing reflectance dependent color variations, and regress the per-frame illumination map using a lightweight UNet $\mathcal{U}_\phi$.

$$\mathbf{L}_t = \sigma\big(\mathcal{U}_\phi(\mathbf{S}_t)\big) \in [0,1]^{H\times W}, \tag{3}$$

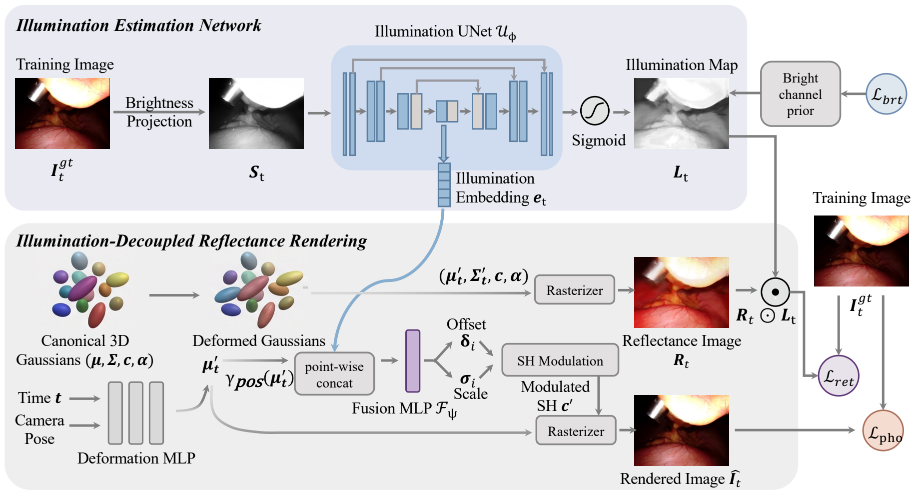

Fig. 1: **Overview of LumiState.** *Top:* An Illumination Estimation Network predicts a spatially smooth illumination map $\mathbf{L}_t$ and a compact embedding $\mathbf{e}_t$ from the input frame. *Bottom:* Per-Gaussian SH modulation decouples reflectance $\mathbf{R}_t$ from illumination, enforcing $\mathbf{R}_t \odot \mathbf{L}_t \approx \mathbf{I}_t^{gt}$.

where σ denotes sigmoid activation. The UNet bottleneck feature serves as a compact illumination embedding $\mathbf{e}_t$, encoding the global illumination state for Gaussian-level modulation below.

Illumination-Decoupled Reflectance Rendering. To recover reflectance image $\mathbf{R}_t$ free of illumination effects at the per-Gaussian level, naive image-space division by $\mathbf{L}_t$ is ill-posed, as it discards multi-view geometric constraints. To address this issue, we introduce an illumination-decoupled reflectance rendering formulation that transfers the estimated 2D illumination field into Gaussian representations. Specifically, each Gaussian is modulated conditioned on both the compact illumination embedding $\mathbf{e}_t$ and its deformed position $\boldsymbol{\mu}_t'$ encoded by sinusoidal positional encoding $\gamma_{\text{pos}}(\cdot)$ [13]. The concatenated feature is fed into a 4-layer fusion MLP $\mathcal{F}_\psi$ to predict a multiplicative scale $\boldsymbol{\sigma}_i$ and an additive offset $\boldsymbol{\delta}_i \in \mathbb{R}^{3C}$:

$$\mathbf{c}_i' = \mathbf{c}_i \odot \exp(\boldsymbol{\sigma}_i) + \frac{\boldsymbol{\delta}_i}{Y_0}, \tag{4}$$

where $Y_0 = 1/\sqrt{4\pi}$. The multiplicative term $\exp(\boldsymbol{\sigma}_i)$ enforces Retinex-consistent illumination scaling, while the additive term $\boldsymbol{\delta}_i$ compensates residual effects beyond pure multiplicative interaction, such as ambient scattering and subsurface contributions. To model non-uniform near-field attenuation and prevent global color shifts, the positional encoding $\gamma_{\text{pos}}(\boldsymbol{\mu}_i)$ enables spatially varying modulation across Gaussians. During rasterization, the modulated coefficients $\mathbf{c}_i'$ are used to render the final image $\hat{\mathbf{I}}_t$, which is directly supervised against the observed image to faithfully reproduce real illumination effects, while the original coefficients $\mathbf{c}_i$ render the reflectance image $\mathbf{R}_t$.

Temporal Illumination Completion. Direct per-frame inference of the illumination estimation network at test time introduces unnecessary computational overhead and temporal instability. To exploit this temporal coherence, we cache the illumination embeddings of training frames and recover the test-time illumination state through linear interpolation:

$$\mathbf{e}_{\text{test}} = (1 - w)\, \mathbf{e}_{t_-} + w\, \mathbf{e}_{t_+}, \tag{5}$$

where $t_\pm$ denote the nearest training timestamps and $w = (t_{\text{test}} - t_-)/(t_+ - t_-)$. This strategy preserves smooth illumination transitions while eliminating the need for per-frame UNet inference during rendering.

2.2 Training Strategy and Loss Functions

The reflectance-illumination decomposition $\mathbf{I}_t = \mathbf{R}_t \odot \mathbf{L}_t$ is inherently ill-posed. Directly optimizing both components often leads to degenerate solutions, where illumination absorbs geometric errors or reflectance becomes contaminated by illumination variations. To stabilize the factorization, we adopt a two-stage training strategy together with Retinex-specific regularization terms that progressively constrain reflectance and illumination during optimization.

Two-stage training. To progressively constrain product ambiguity, we adopt a two-stage training strategy. In *Stage 1*, illumination is not explicitly predicted; only the illumination embedding $\mathbf{e}_t$ modulates Gaussian appearance under photometric supervision, allowing geometry and reflectance to first converge to a stable solution. In *Stage 2*, the full illumination estimation network is activated to produce $\mathbf{L}_t$, and Retinex-specific losses are introduced to explicitly enforce reflectance-illumination consistency.

Loss functions. To supervise the rendered appearance, we employ a photometric loss throughout training that combines L1 distance and D-SSIM [7]:

$$\mathcal{L}_{\text{pho}} = (1 - \lambda_s)\|\hat{\mathbf{I}}_t - \mathbf{I}_t^{\text{gt}}\|_1 + \lambda_s\big(1 - \text{SSIM}(\hat{\mathbf{I}}_t, \mathbf{I}_t^{\text{gt}})\big), \tag{6}$$

with $\lambda_s = 0.2$.

While photometric supervision ensures visual fidelity, it alone cannot resolve the ambiguity of reflectance-illumination factorization. We therefore introduce two Retinex-specific constraints in Stage 2 to explicitly stabilize the decomposition. The Retinex reconstruction consistency loss enforces the forward image formation model, requiring the product of reflectance and illumination to recover the observation:

$$\mathcal{L}_{\text{ret}} = \|\mathbf{R}_t \odot \mathbf{L}_t - \mathbf{I}_t^{\text{gt}}\|_1. \tag{7}$$

However, consistency alone allows trivial solutions where illumination collapses or absorbs most intensity variation. To regularize the illumination field, we adopt a bright-channel prior [1], anchoring $\mathbf{L}_t$ to the local maximum brightness $\mathbf{B}_t = \text{MaxPool}_{2r_0+1}(\mathbf{S}_t)$ with $r_0 = 2$:

$$\mathcal{L}_{\text{brt}} = \|\mathbf{L}_t - \mathbf{B}_t\|_2^2. \tag{8}$$

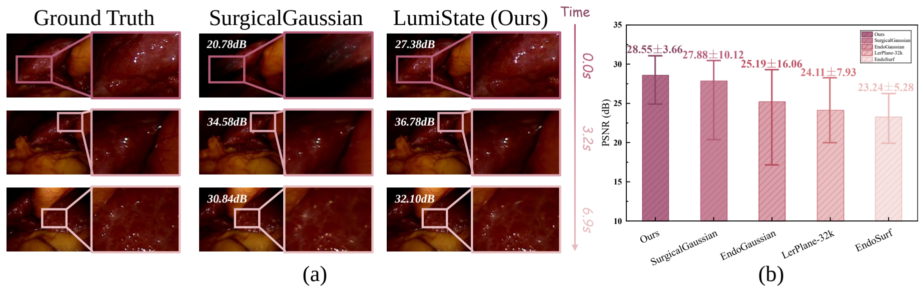

Fig. 2: **Robustness Evaluation on the RLLS Dataset** [25]. **(a)** Visual comparison showing LumiState's superior reconstruction with enhanced details and reduced artifacts over SurgicalGaussian [20]. **(b)** PSNR comparison across patient sequences demonstrates LumiState's consistent performance gains.

Table 2: **Quantitative results on RLLS** [25]. PSNR↑/SSIM↑/LPIPS↓ under low (LIVar) and high (HIVar) illumination variance. The **best** and *second-best* results are highlighted.

	LIVar (*454 fr.*)			**HIVar** (*177 fr.*)			**All** (*631 fr.*)		
Method	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓
EndoSurf† [24]	30.13	0.901	0.186	24.30	0.862	0.255	27.21	0.881	0.221
LerPlane-32k [21]	28.13	0.846	0.218	25.12	0.835	0.205	26.92	0.841	0.213
Endo-4DGX [4]	23.15	0.778	0.278	24.58	0.822	0.213	23.72	0.795	0.252
EndoGaussian [11]	32.20	0.922	0.120	27.78	0.900	0.153	30.43	0.913	0.133
SurgicalGaussian [20]	*33.88*	*0.937*	*0.081*	*29.67*	*0.921*	*0.119*	*32.20*	*0.930*	*0.096*
LumiState (Ours)	**34.83**	**0.943**	**0.071**	**30.58**	**0.933**	**0.091**	**33.13**	**0.939**	**0.079**

Since normalized reflectance satisfies $R(\mathbf{x}) \leq 1$, the bright channel provides an upper bound for illumination, preventing degenerate collapse.

The full objective is:

$$\mathcal{L} = \mathcal{L}_{\text{pho}} + \lambda_{\text{ret}}\mathcal{L}_{\text{ret}} + \lambda_{\text{brt}}\mathcal{L}_{\text{brt}}, \tag{9}$$

with $\lambda_{\text{ret}} = 0.05$ and $\lambda_{\text{brt}} = 0.25$. Only $\mathcal{L}_{\text{pho}}$ is applied in Stage 1.

3 Experiments

Experimental setup. We use two benchmarks. RLLS [25] contains five laparoscopic sequences with frequent camera motion and pronounced illumination fluctuations, split into low illumination variance (LIVar, Patient01-03) and high illumination variance (HIVar, Patient04-05). EndoNeRF-EC [18] provides two robotic surgery scenes (Pulling / Cutting) with tissue deformation under varying illumination. For reconstruction, we compare against EndoSurf [24], LerPlane [21], Endo-4DGX [4], EndoGaussian [11], and SurgicalGaussian [20]; for

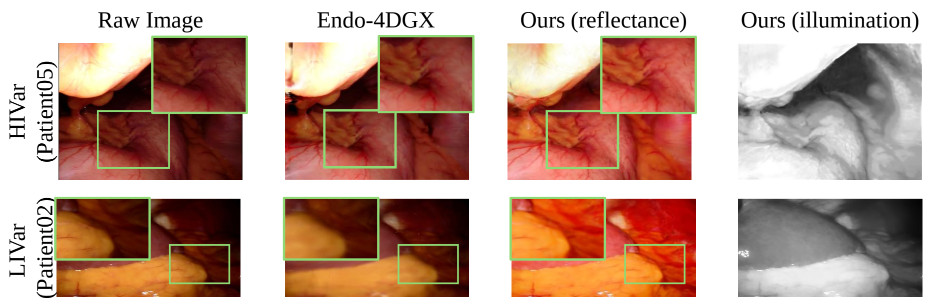

Fig. 3: **Illumination correction on RLLS** [25]. Raw input, Endo-4DGX [4], and our reflectance / illumination on HIVar and LIVar.

Table 3: **No-reference (NR) image quality results on RLLS and EndoNeRF-EC.** The **best** and *second-best* results are highlighted.

Method	RLLS						EndoNeRF-EC					
	LIVar (*454 frames*)			**HIVar** (*177 frames*)			**Pulling**			**Cutting**		
	NIQE↓	BRISQUE↓	LIQE↑	NIQE↓	BRISQUE↓	LIQE↑	NIQE↓	BRISQUE↓	LIQE↑	NIQE↓	BRISQUE↓	LIQE↑
WildGaussians [8]	9.553	76.433	1.068	9.610	80.710	1.089	-	-	-	-	-	-
Endo-4DGX [4]	*7.053*	**53.825**	*1.370*	*7.523*	*56.816*	*1.422*	*6.246*	*41.774*	**1.763**	*6.032*	*42.668*	**2.454**
LumiState (Ours)	**6.911**	*54.433*	**2.529**	**7.007**	**50.639**	**2.601**	**6.043**	**38.502**	*1.392*	**5.767**	**39.924**	*2.320*

illumination correction, against Endo-4DGX and WildGaussians [8]. All experiments run on a single RTX 4090. We report PSNR↑, SSIM↑, LPIPS↓ for reconstruction; NIQE↓, BRISQUE↓, LIQE↑ for illumination correction (no ground-truth illumination available); and R-Var↓, the mean per-pixel temporal standard deviation of reflectance across test timesteps, to quantify illumination leakage into base SH coefficients.

Reconstruction quality. Tab. 2 shows that LumiState achieves the best average metrics on RLLS [25], with consistent gains on both the LIVar subset (+0.95 dB over SurgicalGaussian [20]) and the more challenging HIVar subset (+0.91 dB), demonstrating that explicit decomposition maintains its advantage even under severe photometric inconsistency. Compared with Endo-4DGX [4], LumiState outperforms across all patients, confirming that decomposition-based handling yields better fidelity than affine appearance correction. Fig. 2 shows fewer floating artifacts in highlights and finer detail in shadows.

Illumination correction. Tab. 3 evaluates illumination correction using no-reference metrics. Our method yields competitive or superior perceptual quality across both datasets, with the most consistent improvement on the HIVar subset where illumination variation is most severe. Fig. 3 visualizes the decomposed outputs: our reflectance $\mathbf{R}_t$ exhibits more uniform exposure than Endo-4DGX [4], while the estimated $\mathbf{L}_t$ captures the expected low-frequency radial falloff pattern of near-field illumination.

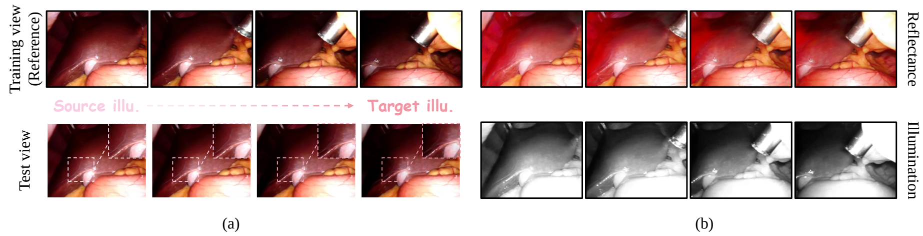

Fig. 4: **Decomposition quality. (a)** Test view rendered under illumination transferred from training frames with progressive variation. **(b)** Reflectance and illumination across different illumination conditions.

Table 4: **Ablation study on RLLS (All Avg.).** R-Var ($\times 10^{-2}$)↓ measures illumination leakage into base SH coefficients (cf. §3). The **best** and *second-best* results are highlighted.

Variant	PSNR↑	LPIPS↓	R-Var↓
(a) **Full model**	**34.83**	**0.070**	**1.08**
(b) → Affine appearance correction ($\boldsymbol{\beta} \odot \mathbf{c} + \boldsymbol{\gamma}$)	34.35	0.079	3.42
(c) → Learnable embedding	*34.58*	*0.075*	2.31
(d) → w/o two-stage training	33.64	0.092	2.94
(e) → w/o $\mathcal{L}_{\text{brt}}$	34.19	0.083	*1.89*

Efficiency. LumiState trains in approximately 50 minutes (vs. 42 min for SurgicalGaussian [20]), with the UNet and fusion MLP adding $\sim 19\%$ overhead. At test time, UNet inference is replaced by embedding interpolation (Eq. 5), achieving 84 fps rendering (vs. 87 fps for SurgicalGaussian) with $4.5GB$ memory on a single RTX 4090.

Ablation and decomposition analysis. Tab. 4 reports reconstruction metrics alongside R-Var on RLLS. The most revealing comparison is (a) vs. (b): replacing our decomposition with affine correction ($\tilde{\mathbf{c}} = \boldsymbol{\beta} \odot \mathbf{c} + \boldsymbol{\gamma}$) reduces PSNR by only 0.48 dB yet *triples* R-Var ($1.08 \to 3.42$), confirming that affine correction improves appearance but leaves illumination entangled in base SH coefficients. Substituting the UNet with a learnable embedding (c) doubles R-Var, showing that image-conditioned estimation provides a stronger inductive bias. Removing Stage 1 (d) causes the largest degradation (PSNR -1.19 dB), confirming that stable geometry is a prerequisite for meaningful decomposition. Dropping $\mathcal{L}_{\text{brt}}$ (e) allows $\mathbf{L}_t$ to collapse toward trivial solutions, moderately degrading both metrics. Fig. 4 provides visual evidence. In (a), a held-out view rendered under illumination transferred from different training frames adapts plausibly while structural detail remains intact, confirming independent controllability. In (b), reflectance maintains consistent appearance across frames with different illumination, while illumination maps capture the expected smooth radial falloff.

4 Conclusion

We presented LumiState, which integrates Retinex-inspired reflectance illumination decomposition into 4D Gaussian Splatting for dynamic endoscopic reconstruction. Experiments on two benchmarks confirm SOTA reconstruction quality with the largest gains under severe illumination variation, while enabling controllable illumination editing and temporally consistent reflectance rendering.

Acknowledgments. Supported by National Natural Science Foundation of China under Grant 62271465, Suzhou Basic Research Program under Grant SYG202338, and Jiangsu Province Science Foundation for Youths (NO. BK20240464).

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Cai, B., Xu, X., Guo, K., Jia, K., Hu, B., Tao, D.: A joint intrinsic-extrinsic prior model for retinex. In: Proceedings of the IEEE international conference on computer vision. pp. 4000–4009 (2017)
2. Chen, Y., Wang, H.: Endogaussians: Single view dynamic gaussian splatting for deformable endoscopic tissues reconstruction. arXiv preprint arXiv:2401.13352 (2024)
3. Horn, B.K., Schunck, B.G.: Determining optical flow. Artificial Intelligence **17**(1), 185–203 (1981). `https://doi.org/https://doi.org/10.1016/0004-3702(81)90024-2`, `https://www.sciencedirect.com/science/article/pii/0004370281900242`
4. Huang, Y., Bai, L., Cui, B., Li, Y., Chen, T., Wang, J., Wu, J., Lei, Z., Liu, H., Ren, H.: Endo-4dgx: Robust endoscopic scene reconstruction and illumination correction with gaussian splatting. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 181–191. Springer (2025)
5. Jiang, Y., Tu, J., Liu, Y., Gao, X., Long, X., Wang, W., Ma, Y.: Gaussianshader: 3d gaussian splatting with shading functions for reflective surfaces. arXiv preprint arXiv:2311.17977 (2023)
6. Kaleta, J., Smolak-Dyżewska, W., Malarz, D., Dall'Alba, D., Korzeniowski, P., Spurek, P.: Pr-endo: Physically based relightable gaussian splatting for endoscopy. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 391–401. Springer (2025)
7. Kerbl, B., Kopanas, G., Leimkühler, T., Drettakis, G.: 3d gaussian splatting for real-time radiance field rendering. ACM Trans. Graph. **42**(4), 139–1 (2023)
8. Kulhanek, J., Peng, S., Kukelova, Z., Pollefeys, M., Sattler, T.: Wildgaussians: 3d gaussian splatting in the wild. arXiv preprint arXiv:2407.08447 (2024)
9. Land, E.H.: The retinex theory of color vision. Scientific american **237**(6), 108–129 (1977)
10. Liang, Z., Zhang, Q., Feng, Y., Shan, Y., Jia, K.: Gs-ir: 3d gaussian splatting for inverse rendering. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 21644–21653 (2024)
11. Liu, Y., Li, C., Yang, C., Yuan, Y.: Endogaussian: Real-time gaussian splatting for dynamic endoscopic scene reconstruction. arXiv preprint arXiv:2401.12561 (2024)

12. Maier-Hein, L., Mountney, P., Bartoli, A., Elhawary, H., Elson, D., Groch, A., Kolb, A., Rodrigues, M., Sorger, J., Speidel, S., et al.: Optical techniques for 3d surface reconstruction in computer-assisted laparoscopic surgery. Medical image analysis **17**(8), 974–996 (2013)
13. Mildenhall, B., Srinivasan, P.P., Tancik, M., Barron, J.T., Ramamoorthi, R., Ng, R.: Nerf: Representing scenes as neural radiance fields for view synthesis. Communications of the ACM **65**(1), 99–106 (2021)
14. Qiu, L., Ren, H.: Endoscope navigation and 3d reconstruction of oral cavity by visual slam with mitigated data scarcity. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops. pp. 2197–2204 (2018)
15. Seitz, S., Curless, B., Diebel, J., Scharstein, D., Szeliski, R.: A comparison and evaluation of multi-view stereo reconstruction algorithms. In: 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06). vol. 1, pp. 519–528 (2006). `https://doi.org/10.1109/CVPR.2006.19`
16. Tang, R., Ma, L.F., Rong, Z.X., Li, M.D., Zeng, J.P., Wang, X.D., Liao, H.E., Dong, J.H.: Augmented reality technology for preoperative planning and intraoperative navigation during hepatobiliary surgery: A review of current methods. Hepatobiliary & Pancreatic Diseases International **17**(2), 101–112 (2018)
17. Wang, Y., Gong, B., Long, Y., Fan, S.H., Dou, Q.: Efficient endonerf reconstruction and its application for data-driven surgical simulation. International Journal of Computer Assisted Radiology and Surgery **19**(5), 821–829 (2024)
18. Wang, Y., Long, Y., Fan, S.H., Dou, Q.: Neural rendering for stereo 3d reconstruction of deformable tissues in robotic surgery. In: International conference on medical image computing and computer-assisted intervention. pp. 431–441. Springer (2022)
19. Wu, G., Yi, T., Fang, J., Xie, L., Zhang, X., Wei, W., Liu, W., Tian, Q., Wang, X.: 4d gaussian splatting for real-time dynamic scene rendering. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 20310–20320 (2024)
20. Xie, W., Yao, J., Cao, X., Lin, Q., Tang, Z., Dong, X., Guo, X.: Surgicalgaussian: Deformable 3d gaussians for high-fidelity surgical scene reconstruction. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 617–627. Springer (2024)
21. Yang, C., Wang, K., Wang, Y., Yang, X., Shen, W.: Neural lerplane representations for fast 4d reconstruction of deformable tissues. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 46–56. Springer (2023)
22. Yang, S., Li, Q., Shen, D., Gong, B., Dou, Q., Jin, Y.: Deform3dgs: Flexible deformation for fast surgical scene reconstruction with gaussian splatting. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 132–142. Springer (2024)
23. Ye, S., Dong, Z.H., Hu, Y., Wen, Y.H., Liu, Y.J.: Gaussian in the dark: Real-time view synthesis from inconsistent dark images using gaussian splatting. In: Computer Graphics Forum. vol. 43, p. e15213. Wiley Online Library (2024)
24. Zha, R., Cheng, X., Li, H., Harandi, M., Ge, Z.: Endosurf: Neural surface reconstruction of deformable tissues with stereo endoscope videos. In: International conference on medical image computing and computer-assisted intervention. pp. 13–23. Springer (2023)
25. Zhao, S., Liu, Y., Wang, Q., Sun, D., Liu, R., Zhou, S.K.: Murphy: relations matter in surgical workflow analysis. arXiv preprint arXiv:2212.12719 (2022)

26. Zhu, L., Wang, Z., Cui, J., Jin, Z., Lin, G., Yu, L.: Endogs: Deformable endoscopic tissues reconstruction with gaussian splatting. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 135–145. Springer (2024)