
NFL-BA: Near-Field Light Bundle Adjustment for SLAM in Dynamic Lighting Supplementary Materials

A Overview of Appendices

Our appendices contain the following additional details:

- Appendix B provides details on data collection and processing
- Appendix C implementation details and Computational Costs
- Appendix D Metrics
- Appendix E we show additional detailed results on the Endoscopy Datasets

To view interactive point clouds and see videos, please see our zipped supplemental. To view the webpage, unzip the folder then click the "RUN_ME" file corresponding to your operating system.

MacOS To run on Mac, you may need to Control-click (or right-click) on `RUN_ME-MacOS.command` and choose *Open*. A Gatekeeper warning appears ("unidentified developer"). Click *Open* again.

Windows After unzipping on Windows: Right-click `RUN_ME-Windows.bat` and select *Run as administrator* (or click *Allow* in the SmartScreen prompt).

Linux If needed, make the script executable once (`chmod +x RUN_ME-Linux.sh`).

B Dataset Processing and Collection

C3VD. We test our method on a subset of 8 videos with at least one video from each section of the colon, with varying camera motion, and anomalies. We choose these 8 videos from the test split of PPSNet [2] to avoid any bias when the SLAM is initialized with PPSNet predicted depth map. The names of the sequences are as follows: `cecum_t1_a`, `cecum_t2_a`, `cecum_t3_a`, `sigmoid_t3_a`, `desc_t4_a_p2`, `trans_t2_a`, `trans_t3_a`, and `trans_t4_a`.

All images were cropped to remove any artifacts resulting from fish-eye correction then downscale and crop the images. Specifically, we resize each image to have a height of 384 pixels while maintaining the aspect ratio, then crop the central region to obtain a 384×384 pixel image.

Colon10k. Since Colon10K provides no ground-truth depths, we compute per-frame estimates with PPSNet. Each image is center-cropped and uniformly resized to 384 × 384 px to match our SLAM input requirements.

Indoor Self Captures. We recorded four indoor scenes using an iPhone 15 with LiDAR, capturing synchronized RGB (1440 × 1920 px) and depth (256 × 192 px) streams. Raw RGB frames are downsampled to 256 × 192 px to align with the depth map resolution. Depth maps are stored as 16-bit values up to 10 m. We logged camera poses via Apple's ARKit framework—code for our custom capture app and preprocessing scripts will be released alongside the dataset. In the zipped portion of the supplementary, you can preview some of the videos used.

32 C Implementation Details

33 As mentioned in the main paper, we ran all models on a single NVIDIA RTX A6000 GPU. The
 34 per-scene optimization takes approximately 1 FPS. We found that NFL-BA only reduces the fps
 35 runtime by a small amount on the C3VD dataset.

Table 1: Runtime: We evaluate the frames per second (**FPS**) for all methods on the C3VD dataset.

Method	Depth	Photo-BA	NFL-BA
NICE-SLAM	Oracle	$\ll 1$	$\ll 1$
	PPSNet	$\ll 1$	$\ll 1$
EndoGSLAM	Oracle	1.79	1.35
	PPSNet	1.53	1.22
	DPT-Hybrid	1.20	0.90
MonoGS	Oracle	1.38	1.09
	PPSNet	1.06	0.93
	DPT-Hybrid	0.99	0.83

36 **NICE-SLAM.** Since the scene is encoded using neural networks, we extract normals from the
 37 occupancy grid, as described in Sec. 4.2 to calculate the shading term. NICE-SLAM requires a
 38 well-defined bounding box which we obtained from the ground truth point clouds (see appendix D).
 39 We also used the default loss weights of NICE-SLAM, setting λ_{ren} to 0.5 during tracking and 0.2
 40 during mapping, and λ_{geo} to 1 in both phases.

41 **EndoGSLAM.** The main difference is the weight map M_t in the bundle adjustment loss (Eq. 3) to
 42 exclude over-exposed pixels that can arise in endoscopy-specific lighting conditions. Furthermore,
 43 given that the shading term $PPS(\cdot)$ is sensitive to depth scales, we rescaled the depth maps so that
 44 their maximum values are approximately 5. Notably, scaling the depth maps did not improve baseline
 45 performance (when using PPS depth maps, the average ATE_t went from 3.03 to 3.38). We used the
 46 default loss weights of EndoGSLAM, λ_{ren} ; λ_{geo} , set to 0.5 and 1 during tracking and 1 and 1 during
 47 mapping, respectively.

48 **MonoGS.** To integrate our method, we treat the Gaussian color features as albedo features and
 49 multiply them with the shading term before rasterization, and then we use the rendered output colors
 50 for bundle adjustment. For all input depths, we set λ_{ren} and λ_{geo} to 0.8 and 0.5, respectively.

51 D Point Clouds and Metrics

52 **Chamfer Distances.** Since ground truth point clouds are unavailable for C3VD, we generate them
 53 by unprojecting 2D images into 3D space with the correct camera configuration and the oracle depth
 54 maps, provided in the C3VD dataset. For neural fields-based SLAM, we use the vertices of the
 55 output meshes as the estimated point clouds, while for Gaussian Splatting-based SLAMs, we use
 56 the Gaussian positions. For point cloud alignment, we use Coherent Point Drift [1] and the Chamfer
 57 distances are also in millimeters.

58 **Coloring Point Clouds.** We use extracted point clouds from 3D Gaussian positions for visualization.
 59 For colors, we directly use Gaussian color features.

60 E Results on endoscopy Datasets

61 We have included a point cloud visualizations for EndoGSLAM below as well as interactive ones in
 62 the zipped portion of the supplementary.

63 **Reproducibility and Statistical Significance** For all experiments for each of the slam systems, we
 64 report the median of three runs for all tables and figures. We have included per-sequence metrics for
 65 the median run below.

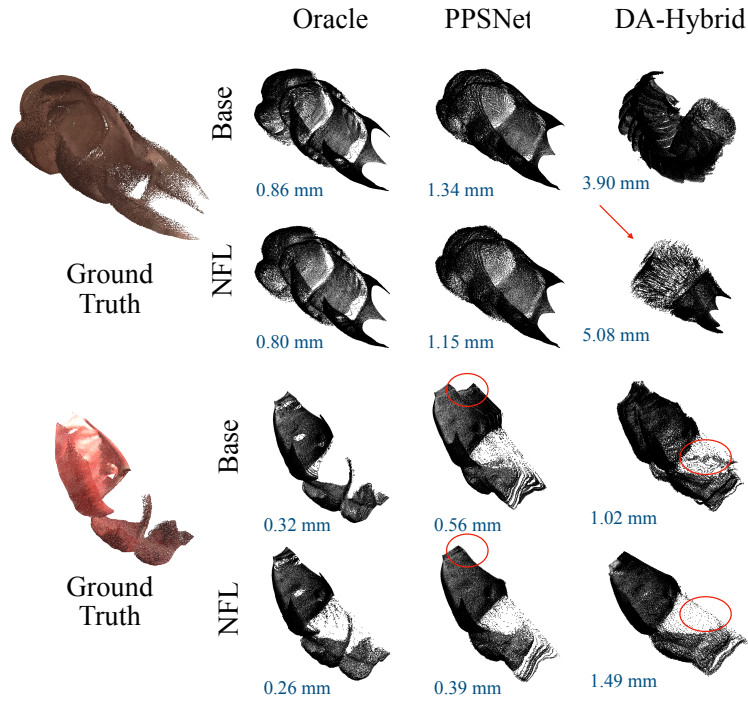


Figure 1: EndoGSLAM Point cloud results for 2 sequences: cecum_t1_a_under_review (top) and desc_t4_a_p2_under_review (bottom).

References

- [1] Andriy Myronenko and Xubo Song. Point set registration: Coherent point drift. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32(12):2262–2275, 2010.
- [2] Akshay Paruchuri, Samuel Ehrenstein, Shuxian Wang, Inbar Fried, Stephen M Pizer, Marc Niethammer, and Roni Sengupta. Leveraging near-field lighting for monocular depth estimation from endoscopy videos. In *Computer Vision – ECCV 2024*, Cham, 2024. Springer Nature Switzerland.

Table 2: Results for sequence **cecum_t1_a_under_review**

Method	Depth	BA	ATE_t (mm)↓	ATE_r (°)↓	Chamfer (mm)↓
NICE-SLAM	Oracle	Photo	2.65	2.82	0.08
		NFL	1.39	2.81	0.15
	PPS-Net	Photo	10.14	2.71	0.51
		NFL	4.14	2.75	0.19
EndoGSLAM	Oracle	Photo	1.39	0.24	0.12
		NFL	0.91	0.31	0.16
	PPS-Net	Photo	2.79	0.60	0.30
		NFL	2.93	0.70	0.35
	DA-Hybrid	Photo	2.64	0.08	0.04
		NFL	8.44	2.42	1.21
MonoGS	Oracle	Photo	1.16	0.39	1.56
		NFL	1.22	0.35	0.95
	PPS-Net	Photo	2.30	0.65	1.19
		NFL	2.45	0.76	1.04
	DA-Hybrid	Photo	5.44	0.30	1.64
		NFL	1.09	0.40	1.65

Table 3: Results for sequence **cecum_t2_a_under_review**

Method	Depth	BA	ATE_t (mm)↓	ATE_r (°)↓	Chamfer (mm)↓
NICE-SLAM	Oracle	Photo	8.13	2.19	0.49
		NFL	1.15	2.82	0.11
	PPS-Net	Photo	1.11	2.13	0.11
		NFL	7.98	2.82	0.50
EndoGSLAM	Oracle	Photo	2.32	1.06	0.53
		NFL	6.75	2.79	1.40
	PPS-Net	Photo	4.55	1.18	0.59
		NFL	4.55	2.82	1.41
	DA-Hybrid	Photo	5.66	0.89	0.45
		NFL	9.47	2.25	1.13
MonoGS	Oracle	Photo	4.04	0.34	1.47
		NFL	6.84	0.48	1.24
	PPS-Net	Photo	6.72	2.71	1.74
		NFL	6.52	2.73	1.70
	DA-Hybrid	Photo	5.20	0.66	2.18
		NFL	4.26	0.32	1.35

Table 4: Results for sequence **cecum_t3_a_under_review**

Method	Depth	BA	ATE_t (mm)↓	ATE_r (°)↓	Chamfer (mm)↓
NICE-SLAM	Oracle	Photo	3.46	2.82	0.11
		NFL	3.42	2.82	0.07
	PPS-Net	Photo	2.80	2.64	0.17
		NFL	3.83	2.59	0.14
EndoGSLAM	Oracle	Photo	0.75	0.15	0.08
		NFL	0.74	0.27	0.14
	PPS-Net	Photo	0.79	0.16	0.08
		NFL	2.18	0.22	0.11
	DA-Hybrid	Photo	1.45	0.62	0.31
		NFL	1.52	0.18	0.09
MonoGS	Oracle	Photo	0.36	0.16	1.23
		NFL	0.87	0.31	0.62
	PPS-Net	Photo	1.07	0.17	1.12
		NFL	1.24	0.14	0.81
	DA-Hybrid	Photo	1.06	0.33	0.95
		NFL	1.16	0.30	0.93

Table 5: Results for sequence **desc_t4_a_p2_under_review**

Method	Depth	BA	ATE_t (mm)↓	ATE_r (°)↓	Chamfer (mm)↓
NICE-SLAM	Oracle	Photo	15.90	2.65	0.59
		NFL	8.88	2.75	0.50
	PPS-Net	Photo	0.87	2.80	0.10
		NFL	0.68	2.80	0.10
EndoGSLAM	Oracle	Photo	0.36	0.56	0.28
		NFL	0.26	0.94	0.47
	PPS-Net	Photo	0.48	0.58	0.29
		NFL	1.00	0.83	0.42
	DA-Hybrid	Photo	0.45	1.16	0.58
		NFL	1.80	2.81	1.40
MonoGS	Oracle	Photo	0.25	1.05	0.76
		NFL	0.13	0.98	0.67
	PPS-Net	Photo	0.49	1.78	0.77
		NFL	0.61	1.67	0.70
	DA-Hybrid	Photo	0.69	2.63	0.80
		NFL	0.20	0.97	0.77

Table 6: Results for sequence **sigmoid_t3_a_under_review**

Method	Depth	BA	ATE_t (mm)↓	ATE_r (°)↓	Chamfer (mm)↓
NICE-SLAM	Oracle	Photo	1.04	2.47	0.05
		NFL	0.84	2.83	0.05
	PPS-Net	Photo	9.47	2.79	0.28
		NFL	6.60	2.55	0.35
EndoGSLAM	Oracle	Photo	4.81	1.14	0.57
		NFL	4.11	2.69	1.35
	PPS-Net	Photo	3.13	1.40	0.70
		NFL	9.82	2.56	1.29
	DA-Hybrid	Photo	8.51	2.36	1.18
		NFL	8.15	2.80	1.41
MonoGS	Oracle	Photo	1.44	0.46	1.17
		NFL	1.19	2.33	0.67
	PPS-Net	Photo	1.20	1.53	4.63
		NFL	1.22	2.22	0.89
	DA-Hybrid	Photo	6.88	2.18	1.42
		NFL	1.10	0.67	1.33

Table 7: Results for sequence **trans_t2_a_under_review**

Method	Depth	BA	ATE_t (mm)↓	ATE_r (°)↓	Chamfer (mm)↓
NICE-SLAM	Oracle	Photo	0.77	2.83	0.05
		NFL	0.94	2.80	0.05
	PPS-Net	Photo	9.18	2.82	0.26
		NFL	9.85	2.81	0.31
EndoGSLAM	Oracle	Photo	4.21	1.80	0.90
		NFL	0.49	2.82	1.41
	PPS-Net	Photo	10.05	1.66	0.84
		NFL	0.90	1.46	0.73
	DA-Hybrid	Photo	9.14	2.77	1.39
		NFL	14.59	2.18	1.10
MonoGS	Oracle	Photo	14.38	1.42	1.49
		NFL	1.51	2.72	0.97
	PPS-Net	Photo	14.64	1.91	1.52
		NFL	3.08	1.02	1.09
	DA-Hybrid	Photo	15.18	2.12	1.68
		NFL	9.51	1.41	1.45

Table 8: Results for sequence **trans_t3_a_under_review**

Method	Depth	BA	ATE_t (mm)↓	$ATE_r(^{\circ})$ ↓	Chamfer (mm)↓
NICE-SLAM	Oracle	Photo	0.97	2.81	0.15
		NFL	6.26	2.79	0.68
	PPS-Net	Photo	3.78	2.80	0.22
		NFL	1.12	2.65	0.13
EndoGSLAM	Oracle	Photo	0.17	2.70	1.35
		NFL	0.20	2.70	1.35
	PPS-Net	Photo	0.30	2.80	1.40
		NFL	0.50	2.81	1.41
	DA-Hybrid	Photo	0.46	2.81	1.41
		NFL	0.33	2.74	1.37
MonoGS	Oracle	Photo	0.31	2.67	0.76
		NFL	0.26	2.66	0.60
	PPS-Net	Photo	0.33	2.79	0.89
		NFL	0.61	2.83	0.82
	DA-Hybrid	Photo	0.29	2.81	0.91
		NFL	0.38	2.71	0.71

Table 9: Results for sequence **trans_t4_a_under_review**

Method	Depth	BA	ATE_t (mm)↓	$ATE_r(^{\circ})$ ↓	Chamfer (mm)↓
NICE-SLAM	Oracle	Photo	0.39	2.81	0.05
		NFL	0.23	2.83	0.05
	PPS-Net	Photo	7.26	2.77	0.14
		NFL	7.88	2.73	0.23
EndoGSLAM	Oracle	Photo	2.64	1.47	0.74
		NFL	1.99	2.02	1.01
	PPS-Net	Photo	2.02	1.73	0.86
		NFL	2.58	1.78	0.89
	DA-Hybrid	Photo	2.99	2.08	1.04
		NFL	10.67	2.72	1.37
MonoGS	Oracle	Photo	1.23	2.41	0.83
		NFL	0.76	2.06	0.62
	PPS-Net	Photo	1.12	2.10	0.89
		NFL	1.70	1.80	0.86
	DA-Hybrid	Photo	2.30	2.46	1.10
		NFL	1.14	2.36	0.85