Supplementary Materials for ABLE-NeRF

Appendix A. Optimisation Details

Our implementation is based on PyTorch Lightning. We optimise our model with Adam and we set the learning rate as 5×10^{-4} which is annealed log-linearly to 1×10^{-4}, with a warm-up phase of 1250 iterations. We set the hyper-parameters as $\beta_1 = 0.9$ and $\beta_2 = 0.999$. Each scene in the Blender and Shiny Blender dataset takes 1.5 days to train on 8 Tesla v100 GPUs with a total batch size of 8192 rays. To render an image on a single Tesla v100 GPU takes 110s.

Appendix B. Blender Dataset Details

We report full breakdown of individual scenes in the Blender Dataset in Tables A.1, A.2, A.3.

We have included the test scene images, compiled into videos in the zipped folder, and we compare our results with the rendered test set images from Ref-NeRF[1].

	chair	lego	materials	mic	hotdog	ficus	drums	ship
PhySG [4]	24.00	20.19	18.86	22.33	24.08	19.02	20.99	15.35
VolSDF [3]	30.57	29.46	29.13	30.53	35.11	22.91	20.43	25.51
Mip-NeRF [1]	35.12	35.92	30.62	36.76	37.34	33.19	25.36	30.52
Ref-NeRF [2]	35.83	36.25	35.41	36.76	37.72	33.91	25.79	30.28
Ours,No LE	35.76	36.62	34.57	35.90	38.68	34.28	25.98	30.60
Ours	36.25	38.03	35.46	37.11	39.07	35.69	26.84	31.75

A.1. Per-scene test set PSNRs on Blender Dataset. Results retrieved from [2].

	chair	lego	materials	mic	hotdog	ficus	drums	ship
PhySG [4]	0.898	0.821	0.838	0.933	0.912	0.873	0.884	0.727
VolSDF [3]	0.949	0.951	0.954	0.969	0.972	0.929	0.893	0.842
Mip-NeRF [1]	0.981	0.980	0.959	0.992	0.982	0.980	0.933	0.885
Ref-NeRF [2]	0.984	0.981	0.983	0.992	0.984	0.983	0.937	0.880
Ours, No LE	0.941	0.985	0.983	0.991	0.987	0.986	0.945	0.893
Ours	0.987	0.987	0.985	0.993	0.988	0.989	0.951	0.919

A.2. Per-scene test set SSIMs on Blender Dataset. Results retrieved from [2].

Appendix C. Shiny Blender Dataset Details

We report full breakdown of individual scenes in the Shiny Blender Dataset in Tables A.4, A.5, A.6.

[1]We thank Dor Verbin for providing us with the test images from Ref-NeRF model output

f	chair	lego	materials	mic	hotdog	ficus	drums	ship
PhySG [4]	0.093	0.172	0.142	0.082	0.117	0.112	0.113	0.322
VolSDF [3]	0.056	0.054	0.048	0.191	0.043	0.068	0.119	0.191
Mip-NeRF [1]	0.020	0.018	0.040	0.008	0.026	0.021	0.064	0.135
Ref-NeRF [2]	0.017	0.018	0.022	0.007	0.022	0.019	0.059	0.139
Ours, No LE	0.021	0.018	0.031	0.010	0.024	0.017	0.065	0.147
Ours	0.018	0.015	0.029	0.008	0.021	0.014	0.057	0.122

A.3. Per-scene test set LPIPS on Blender Dataset. Results retrieved from [2].

We have included the test scene images, compiled into videos in the zipped folder, and we compare our results with the rendered test set images from Ref-NeRF.

	teapot	toaster	car	ball	coffee	helmet
PhySG [4]	35.83	18.59	24.40	27.24	23.71	27.51
Mip-NeRF [1]	46.00	22.37	26.50	25.94	30.36	27.39
Ref-NeRF, no pred. normals [2]	47.09	23.32	27.19	26.09	31.79	30.54
Ref-NeRF [2]	47.90	25.70	30.82	47.46	34.21	29.68
Ours	47.30	26.52	28.76	36.62	33.01	31.04

A.4. Per-scene test set PSNRs on Shiny Blender Dataset. Results retrieved from [2].

	teapot	toaster	car	ball	coffee	helmet
PhySG [4]	0.990	0.805	0.910	0.947	0.922	0.953
Mip-NeRF [1]	0.997	0.891	0.922	0.935	0.966	0.939
Ref-NeRF, no pred. normals [2]	0.997	0.898	0.926	0.865	0.967	0.962
Ref-NeRF [2]	0.998	0.922	0.955	0.995	0.974	0.958
Ours	0.998	0.949	0.942	0.984	0.975	0.968

A.5. Per-scene test set SSIMs on Shiny Blender Dataset. Results retrieved from [2].

	teapot	toaster	car	ball	coffee	helmet
PhySG [4]	0.022	0.194	0.091	0.179	0.150	0.089
Mip-NeRF [1]	0.008	0.123	0.059	0.168	0.086	0.108
Ref-NeRF, no pred. normals [2]	0.006	0.134	0.064	0.272	0.087	0.068
Ref-NeRF [2]	0.004	0.095	0.041	0.059	0.078	0.075
Ours	0.007	0.092	0.052	0.107	0.114	0.080

A.6. Per-scene test set LPIPS on Shiny Blender Dataset. Results retrieved from [2].

References

[1] Jonathan T Barron, Ben Mildenhall, Matthew Tancik, Peter Hedman, Ricardo Martin-Brualla, and Pratul P Srinivasan. Mip-nerf: A multiscale representation for anti-aliasing neural radiance fields. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 5855–5864, 2021. 1

[2] Dor Verbin, Peter Hedman, Ben Mildenhall, Todd Zickler, Jonathan T Barron, and Pratul P Srinivasan. Ref-nerf: Structured view-dependent appearance for neural radiance fields. In *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 5481–5490. IEEE, 2022. 1

[3] Lior Yariv, Jiatao Gu, Yoni Kasten, and Yaron Lipman. Volume rendering of neural implicit surfaces. *Advances in Neural Information Processing Systems*, 34:4805–4815, 2021. 1

[4] Kai Zhang, Fujun Luan, Qianqian Wang, Kavita Bala, and Noah Snavely. Physg: Inverse rendering with spherical gaussians for physics-based material editing and relighting. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 5453–5462, 2021. 1