GSNB: Gaussian Splatting With Neural Basis Extension
Supplementary Material

Anonymous Author(s)

A MORE ABLATION STUDIES
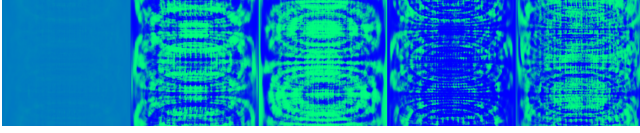
A.1 Ablation Study On Baking Image

Figure 1: Visualization of the baking image in the CD scene from the Shiny [7] dataset. For the sake of brevity, we show only the first 5 channels of the image, which should have 16 channels in our implementation.

As shown in Figure 1, We visualize one of the bakery images used in our real-time rendering pipeline and the image shows rich high frequency information. We also perform an ablation study on the baking image size to determine the appropriate size for preserving the Neural Basis Extension network result. The experiment is conducted on 8 scenes from the Spaces [2] dataset, we compare the PSNR score for the different resolution images and find out that although the performance is improved with the resolution increase, the improvement is limited compared to the disk space inflation. We choose to use 400×400 as the proper size of the baking in most of the test scene, while for a few scenes we use 800×800. This experiment is performed on the pruned model described in the main body.
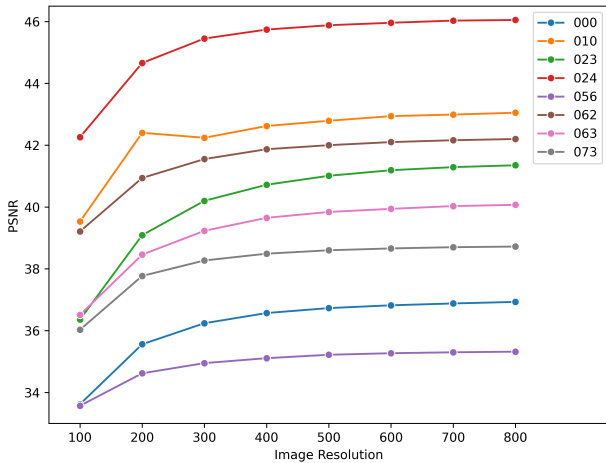

Figure 2: Effect of baking resolution on performance in 8 scenes from the Spaces [2] dataset

Table 1: Pruning strategy influence on PSNR performance

Method | Scene	Mic	Chair	Ship	Materials	Lego	Drums	Ficus	Hotdog	Avg.
w/o pruning	36.90	35.68	31.81	30.65	36.40	26.68	35.66	38.08	33.98
SP	37.04	36.09	31.56	30.35	36.50	26.59	35.61	38.13	33.98
HFP	36.47	35.54	31.35	29.79	35.94	26.44	35.16	37.81	33.56
SFP	37.12	36.04	31.87	30.68	36.75	26.63	35.62	38.28	34.12

Table 2: Pruning strategy influence on SSIM performance

Method | Scene	Mic	Chair	Ship	Materials	Lego	Drums	Ficus	Hotdog	Avg.
w/o pruning	0.993	0.988	0.907	0.961	0.983	0.955	0.987	0.985	0.970
SP	0.993	0.988	0.905	0.959	0.982	0.953	0.987	0.985	0.969
HFP	0.992	0.988	0.906	0.957	0.982	0.954	0.986	0.985	0.969
SFP	0.993	0.988	0.907	0.961	0.983	0.954	0.987	0.986	0.970

Table 3: Pruning strategy influence on LPIPS performance

Method | Scene	Mic	Chair	Ship	Materials	Lego	Drums	Ficus	Hotdog	Avg.
w/o pruning	0.006	0.010	0.108	0.037	0.016	0.035	0.012	0.020	0.030
SP	0.006	0.010	0.113	0.039	0.018	0.037	0.012	0.021	0.032
HFP	0.007	0.011	0.112	0.042	0.018	0.038	0.013	0.022	0.033
SFP	0.006	0.010	0.109	0.037	0.016	0.036	0.012	0.020	0.031

A.2 Ablation Study On Pruning Strategy

To decide the best pruning strategy to apply to the trained GSNB, we test three different pruning strategies. The first strategy, we call Single Pruning (SP), filters out the less important 3D Gaussians at the beginning of the pruning process and performs re-optimization on the pruned model. The second strategy, Hard Filter Pruning (HFP), is based on Structured Pruning [5], which prunes the model several times until we reach the specified pruning ratio. Each pruning removes a portion of the 3D Gaussians that are considered less important by our importance score. The third strategy is based on Soft Filter Pruning (SFP) [3]. The method also performs multiple prunings but instead of directly bumping the redundant Gaussians, the method simply resets the Gaussians' feature coefficient to be 0. The model is not actually pruned until the last pruning. We test the three methods on 8 scenes from the Synthetic NeRF [6] dataset and set the pruning ratio to 0.6.

Ablation studies reveal that the three strategies show no obvious difference in terms of pruning performance. Although the SFP method is slightly better in most of the scene, we adopt the Single Pruning method for simplicity.

B MORE EXPERIMENT RESULTS

To show the specific performance of our method, we provide our complete experimental results on each scene of the test dataset. Our method shows state of the art result in most of the scene.We color each cell as best , second best and third best .

Table 4: Quantitative comparison of PSNR values on Shiny [7] dataset

Method | Scene	Shiny	Lab	Tools	Giants	Avg.
3D-GS	30.93	30.36	27.68	25.09	28.52
Nex	31.43	30.43	28.16	26.00	29.01
Scaffold-GS	33.35	32.17	28.06	25.18	29.69
Ours	34.35	33.04	27.78	25.35	30.13
Ours w/ pruning	34.45	33.11	27.87	25.37	30.20

Table 5: Quantitative comparison of SSIM values on Shiny [7] dataset

Method | Scene	Shiny	Lab	Tools	Giants	Avg.
3D-GS	0.939	0.927	0.921	0.866	0.913
Nex	0.958	0.949	0.953	0.898	0.940
Scaffold-GS	0.949	0.945	0.929	0.851	0.918
Ours	0.954	0.952	0.924	0.868	0.925
Ours w/ pruning	0.954	0.952	0.925	0.859	0.923

Table 6: Quantitative comparison of LPIPS values on Shiny [7] dataset

Method | Scene	Shiny	Lab	Tools	Giants	Avg.
3D-GS	0.118	0.139	0.153	0.103	0.128
Nex	0.129	0.146	0.151	0.147	0.143
Scaffold-GS	0.100	0.103	0.127	0.141	0.118
Ours	0.098	0.110	0.148	0.105	0.115
Ours w/ pruning	0.103	0.116	0.149	0.116	0.121

Table 7: Quantitative comparison of PSNR values on Mip-NeRF360 [1] dataset

Method | Scene	Bicycle	Bonsai	Counter	Stump	Garden	Kitchen	Room
Plenoxels	21.91	24.67	23.62	20.66	23.49	23.42	27.59
iNGP	22.17	30.69	26.69	23.47	25.07	29.48	29.69
Mip-NeRF	24.30	33.39	29.44	26.17	26.87	31.98	31.46
3DGS	25.09	32.32	29.07	26.65	27.27	31.54	31.66
Scaffold-GS	25.05	32.49	29.44	26.49	27.33	31.59	32.13
Ours	25.13	33.01	29.54	26.58	27.50	31.99	31.79
Ours w/ pruning	25.07	33.58	29.87	26.49	27.53	32.30	32.23

Table 8: Quantitative comparison of SSIM values on Mip-NeRF360 [1] dataset

Method | Scene	Bicycle	Bonsai	Counter	Stump	Garden	Kitchen	Room
Plenoxels	0.496	0.814	0.759	0.523	0.606	0.648	0.842
iNGP	0.512	0.906	0.817	0.594	0.701	0.858	0.871
Mip-NeRF360	0.685	0.938	0.892	0.745	0.809	0.917	0.910
3DGS	0.746	0.946	0.913	0.768	0.855	0.931	0.925
Scaffold-GS	0.738	0.946	0.914	0.754	0.846	0.929	0.927
Ours	0.745	0.947	0.917	0.767	0.858	0.933	0.926
Ours w/ pruning	0.743	0.949	0.919	0.760	0.853	0.935	0.929

Table 9: Quantitative comparison of LPIPS values on Mip-NeRF360 [1] dataset

Method | Scene	Bicycle	Bonsai	Counter	Stump	Garden	Kitchen	Room
Plenoxels	0.506	0.398	0.441	0.503	0.386	0.447	0.419
iNGP	0.446	0.205	0.306	0.421	0.257	0.195	0.261
Mip-NeRF	0.305	0.179	0.207	0.265	0.171	0.128	0.213
3DGS	0.244	0.180	0.183	0.243	0.122	0.116	0.197
Scaffold-GS	0.266	0.186	0.195	0.276	0.143	0.121	0.200
Ours	0.243	0.180	0.182	0.243	0.120	0.114	0.198
Ours w/ pruning	0.242	0.184	0.185	0.246	0.130	0.118	0.199

Table 10: Quantitative comparison of metrics on Tanks & Temples [4] dataset

Dataset	Truck			Train		
Method | Metric	PSNR	SSIM	LPIPS	PSNR	SSIM	LPIPS
Plenoxels	23.22	0.774	0.335	18.93	0.663	0.422
iNGP	23.38	0.800	0.249	20.46	0.689	0.360
Mip-NeRF360	24.91	0.857	0.159	19.52	0.660	0.354
3D-GS	25.15	0.864	0.194	21.99	0.809	0.241
Scaffold-GS	25.77	0.883	0.147	22.15	0.822	0.206
Ours	25.54	0.880	0.146	22.44	0.816	0.205
Ours w/ pruning	25.79	0.882	0.145	22.70	0.814	0.219

Table 11: Quantitative comparison of PSNR values on Synthetic NeRF [6] dataset

Method | Scene	Mic	Chair	Ship	Materials	Lego	Drums	Ficus	Hotdog	Avg.
Mip-NeRF	36.51	35.14	30.41	30.71	35.70	25.48	33.29	37.48	33.09
Tri-MipRF	37.75	36.10	28.78	30.73	36.15	26.59	34.51	38.54	33.64
GS-Shader	35.23	35.83	30.82	30.07	35.87	26.36	34.97	37.85	33.38
3DGS	36.67	35.52	31.67	30.49	36.08	26.28	35.49	38.10	33.79
Scaffold-GS	37.25	35.28	31.17	30.65	35.69	26.44	35.21	37.73	33.68
Ours	36.90	35.68	31.81	30.65	36.40	26.68	35.66	38.08	33.98
Ours w/ pruning	37.04	36.09	31.56	30.35	36.50	26.59	35.61	38.13	33.98

Table 12: Quantitative comparison of SSIM values on Synthetic NeRF [6] dataset

Method | Scene	Mic	Chair	Ship	Materials	Lego	Drums	Ficus	Hotdog	Avg.
Mip-NeRF	0.991	0.981	0.882	0.959	0.978	0.932	0.980	0.982	0.961
Tri-MipRF	0.992	0.985	0.879	0.953	0.982	0.939	0.983	0.984	0.962
GS-Shader	0.991	0.987	0.905	0.960	0.983	0.949	0.985	0.985	0.968
3DGS	0.993	0.988	0.906	0.960	0.983	0.950	0.987	0.985	0.970
Scaffold-GS	0.992	0.985	0.898	0.960	0.980	0.950	0.985	0.983	0.967
Ours	0.993	0.988	0.907	0.961	0.983	0.955	0.987	0.985	0.970
Ours w/ pruning	0.993	0.988	0.905	0.959	0.982	0.953	0.987	0.985	0.969

Table 13: Quantitative comparison of LPIPS values on Synthetic NeRF [6] dataset

Method | Scene	Mic	Chair	Ship	Materials	Lego	Drums	Ficus	Hotdog	Avg.
Mip-NeRF	0.009	0.021	0.138	0.040	0.021	0.065	0.020	0.027	0.043
Tri-MipRF	0.008	0.016	0.136	0.052	0.016	0.066	0.020	0.021	0.042
GS-Shader	0.006	0.012	0.098	0.033	0.014	0.040	0.013	0.019	0.029
3DGS	0.006	0.010	0.106	0.037	0.016	0.037	0.012	0.020	0.031
Scaffold-GS	0.008	0.013	0.114	0.040	0.019	0.042	0.013	0.023	0.034
Ours	0.006	0.010	0.108	0.037	0.016	0.035	0.012	0.020	0.030
Ours w/ pruning	0.006	0.010	0.113	0.039	0.018	0.037	0.012	0.021	0.032

Table 14: Quantitative comparison of PSNR on Spaces [2] dataset

Method	Scene	Scene 0	Scene 10	Scene 23	Scene 24	Scene 56	Scene 62	Scene 63	Scene 73	Avg.
3D-GS	30.69	27.98	26.01	43.98	29.91	32.17	32.07	30.77	31.70
Nex	37.61	37.61	35.69	37.77	34.77	35.34	35.44	34.81	36.13
Scaffold-GS	34.35	40.23	34.80	41.51	34.53	39.48	37.66	36.70	37.41
Ours	34.61	42.42	39.09	43.57	34.28	41.33	38.64	37.58	38.94
Ours w/ pruning	37.05	43.10	41.63	46.33	35.43	42.39	40.30	38.82	40.63

Table 15: Quantitative comparison of SSIM on Spaces [2] dataset

Method	Scene	Scene 0	Scene 10	Scene 23	Scene 24	Scene 56	Scene 62	Scene 63	Scene 73	Avg.
3D-GS	0.908	0.900	0.864	0.988	0.922	0.930	0.968	0.892	0.922
Nex	0.989	0.989	0.986	0.989	0.981	0.984	0.987	0.986	0.986
Scaffold-GS	0.959	0.981	0.969	0.987	0.959	0.979	0.980	0.977	0.974
Ours	0.962	0.982	0.977	0.989	0.954	0.979	0.981	0.979	0.975
Ours w/ pruning	0.974	0.984	0.982	0.991	0.961	0.980	0.984	0.981	0.980

Table 16: Quantitative comparison of LPIPS on Spaces [2] dataset

Method	Scene	Scene 0	Scene 10	Scene 23	Scene 24	Scene 56	Scene 62	Scene 63	Scene 73	Avg.
3D-GS	0.203	0.203	0.270	0.025	0.137	0.202	0.060	0.116	0.152
Nex	0.049	0.095	0.098	0.090	0.087	0.121	0.073	0.065	0.085
Scaffold-GS	0.055	0.031	0.058	0.027	0.072	0.061	0.031	0.035	0.046
Ours	0.059	0.039	0.066	0.026	0.083	0.076	0.038	0.038	0.053
Ours w/ pruning	0.046	0.042	0.061	0.027	0.079	0.080	0.037	0.036	0.051

Table 17: Quantitative comparison of FPS and Memory usage on Spaces [2] dataset

Dataset	Scene 0		Scene 10		Scene 23		Scene 24		Scene 56		Scene 62		Scene 63		Scene 73		Avg.	
Method | Metric	FPS↑	Mem↓	FPS↑	Mem↓	FPS↑	Mem↓	FPS↑	Mem↓	FPS↑	Mem↓	FPS↑	Mem↓	FPS↑	Mem↓	FPS↑	Mem↓	FPS↑	Mem↓
3D-GS	490	84MB	542	70MB	442	75MB	493	66MB	481	104MB	614	37MB	426	124MB	468	125MB	495	86MB
Scaffold-GS	102	85MB	144	54MB	120	63MB	179	43MB	146	108MB	120	62MB	91	95MB	107	101MB	126	76MB
Ours	314	122MB	465	42MB	306	72MB	537	51MB	414	122MB	504	44MB	400	77MB	427	83MB	421	77MB
Ours w/ pruning	484	50MB	592	18MB	514	30MB	699	22MB	497	50MB	599	19MB	508	32MB	513	34MB	551	32MB

Table 18: Quantitative comparison of FPS and Memory usage on Shiny [7] dataset

Dataset	Shiny		Lab		Tools		Giants		Avg.	
Method | Metric	FPS↑	MEM↓	FPS↑	MEM↓	FPS↑	MEM↓	FPS↑	MEM↓	FPS↑	MEM↓
3D-GS	239	159MB	256	133MB	234	310MB	156	680MB	221	321MB
Scaffold-GS	97	120MB	66	184MB	51	284MB	52	308MB	66.47	224MB
Ours	294	114MB	208	101MB	229	300MB	167	547MB	225	266MB
Ours w/ pruning	345	47MB	481	42MB	337	113MB	250	220MB	353	105MB

REFERENCES

[1] Jonathan T. Barron, Ben Mildenhall, Dor Verbin, Pratul P. Srinivasan, and Peter Hedman. 2022. Mip-NeRF 360: Unbounded Anti-Aliased Neural Radiance Fields. *CVPR* (2022).

[2] John Flynn, Michael Broxton, Paul Debevec, Matthew DuVall, Graham Fyffe, Ryan Overbeck, Noah Snavely, and Richard Tucker. 2019. Deepview: View synthesis with learned gradient descent. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2367–2376.

[3] Yihui He, Xiangyu Zhang, and Jian Sun. 2017. Channel Pruning for Accelerating Very Deep Neural Networks. In *2017 IEEE International Conference on Computer Vision (ICCV)*. 1398–1406. https://doi.org/10.1109/ICCV.2017.155

[4] Arno Knapitsch, Jaesik Park, Qian-Yi Zhou, and Vladlen Koltun. 2017. Tanks and temples: Benchmarking large-scale scene reconstruction. *ACM Transactions on Graphics (ToG)* 36, 4 (2017), 1–13.

[5] Z. Liu, J. Li, Z. Shen, G. Huang, S. Yan, and C. Zhang. 2017. Learning Efficient Convolutional Networks through Network Slimming. In *2017 IEEE International Conference on Computer Vision (ICCV)*. IEEE Computer Society, Los Alamitos, CA, USA, 2755–2763. https://doi.org/10.1109/ICCV.2017.298

[6] Ben Mildenhall, Pratul P. Srinivasan, Matthew Tancik, Jonathan T. Barron, Ravi Ramamoorthi, and Ren Ng. 2021. NeRF: representing scenes as neural radiance fields for view synthesis. *Commun. ACM* 65, 1 (dec 2021), 99–106. https://doi.org/10.1145/3503250

[7] Suttisak Wizadwongsa, Pakkapon Phongthawee, Jiraphon Yenphraphai, and Supasorn Suwajanakorn. 2021. NeX: Real-time View Synthesis with Neural Basis Expansion. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.