Mip-Splatting: Alias-free 3D Gaussian Splatting
Supplementary Material

Zehao Yu[1,2] Anpei Chen[1,2] Binbin Huang[3] Torsten Sattler[4] Andreas Geiger[1,2]

[1]University of Tübingen [2]Tübingen AI Center [3]ShanghaiTech University
[4]Czech Technical University in Prague

https://niujinshuchong.github.io/mip-splatting

In this **supplementary document**, we first present ablation studies of Mip-Splatting in Section 1. Next, we report additional quantitative and quality results in Section 3.

1. Ablation

In this section, we evaluate the effectiveness of our 3D smoothing filter and 2D Mip filter in Section 1.1 and Section 1.2. Then, we present an additional experiment to evaluate both zoom-in and zoom-out effects in the same dataset in Section 1.3.

1.1. Effectiveness of the 3D Smoothing Filter

To evaluate the effectiveness of the 3D smoothing filter, we conduct an ablation with the single-scale training and multi-scale testing setting to simulate zoom-in effects in the Mip-NeRF 360 dataset [3]. The quantitative result is presented in Table 1. Omitting the 3D smoothing filter results in high-frequency artifacts when rendering higher resolution image, as depicted in Figure 1. Excluding the 2D Mip filter causes a slight decline in performance as this filter's role is mainly for mitigating zoom-out artifacts, as we will shown next. The absence of both the 3D smoothing filter and the 2D Mip filter leads to an excessive generation of small Gaussian primitives, due to the density control mechanism, resulting in out of memory error even on an A100 GPU with 40GB memory. Hence, we don't report the result.

1.2. Effectiveness of the 2D Mip Filter

To evaluate the effectiveness of the 2D Mip filter, we perform an ablation study with the single-scale training and multi-scale testing setting to simulate zoom-out effects in the Blender dataset [10]. The quantitative results are shown in Table 2. Upon removing the dilation operation from 3DGS [9] (*3DGS - Dilation*), the dilation effects are eliminated, outperforming 3DGS in this context. However, it also results in aliasing artifacts due to a lack of anti-aliasing. Mip-Splatting outperforms all baseline methods by a large margin. Removing the 2D Mip filter results in a notable decline in performance, validating its critical role in anti-aliasing. Without the 3D smoothing filter, it still produces alias-free rendering as the 3D filter aims at addressing the high-frequency artifacts when zooming in.

1.3. Single-scale Training and Multi-scale Testing

In the main paper, we evaluate the zoom-out effects by rendering lower resolution images on the Blender dataset [10] following [2, 8] and simulating the zoom-in effects by rendering higher resolution images on the Mip-NeRF 360 dataset [3]. Here we present an addition experiment evaluating both zoom-out and zoom-in effects on the Mip-NeRF 360 dataset [3]. We use the images downsampled by a factor of 4 for training and evaluate it at multiple resolutions ($1/4\times$, $1/2\times$, $1\times$, $2\times$, $4\times$). The quantitative results are presented in Table 3 and the qualitative comparison is shown in Figure 2. Mip-Splatting significantly outperforms 3DGS [9] and 3DGS + EWA [16] in rendering quality when zooming in and out, which is consistent to our main results. Further, removing our 3D smoothing filter leads to high-frequency artifacts, while removing our 2D Mip-filter results in aliasing artifacts, as evidenced in Figure 2.

2. Relation with Airy disk and 2D Mip filter

The Airy disk is a result of diffraction in optical systems. It describes how a point light in the scene is blurred by the imaging system before being recorded by the sensor. It represents the smallest area in the image (or maximum frequency in the scene) that can be resolved as a distinct feature [12]. In other words, it determines the smallest theoretical "pixel" in the image [1].

A pixel integrates over all incoming light rays that fall onto the pixel area and hence acts as a "2D box filter". In contrast to the Airy disk which is defined by the camera lens, the pixel size is determined by the image sensor. For the datasets (e.g., Mip-NeRF 360) and downsampling rate

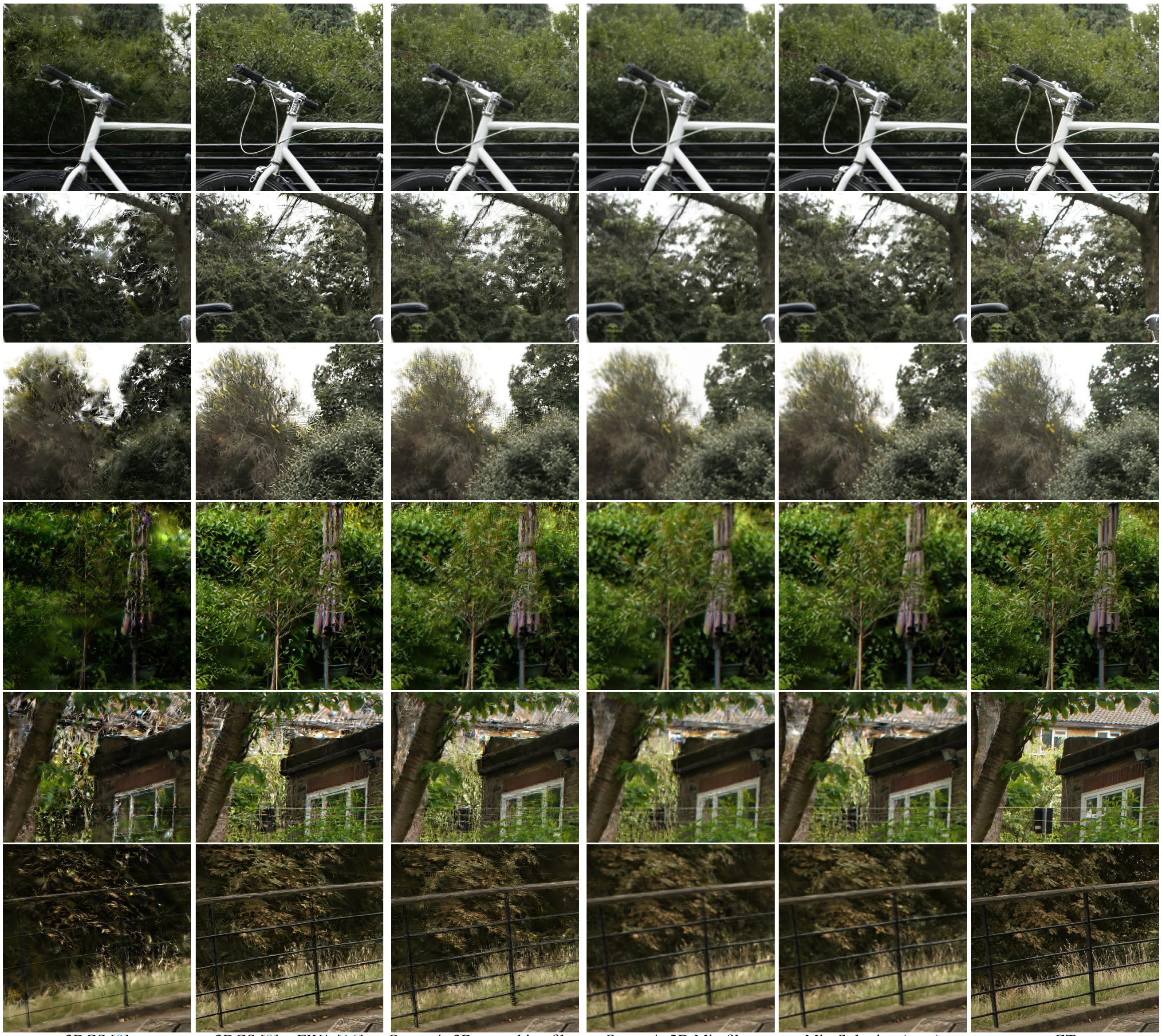

Figure 1. **Single-scale Training and Multi-scale Testing on the Mip-NeRF 360 Dataset [3].** All models are trained on images downsampled by a factor of 8 and rendered at full resolution to demonstrate zoom-in/moving closer effects. Removing the 3D smoothing filter results in high-frequency artifacts. Mip-Splatting renders images that closely approximate ground truth. Zoom in for a better view.

	PSNR ↑					SSIM ↑					LPIPS ↓				
	1× Res.	2× Res.	4× Res.	8× Res.	Avg.	1× Res.	2× Res.	4× Res.	8× Res.	Avg.	1× Res.	2× Res.	4× Res.	8× Res.	Avg.
3DGS [9]	29.19	23.50	20.71	19.59	23.25	0.880	0.740	0.619	0.619	0.715	0.107	0.243	0.394	0.476	0.305
3DGS [9] + EWA [16]	29.30	25.90	23.70	22.81	25.43	0.880	0.775	0.667	0.643	0.741	0.114	0.236	0.369	0.449	0.292
Mip-Splatting (ours)	29.39	27.39	26.47	26.22	27.37	0.884	0.808	0.754	0.765	0.803	0.108	0.205	0.305	0.392	0.252
Mip-Splatting (ours) w/o 3D smoothing filter	29.41	27.09	25.83	25.38	26.93	0.881	0.795	0.722	0.713	0.778	0.107	0.214	0.342	0.424	0.272
Mip-Splatting (ours) w/o 2D Mip filter	29.29	27.22	26.31	26.08	27.23	0.882	0.798	0.742	0.759	0.795	0.107	0.214	0.319	0.407	0.262

Table 1. **Single-scale Training and Multi-scale Testing on the Mip-NeRF 360 Dataset [3].** All methods are trained on the smallest scale (1×) and evaluated across four scales (1×, 2×, 4×, and 8×), with evaluations at higher sampling rates simulating zoom-in effects. While our method yields comparable results at the training resolution, it significantly surpasses all previous work at all other scales. Omitting the 3D smoothing filter results in high-frequency artifacts when rendering higher resolution image as shown in 1, while the excluding the 2D Mip filter only causes a slight decline in performance as this filter's role is mainly for mitigating zoom-out artifacts.

	PSNR ↑					SSIM ↑					LPIPS ↓				
	Full Res.	1/2 Res.	1/4 Res.	1/8 Res.	Avg.	Full Res.	1/2 Res.	1/4 Res.	1/8 Res.	Avg.	Full Res.	1/2 Res.	1/4 Res.	1/8 Res	Avg.
3DGS [9]	33.33	26.95	21.38	17.69	24.84	0.969	0.949	0.875	0.766	0.890	0.030	0.032	0.066	0.121	0.063
3DGS [9] + EWA [16]	33.51	31.66	27.82	24.63	29.40	0.969	0.971	0.959	0.940	0.960	0.032	0.024	0.033	0.047	0.034
3DGS [9] - Dilation	33.38	33.06	29.68	26.19	30.58	0.969	0.973	0.964	0.945	0.963	0.030	0.024	0.041	0.075	0.042
Mip-Splatting (ours)	33.36	34.00	31.85	28.67	31.97	0.969	0.977	0.978	0.973	0.974	0.031	0.019	0.019	0.026	0.024
Mip-Splatting (ours) w/o 3D smoothing filter	33.67	34.16	31.56	28.20	31.90	0.970	0.977	0.978	0.971	0.974	0.030	0.018	0.019	0.027	0.024
Mip-Splatting (ours) w/o 2D Mip filter	33.51	33.38	29.87	26.28	30.76	0.970	0.975	0.966	0.946	0.964	0.031	0.022	0.039	0.073	0.041

Table 2. **Single-scale Training and Multi-scale Testing on the Blender Dataset [10].** All methods are trained on full-resolution images and evaluated at four different (smaller) resolutions, with lower resolutions simulating zoom-out effects. While Mip-Splatting yields comparable results at training resolution, it significantly surpasses previous work at all other scales. Removing the 2D Mip filter results in a notable decline in performance at lower resolutions, validating its critical role in anti-aliasing. Removing the 3D smoothing filter achieves similar performance since the 3D filter aims at addressing the high-frequency artifacts when zooming in.

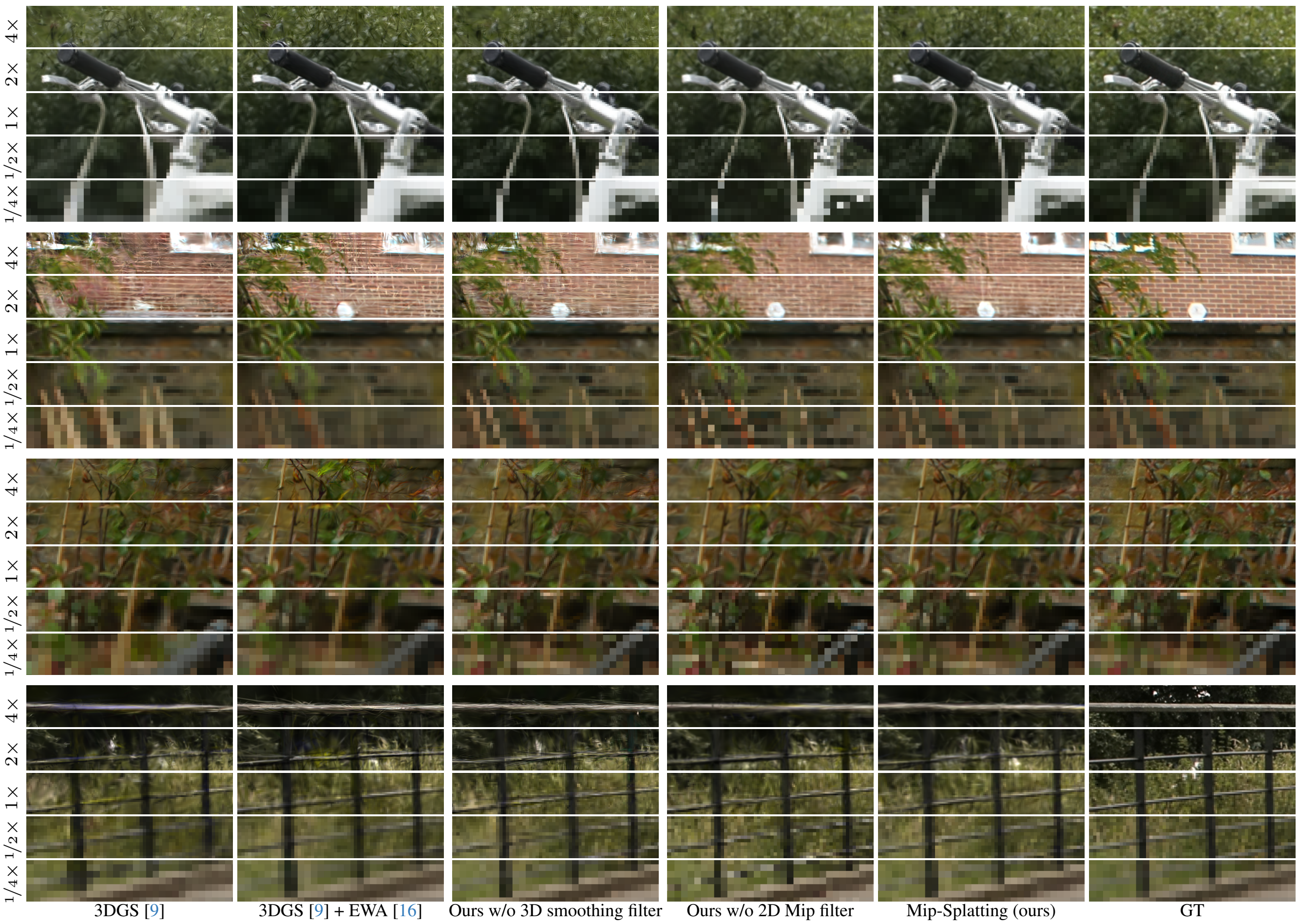

Figure 2. **Single-scale Training and Multi-scale Testing on the Mip-NeRF 360 Dataset [3].** All methods are trained at 1× resolution and evaluated at different resolutions to mimic zoom-out (1/4× and 1/2×) and zoom-in (2× and 4×). Mip-Splatting surpasses both 3DGS [9] and 3DGS + EWA [16] across different resolutions. Removing 3D smoothing filter leads to high-frquency artifacts when zooming in, while omitting 2D Mip filter results in aliasing artifacts when zooming out.

(4x) we consider in our experiments, the Airy disk is smaller than the pixel size and hence does not play a significant role.

Modeling the Airy disk or PSF is a promising future direction, especially when using high resolution images where the diffraction limits are reached. In this case, a more accurate model of the imaging process might lead to more accurate reconstructions. However, additional model parameters might also lead to an increase in optimization overhead. Note that other factors such as focus accuracy, motion blur and imperfect lenses also affect the results [1].

3. Additional Results

In this section, we provide more qualitative and quantitative results on the Blender dataset [10] in Section 3.1 and the

	PSNR ↑						SSIM ↑						LPIPS ↓					
	1/4 Res.	1/2 Res.	1× Res.	2× Res.	4× Res.	Avg.	1/4 Res.	1/2 Res.	1× Res.	2× Res.	4× Res.	Avg.	1/4 Res.	1/2 Res.	1× Res.	2× Res.	4× Res.	Avg.
3DGS [9]	20.85	24.66	28.01	25.08	23.37	24.39	0.681	0.812	0.834	0.766	0.735	0.765	0.203	0.158	0.166	0.275	0.383	0.237
3DGS [9] + EWA [16]	27.40	28.39	28.09	26.43	25.30	27.12	0.888	0.871	0.833	0.774	0.738	0.821	0.103	0.126	0.171	0.276	0.385	0.212
Mip-Splatting (ours)	28.98	29.02	28.09	27.25	26.95	28.06	0.908	0.880	0.835	0.798	0.800	0.844	0.086	0.114	0.168	0.248	0.331	0.189
Mip-Splatting (ours) w/o 3D smoothing filter	28.69	28.94	28.05	27.06	26.61	27.87	0.905	0.879	0.833	0.790	0.780	0.837	0.088	0.115	0.168	0.261	0.359	0.198
Mip-Splatting (ours) w/o 2D Mip filter	26.09	28.04	28.05	27.27	27.00	27.29	0.815	0.856	0.834	0.798	0.802	0.821	0.167	0.132	0.167	0.249	0.335	0.210

Table 3. **Single-scale Training and Multi-scale Testing on the Mip-NeRF 360 Dataset [3].** All methods are trained on the middle scale (1×) and evaluated across four scales (1/4×, 1/2×, 1×, 2×, and 4×), with evaluations at higher sampling rates simulating zoom-in effects. While our method yields comparable results at the training resolution, it significantly surpasses all previous work at all other scales. Omitting the 3D smoothing filter results in high-frequency artifacts when rendering higher resolution images, while removing the 2D Mip filter results in aliasing artifacts when rendering lower resolution images, as shown in Figure 2.

Mip-NeRF 360 dataset [3] in Section 3.2.

3.1. Blender Dataset

We evaluate Mip-Splatting under two different settings in the Blender dataset [10]. For *multi-scale training and multi-scale testing*, the quantitative results are compiled in Table 4, where Mip-Splatting achieves state-of-the-art performance. Additionally, per-scene metrics for *single-scale training and multi-scale testing* are presented in Table 5. A qualitative comparison against leading methods is shown in Figure 3. Mip-Splatting outperforms both 3DGS [9] and 3DGS + EWA [16], particularly noticeable when zooming out, i.e. at lower resolutions.

3.2. Mip-NeRF 360 Dataset

We further evaluate Mip-Splatting on the Mip-NeRF 360 dataset [3] across two experimental setups. In the first setup, we follow the standard approach where models are trained and evaluated at the same scale, with indoor scenes downsampled by a factor of two and outdoor scenes by four. Quantitative results with per-scene metrics are shown in Table 6, our method performs on par with 3DGS [9] and 3DGS + EWA [16] in this challenging benchmark, without any decrease in performance.

In the second setup, models are trained on data downsampled by a factor of 8 and rendered at successively higher resolutions (1×, 2×, 4×, and 8×) to simulate zoom-in effects. The quantitative results with per-scene metrics can be found in Table 7. Qualitative comparison with state-of-the-art methods are provided in Figure 4. Mip-Splatting effectively eliminates high-frequency artifacts, yielding high quality renderings that more closely resemble ground truth.

References

[1] Lens diffraction & photography. https://www.cambridgeincolour.com/tutorials/diffraction-photography.htm. 1, 3

[2] 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. *ICCV*, 2021. 1, 5, 6, 7, 9

[3] Jonathan T. Barron, Ben Mildenhall, Dor Verbin, Pratul P. Srinivasan, and Peter Hedman. Mip-nerf 360: Unbounded anti-aliased neural radiance fields. *CVPR*, 2022. 1, 2, 3, 4, 7, 8, 10

[4] Jonathan T. Barron, Ben Mildenhall, Dor Verbin, Pratul P. Srinivasan, and Peter Hedman. Zip-nerf: Anti-aliased grid-based neural radiance fields. *Proc. of the IEEE International Conf. on Computer Vision (ICCV)*, 2023. 7, 8, 10

[5] Anpei Chen, Zexiang Xu, Andreas Geiger, Jingyi Yu, and Hao Su. Tensorf: Tensorial radiance fields. 2022. 5, 6

[6] Boyang Deng, Jonathan T. Barron, and Pratul P. Srinivasan. JaxNeRF: an efficient JAX implementation of NeRF, 2020. 7

[7] Sara Fridovich-Keil, Alex Yu, Matthew Tancik, Qinhong Chen, Benjamin Recht, and Angjoo Kanazawa. Plenoxels: Radiance fields without neural networks. In *CVPR*, 2022. 5, 7

[8] Wenbo Hu, Yuling Wang, Lin Ma, Bangbang Yang, Lin Gao, Xiao Liu, and Yuewen Ma. Tri-miprf: Tri-mip representation for efficient anti-aliasing neural radiance fields. In *Proc. of the IEEE International Conf. on Computer Vision (ICCV)*, 2023. 1, 5, 6, 9

[9] Bernhard Kerbl, Georgios Kopanas, Thomas Leimkühler, and George Drettakis. 3d gaussian splatting for real-time radiance field rendering. *ACM Transactions on Graphics*, 42 (4), 2023. 1, 2, 3, 4, 5, 6, 7, 8, 9, 10

[10] Ben Mildenhall, Pratul P. Srinivasan, Matthew Tancik, Jonathan T. Barron, Ravi Ramamoorthi, and Ren Ng. Nerf: Representing scenes as neural radiance fields for view synthesis. In *ECCV*, 2020. 1, 3, 4, 5, 6, 7, 9

[11] Ben Mildenhall, Dor Verbin, Pratul P. Srinivasan, Peter Hedman, Ricardo Martin-Brualla, and Jonathan T. Barron. MultiNeRF: A Code Release for Mip-NeRF 360, Ref-NeRF, and RawNeRF, 2022. 7

[12] Ansh Mittal. Neural radiance fields: Past, present, and future. *arXiv preprint arXiv:2304.10050*, page 3, 2023. 1

[13] Thomas Müller, Alex Evans, Christoph Schied, and Alexander Keller. Instant neural graphics primitives with a multiresolution hash encoding. *ACM Trans. Graph.*, 41(4):102:1–102:15, 2022. 5, 6, 7, 8

[14] Lior Yariv, Peter Hedman, Christian Reiser, Dor Verbin, Pratul P. Srinivasan, Richard Szeliski, Jonathan T. Barron, and Ben Mildenhall. Bakedsdf: Meshing neural sdfs for real-time view synthesis. *arXiv*, 2023. 7

[15] Kai Zhang, Gernot Riegler, Noah Snavely, and Vladlen

	PSNR								
	chair	*drums*	*ficus*	*hotdog*	*lego*	*materials*	*mic*	*ship*	*Average*
NeRF w/o $\mathcal{L}_{\text{area}}$ [2, 10]	29.92	23.27	27.15	32.00	27.75	26.30	28.40	26.46	27.66
NeRF [10]	33.39	25.87	30.37	35.64	31.65	30.18	32.60	30.09	31.23
MipNeRF [2]	37.14	27.02	33.19	39.31	35.74	32.56	38.04	33.08	34.51
Plenoxels [7]	32.79	25.25	30.28	34.65	31.26	28.33	31.53	28.59	30.34
TensoRF [5]	32.47	25.37	31.16	34.96	31.73	28.53	31.48	29.08	30.60
Instant-ngp [13]	32.95	26.43	30.41	35.87	31.83	29.31	32.58	30.23	31.20
Tri-MipRF [8]*	37.67	27.35	33.57	38.78	35.72	31.42	37.63	32.74	34.36
3DGS [9]	32.73	25.30	29.00	35.03	29.44	27.13	31.17	28.33	29.77
3DGS [9] + EWA [16]	35.77	27.14	33.65	37.74	32.75	30.21	35.21	31.63	33.01
Mip-Splatting (ours)	37.48	27.74	34.71	39.15	35.07	31.88	37.68	32.80	34.56

	SSIM								
	chair	*drums*	*ficus*	*hotdog*	*lego*	*materials*	*mic*	*ship*	*Average*
NeRF w/o $\mathcal{L}_{\text{area}}$ [2, 10]	0.944	0.891	0.942	0.959	0.926	0.934	0.958	0.861	0.927
NeRF [10]	0.971	0.932	0.971	0.979	0.965	0.967	0.980	0.900	0.958
MipNeRF [2]	0.988	0.945	0.984	0.988	0.984	0.977	0.993	0.922	0.973
Plenoxels [7]	0.968	0.929	0.972	0.976	0.964	0.959	0.979	0.892	0.955
TensoRF [5]	0.967	0.930	0.974	0.977	0.967	0.957	0.978	0.895	0.956
Instant-ngp [13]	0.971	0.940	0.973	0.979	0.966	0.959	0.981	0.904	0.959
Tri-MipRF [8]*	0.990	0.951	0.985	0.988	0.986	0.969	0.992	0.929	0.974
3DGS [9]	0.976	0.941	0.968	0.982	0.964	0.956	0.979	0.910	0.960
3DGS [9] + EWA [16]	0.986	0.958	0.988	0.988	0.979	0.972	0.990	0.929	0.974
Mip-Splatting (ours)	0.991	0.963	0.990	0.990	0.987	0.978	0.994	0.936	0.979

	LPIPS								
	chair	*drums*	*ficus*	*hotdog*	*lego*	*materials*	*mic*	*ship*	*Average*
NeRF w/o $\mathcal{L}_{\text{area}}$ [2, 10]	0.035	0.069	0.032	0.028	0.041	0.045	0.031	0.095	0.052
NeRF [10]	0.028	0.059	0.026	0.024	0.035	0.033	0.025	0.085	0.044
MipNeRF [2]	0.011	0.044	0.014	0.012	0.013	0.019	0.007	0.062	0.026
Plenoxels [7]	0.040	0.070	0.032	0.037	0.038	0.055	0.036	0.104	0.051
TensoRF [5]	0.042	0.075	0.032	0.035	0.036	0.063	0.040	0.112	0.054
Instant-ngp [13]	0.035	0.066	0.029	0.028	0.040	0.051	0.032	0.095	0.047
Tri-MipRF [8]*	0.011	0.046	0.016	0.014	0.013	0.033	0.008	0.069	0.026
3DGS [9]	0.025	0.056	0.030	0.022	0.038	0.040	0.023	0.086	0.040
3DGS [9] + EWA [16]	0.017	0.039	0.013	0.016	0.024	0.026	0.011	0.070	0.027
Mip-Splatting (ours)	0.010	0.031	0.009	0.011	0.012	0.018	0.005	0.059	0.019

Table 4. **Multi-scale Training and Multi-scale Testing on the the Blender dataset [10]**. For each scene, we report the arithmetic mean of each metric averaged over the 4 scales used in the dataset.

Koltun. Nerf++: Analyzing and improving neural radiance fields. *arXiv:2010.07492*, 2020. 7

[16] Matthias Zwicker, Hanspeter Pfister, Jeroen Van Baar, and Markus Gross. Ewa volume splatting. In *Proceedings Visualization, 2001. VIS'01.*, pages 29–538. IEEE, 2001. 1, 2, 3, 4, 5, 6, 7, 8, 9, 10

	PSNR								
	chair	*drums*	*ficus*	*hotdog*	*lego*	*materials*	*mic*	*ship*	*Average*
NeRF [10]	31.99	25.31	30.74	34.45	30.69	28.86	31.41	28.36	30.23
MipNeRF [2]	32.89	25.58	31.80	35.40	32.24	29.46	33.26	29.88	31.31
TensoRF [5]	32.17	25.51	31.19	34.69	31.46	28.60	31.50	28.71	30.48
Instant-ngp [13]	32.18	25.05	31.32	34.85	31.53	28.59	32.15	28.84	30.57
Tri-MipRF [8]	32.48	24.01	28.41	34.45	30.41	27.82	31.19	27.02	29.47
3DGS [9]	26.81	21.17	26.02	28.80	25.36	23.10	24.39	23.05	24.84
3DGS [9] + EWA [16]	32.85	24.91	31.94	33.33	29.76	27.36	27.68	27.41	29.40
Mip-Splatting (ours)	35.69	26.50	32.99	36.18	32.76	30.01	31.66	29.98	31.97

	SSIM								
	chair	*drums*	*ficus*	*hotdog*	*lego*	*materials*	*mic*	*ship*	*Average*
NeRF [10]	0.968	0.936	0.976	0.977	0.963	0.964	0.980	0.887	0.956
MipNeRF [2]	0.974	0.939	0.981	0.982	0.973	0.969	0.987	0.915	0.965
TensoRF [5]	0.970	0.938	0.978	0.979	0.970	0.963	0.981	0.906	0.961
Instant-ngp [13]	0.970	0.935	0.977	0.980	0.969	0.962	0.982	0.909	0.961
Tri-MipRF [8]	0.971	0.908	0.957	0.975	0.957	0.953	0.975	0.883	0.947
3DGS [9]	0.915	0.851	0.921	0.930	0.882	0.882	0.909	0.827	0.890
3DGS [9] + EWA [16]	0.978	0.942	0.983	0.977	0.964	0.958	0.963	0.912	0.960
Mip-Splatting (ours)	0.988	0.958	0.988	0.987	0.982	0.974	0.986	0.930	0.974

	LPIPS								
	chair	*drums*	*ficus*	*hotdog*	*lego*	*materials*	*mic*	*ship*	*Average*
NeRF [10]	0.040	0.067	0.027	0.034	0.043	0.049	0.035	0.132	0.053
MipNeRF [2]	0.033	0.062	0.022	0.025	0.030	0.041	0.023	0.092	0.041
TensoRF [5]	0.036	0.066	0.027	0.030	0.035	0.052	0.034	0.102	0.048
Instant-ngp [13]	0.036	0.074	0.035	0.030	0.035	0.054	0.034	0.096	0.049
Tri-MipRF [8]	0.026	0.086	0.041	0.023	0.036	0.048	0.023	0.117	0.050
3DGS [9]	0.047	0.087	0.055	0.034	0.064	0.055	0.046	0.113	0.063
3DGS [9] + EWA [16]	0.023	0.051	0.017	0.018	0.033	0.027	0.024	0.077	0.034
Mip-Splatting (ours)	0.014	0.035	0.012	0.014	0.016	0.019	0.015	0.066	0.024

Table 5. **Single-scale Training and Multi-scale Testing on the the Blender dataset [10]**. For each scene, we report the arithmetic mean of each metric averaged over the four scales used in the dataset.

	PSNR								
	bicycle	*flowers*	*garden*	*stump*	*treehill*	*room*	*counter*	*kitchen*	*bonsai*
NeRF [6, 10]	21.76	19.40	23.11	21.73	21.28	28.56	25.67	26.31	26.81
mip-NeRF [2]	21.69	19.31	23.16	23.10	21.21	28.73	25.59	26.47	27.13
NeRF++ [15]	22.64	20.31	24.32	24.34	22.20	28.87	26.38	27.80	29.15
Plenoxels [7]	21.91	20.10	23.49	20.661	22.25	27.59	23.62	23.42	24.67
Instant NGP [13, 14]	22.79	19.19	25.26	24.80	22.46	30.31	26.21	29.00	31.08
mip-NeRF 360 [3, 11]	24.40	21.64	26.94	26.36	22.81	31.40	29.44	32.02	33.11
Zip-NeRF [4]	25.80	22.40	28.20	27.55	23.89	32.65	29.38	32.50	34.46
3DGS [9]	25.25	21.52	27.41	26.55	22.49	30.63	28.70	30.32	31.98
3DGS [9]*	25.63	21.77	27.70	26.87	22.75	31.69	29.08	31.56	32.29
3DGS [9] + EWA [16]	25.64	21.86	27.65	26.87	22.91	31.68	29.21	31.59	32.51
Mip-Splatting (ours)	25.72	21.93	27.76	26.94	22.98	31.74	29.16	31.55	32.31

	SSIM								
	bicycle	*flowers*	*garden*	*stump*	*treehill*	*room*	*counter*	*kitchen*	*bonsai*
NeRF [6, 10]	0.455	0.376	0.546	0.453	0.459	0.843	0.775	0.749	0.792
mip-NeRF [2]	0.454	0.373	0.543	0.517	0.466	0.851	0.779	0.745	0.818
NeRF++ [15]	0.526	0.453	0.635	0.594	0.530	0.852	0.802	0.816	0.876
Plenoxels [7]	0.496	0.431	0.606	0.523	0.509	0.842	0.759	0.648	0.814
Instant NGP [13, 14]	0.540	0.378	0.709	0.654	0.547	0.893	0.845	0.857	0.924
mip-NeRF 360 [3, 11]	0.693	0.583	0.816	0.746	0.632	0.913	0.895	0.920	0.939
Zip-NeRF [4]	0.769	0.642	0.860	0.800	0.681	0.925	0.902	0.928	0.949
3DGS [9]	0.771	0.605	0.868	0.775	0.638	0.914	0.905	0.922	0.938
3DGS [9]*	0.777	0.622	0.873	0.783	0.652	0.928	0.916	0.933	0.948
3DGS [9] + EWA [16]	0.777	0.620	0.871	0.784	0.655	0.927	0.916	0.933	0.948
Mip-Splatting (ours)	0.780	0.623	0.875	0.786	0.655	0.928	0.916	0.933	0.948

	LPIPS								
	bicycle	*flowers*	*garden*	*stump*	*treehill*	*room*	*counter*	*kitchen*	*bonsai*
NeRF [6, 10]	0.536	0.529	0.415	0.551	0.546	0.353	0.394	0.335	0.398
mip-NeRF [2]	0.541	0.535	0.422	0.490	0.538	0.346	0.390	0.336	0.370
NeRF++ [15]	0.455	0.466	0.331	0.416	0.466	0.335	0.351	0.260	0.291
Plenoxels [7]	0.506	0.521	0.3864	0.503	0.540	0.419	0.441	0.447	0.398
Instant NGP [13, 14]	0.398	0.441	0.255	0.339	0.420	0.242	0.255	0.170	0.198
mip-NeRF 360 [3, 11]	0.289	0.345	0.164	0.254	0.338	0.211	0.203	0.126	0.177
Zip-NeRF [4]	0.208	0.273	0.118	0.193	0.242	0.196	0.185	0.116	0.173
3DGS [9]	0.205	0.336	0.103	0.210	0.317	0.220	0.204	0.129	0.205
3DGS [9]*	0.205	0.329	0.103	0.208	0.318	0.192	0.178	0.113	0.174
3DGS [9] + EWA [16]	0.213	0.335	0.111	0.210	0.325	0.192	0.179	0.113	0.173
Mip-Splatting (ours)	0.206	0.331	0.103	0.209	0.320	0.192	0.179	0.113	0.173

Table 6. **Single-scale Training and Single-scale Testing on the Mip-NeRF 360 dataset [3].** Indoor scenes are downsampled by a factor of 2 and outdoor scenes by 4.

	PSNR								
	bicycle	*flowers*	*garden*	*stump*	*treehill*	*room*	*counter*	*kitchen*	*bonsai*
Instant-NGP [13]	22.51	20.25	24.65	23.15	22.24	29.48	26.18	27.10	29.66
mip-NeRF 360 [3]	24.21	21.60	25.82	25.59	22.78	22.95	27.72	28.78	31.63
zip-NeRF [4]	23.05	20.05	18.07	23.94	22.53	20.51	26.08	27.37	30.05
3DGS [9]	21.34	19.43	21.94	22.63	20.91	28.10	25.33	23.68	25.89
3DGS [9] + EWA [16]	23.74	20.94	24.69	24.81	21.93	29.80	27.23	27.07	28.63
Mip-Splatting (ours)	25.26	22.02	26.78	26.65	22.92	31.56	28.87	30.73	31.49

	SSIM								
	bicycle	*flowers*	*garden*	*stump*	*treehill*	*room*	*counter*	*kitchen*	*bonsai*
Instant-NGP [13]	0.538	0.473	0.647	0.590	0.544	0.868	0.795	0.764	0.877
mip-NeRF 360 [3]	0.662	0.567	0.716	0.715	0.628	0.795	0.845	0.828	0.910
zip-NeRF [4]	0.640	0.521	0.548	0.661	0.590	0.655	0.784	0.800	0.865
3DGS [9]	0.638	0.536	0.675	0.662	0.591	0.878	0.826	0.789	0.838
3DGS [9] + EWA [16]	0.671	0.563	0.718	0.693	0.608	0.889	0.843	0.813	0.874
Mip-Splatting (ours)	0.738	0.613	0.786	0.776	0.659	0.921	0.897	0.903	0.933

	LPIPS								
	bicycle	*flowers*	*garden*	*stump*	*treehill*	*room*	*counter*	*kitchen*	*bonsai*
Instant-NGP [13]	0.500	0.486	0.372	0.469	0.511	0.270	0.310	0.286	0.229
mip-NeRF 360 [3]	0.358	0.400	0.296	0.333	0.391	0.256	0.228	0.210	0.182
zip-NeRF [4]	0.353	0.397	0.346	0.349	0.366	0.302	0.277	0.232	0.236
3DGS [9]	0.336	0.406	0.295	0.353	0.406	0.223	0.239	0.245	0.242
3DGS [9] + EWA [16]	0.322	0.395	0.281	0.334	0.405	0.217	0.231	0.216	0.227
Mip-Splatting (ours)	0.281	0.373	0.233	0.281	0.369	0.193	0.199	0.165	0.176

Table 7. **Single-scale Training and Multi-scale Testing on the Mip-NeRF 360 dataset [3]**. All models are trained on images downsampled by a factor of 8 and rendered at higher resolutions to simulates zoom-in effects.

Figure 3. **Single-scale Training and Multi-scale Testing on the Blender Dataset [10].** All methods are trained at full resolution and evaluated at different (smaller) resolutions to mimic zoom-out. Methods based on 3DGS capture fine details better than Mip-NeRF [2] and Tri-MipRF [8] at training resolution. Mip-Splatting surpasses both 3DGS [9] and 3DGS + EWA [16] at lower resolutions.

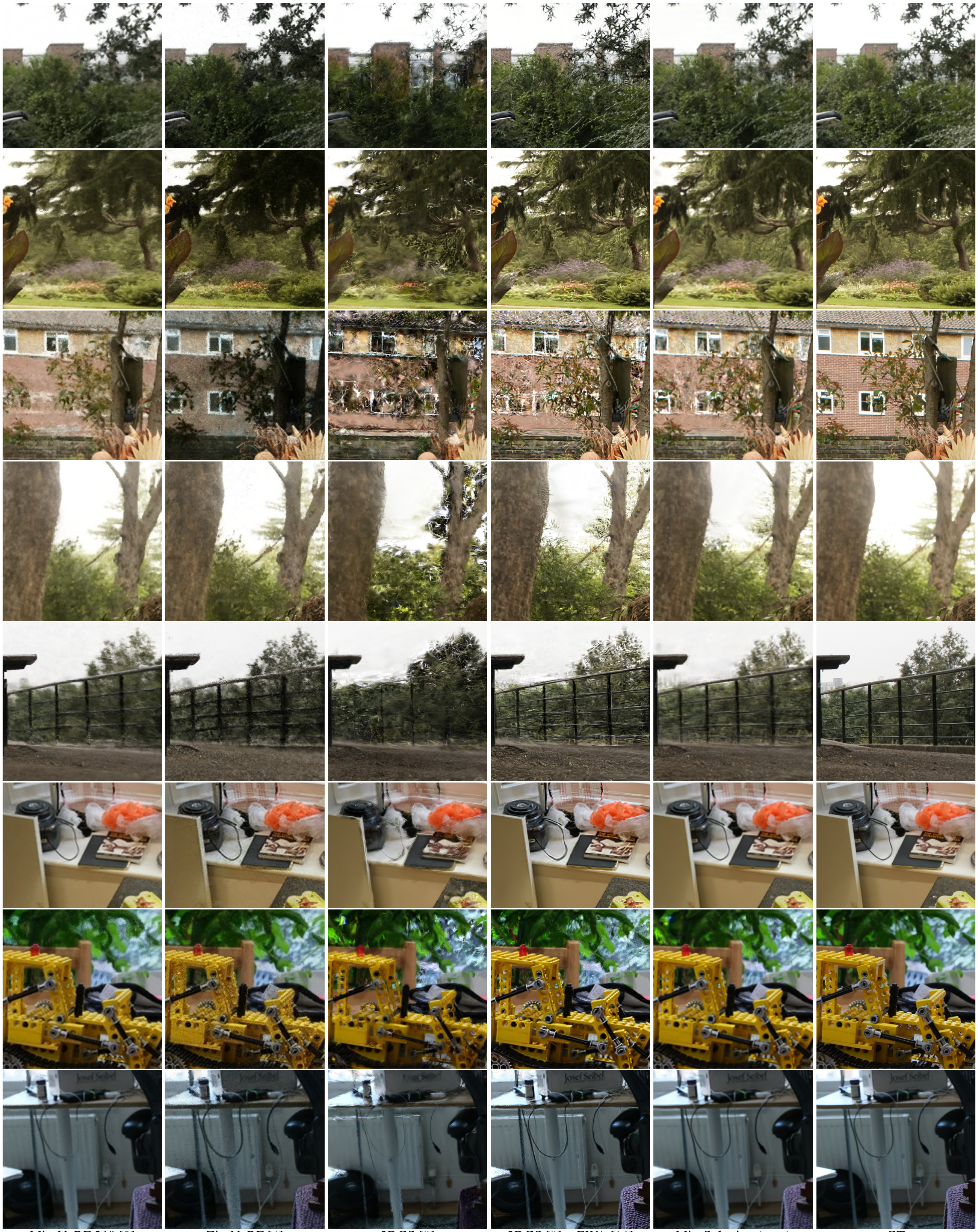

Figure 4. **Single-scale Training and Multi-scale Testing on the Mip-NeRF 360 Dataset [3].** All models are trained on images downsampled by a factor of eight and rendered at full resolution to demonstrate zoom-in/moving closer effects. In contrast to prior work, Mip-Splatting renders images that closely approximate ground truth. Please also note the high-frequency artifacts of 3DGS + EWA [16].