EAP-GS: Efficient Augmentation of Pointcloud for 3D Gaussian Splatting in Few-shot Scene Reconstruction

Supplementary Material

1. Implementation Details

For initialization, camera poses are assumed to be known based on the full-view SfM methods, following the conventions of few-shot settings. Aligning with the configuration used in FSGS, we executed COLMAP by the parameters shown in Tab. S.1 to produce a coarse pointcloud. To generate a fine pointcloud using DetectorfreeSfM, we configured *img_resize* to 2000 for the LLFF dataset and 1500 for the Mip-NeRF360 dataset, while keeping other parameters at their default values.

Table S.1. Initialization parameters for COLMAP.

Step	Parameter	Value
Feature Extraction	max image size	4032
	max num features	32768
	estimate affine shape	1
	domain size pooling	1
Feature Matching	max num matches	32768
	guided matching	1
Triangulation	ba local max num iterations	40
	ba local max refinements	3
	ba global max num iterations	100

Considering the converge speed of few-shot renconstruction, we trained all scenarios for 5,000 iterations. As a start, we focused on fully reconstructing the scene by the initial pointcloud for the first 2000 iterations since the initial pointcloud is relatively accurate. In the subsequent 2500 iterations, adaptive density control (ADC) is utilized to progressively increase the number of Gaussians every 200 iterations. During the final 500 iterations, we fine-tune the reconstruction with the number of Gaussians frozen. The opacity reset trick is not applied for the number of Guassians in the few-shot reconstruction usually won't be overly abundant. The training strategies for baseline methods remain unchanged. Learning rates are set to their default values.

2. Comparison of Initial Pointcloud

As shown in Fig. S.1, we visualize the coarse pointcloud $\mathcal{X}_{\mathrm{c}}$ and fine pointcloud $\mathcal{X}_{\mathrm{f}}$. $\mathcal{X}_{\mathrm{c}}$ is generated directly by the COLMAP and serves as the initial input for reconstructions of 3DGS and other baseline methods, while $\mathcal{X}_{\mathrm{f}}$ is derived from the Attentional Pointcloud Augmentation (APA) technique based on $\mathcal{X}_{\mathrm{c}}$ (more details can be found in Sec. 3.2 of main manuscript). Our technique significantly increases the number of initial points, with some scene points boosting nearly 10 times (e.g., *Bonsai*, *Garden*, *Kitchen*). Moreover, the points of $\mathcal{X}_{\mathrm{f}}$ have a better density distribution that matches the scene complexity. $\mathcal{X}_{\mathrm{f}}$ successfully generates 3D points in regions with weakly pronounced texture (e.g., the desktop in *Fortress*, the carpet in *Horns*, the tablecloths in *Kitchen*) and peripheral areas (e.g., the ceiling in *Fern*, the brown cabinet in *Counter*, the surrounding trees in *Stump*), which paves the way for high-quality reconstruction.

Comparison of the results reveals that a lot of information exists within the sparse views. The fine pointcloud $\mathcal{X}_{\mathrm{f}}$ can even reflect the rough structure of some scenes. For instance, in *Fortress* and *Kitchen*, the basic content and scene geometry are already well defined by the pointcloud $\mathcal{X}_{\mathrm{f}}$, with the exception of partially occluded areas. However, the COLMAP produces only a small number of accurate 3D feature points because of the limitations of view tracks, so that filters out significant amounts of information. Due to the view information is not fully explored during the initialization stage, it becomes very challenging to find an accurate optimization path in subsequent reconstruction. Therefore, by retaining two-view tracks as an option for pointcloud generation, we provide a high-quality and efficient pointcloud initialization for optimization process, ultimately leading to more accurate and stable reconstruction results.

3. Additional Results

Additional qualitative results for the LLFF and Mip-NeRF360 datasets are shown in Fig. S.2 and Fig. S.3. All methods successfully reconstruct the core scene geometry, which is consistent with the conclusions in Sec. 4.2 of main manuscript. However, 3DGS failed to produce acceptable results for some regions with weakly pronounced texture (marked by red boxes) and peripheral areas (marked by yellow boxes) that are typicially with sparsely distributed initial pointclouds (as shown in Fig. S.1), leading to unrealistic artifacts such as blur and floaters. The improvements are not obvious by other baseline methods. In contrast, our method effectively addresses this issue, and gives a reasonable interpretation of regions with weakly pronounced texture and peripheral areas, resulting in the best reconstruction performance. That is because the APA technique generates more points in regions with complex structure but sparse density distribution of the scene. This provides a good guidance to Gaussians generation in the reconstruction stage, allowing for capturing richer details, e.g., the carpet in *Horn*, the

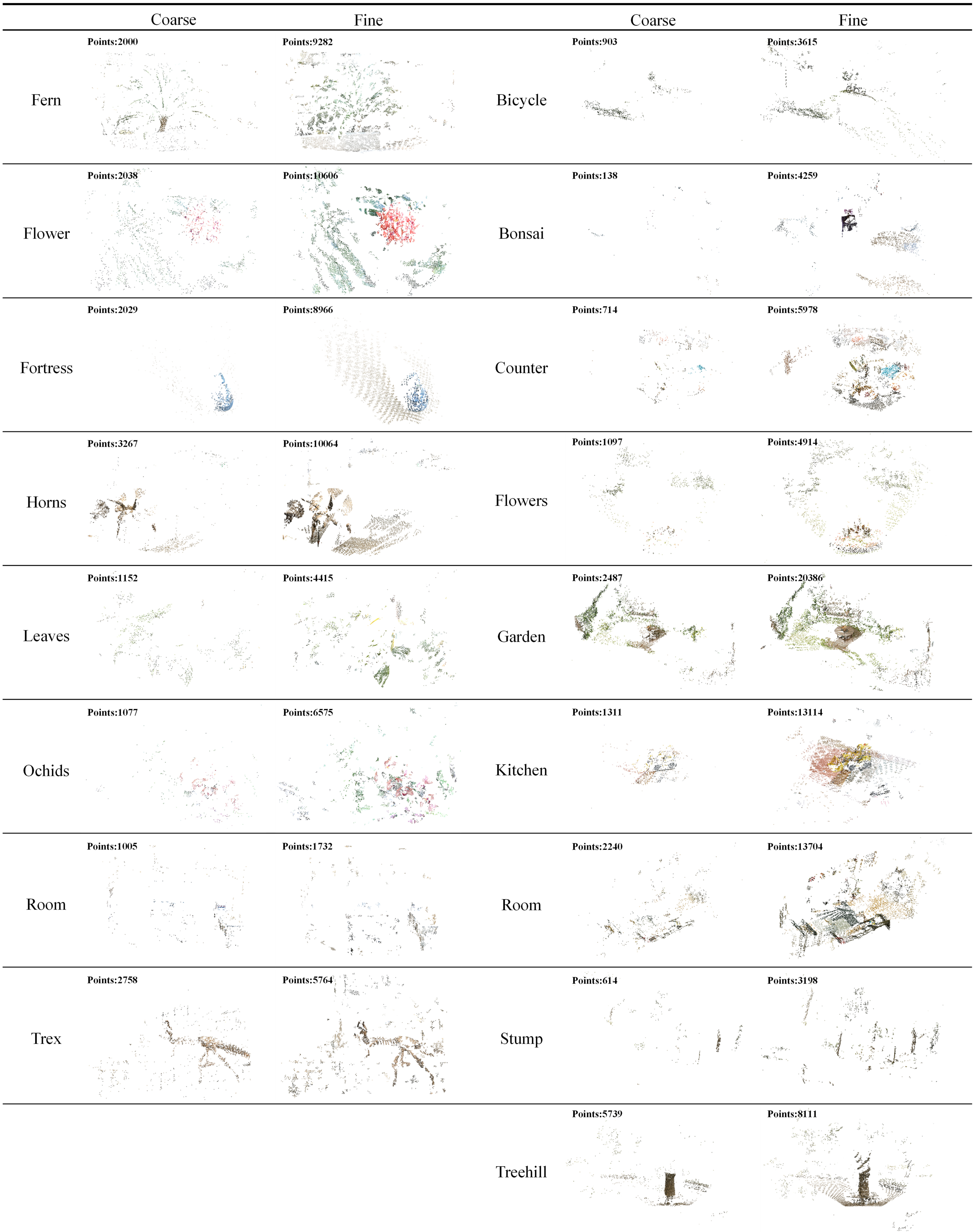

Figure S.1. Comparison of coarse and fine pointcloud. The coarse pointcloud serves as the initial condition for 3DGS and other baseline methods, while the fine pointcloud is utilized by our proposed EAP-GS.

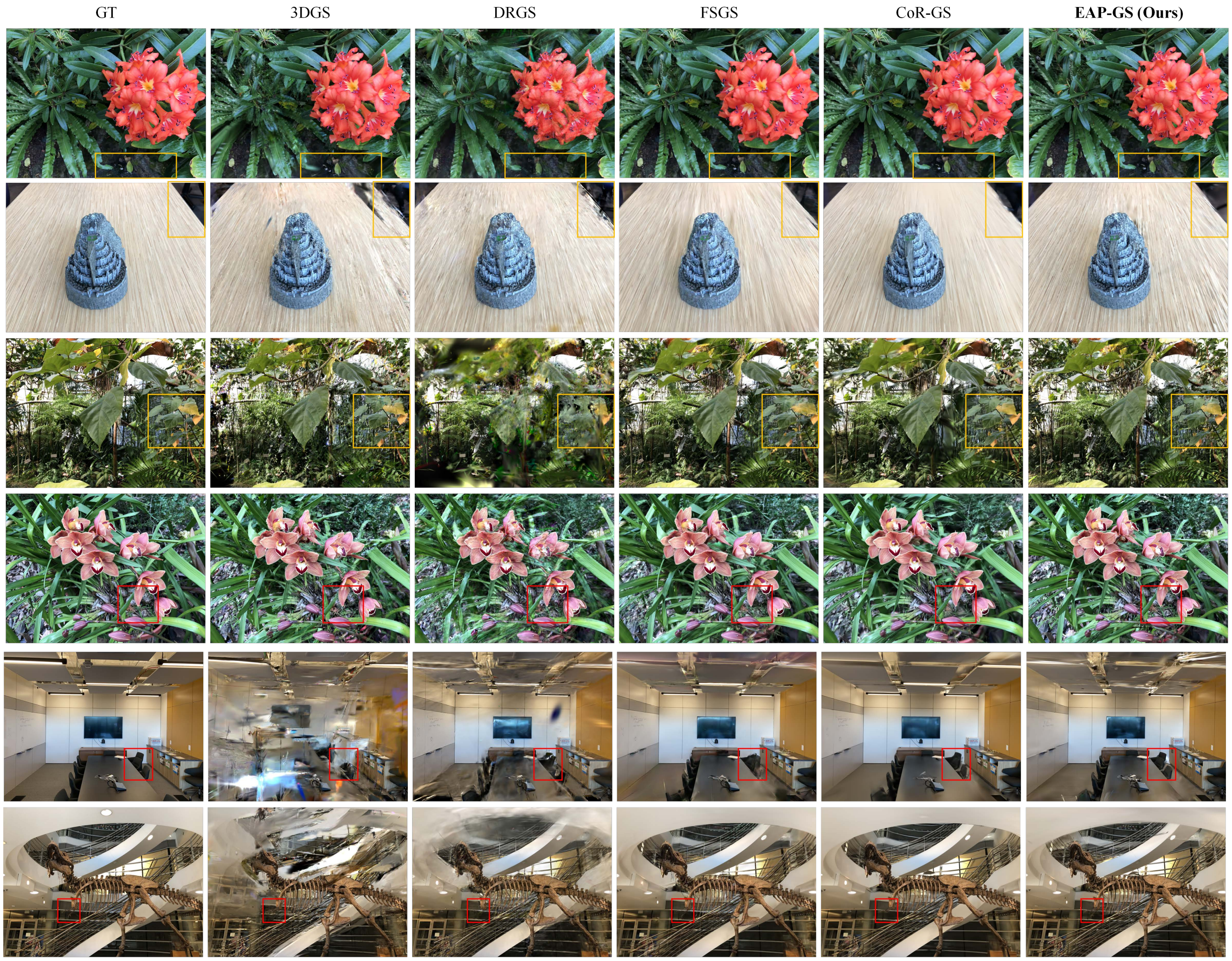

Figure S.2. Qualitative Comparison on the rest of the LLFF datasets. We demonstrate the qualitative comparison results with our main competitors 3DGS [13], DRGS [5], FSGS [40] and CoR-GS [35] on a certain testing view. The red and yellow boxes indicate the significant difference in reconstruction of the regions with weakly pronounced texture and peripheral areas, respectively.

grassland in *Bicycle* and *Flowers*, the glass lid in *Counter* and the chair in *Kitchen*.

In Tab. S.2 and Tab. S.3, we present quantitative results for each scene in the LLFF and Mip-NeRF360 datasets. Our method achieves leading scores in most scenarios while using fewer Gaussians and requiring less computation time. It is worth noting that our method obtains a relatively smaller improvement in PSNR compared to SSIM and LPIPS. That is because PSNR measures the mean squared error between two images and focus on smoothness, whereas SSIM and LPIPS emphasize the similarity of structure and details. As discussed, the APA technique significantly enhances the representation of scene details. As for the main parts of scene, all methods produce reasonable reconstruction results in most cases.

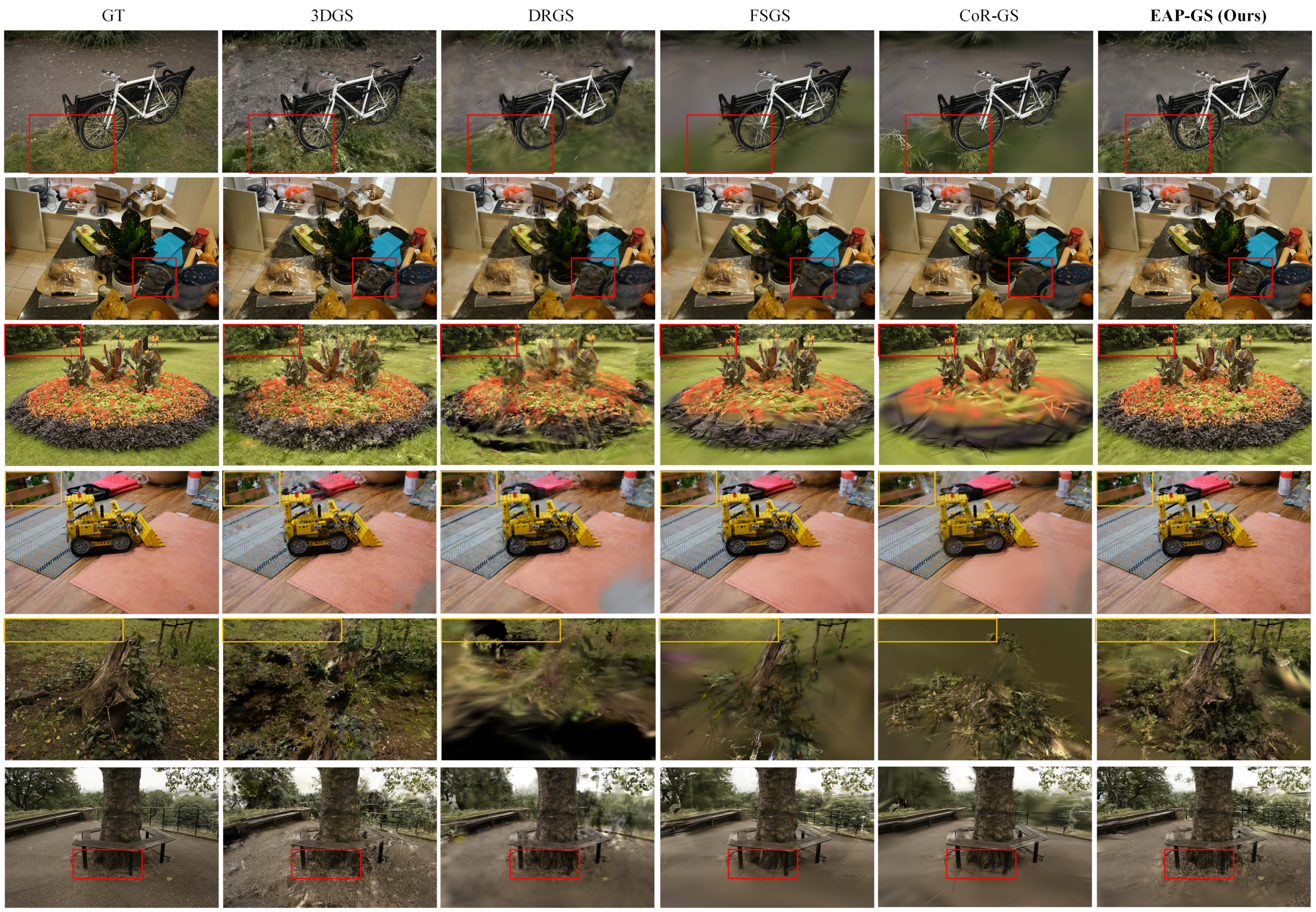

Figure S.3. Qualitative Comparison on the rest of the Mip-NeRF360 datasets. We demonstrate the qualitative comparison results with our main competitors 3DGS [13], DRGS [5], FSGS [40] and CoR-GS [35] on a certain testing view. The red and yellow boxes indicate the significant difference in reconstruction of the regions with weakly pronounced texture and peripheral areas, respectively.

Table S.2. Quantitative results on the LLFF datasets. The overall perforamnce of proposed EAP-GS and existing 3DGS-based methods are compared on the LLFF dataset. The best and second-best scores are in red and orange respectively.

		3DGS	DRGS	FSGS	CoR-GS	**EAP-GS(Ours)**
Fern	PSNR	15.53	19.04	19.47	19.60	19.79
	SSIM	0.4592	0.6033	0.6435	0.6417	0.6610
	LPIPS	0.3560	0.2260	0.2089	0.2293	0.1773
	Time (min)	12.85	0.92	22.58	16.32	1.73
	Number	507k	345k	194k	113k	168k
Flower	PSNR	16.00	17.93	19.34	19.69	19.65
	SSIM	0.4300	0.4709	0.6117	0.6339	0.6153
	LPIPS	0.3459	0.3401	0.2313	0.2296	0.2082
	Time (min)	10.32	1.23	34.35	15.55	1.75
	Number	298k	1239k	1346k	89k	128k
Fortress	PSNR	18.58	20.88	22.67	22.80	22.65
	SSIM	0.5076	0.5873	0.6034	0.7010	0.7167
	LPIPS	0.1917	0.1677	0.1404	0.1546	0.1106
	Time (min)	11.85	0.95	33.07	15.37	1.65
	Number	165k	186k	642k	59k	86k
Horns	PSNR	15.10	17.09	19.14	18.15	18.08
	SSIM	0.4590	0.5444	0.6505	0.6276	0.6351
	LPIPS	0.3459	0.3491	0.2175	0.2864	0.2192
	Time (min)	10.60	0.78	23.82	14.45	1.83
	Number	273k	214k	163k	73k	123k
Leaves	PSNR	11.80	13.37	15.83	14.40	14.55
	SSIM	0.2471	0.3014	0.4831	0.4235	0.4362
	LPIPS	0.3313	0.4904	0.1778	0.3017	0.2154
	Time (min)	16.50	0.77	23.58	15.05	1.90
	Number	1029k	746k	457k	194k	227k
Orchids	PSNR	14.16	14.80	15.44	14.77	15.89
	SSIM	0.3675	0.3532	0.4305	0.4246	0.4557
	LPIPS	0.2652	0.2835	0.2310	0.2713	0.2036
	Time (min)	10.72	1.05	21.65	14.53	1.77
	Number	255k	487k	188k	83k	131k
Room	PSNR	13.46	19.18	20.64	20.42	20.53
	SSIM	0.5601	0.7748	0.8316	0.8454	0.8305
	LPIPS	0.4999	0.2172	0.1584	0.1511	0.1662
	Time (min)	10.65	0.87	20.98	13.40	1.58
	Number	170k	144k	104k	35k	38k
Trex	PSNR	12.37	17.52	20.24	20.01	20.31
	SSIM	0.4686	0.6423	0.7422	0.7559	0.7684
	LPIPS	0.4040	0.2637	0.1454	0.1473	0.1329
	Time (min)	12.32	1.00	23.90	14.97	1.75
	Number	337k	346k	196k	80k	97k
Average	PSNR	14.63	17.48	19.10	18.73	18.93
	SSIM	0.4374	0.5347	0.6246	0.6317	0.6399
	LPIPS	0.3425	0.2922	0.1888	0.2214	0.1792
	Time (min)	11.98	0.95	25.49	14.96	1.75
	Number	379k	463k	411k	91k	125k

Table S.3. Quantitative results on the Mip-NeRF360 datasets. The overall perforamnce of proposed EAP-GS and existing 3DGS-based methods are compared on the Mip-NeRF360 dataset. The best and second-best scores are in red and orange respectively.

		3DGS	DRGS	FSGS	CoR-GS	**EAP-GS(Ours)**
Bicycle	PSNR	16.23	17.37	18.73	18.22	17.71
	SSIM	0.2664	0.3443	0.3760	0.3775	0.3561
	LPIPS	0.4424	0.6413	0.5410	0.5690	0.4779
	Time (min)	30.68	1.37	7.65	70.25	2.82
	Number	2239k	145k	254k	240k	599k
Bonsai	PSNR	15.80	18.25	18.51	18.35	19.08
	SSIM	0.5024	0.5502	0.6028	0.6188	0.6468
	LPIPS	0.3599	0.4151	0.3106	0.3380	0.2441
	Time (min)	9.60	1.00	4.83	28.13	1.42
	Number	460k	152k	127k	123k	275k
Counter	PSNR	16.82	17.71	17.91	18.69	18.12
	SSIM	0.5059	0.5467	0.5761	0.6116	0.5962
	LPIPS	0.3253	0.4062	0.3508	0.3332	0.2852
	Time (min)	11.08	0.93	4.50	26.23	1.27
	Number	486k	114k	67k	82k	216k
Flowers	PSNR	12.57	13.11	14.36	14.20	13.83
	SSIM	0.1609	0.2117	0.2365	0.2426	0.2437
	LPIPS	0.5257	0.7378	0.6347	0.7171	0.5513
	Time (min)	27.42	1.40	8.28	60.70	2.48
	Number	1954k	232k	214k	137k	719k
Garden	PSNR	17.58	18.56	19.06	19.56	19.54
	SSIM	0.4421	0.4351	0.4564	0.4734	0.5268
	LPIPS	0.3269	0.5123	0.4515	0.4711	0.3198
	Time (min)	29.98	1.53	8.52	66.92	2.85
	Number	1940k	336k	202k	220k	840k
Kitchen	PSNR	17.89	17.62	18.86	18.22	20.43
	SSIM	0.6116	0.5133	0.6501	0.6395	0.7179
	LPIPS	0.2757	0.4530	0.2865	0.2925	0.1871
	Time (min)	14.27	0.98	5.00	28.92	1.52
	Number	650k	88k	96k	112k	336k
Room	PSNR	18.53	20.07	20.27	20.97	21.16
	SSIM	0.6478	0.6933	0.7121	0.7363	0.7488
	LPIPS	0.2635	0.2963	0.2517	0.2455	0.2127
	Time (min)	12.05	1.02	5.12	31.97	1.42
	Number	652k	201k	117k	166k	322k
Stump	PSNR	15.32	16.35	16.96	16.70	17.59
	SSIM	0.1776	0.3050	0.3145	0.3117	0.3028
	LPIPS	0.4673	0.7353	0.5766	0.5986	0.4890
	Time (min)	25.28	1.17	7.00	66.22	2.50
	Number	2269k	337k	267k	192k	548k
Treehill	PSNR	13.80	14.91	16.75	16.48	15.25
	SSIM	0.2822	0.3660	0.3972	0.3916	0.3591
	LPIPS	0.5164	0.6732	0.5755	0.6105	0.5571
	Time (min)	30.90	1.37	7.67	62.28	2.27
	Number	2308k	255k	259k	265k	254k
Average	PSNR	16.06	17.11	17.93(4)	17.93(2)	18.08
	SSIM	0.3997	0.4406	0.4802	0.4892	0.4998
	LPIPS	0.3892	0.5412	0.4421	0.4639	0.3696
	Time (min)	21.25	1.20	6.51	49.07	2.06
	Number	1440k	207k	178k	171k	457k